Advances in Intelligent Systems and Computing 251

Van-Nam Huynh
Vladik Kreinovich
Songsak Sriboonchitta *Editors*

Modeling Dependence in Econometrics

Springer

Advances in Intelligent Systems and Computing

Volume 251

Series Editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland
e-mail: kacprzyk@ibspan.waw.pl

For further volumes:
http://www.springer.com/series/11156

About this Series

The series "Advances in Intelligent Systems and Computing" contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing.

The publications within "Advances in Intelligent Systems and Computing" are primarily textbooks and proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

Advisory Board

Van-Nam Huynh · Vladik Kreinovich
Songsak Sriboonchitta
Editors

Modeling Dependence in Econometrics

Springer

Editors
Van-Nam Huynh
Japan Advanced Institute of Science and Technology
Ishikawa
Japan

Songsak Sriboonchitta
Faculty of Economics
Chiangmai University
Chiangmai
Thailand

Vladik Kreinovich
Department of Computer Science
University of Texas at El Paso
El Paso, Texas
USA

ISSN 2194-5357 ISSN 2194-5365 (electronic)
ISBN 978-3-319-03394-5 ISBN 978-3-319-03395-2 (eBook)
DOI 10.1007/978-3-319-03395-2
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013954015

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The Seventh International Conference of the Thailand Econometric Society (**TES 2014**) is held in Chiang Mai, Thailand, during January 8^{th}–10^{th}, 2014, and hosted by the School of Economics, Chiang Mai University, Thailand.

The aim of this conference is to provide an international forum for researchers and practitioners in the areas of econometrics and quantitative analysis in economics to stimulate new ideas, present and discuss fundamental and applied research problems, as well as to foster research collaborations. The main theme of the TES 2014 conference is: "Modeling Dependence in Econometrics".

This volume contains papers presented at the TES 2014 conference, which were carefully evaluated and recommended for publication by the Scientific Committee. We appreciate the efforts of all the authors who submitted papers and regret that not all of them can be included. The volume begins with a keynote paper by Christian Francq and Jean-Michel Zakoïan, and is followed by invited and contributed papers arranged into two parts: Fundamental Theory and Applications.

The TES 2014 conference is financially supported by the Chiang Mai School of Economics (CMSE), Thailand. Thanks to Dean Pisit Leeahtam and CMSE for providing crucial support throughout the organization of TES 2014. Special thanks to Prof. Hung T. Nguyen for his valuable advice and constant support.

We sincerely wish to express our appreciation to all the members of Advisory Board, Administrative Committee, Scientific Committee and Local Organizing Committee for their great help and support. We would also like to thank Prof. Janusz Kacprzyk (Series Editor) and Dr. Thomas Ditzinger (Senior Editor, Engineering/Applied Sciences) for their support and cooperation in this publication.

Last, but not the least, we wish to thank all the authors and participants for their contributions and fruitful discussions that made this conference a success.

Chiang Mai, Thailand
January 2014

Van-Nam Huynh
Vladik Kreinovich
Songsak Sriboonchitta

Contents

Part I
Keynote Paper

Multi-level Conditional VaR Estimation in Dynamic Models*

Christian Francq and Jean-Michel Zakoïan**

Abstract. We consider joint estimation of conditional Value-at-Risk (VaR) at several levels, in the framework of general conditional heteroskedastic models. The volatility is estimated by Quasi-Maximum Likelihood (QML) in a first step, and the residuals are used to estimate the innovations quantiles in a second step. The joint limiting distribution of the volatility parameter and a vector of residual quantiles is derived. We deduce confidence intervals for general Distortion Risk Measures (DRM) which can be approximated by a finite number of VaR's. We also propose an alternative approach based on non Gaussian QML which, although numerically more cumbersome, has interest when the innovations distribution is fat tailed. An empirical study based on stock indices illustrates the theoretical findings.

1 Introduction

Under the regulations introduced in Finance since Basel 2, bank capital is risk-sensitive. Financial institutions are required to measure the riskiness of their assets and, for instance, to hold more capital to compensate more risk. While the Value-at-Risk (VaR), defined as a quantile of some loss distribution, continues to play

Christian Francq
CREST and University Lille 3 (EQUIPPE),
BP 60149, 59653 Villeneuve d'Ascq cedex, France
e-mail: christian.francq@univ-lille3.fr

Jean-Michel Zakoïan
EQUIPPE (University Lille 3) and CREST,
15 boulevard Gabriel Péri, 92245 Malakoff Cedex, France
e-mail: zakoian@ensae.fr

* The authors gratefully acknowledge financial support of the ANR via the Project ECONOM&RISK (ANR 2010 blanc 1804 03). The second author gratefully thanks the IDR "Risques systemiques" for financial support.

** Corresponding author.

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_1,

a prominent role in the mainstream financial risk management, a variety of alternative risk measures have been introduced and studied in recent years. The Expected Shortfall (ES), and more generally the Distortion Risk Measures (DRM), are quantile-based measures which, by comparison with the VaR at a given level, give further insight on the shape of the loss distribution[1].

Whatever the choice of a risk measure, it depends on unknown characteristics of the loss distribution which, for practical use, have to be estimated. In the so-called standard approach, the quantity of interest is a parameter, defined as a characteristic of the *marginal* loss distribution. In the so-called advanced approaches, the focus is on *conditional* characteristics of the loss distributions, that is, characteristics which, at the current date, take into account the available past information. The conditional VaR, and more generally conditional risk measures, are stochastic processes which are not directly observable, just like volatility. This complicates the statistical inference of risk measures. The problem is not only to get consistent estimators of conditional risks but also to evaluate the accuracy of such estimators[2].

Confidence intervals for conditional VaR's were derived, in the recent econometric literature, using different approaches. Chan, Deng, Peng, Xia (2007) constructed confidence intervals under the assumption that the errors have heavy tails, using the Extreme-Value Theory, while Spierdijk (2013) proposed a residual subsample bootstrap approach. Francq and Zakoïan (2012) used a QML approach. They showed that the problem of estimating a conditional risk measure, for instance a VaR at a given level, in GARCH-type models reduced to the estimation of a parameter, called risk parameter.

In the present article we extend those results to the joint estimation of several conditional risks. In practice, it is often important to handle several risk levels, in order to have a better view on the tail properties of the conditional distribution. We will provide statistical tools for jointly estimating conditional VaR's corresponding to different levels, in a general GARCH-type framework which does not impose a specific form for the volatility, and for estimating the accuracy of such VaR estimates. Our approach is aimed at, not only providing VaR estimates, but also confidence intervals based on asymptotic results. A tractable risk measure based on a vector of risk levels can be defined by weighting the corresponding VaR's, that is, by defining a *portfolio* of VaR's. This approach can be connected with DRMs through an appropriate choice of the weights. For a given DRM, our asymptotic results allow us to construct upper and lower bounds based on a finite number of VaR's.

[1] These measures are also advocated because, contrary to the VaR, they satisfy a set of "coherence requirements" for a large family of distributions.

[2] In July 2009, the Basel Committee issued a directive requiring that financial institutions quantify "model risk". The Committee states that *"Banks must explicitly assess the need for valuation adjustments to reflect two forms of model risk: the model risk associated with using a possibly incorrect valuation methodology; and the risk associated with using unobservable (and possibly incorrect) calibration parameters in the valuation model."* For instance, an important issue in determining the reserves of a financial institution is whether VaR estimates remain reliable in very hectic periods.

This paper is organized as follows. In Section 2, we start by introducing a general class of GARCH-type models. Then we derive the asymptotic joint distribution of the Quasi-Maximum Likelihood Estimator (QMLE) and a vector of empirical quantiles of the residuals. We deduce asymptotic confidence intervals for the VaR's and for VaR portfolios. Section 3 proposes another approach for conditional VaR estimation based on non Gaussian QMLEs. An empirical illustration based on major stock indices is proposed in Section 4. Section 5 concludes.

2 Two-Step VaR Estimation in Volatility Models

2.1 Conditional VaR in a General Model

Consider a GARCH-type model of the form

$$\begin{cases} \varepsilon_t = \sigma_t \eta_t \\ \sigma_t = \sigma(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots; \boldsymbol{\theta}_0) \end{cases} \tag{1}$$

where (η_t) is a sequence of iid random variables, η_t is independent of $\{\varepsilon_u, u < t\}$, $\boldsymbol{\theta}_0 \in \mathbb{R}^d$ is a parameter belonging to a parameter space Θ, and $\sigma : \mathbb{R}^\infty \times \Theta \to (0, \infty)$. The most widely used specifications of volatility belong to this class, in particular the GARCH(p,q) model of Engle (1982) and Bollerslev (1986),

$$\begin{cases} \varepsilon_t = \sigma_t \eta_t, \\ \sigma_t^2 = \omega_0 + \sum_{i=1}^q a_{0i} \varepsilon_{t-i}^2 + \sum_{j=1}^p b_{0j} \sigma_{t-j}^2, \end{cases} \tag{2}$$

where $\boldsymbol{\theta}_0 = (\omega_0, a_{01}, \ldots, b_{0p})'$ satisfies $\omega_0 > 0, a_{0i} \geq 0$, $b_{0j} \geq 0$. For this model, if the lag polynomial $\beta(L) = 1 - \sum_{j=1}^p b_{0j} L^j$ has its roots outside the unit disk, we have a representation of the form (1) given by

$$\sigma_t^2 = \beta(1)^{-1} \omega_0 + \sum_{i=1}^{\infty} \gamma_i \varepsilon_{t-i}^2,$$

where $\beta(L)^{-1} \sum_{i=1}^q a_i L^i = \sum_{i=1}^\infty \gamma_i L^i$. Other classical examples of models belonging to the class (1) are the EGARCH, GJR-GARCH, TGARCH, QGARCH, APARCH, Log-GARCH, models introduced, respectively, by Nelson (1991), Glosten, Jagannathan and Runkle (1993), Zakoïan (1994), Sentana (1995), Ding, Granger and Engle (1993), and for the log-GARCH, under slightly different forms, by Geweke (1986), Pantula (1986) and Milhøj (1987). See Francq and Zakoïan (2010) for an overview on GARCH models.

The *conditional* VaR of a process (ε_t) at risk level $\alpha \in (0,1)$, denoted by $\text{VaR}_t(\alpha)$, is defined by

$$P_{t-1}[\varepsilon_t < -\text{VaR}_t(\alpha)] = \alpha,$$

where P_{t-1} denotes the historical distribution conditional on $\{\varepsilon_u, u < t\}$. When (ε_t) satisfies (1), the theoretical VaR is then given by

$$\mathrm{VaR}_t(\alpha) = -\sigma(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots; \boldsymbol{\theta}_0)\xi_\alpha \tag{3}$$

where ξ_α is the α-quantile of η_t.

Remark 1. It can be noted that in the standard GARCH(p,q) model, the conditional VaR at level α automatically satisfies the stochastic recurrence equation

$$\mathrm{VaR}_t^2(\alpha) = \omega_0\xi_\alpha^2 + \sum_{i=1}^{q} a_{0i}\xi_\alpha^2\varepsilon_{t-i}^2 + \sum_{j=1}^{p} b_{0j}\mathrm{VaR}_{t-j}^2(\alpha).$$

Direct modelling of the conditional VaR has been proposed in several papers, for instance Engle and Manganelli (2004), Koenker and Xiao (2006), Gouriéroux and Jasiak (2008). A difficulty in this approach is to constrain the model so as to guarantee the monotonicity of the conditional VaR as a function of the risk level. Monotonicity is automatically satisfied in our approach.

2.2 *Asymptotic Properties of the Multi-level Two-Step VaR Estimator*

A two-step standard method for evaluating the VaR at different levels $\alpha_i \in (0,1)$, for $i = 1, \ldots, m$ consists in estimating the volatility parameter $\boldsymbol{\theta}_0$ by Gaussian QMLE, and then estimating the ξ_{α_i} by the corresponding empirical quantiles of the residuals; see, for instance, Chapter 2 in McNeil, Frey and Embrechts (2005). For a comparison of alternative strategies based on residuals following a preliminary volatility estimation, see Kuester, Mittnik and Paolella (2006).

Given observations $\varepsilon_1, \ldots, \varepsilon_n$, and arbitrary initial values $\tilde{\varepsilon}_i$ for $i \leq 0$, we define, under assumptions given below,

$$\tilde{\sigma}_t(\boldsymbol{\theta}) = \sigma(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots, \varepsilon_1, \tilde{\varepsilon}_0, \tilde{\varepsilon}_{-1}, \ldots; \boldsymbol{\theta}),$$

which is used to approximate $\sigma_t(\boldsymbol{\theta}) = \sigma(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots, \varepsilon_1, \varepsilon_0, \varepsilon_{-1}, \ldots; \boldsymbol{\theta})$. A QMLE of $\boldsymbol{\theta}_0$ in Model (1) is defined as any measurable solution $\widehat{\boldsymbol{\theta}}_n$ of

$$\widehat{\boldsymbol{\theta}}_n = \arg\min_{\boldsymbol{\theta}\in\Theta} \widetilde{Q}_n(\boldsymbol{\theta}), \tag{4}$$

with

$$\widetilde{Q}_n(\boldsymbol{\theta}) = n^{-1}\sum_{t=1}^{n} \widetilde{\ell}_t(\boldsymbol{\theta}), \qquad \widetilde{\ell}_t(\boldsymbol{\theta}) = \frac{\varepsilon_t^2}{\widetilde{\sigma}_t^2(\boldsymbol{\theta})} + \log\widetilde{\sigma}_t^2(\boldsymbol{\theta}).$$

The following assumptions are required to derive the asymptotic properties of the QMLE $\widehat{\boldsymbol{\theta}}_n$.

A1: (ε_t) is a strictly stationary and ergodic solution of Model (1). Moreover, $E|\varepsilon_0|^s < \infty$ for some $s > 0$.

A2: For any real sequence (x_i), the function $\boldsymbol{\theta} \mapsto \sigma(x_1, x_2, \ldots; \boldsymbol{\theta})$ is continuous. Almost surely, $\sigma_t(\boldsymbol{\theta}) \in (\underline{\omega}, \infty]$ for any $\boldsymbol{\theta} \in \Theta$ and for some $\underline{\omega} > 0$.

A3: The function $\boldsymbol{\theta} \mapsto \sigma(x_1, x_2, \ldots; \boldsymbol{\theta})$ has continuous second-order derivatives, and

$$\sup_{\boldsymbol{\theta}\in\Theta} \left\{ |\sigma_t(\boldsymbol{\theta}) - \tilde{\sigma}_t(\boldsymbol{\theta})| + \left\| \frac{\partial \sigma_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} - \frac{\partial \tilde{\sigma}_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right\| + \left\| \frac{\partial^2 \sigma_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} - \frac{\partial^2 \tilde{\sigma}_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right\| \right\} \leq C_1 \rho^t,$$

where C_1 is a random variable which is measurable with respect to $\{\varepsilon_u, u < 0\}$ and $\rho \in (0,1)$ is a constant.

A4 $(\boldsymbol{\theta}_0^*)$**:** $\boldsymbol{\theta}_0^*$ belongs to the interior of Θ and $\sigma_t(\boldsymbol{\theta}_0^*)/\sigma_t(\boldsymbol{\theta}) = 1$ *a.s.* iff $\boldsymbol{\theta} = \boldsymbol{\theta}_0^*$.

A5 $(\boldsymbol{\theta}_0^*)$**:** There exist no non-zero $x \in \mathbb{R}^d$ such that $x' \frac{\partial \sigma_t(\boldsymbol{\theta}_0^*)}{\partial \boldsymbol{\theta}} = 0, \quad a.s.$

A6$(\boldsymbol{\theta}_0^*)$**:** There exists a neighborhood $V(\boldsymbol{\theta}_0^*)$ of $\boldsymbol{\theta}_0^*$ such that the following variables have finite expectation:

$$\sup_{\boldsymbol{\theta}\in V(\boldsymbol{\theta}_0^*)} \left\| \frac{1}{\sigma_t(\boldsymbol{\theta})} \frac{\partial \sigma_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right\|^4, \quad \sup_{\boldsymbol{\theta}\in V(\boldsymbol{\theta}_0^*)} \left\| \frac{1}{\sigma_t(\boldsymbol{\theta})} \frac{\partial^2 \sigma_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right\|^2, \quad \sup_{\boldsymbol{\theta}\in V(\boldsymbol{\theta}_0^*)} \left| \frac{\sigma_t(\boldsymbol{\theta}_0^*)}{\sigma_t(\boldsymbol{\theta})} \right|^{2\delta}.$$

Note that Assumptions **A2**, **A3**, **A5** and **A6** can be simplified for specific forms of σ_t: for instance if the model is the GARCH (p,q) Model (2), **A2** reduces to standard assumptions on the lag polynomials of the volatility and **A3**, **A5**, **A6** can be directly verified. Note also that the only moment assumption on the observed process is the existence of a small moment in **A1**, which is automatically satisfied for standard models such as the classical GARCH(p,q).

Now let the residuals of the QML estimation

$$\hat{\eta}_t = \frac{\varepsilon_t}{\tilde{\sigma}_t(\widehat{\boldsymbol{\theta}}_n)}, \qquad t = 1, \ldots, n,$$

and let ξ_{n,α_i} denote the empirical α_i-quantile of $\hat{\eta}_1, \ldots, \hat{\eta}_n$. Let $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_m)'$, $\boldsymbol{\xi}_{n,\boldsymbol{\alpha}} = (\xi_{n,\alpha_1}, \ldots, \xi_{n,\alpha_m})'$ and let $\boldsymbol{\xi}_{\boldsymbol{\alpha}} = (\xi_{\alpha_1}, \ldots, \xi_{\alpha_m})'$ denote the vector of population quantiles.

Remark 2. The derivation of the joint asymptotic properties of sample quantiles goes back to Cramér (1946) in the iid case. Different articles have extended these results for the marginal quantiles of stationary processes, under different dependence assumptions. See Dominicy, Hörmann, Ogata and Veredas (2013) and the references therein. We cannot apply their results because $(\hat{\eta}_t)$ is not a stationary process.

The next result gives the joint asymptotic distributions of $(\widehat{\boldsymbol{\theta}}_n', \boldsymbol{\xi}_{n,\boldsymbol{\alpha}}')$. Let $\boldsymbol{D}_t(\boldsymbol{\theta}) = \sigma_t^{-1}(\boldsymbol{\theta}) \partial \sigma_t(\boldsymbol{\theta}) / \partial \boldsymbol{\theta}$.

Theorem 1. *Assume* $\xi_{\alpha_i} < 0$*, for* $i = 1, \ldots, m$*,* $E\eta_t^2 = 1$ *and* $\kappa_4 := E\eta_t^4 < \infty$*. Suppose that* η_1 *admits a density* f *which is continuous and strictly positive in a neighborhood of* ξ_{α_i}*, for* $i = 1, \ldots, m$*. Let* **A1**-**A3** *and* **A4**$(\boldsymbol{\theta}_0)$-**A6**$(\boldsymbol{\theta}_0)$ *hold. Then*

$$\begin{pmatrix} \sqrt{n}\left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0\right) \\ \sqrt{n}(\boldsymbol{\xi}_{n,\boldsymbol{\alpha}} - \boldsymbol{\xi}_{\boldsymbol{\alpha}}) \end{pmatrix} \xrightarrow{\mathscr{L}} \mathscr{N}(0, \boldsymbol{\Sigma}_\alpha), \qquad \boldsymbol{\Sigma}_\alpha = \begin{pmatrix} \frac{\kappa_4 - 1}{4} \boldsymbol{J}^{-1} & \boldsymbol{\lambda}_{\boldsymbol{\alpha}}' \otimes \boldsymbol{J}^{-1} \boldsymbol{\Omega} \\ \boldsymbol{\lambda}_{\boldsymbol{\alpha}} \otimes \boldsymbol{\Omega}' \boldsymbol{J}^{-1} & \boldsymbol{\zeta}_{\boldsymbol{\alpha}} \end{pmatrix},$$

where $\boldsymbol{\Omega} = E(\boldsymbol{D}_t)$, $\boldsymbol{J} = E(\boldsymbol{D}_t\boldsymbol{D}_t')$ *with* $\boldsymbol{D}_t = \boldsymbol{D}_t(\boldsymbol{\theta}_0)$, $\boldsymbol{\lambda}_{\boldsymbol{\alpha}} = (\lambda_{\alpha_1}, \ldots, \lambda_{\alpha_m})'$, $\boldsymbol{\zeta}_{\boldsymbol{\alpha}} = (\zeta_{ij})_{1\leq i,j\leq m}$ *and*

$$\lambda_{\alpha_i} = \xi_{\alpha_i}\frac{\kappa_4 - 1}{4} + \frac{p_{\alpha_i}}{2f(\xi_{\alpha_i})},$$
$$\zeta_{ij} = \xi_{\alpha_i}\xi_{\alpha_j}\frac{\kappa_4 - 1}{4} + \frac{\xi_{\alpha_i}p_{\alpha_j}}{2f(\xi_{\alpha_j})} + \frac{\xi_{\alpha_j}p_{\alpha_i}}{2f(\xi_{\alpha_i})} + \frac{(\alpha_i \wedge \alpha_j)(1 - \alpha_i \vee \alpha_j)}{f(\xi_{\alpha_i})f(\xi_{\alpha_j})},$$

with $p_\alpha = E\left(\eta_1^2 \mathbf{1}_{\{\eta_1 < \xi_\alpha\}}\right) - \alpha$.

Proof. In view of Francq and Zakoïan (Proof of Theorem 4, 2012), we have, for $i = 1, \ldots, m$,

$$\sqrt{n}(\xi_{\alpha_i} - \xi_{n,\alpha_i}) = \xi_{\alpha_i}\boldsymbol{\Omega}'\sqrt{n}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) + \frac{1}{f(\xi_{\alpha_i})}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}(\mathbf{1}_{\{\eta_t < \xi_{\alpha_i}\}} - \alpha_i) + o_P(1),$$

and

$$\sqrt{n}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) = \frac{-\boldsymbol{J}^{-1}}{2\sqrt{n}}\sum_{t=1}^{n}(1 - \eta_t^2)\boldsymbol{D}_t + o_P(1).$$

Hence

$$\mathrm{Cov}_{as}\left(\sqrt{n}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0), \frac{1}{\sqrt{n}}\sum_{t=1}^{n}(\mathbf{1}_{\{\eta_t < \xi_{\alpha_i}\}} - \alpha_i)\right) = \frac{1}{2}p_{\alpha_i}\boldsymbol{J}^{-1}\boldsymbol{\Omega}.$$

It follows that, for $i \leq j$,

$$\begin{aligned}
&\mathrm{Cov}_{as}\{\sqrt{n}(\xi_{\alpha_i} - \xi_{n,\alpha_i}), \sqrt{n}(\xi_{\alpha_j} - \xi_{n,\alpha_j})\} \\
&= \left\{\xi_{\alpha_i}\xi_{\alpha_j}\frac{\kappa_4 - 1}{4} + \frac{\xi_{\alpha_i}p_{\alpha_j}}{2f(\xi_{\alpha_j})} + \frac{\xi_{\alpha_j}p_{\alpha_i}}{2f(\xi_{\alpha_i})}\right\}\boldsymbol{\Omega}'J^{-1}\boldsymbol{\Omega} + \frac{\alpha_i(1 - \alpha_j)}{f(\xi_{\alpha_i})f(\xi_{\alpha_j})}, \\
&\mathrm{Cov}_{as}\left(\sqrt{n}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0), \sqrt{n}(\xi_{\alpha_i} - \xi_{n,\alpha_i})\right) \\
&= \lambda_{\alpha_i}\boldsymbol{J}^{-1}\boldsymbol{\Omega}.
\end{aligned}$$

We have $\boldsymbol{\Omega}'\boldsymbol{J}^{-1}\boldsymbol{\Omega} = 1$ (see Remark 3.1 in Francq and Zakoïan, 2013) and thus we obtain

$$\mathrm{Cov}_{as}\{\sqrt{n}(\xi_{\alpha_i} - \xi_{n,\alpha_i}), \sqrt{n}(\xi_{\alpha_j} - \xi_{n,\alpha_j})\} = \zeta_{ij}.$$

By the CLT for martingale differences, we get the announced result. □

Let $\mathbf{VaR}_t(\boldsymbol{\alpha}) = (\mathrm{VaR}_t(\alpha_1), \ldots, \mathrm{VaR}_t(\alpha_m))'$, the vector of VaR's at levels α_i. We have

$$\mathbf{VaR}_t(\boldsymbol{\alpha}) = -\sigma(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots; \boldsymbol{\theta}_0)\boldsymbol{\xi}_{\boldsymbol{\alpha}}. \tag{5}$$

A natural estimator of $\mathbf{VaR}_t(\boldsymbol{\alpha})$ is thus

$$\widehat{\mathbf{VaR}}_t(\boldsymbol{\alpha}) = -\tilde{\sigma}_t(\widehat{\boldsymbol{\theta}}_n)\boldsymbol{\xi}_{n,\boldsymbol{\alpha}}.$$

Remark 3. A classical problem, called quantile crossing, in quantile regression is that two or more estimated conditional quantile functions can cross or overlap. This drawback occurs because each conditional quantile function is independently estimated (see Koenker (2005)). It is thus worth noting that our estimation procedure does not face this problem. By construction, the estimated conditional VaR are monotonous functions of the α's.

Remark 4. For the standard GARCH(p,q) model, we have $\boldsymbol{J}^{-1}\boldsymbol{\Omega} = 2\overline{\boldsymbol{\theta}}_0$, where

$$\overline{\boldsymbol{\theta}}_0 = \begin{pmatrix} \boldsymbol{\theta}_0^{[1:q+1]} \\ 0_p \end{pmatrix}, \quad \boldsymbol{\theta}_0^{[1:q+1]} = (\omega_0, a_{01}, \ldots, a_{0q})',$$

(see Francq and Zakoïan (2013)), and the asymptotic variance in Theorem 1 takes the more explicit form

$$\boldsymbol{\Sigma}_{\boldsymbol{\alpha}} = \begin{pmatrix} \frac{\kappa_4 - 1}{4}\boldsymbol{J}^{-1} & 2\boldsymbol{\lambda}'_{\boldsymbol{\alpha}} \otimes \overline{\boldsymbol{\theta}}_0 \\ 2\boldsymbol{\lambda}_{\boldsymbol{\alpha}} \otimes \overline{\boldsymbol{\theta}}'_0 & \boldsymbol{\zeta}_{\boldsymbol{\alpha}} \end{pmatrix}.$$

2.3 Constructing Confidence Intervals for the VaR's

Let $\widehat{\boldsymbol{\Sigma}}_\alpha$ denote a consistent estimator of the asymptotic variance $\boldsymbol{\Sigma}_\alpha$. Such an estimator can be constructed by i) replacing $\boldsymbol{J}$ by $\hat{\boldsymbol{J}} = n^{-1}\sum_{t=1}^n \boldsymbol{D}_t(\widehat{\boldsymbol{\theta}}_n)\boldsymbol{D}_t(\widehat{\boldsymbol{\theta}}_n)'$; ii) using the residuals $\hat{\eta}_t$ to construct an estimator $\hat{f}$ of the density function f of the innovation, and to replace the theoretical moments of the process (η_t) by their empirical counterpart.

The delta method thus suggests a $(1-\alpha_0)\%$ confidence interval (CI) for the $\mathrm{VaR}_t(\alpha_i)$ whose bounds are

$$-\tilde{\sigma}_t(\widehat{\boldsymbol{\theta}}_{n,\alpha_i})\xi_{n,\alpha_i} \pm \frac{\Phi^{-1}_{1-\alpha_0/2}}{\sqrt{n}} \left\{ \left(\widehat{\boldsymbol{\Delta}}_\alpha \widehat{\boldsymbol{\Sigma}}_\alpha \widehat{\boldsymbol{\Delta}}'_\alpha \right)_{ii} \right\}^{1/2}, \tag{6}$$

where

$$\widehat{\boldsymbol{\Delta}}_\alpha = \left(\boldsymbol{\xi}_{n,\boldsymbol{\alpha}} \frac{\partial \tilde{\sigma}_t(\widehat{\boldsymbol{\theta}}_{n,\alpha})}{\partial \boldsymbol{\theta}'}, \tilde{\sigma}_t(\widehat{\boldsymbol{\theta}}_n)\boldsymbol{I}_m \right),$$

$\Phi^{-1}_{\alpha_0}$ denotes the α_0-quantile of the standard Gaussian distribution, and $\boldsymbol{I}_m$ denotes the $m \times m$ identity matrix. Note that the choice of α_0 (the risk estimation level) is independent from that of the α_i's (the financial risk levels). Drawing such CI allows to underline the importance of the estimation risk for VaR evaluation.

2.4 A Portfolio of VaR's

Focusing only on VaR at a given level for measuring risk can be misleading since it gives a limited view of the distribution, which may result in lack of robustness for risk management and risk control. To circumvent this problem, several risk measures have to be jointly considered in practice. To this aim, Distortion Risk Measures (DRM) have been introduced in the insurance literature, in a series of papers by Wang and coauthors [see Wang (2000) and the references therein]. A particular case is the conditional expected shortfall (ES) which, at level $\alpha \in (0,1)$, can be written as

$$\mathrm{ES}_t(\alpha) = -E_{t-1}[\varepsilon_t \mid \varepsilon_t < -\mathrm{VaR}_t(\alpha)] = \frac{1}{\alpha}\int_0^\alpha \mathrm{VaR}_t(u)du.$$

More general DRM take the form

$$\mathrm{DRM}_t = \int_0^1 \mathrm{VaR}_t(u)dG(u), \tag{7}$$

where the distortion function, G, is a given cumulative distribution function (cdf) on $[0,1]$[3]. It follows from (3) that, for Model (1),

$$\mathrm{DRM}_t = -\sigma(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots; \boldsymbol{\theta}_0)\int_0^1 \xi_u dG(u). \tag{8}$$

In the spirit of DRM, a risque measure which can be interpreted as a portfolio of VaR's at different levels is defined by

$$\boldsymbol{p}'\mathbf{VaR}_t(\boldsymbol{\alpha}) = \sum_{i=1}^m p_i \mathrm{VaR}_t(\alpha_i)$$

where $\boldsymbol{p} = (p_1, \ldots, p_m)$ with $p_i \geq 0$ for $i = 1, \ldots, m$ and $\sum_{i=1}^m p_i = 1$. This risk measure can be interpreted as a special DRM with associated distortion function corresponding to Dirac masses at the points α_i. In view of (6), an asymptotic CI at level α_0 for this risk measure is

$$-\tilde{\sigma}_t(\widehat{\boldsymbol{\theta}}_{n,\alpha_i})\boldsymbol{p}'\boldsymbol{\xi}_{n,\boldsymbol{\alpha}} \pm \frac{\Phi^{-1}_{1-\alpha_0/2}}{\sqrt{n}}\left\{\boldsymbol{p}'\widehat{\boldsymbol{\Delta}}_\alpha \widehat{\boldsymbol{\Sigma}}_\alpha \widehat{\boldsymbol{\Delta}}'_\alpha \boldsymbol{p}\right\}^{1/2}. \tag{9}$$

2.5 Choosing the Weights to Approximate DRMs

An estimator of the DRM in (8) can be constructed as follows:

[3] Examples of DRM are the Proportional Hazard DRM, defined with $G(u) = u^r$, and the Exponential DRM defined with $G(u) = (1-e^{ru})/(1-e^r)$, both of them defined for $r > 0$. These distortion functions are concave for $0 < r < 1$ and $r > 0$, respectively, which corresponds to coherent risk measure in the sense of Artzner, Delbaen, Eber and Heath (1999) [see e.g. Wirch and Hardy (1999)].

$$\widehat{\mathrm{DRM}}_t = -\sigma(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots; \widehat{\boldsymbol{\theta}}_n) \sum_{i=1}^{n} \left\{ G\left(\frac{i}{n}\right) - G\left(\frac{i-1}{n}\right) \right\} \hat{\eta}_{n,i}, \tag{10}$$

where $(\hat{\eta}_{n,n-i})$ denotes the order statistics, obtained by ranking the $\hat{\eta}_t$ in ascending order: $\hat{\eta}_{n,1} < \cdots < \hat{\eta}_{n,n}$.

However, deriving the asymptotic distribution of this estimator, might be a formidable task. To our knowledge such results do not exist in the literature. In this section, we use VaR portfolios to obtain lower and upper bounds for a class of DRM, leading to (approximate) asymptotic CI's for such DRM.

It is not restrictive to assume $\alpha_1 < \alpha_2 < \ldots < \alpha_m$. Suppose that the support of the distortion cdf G is $[\alpha_1, \alpha_m]$, that is

$$\mathrm{DRM}_t = \int_{\alpha_1}^{\alpha_m} \mathrm{VaR}_t(u) dG(u). \tag{11}$$

In other words, we focus on "moderate risks": we do not consider extreme risks, corresponding to values of α approaching 0. An example of class of such DRM, parameterized by the coefficient $r > 0$ and adapted from the so-called "proportional hazard" DRM, is defined by

$$G(u) = \left(\frac{u - \alpha_1}{\alpha_m - \alpha_1} \right)^r \mathbf{1}_{u \in (\alpha_1, \alpha_m)} + \mathbf{1}_{u \in (\alpha_m, 1)}, \tag{12}$$

where $\mathbf{1}_A$ denotes the indicator function of any set A.

Lower and upper bounds for the DRM in (11), can be constructed as follows. Because $u \mapsto \mathrm{VaR}_t(u)$ is decreasing we have, noting that $G(\alpha_1) = 0$ and $G(\alpha_m) = 1$,

$$\boldsymbol{p}_L' \mathbf{VaR}_t(\boldsymbol{\alpha}) \leq \mathrm{DRM}_t(\alpha) \leq \boldsymbol{p}_U' \mathbf{VaR}_t(\boldsymbol{\alpha})$$

where

$$\begin{aligned} \boldsymbol{p}_L &= (0, G(\alpha_2), G(\alpha_3) - G(\alpha_2), \ldots, 1 - G(\alpha_{m-1})), \\ \boldsymbol{p}_U &= (G(\alpha_2), G(\alpha_3) - G(\alpha_2), \ldots, 1 - G(\alpha_{m-1}), 0). \end{aligned}$$

It follows that a CI at significance level $\alpha_0^* \leq \alpha_0$ for this risk measure is

$$\left[-\tilde{\sigma}_t(\widehat{\boldsymbol{\theta}}_{n,\alpha_i}) \boldsymbol{p}_L' \boldsymbol{\xi}_{n,\boldsymbol{\alpha}} - \frac{\Phi^{-1}_{1-\alpha_0/2}}{\sqrt{n}} \left\{ \boldsymbol{p}_L' \widehat{\boldsymbol{\Delta}}_\alpha \widehat{\boldsymbol{\Sigma}}_\alpha \widehat{\boldsymbol{\Delta}}_\alpha' \boldsymbol{p}_L \right\}^{1/2}, \right.$$
$$\left. -\tilde{\sigma}_t(\widehat{\boldsymbol{\theta}}_{n,\alpha_i}) \boldsymbol{p}_U' \boldsymbol{\xi}_{n,\boldsymbol{\alpha}} + \frac{\Phi^{-1}_{1-\alpha_0/2}}{\sqrt{n}} \left\{ \boldsymbol{p}_U' \widehat{\boldsymbol{\Delta}}_\alpha \widehat{\boldsymbol{\Sigma}}_\alpha \widehat{\boldsymbol{\Delta}}_\alpha' \boldsymbol{p}_U \right\}^{1/2} \right]. \tag{13}$$

3 NonGaussian QML Estimation of VaR's

In this section we develop an alternative method for estimating the conditional VaR's. This method is based on a reparameterization of model (1). QML inferences based on similar reparameterizations were proposed by Francq, Lepage and Zakoïan (2011), Fan, Qi and Xiu (2012), Francq and Zakoïan (2013).

3.1 Reparameterization and VaR Parameter

The approach of this section requires the following assumption

A7: There exists a function H such that for any $\boldsymbol{\theta} \in \Theta$, for any $K > 0$, and any sequence $(x_i)_i$

$$K\sigma(x_1, x_2, \ldots; \boldsymbol{\theta}) = \sigma(x_1, x_2, \ldots; \boldsymbol{\theta}^*), \quad \text{where } \boldsymbol{\theta}^* = H(\boldsymbol{\theta}, K).$$

This assumption is not very restrictive as it is satisfied by all commonly used GARCH-type formulations, in particular those mentioned in Section 2. It means that scaling the volatility is equivalent to a change of parameter. In general, the new parameter satisfies $\boldsymbol{\theta}^* \geq \boldsymbol{\theta}$, componentwise, when $K \geq 1$. For instance, in the GARCH(p,q) model (2) we have $\boldsymbol{\theta}^* = (K^2\omega, K^2 a_1, \ldots, K^2 a_q, b_1, \ldots, b_p)'$.

In view of (3), we have under **A7**, provided α_i is small enough so that $-\xi_{\alpha_i} > 0$,

$$\text{VaR}_t(\alpha_i) = \sigma(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots; \boldsymbol{\theta}_{0\alpha_i}) \tag{14}$$

where $\boldsymbol{\theta}_{0,\alpha_i} = H(\boldsymbol{\theta}_0, -\xi_{\alpha_i})$. This parameter depends on both the dynamics of the GARCH process, through the volatility parameters, and the innovations distribution through the α-quantile. It is called VaR-parameter in Francq and Zakoïan (2012) (hereafter FZ). Similarly, if $-\int_0^1 \xi_u dG(u) > 0$, the DRM in (8) can be written as

$$\text{DRM}_t = \sigma(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots; \boldsymbol{\theta}_0^G) \tag{15}$$

where $\boldsymbol{\theta}_0^G$ is a DRM-parameter defined by

$$\boldsymbol{\theta}_0^G = H\left(\boldsymbol{\theta}_0, -\int_0^1 \xi_u dG(u)\right). \tag{16}$$

It follows from (14) that, with the notation used in (5),

$$\mathbf{VaR}_t(\boldsymbol{\alpha}) = \begin{pmatrix} \sigma(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots; \boldsymbol{\theta}_{0\alpha_1}) \\ \vdots \\ \sigma(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots; \boldsymbol{\theta}_{0\alpha_m}) \end{pmatrix}.$$

The approach, in this section, consists in estimating by QML the $\boldsymbol{\theta}_{0\alpha_i}$'s instead of $\boldsymbol{\theta}_0$. The idea is to interpret, for $i = 1, \ldots, m$, the VaR-parameter $\boldsymbol{\theta}_{0\alpha_i}$ as a volatility parameter in a reparameterized model. We note that

$$\varepsilon_t = \sigma(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots; \boldsymbol{\theta}_{0\alpha_i})\eta_{i,t}, \qquad \text{where } \eta_{i,t} = \frac{\eta_t}{-\xi_{\alpha_i}}.$$

The problem is thus to estimate by QML the model

$$\begin{cases} \varepsilon_t = \sigma_{i,t}\eta_{i,t}, \quad P[\eta_{i,t} < -1] = \alpha_i, \\ \sigma_{i,t} = \sigma(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots; \boldsymbol{\theta}_{0,\alpha_i}). \end{cases} \tag{17}$$

Note that the Gaussian QML cannot be employed because it requires the assumption that $E\eta_t^2 = 1$. FZ derived the asymptotic distribution of the non-Gaussian QMLE of $\boldsymbol{\theta}_{0,\alpha_i}$ defined by

$$\widehat{\boldsymbol{\theta}}_{n,\alpha_i} = \arg\max_{\boldsymbol{\theta}\in\Theta} \sum_{t=1}^{n} \log \frac{1}{\tilde{\sigma}_t(\boldsymbol{\theta})} h_{\alpha_i}\left(\frac{\varepsilon_t}{\tilde{\sigma}_t(\boldsymbol{\theta})}\right) \tag{18}$$

where h_{α_i} is given by

$$h_{\alpha_i}(x) = \lambda\,\alpha_i(1-2\alpha_i)|x|^{2\lambda\alpha_i - 1}\{|x|^{-\lambda}\mathbf{1}_{\{|x|>1\}} + \mathbf{1}_{\{|x|\le 1\}}\} \tag{19}$$

for some (unimportant) positive constant λ.

As noted by FZ, the non-Gaussian QML estimator in (18) can be interpreted as a nonlinear quantile regression estimator. Letting $\rho_\alpha(u) = u(\alpha - \mathbf{1}_{\{u\le 0\}})$, for $\alpha \in (0,1)$, we have

$$\widehat{\boldsymbol{\theta}}_{n,\alpha_i} = \arg\min_{\boldsymbol{\theta}\in\Theta} \frac{1}{n}\sum_{t=1}^{n} \rho_{1-2\alpha_i}\left\{\log\left(\frac{|\varepsilon_t|}{\tilde{\sigma}_t(\boldsymbol{\theta})}\right)\right\}.$$

In the next section, we derive the joint distribution of the $\widehat{\boldsymbol{\theta}}_{n,\alpha_i}$'s.

3.2 Asymptotic Joint Distribution of the VaR Parameter Estimators

We introduce the following additional assumption.

A8: The density f of η_0 is symmetric, continuous and strictly positive at the points ξ_{α_i}, for $i = 1, \ldots, m$, and satisfies $M = \sup_{x\in\mathbb{R}} |x| f(x) < \infty$. Moreover $E|\log|\eta_0|| < \infty$.

Let $\boldsymbol{\theta}_{0\boldsymbol{\alpha}} = (\boldsymbol{\theta}'_{0\alpha_1}, \ldots, \boldsymbol{\theta}'_{0,\alpha_m})'$ and let $\widehat{\boldsymbol{\theta}}_{n,\boldsymbol{\alpha}} = (\widehat{\boldsymbol{\theta}}'_{n,\alpha_1}, \ldots, \widehat{\boldsymbol{\theta}}'_{n,\alpha_m})'$.

Theorem 2. *Under the assumptions* **A1-A3**, **A7**, **A8** *and if, for* $i = 1, \ldots, m$, $\alpha_i \in (0, 1/2)$ *and* **A4**$(\boldsymbol{\theta}_{0\alpha_i})$-**A6**$(\boldsymbol{\theta}_{0\alpha_i})$ *hold, there exists a sequence of local minimizers* $\widehat{\boldsymbol{\theta}}_{n,\boldsymbol{\alpha}}$ *of the QML criterion satisfying*

$$\sqrt{n}(\widehat{\boldsymbol{\theta}}_{n,\boldsymbol{\alpha}} - \boldsymbol{\theta}_{0,\boldsymbol{\alpha}}) \xrightarrow{d} \mathcal{N}(0, \boldsymbol{\Xi}_{\boldsymbol{\alpha}}),$$

where $\boldsymbol{\Xi}_{\boldsymbol{\alpha}}$ *is a* $md \times md$ *matrix whose* (i,j)*-block of size* $d \times d$ *is*

$$\boldsymbol{\Xi}_{\boldsymbol{\alpha}}[i,j] = \frac{2\alpha_i \wedge \alpha_j (1 - 2\alpha_i \vee \alpha_j)}{4 f(\xi_{\alpha_i}) f(\xi_{\alpha_j})} \boldsymbol{J}^{-1}_{\alpha_i \alpha_i} \boldsymbol{J}_{\alpha_i \alpha_j} \boldsymbol{J}^{-1}_{\alpha_j \alpha_j}$$

with $\boldsymbol{J}_{\alpha_i \alpha_j} = E \boldsymbol{D}_t(\boldsymbol{\theta}_{0,\alpha_i}) \boldsymbol{D}_t'(\boldsymbol{\theta}_{0,\alpha_j})$.

Proof. Note that $\hat{\boldsymbol{v}}_{n,\alpha_i} := \sqrt{n}(\widehat{\boldsymbol{\theta}}_{n,\alpha_i} - \boldsymbol{\theta}_{0,\alpha_i})$ is such that

$$\hat{\boldsymbol{v}}_{n,\alpha_i} = \arg \min_{\boldsymbol{v} \in \Lambda_{n,\alpha_i}} \tilde{S}_{n,\alpha_i}(\boldsymbol{v}),$$

where $\Lambda_{n,\alpha_i} := \sqrt{n}(\Theta - \boldsymbol{\theta}_{0,\alpha_i})$ and

$$\tilde{S}_{n,\alpha_i}(\boldsymbol{v}) = \sum_{t=1}^{n} \rho_{1-2\alpha_i} \left\{ \log \left(\frac{|\varepsilon_t|}{\tilde{\sigma}_t(\boldsymbol{\theta}_{0,\alpha_i} + n^{-1/2}\boldsymbol{v})} \right) \right\} - \rho_{1-2\alpha_i} \left\{ \log \left(\frac{|\varepsilon_t|}{\tilde{\sigma}_t(\boldsymbol{\theta}_{0,\alpha_i})} \right) \right\}.$$

For notational convenience, write $a \overset{c}{=} b$ when $a = b + c$. Showing that the initial values are asymptotically negligible, and noting that $\varepsilon_t / \sigma_t(\boldsymbol{\theta}_{0,\alpha_i}) = -\eta_t / \xi_{\alpha_i}$, it can be proven that, uniformly in v belonging to an arbitrary compact set (see Lemma 2 in FZ),

$$\begin{aligned} \tilde{S}_{n,\alpha_i}(\boldsymbol{v}) \overset{o_P(1)}{=} S_{n,\alpha_i}(\boldsymbol{v}) := & \sum_{t=1}^{n} \rho_{1-2\alpha_i} \left\{ \log \left(\frac{|\varepsilon_t|}{\sigma_t(\boldsymbol{\theta}_{0,\alpha_i} + n^{-1/2}\boldsymbol{v})} \right) \right\} \\ & - \rho_{1-2\alpha_i} \left\{ \log \left| \frac{\eta_t}{\xi_i} \right| \right\}. \end{aligned}$$

Doing a Taylor expansion of $\log \sigma_t(\boldsymbol{\theta}_{0,\alpha_i} + n^{-1/2}\boldsymbol{v})$ around $\boldsymbol{v} = \boldsymbol{0}$, and using Lemma 2 in FZ, we obtain

$$\begin{aligned} S_{n,\alpha_i}(\boldsymbol{v}) \overset{o_P(1)}{=} S^*_{n,\alpha_i}(\boldsymbol{v}) := & \sum_{t=1}^{n} \rho_{1-2\alpha_i} \left\{ \log \left| \frac{\eta_t}{\xi_{\alpha_i}} \right| - \frac{1}{\sqrt{n}} \boldsymbol{v}' \boldsymbol{D}_t(\boldsymbol{\theta}_{0,\alpha_{\alpha_i}}) \right\} \\ & - \rho_{1-2\alpha_i} \left\{ \log \left| \frac{\eta_t}{\xi_{\alpha_i}} \right| \right\}. \end{aligned}$$

Note that $S^*_{n,\alpha_i}(\cdot)$ is equal to the function $Z_n(\cdot)$ defined by Equation (17) in Koenker and Xiao (2006), when applied to the quantile regression of $\log|\eta_t/\xi_{\alpha_i}|$ on $\boldsymbol{D}_t(\boldsymbol{\theta}_{0,\alpha_{\alpha_i}})$ at the level $1 - 2\alpha_i$. Even if our framework is not that of the above-mentioned paper, similar results hold true. More precisely, FZ show that the finite-dimensional distributions of $S^*_{n,\alpha_i}(\boldsymbol{v})$ and

$$S^{**}_{n,\alpha_i}(\boldsymbol{v}) := -\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \boldsymbol{v}' \boldsymbol{D}_t(\boldsymbol{\theta}_{0,\alpha_i}) \left(1 - 2\alpha_i - \mathbf{1}_{\{|\eta_t| < -\xi_{\alpha_i}\}} \right) + f(\xi_{\alpha_i}) \boldsymbol{v}' \boldsymbol{J}_{\alpha_i \alpha_i} \boldsymbol{v}$$

converge to those of the same Gaussian process. Noting that the trajectories of $S^*_{n,\alpha_i}(\cdot)$ and $S^{**}_{n,\alpha_i}(\cdot)$ are convex, we also have uniform convergence over every compact set in the space of the continuous function on $\mathbb{R}^d$. By Lemma 2.2 in Davis,

Knight and Liu (1992) the minima of $S^*_{n,\alpha_i}(\cdot)$ and $S^{**}_{n,\alpha_i}(\cdot)$ are asymptotically the same. By Remark 1 of the above-mentioned paper, we finally obtain

$$\hat{\boldsymbol{v}}_{n,\alpha_i} \overset{o_P(1)}{=} \frac{1}{2f(\xi_{\alpha_i})} \boldsymbol{J}^{-1}_{\alpha_i\alpha_i} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \boldsymbol{D}_t(\boldsymbol{\theta}_{0,\alpha_i}) \left(1 - 2\alpha_i - \mathbf{1}_{\{|\eta_t| < -\xi_{\alpha_i}\}}\right).$$

The conclusion follows easily. □

Remark 5. Theorems 1 and 2 provide the asymptotic distributions of two estimators for VaR portfolios. At first sight, the method of this section is not attractive because it is more cumbersome, from a numerical point of view, than that of the previous section. Indeed, it requires the optimization of m QML criteria, whereas the first method requires one. However, it is important to note that the assumptions required for the asymptotic results are different. In particular, the fourth moment assumption $E\eta_t^4 < \infty$ of the first method, is not required in Theorem 2. On the other hand, the latter theorem is valid under a symmetry assumption on the noise distribution. To conclude, the method of this section can only be recommended in presence of very heavy-tailed errors distribution.

4 Empirical Illustration

In this section we present empirical results using returns of nine major stock indices: CAC (Paris), DAX (Frankfurt), FTSE (London), Nikkei (Tokyo), NSE (Bombay), SMI (Switzerland), SP500 (New York), SPTSX (Toronto), and SSE (Shanghai). Our sample spans the period from January, 2 1991 to August, 26 2011 (but all series are not available for the whole period, see Table 1 for the sample sizes). For each series of log-returns, $\varepsilon_t = \log(p_t/p_{t-1})$ where p_t denotes the value of the index, we used a GARCH(1,1) model for the volatility dynamics. We estimated the DRM parameter $\boldsymbol{\theta}_0^G = (\omega^G, a^G, b^G)$, defined in (16), with $r = 1/2$, $\alpha_1 = 0.01$ and $\alpha_m = 0.1$ for the DRM function G defined in (12). In view of (10) and (16), the DRM-parameter estimator is given by

$$\widehat{\boldsymbol{\theta}}_n^G = H\left(\widehat{\boldsymbol{\theta}}_n, -\sum_{i=1}^{n} \left\{ G\left(\frac{i}{n}\right) - G\left(\frac{i-1}{n}\right) \right\} \hat{\eta}_{n,i} \right),$$

where $H(\omega, \alpha, \beta; K) = (K^2\omega, K^2\alpha, \beta)$. The CI's are obtained using (13) with $m = 20$ and $\alpha_0 = 5\%$.

We report in Table 1 our estimates of the conditional DRM parameter and the corresponding CI's. Caution is needed in the interpretation of this table because the DRM parameter is not the usual volatility parameter. In particular, the fact that $a^G + b^G > 1$ is not in contradiction with the usual empirical finding, $a + b \approx 1$, for GARCH(1,1) models. Noticeable differences appear between these series, particularly for the coefficients ω^G and a^G and their CI's. Replacing the number $m = 20$

Table 1 Estimation of the conditional DRM parameter for 9 stock market indices. The approximate 95% confidence intervals are displayed into brackets.

Index	n	ω^G	a^G	b
CAC	5229	0.11 [0.05,0.17]	0.31[0.22,0.41]	0.90 [0.88,0.92]
DAX	5226	0.12 [0.04,0.20]	0.31[0.18,0.45]	0.90 [0.86,0.93]
FTSE	5217	0.04 [0.02,0.07]	0.32[0.24,0.41]	0.91 [0.89,0.92]
Nikkei	5078	0.20 [0.11,0.30]	0.37[0.26,0.48]	0.88 [0.85,0.91]
NSE	2265	0.25 [0.06,0.46]	0.40[0.20,0.65]	0.87 [0.82,0.92]
SMI	5209	0.17 [0.08,0.27]	0.46[0.27,0.65]	0.84 [0.79,0.89]
SP500	5206	0.03 [0.01,0.05]	0.27[0.19,0.36]	0.92 [0.90,0.94]
SPTSX	2934	0.03 [0.01,0.06]	0.27[0.17,0.38]	0.93 [0.91,0.95]
SSE	2982	0.11 [0.03,0.20]	0.25[0.15,0.37]	0.93 [0.90,0.95]

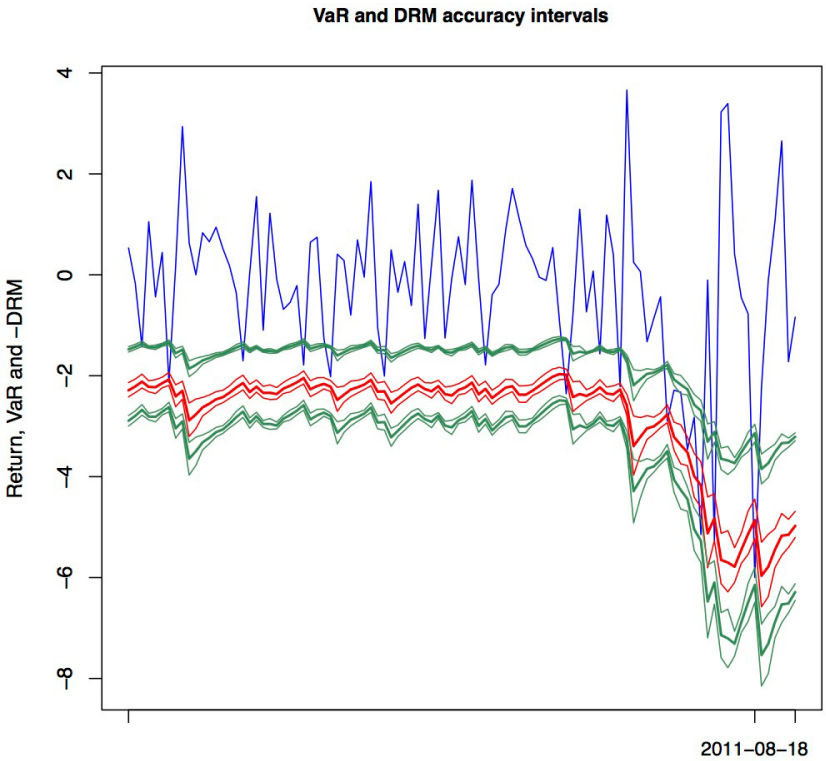


Fig. 1 Returns (in blue), estimated -VaR (at the 10% and 1% levels, in green), -DRM (in red), and CI's of the VaR's and DRMs, for the DAX index from April, 8, 2011 to August, 26, 2011. Estimation of the volatility and risk parameters is based on the 1000 previous values.

of α_i's used for the discretization of the DRM by $m = 10$ or $m = 30$ left almost unchanged the CI's, so we did not report the results. We can depict three categories of assets: i) the FTSE, SP500 and SPTSX display similar coefficients, relatively small ω^G's, large persistence parameter b's, small CI's; ii) Nikkei, NSE and SMI provide, by comparison, larger ω^G's and a^G's, smaller persistence and much larger CI's; iii) the CAC, DAX and SSE display intermediate results. Examination of the CI's shows that the differences between parameters of series in groups i) and ii) are statistically significant. Note also that larger CI's are not always due to smaller sample size.

Figure 1 displays the returns, estimated -VaR (at the 10% and 1% levels), -DRM, and their accuracy intervals for the DAX index from April, 8, 2011 to August, 26, 2011. The $(1-\alpha_0)\%$ confidence intervals (for $\alpha_0 = 5\%$) are obtained from formula (13). We reported the opposite of the conditional risks (VaR and DRM), because in terms of capital reserves, only large negative returns matter. As expected, the accuracy on VaR estimation decreases when the risk α approaches 0. Interestingly, the accuracy of the DRM is comparable to that of the VaR's, despite the more sophisticated construction of this measure of risk. Note also that, in turbulent periods, both the market risks, as measured by the VaR's or the DRM, and the estimation risks, as measured by the CI's, increase.

5 Conclusion

In this paper, we proposed procedures for joint statistical inference on the VaR's at different levels, in the framework of conditionally heteroskedastic models. We also introduced an approximation of general DRM based on a finite number of VaR's. Our empirical analysis showed that confidence intervals based on this measure of risk have similar magnitude as those obtained for VaR's.

One alternative for deriving the asymptotic distribution of the DRM estimator would be to establish a functional CLT, in function of α, for the vector of the volatility parameter estimator and the empirical quantile of the residuals. Deriving this asymptotic distribution could be a formidable challenge. Moreover, the asymptotic distribution would certainly be non explicit. The approximation proposed in this article, which provides an explicit and easily estimable asymptotic distribution, thus has the advantage of simplicity.

One object of this study was also to draw attention on the estimation risk, in other words the effects of parameter estimation on the accuracy of VaR's evaluations. We showed that estimation risk can be explicitly taken into account, leading to confidence bounds for portfolios, or more generally any smooth function, of VaR's. For risk management purposes, or from a regulation point of view, such confidence intervals could be used to increase the capital reserve in order to account for the underlying estimation uncertainty.

References

1. Artzner, P., Delbaen, F., Eber, J.-M., Heath, D.: Coherent measures of risk. Mathematical Finance 9, 203–228 (1999)
2. Bollerslev, T.: Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307–327 (1986)
3. Chan, N.H., Deng, S.J., Peng, L., Xia, Z.: Interval estimation of value-at-risk based on GARCH models with heavy-tailed innovations. Journal of Econometrics 137, 556–576 (2007)
4. Cramer, H.: Mathematical Methods of Statistics. Princeton Mathematical Series, vol. 9. Princeton University Press, Princeton (1946)
5. Davis, R.A., Knight, K., Liu, J.: M-estimation for autoregressions with infinite variance. Stochastic Processes and their Applications 40, 145–180 (1992)
6. Ding, Z., Granger, C., Engle, R.F.: A long memory property of stock market returns and a new model. Journal of Empirical Finance 1, 83–106 (1993)
7. Dominicy, Y., Hörmann, S., Ogata, H., Veredas, D.: Marginal Quantiles for StationaryProcesses. Statistics and Probability Letters 83, 28–36 (2013)
8. Engle, R.F.: Autoregressive conditional heteroskedasticity with estimates of the variance of the United Kingdom inflation. Econometrica 50, 987–1007 (1982)
9. Engle, R.F., Manganelli, S.: CAViaR: Conditional Value at risk by Quantile Regression. Journal of Business and Economic Statistics 22, 367–381 (2004)
10. Fan, J., Qi, L., Xiu, D.: Quasi Maximum Likelihood Estimation of GARCH Models with Heavy-Tailed Likelihoods. Discussion Paper, `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1540363`
11. Francq, C., Lepage, G., Zakoïan, J.-M.: Two-stage non Gaussian QML estimation of GARCH Models and testing the efficiency of the Gaussian QMLE. Journal of Econometrics 165, 246–257 (2011)
12. Francq, C., Zakoïan, J.-M.: GARCH Models: Structure, Statistical Inference and Financial Applications. John Wiley (2010)
13. Francq, C., Zakoïan, J.-M.: Risk-parameter estimation in volatility models. MPRA Preprint No. 41713 (2012)
14. Francq, C., Zakoïan, J.M.: Optimal predictions of powers of conditionally heteroskedastic processes. Journal of the Royal Statistical Society - Series B 75, 345–367 (2013)
15. Geweke, J.: Modeling the persistence of conditional variances: a comment. Econometric Review 5, 57–61 (1986)
16. Glosten, L.R., Jaganathan, R., Runkle, D.: On the relation between the expected values and the volatility of the nominal excess return on stocks. Journal of Finance 48, 1779–1801 (1993)
17. Gouriéroux, C., Jasiak, J.: Dynamic Quantile Models. Journal of Econometrics 147, 198–205 (2008)
18. Knight, K.: Limiting distributions for L_1 regression estimators under general conditions. The Annals of Statistics 26, 755–770 (1998)
19. Koenker, R.: Quantile Regression. Cambridge University Press, Cambridge (2005)
20. Koenker, R., Xiao, Z.: Quantile autoregression. Journal of the American Statistical Association 101, 980–990 (2006)
21. Kuester, K., Mittnik, S., Paolella, M.S.: Value-at-Risk predictions: A comparison of alternative strategies. Journal of Financial Econometrics 4, 53–89 (2006)
22. McNeil, A.J., Frey, R., Embrechts, P.: Quantitative Risk Management. Princeton University Press (2005)

23. Milhøj, A.: A multiplicative parameterization of ARCH Models. Working Paper, Department of Statistics, University of Copenhagen (1987)
24. Nelson, D.B.: Conditional Heteroskedasticity in Asset Returns: a New Approach. Econometrica 59, 347–370 (1991)
25. Pantula, S.G.: Modeling the persistence of conditional variances: a comment. Econometric Review 5, 71–74 (1986)
26. Pollard, D.: Asymptotics for Least Absolute Deviation Regression Estimators. Econometric Theory 7, 186–199 (1991)
27. Sentana, E.: Quadratic ARCH Models. Review of Economic Studies 62, 639–661 (1995)
28. Spierdijk, L.: Confidence intervals for ARMA-GARCH Value-at-Risk. Working paper, University of Groningen (2013)
29. Wang, S.: A class of distortion operators for pricing financial and insurance risks. Journal of Risk and Insurance 67, 15–36 (2000)
30. Wirch, J.L., Hardy, M.R.: A Synthesis of Risk Measures for Capital Adequacy. Insurance: Mathematics and Economics 25, 337–347 (1999)
31. Zakoïan, J.M.: Threshold Heteroskedastic Models. Journal of Economic Dynamics and Control 18, 931–955 (1994)

Part II
Fundamental Theory

The Effects of Management and Provision Accounts on Hedge Fund Returns – Part I: The High Water Mark Scheme*

Serge Darolles and Christian Gourieroux

Abstract. A characteristic of hedge funds is not only an active portfolio management, but also the allocation of portfolio performance between different accounts, which are the accounts for the external investors and an account for the management firm, respectively. Despite a lack of transparency in hedge fund market, the strategy of performance allocation is publicly available. This paper shows that, for the High Water Mark Scheme, these complex performance allocation strategies might explain empirical facts observed in hedge fund returns, such as return persistence, skewed return distribution, bias ratio, or implied increasing risk appetite.

1 Introduction

The applied literature has shown that the return dynamics of individual hedge funds[1] (HF) are very different from the return dynamics of more standard assets such as stocks, currencies, or mutual funds. The HF return dynamics can depend on the management style, but generally, feature persistence, especially at short term and in extreme returns [Agarwal, Naik (2000), Koh, Koh, Teo (2003), Getmanski, Lo, Makarov (2004)], local asymmetries around zero, called bias ratio in the literature [Abdulali (2006), Bollen, Pool (2009), Darolles, Gourieroux, Jasiak (2009)], very heavy tails, for instance for Convertible Arbitrage or Fixed Income Arbitrage funds; moreover, some HF returns are weakly correlated with major asset market returns

Serge Darolles
Universite Paris-Dauphine and CREST
e-mail: serge.darolles@dauphine.fr

Christian Gourieroux
CREST and University of Toronto
e-mail: christian.gourieroux@ensae.fr

* The authors gratefully acknowledge financial support of the chair QuantValley/Risk Foundation "Quantitative Management Initiative". The second author gratefully acknowledges financial support of NSERC Canada.

[1] Note that we are interested in the return dynamics of individual hedge funds, not in hedge funds indices.

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_2,

[Fung, Hsieh(1999)]. These empirical facts reflect an underlying nonlinear dynamic of HF return, which can be explained by:

i) The frequent path dependent updating of the portfolio associated with the fund [see e.g. Lo (2008)];
ii) The procedure used to allocate the performance between different accounts, that are the investor's account and the account of the management firm.

Since the sequence of portfolio updatings and allocations are not observable by the econometrician and the standard investor[2], the management style and its effect on returns are difficult to analyse. On the other hand, the procedures used to allocate the total performance between the different accounts are precisely described in the prospectus written at the creation of the fund and validated by the appropriate authorities. The aim of this paper and of its companion paper [Darolles, Gourieroux (2013)] is to discuss the possible effects of these rather complex procedures and to see if they can partly explain empirical facts observed on individual HF returns[3].

In Section 2, we provide an example of allocations between accounts used in practice. We consider the rather standard high-water mark (HWM) scheme. The presence of several accounts can imply significant differences between the return of the managed porfolio and the published HF return. We describe in detail the nonlinear filter to pass from the portfolio return to the published HF return.

Section 3 compares the portfolio and fund returns when the portfolio returns are independent and identically Gaussian distributed. The i.i.d. Gaussian assumption on portfolio returns corresponds to a rather exogenous portfolio management, whereas the hedge fund manager will account for the existence of multiple accounts in his/her management strategy.

In Section 4, we discuss the mean-variance efficient portfolio management according to the account of interest. If the fund performance has to be maximized, the management differs from the standard mean-variance management of the global portfolio. More precisely, the allocation scheme between accounts has a significant

[2] They are known by the fund manager and partly known by large investors, who profit of due diligence, or investors in US funds reporting their holdings on Form 13F with the Security Exchange Commission (SEC). This creates asymmetric information on HF markets.

[3] Performance based fees (also called incentive fees) are characteristics of hedge funds; they are much less frequent for mutual funds. For instance, in 1999 only 108 out of a total 6.716 bond and stock mutual funds used incentive fees [Elton, Gruber, Blake(2001)]. Moreover by law the mutual funds must use a special form of incentive fees known as fulcrum fee (see the 1970 amendment to the Investment Company Act of 1940). Typically, the fulcrum fees are centered around an index[4] and have upper and lower limits in size. Such constraints do not exist for HF.

[4] As noted in Elton, Gruber, Blake (2003), 43 different indices were used as benchmark in 1999, as the S&P 500 index, the Russell 2000, Morgan Stanley's EAFE, Lipper Growth, or Income Fund Index.

impact on the optimal portfolio management. There exists a theoretical literature in the introduction of multiple accounts as an incentive for the hedge fund manager[5]. However, this question is often considered under rather irrealistic assumptions such as continuous time incentives, whereas the barrier effects apply monthly [see e.g. Goetzmann, Ingersoll, Ross (2003), Kouwenberg, Ziemba (2007)], competitive hedge fund market, whereas each hedge fund has a specific design and its secondary market is not very active [see e.g. Christoffersen, Musto, Yilmaz (2013)], two periods instead of multiperiod optimization [see e.g. Christoffersen, Musto, Yilmaz (2013)], risk-neutral manager [Paganeas, Westerfield (2009)], binary returns [Christoffersen, Musto, Yilmaz (2013)], or rather ad-hoc account description, which does not correspond to the account allocations proposed in the hedge fund industry [Kazemi, Li (2009), Aragon, Nanda (2012)]. We try in this section to stay as close as possible to the actual hedge fund designs and to focus on their dynamic features. Section 5 contains conclusions. Proofs are gathered in Appendices.

2 High-Water Mark Allocation Scheme

There exist almost as many account allocation schemes as hedge funds shares, which explains why any precautionary investor, regulator, or researcher[6] should study in details the prospectus of the funds. We describe below a standard scheme used to allocate the performance between the account invested by external clients, called class A units in the following, and the account invested by the management firm, called class B units[7].

This allocation scheme is parametrized by an allocation rate, called performance fee rate, a return benchmark, called hurdle rate, and a validity period corresponding to the duration between consecutive resets of class B account. These parameters differ according to the fund share.

2.1 Allocation between A and B Accounts

Let us first consider two accounts, with respective values A_t, B_t at month t, $t = 0, \ldots, T$. The contractual hurdle rate is denoted by $y_{h,t}, y_{h,t} \geq 0$, and is assumed to be

[5] There exists also a more empirical literature studying the links between the risk taken by the hedge fund manager, often summarized by means of the HF return volatility, and some characteristics of the HF, such as proxies for the optional feature of the compensation scheme [see e.g. Kazemi, Li (2009)]. These analysis are often based on the rather simple static linear regression techniques and thus neglect the complexity of the compensation scheme, especially its dynamics and nonlinear features.

[6] Typically, it is misleading to consider as an homogenous class the set of funds reporting a high-water mark benchmark in the standard Lipper/TASS database [see e.g. Aragon, Nanda (2012)].

[7] To simplify, we assume that there is neither redemption, nor subscription after the inception date and no misreporting of the data. The changes observed in the values of the different accounts come from the evolution of the portfolio return only.

predetermined and observable at date t. The contractual hurdle rate is a benchmark introduced to define the performance allocations. This hurdle can be set to zero [see e.g. Panageas, Westerfield (2009)], or to a cash return like the 1-month London Interbank Offered Rate (LIBOR)[8]. The maximal value reached on the past by account A is discounted at rate $y_{h,t}$ and called the high-water mark (HWM). This HWM is first computed at date t by:

$$HWM_t = \max_{0\leq\tau\leq t}\left[A_\tau \prod_{\tau^*=\tau}^{t}(1+y_{h,\tau^*})\right],\ t=0,...,T-1, \tag{1}$$

and then compare to A_{t+1} at date $t+1$. The fee schedule is endogenous[9] as a function of past successes, but is entirely defined at date t, due to the choice of the predetermined hurdle rate. We deduce that:

$$HWM_t = \max\left[HWM_{t-1}, A_t\right](1+y_{h,t}),\ t=1,...,T-1, \tag{2}$$

with initial condition $HWM_0 = A_0(1+y_{h,0})$.

At period t, the global portfolio value $A_t + B_t$ is invested and provides at the end of the period a return net of base management fees[10] denoted by y_{t+1}. Then, the change in total portfolio value $(A_t + B_t)y_{t+1}$ is allocated between the two accounts. The performance fee is not charged if the fund is globally in a deficit of performance with respect to the high-water mark. Thus, this allocation depends on the location of:

$$A_t(1+y_{t+1}), \tag{3}$$

with respect to the predetermined HWM_t as follows:

1. if $HWM_t \geq A_t(1+y_{t+1})$,

$$\begin{cases} A_{t+1} = A_t(1+y_{t+1}), \\ B_{t+1} = B_t(1+y_{t+1}). \end{cases} \tag{4}$$

2. If $HWM_t < A_t(1+y_{t+1})$,

$$\begin{cases} A_{t+1} = A_t(1+y_{t+1}) - \alpha[A_t(1+y_{t+1}) - HWM_t], \\ B_{t+1} = B_t(1+y_{t+1}) + \alpha[A_t(1+y_{t+1}) - HWM_t], \end{cases} \tag{5}$$

[8] The hurdle rate has to be defined in the same currency as the fund reference currency, e.g. US Dollar, Euro, Yen, ...

[9] Exogenous HWM of the type $HWM_t = HWM_0\prod_{\tau=1}^{t}(1+y_{h,\tau})$ are often assumed in the HF literature [see e.g. Hodder, Jackwerth (2007)]. Such HWM schemes correspond to the fulcrum scheme for mutual funds, but are very different from the actual HWM for hedge funds.

[10] The base management fee is generally proportional to the asset value managed by the fund. Without loss of generality, we take them into account by considering portfolio return net of base management fee.

where $0 < \alpha < 1$ is the (high-water mark) performance fee rate. This performance fee rate varies from 15% to 50%, with an increase in recent years [see e.g. Fung, Hsieh(1999), Table 2, Zuckerman (2004)]. It is equal to 20% for the Quantum Fund reported in Goetzmannn, Ingersoll, Ross (2003).

The updating equations (2.4)-(2.5) can also be written as:

$$\begin{cases} A_{t+1} = A_t(1+y_{t+1}) - \alpha A_t \left[y_{t+1} - (HWM_t - A_t)/A_t\right]^+, \\ B_{t+1} = B_t(1+y_{t+1}) + \alpha A_t \left[y_{t+1} - (HWM_t - A_t)/A_t\right]^+, \end{cases} \tag{6}$$

where $X^+ = \max(X,0)$, to highlight the presence of an option component. When the fund gains enough value, the manager is paid and the strike price increases, but when the fund loses money, the strike price remains unchanged and the manager retains his/her option at the old strike price.

At short term horizon equal to 1, the future account values involve the payoff of a European call option[11] written on y_{t+1}, with predetermined path dependent strike equal to $y_{0,t} = (HWM_t - A_t)/A_t$. At larger horizon, we get a sequence of European calls with changing strike prices. Both rolling effect and path dependent strike show that the option interpretation of the account allocation is significantly different from the simplified European call interpretations introduced for instance in Kouwenberg and Ziemba (2007) or Diez de los Rios, Garcia (2008), eq. (2.5), which neglects path dependence.

For a zero hurdle rate, the recursive equation for account A can also be written as:

$$A_{t+1} = A_t(1+y_{t+1}) - \alpha(HWM_{t+1} - HWM_t)^+, \tag{7}$$

which shows that the fund manager receives a fraction of the increase in HWM as conpensation.

In practice, the management firm is periodically paid by means of the management account, generally at the end of the year. The recursive equations are valid on a period $\{0, T-1\}$ of a given length T corresponding to the duration between consecutive resets, i.e. 0 and T. At time T, the management account is reset to the initial fixed[12] contractual value B_0 and the HWM reset[13] to $A_T(1+y_{h,T})$. Since the allocation scheme may create nonstationary features, this practice breaks down possible explosive behavior.

If the reset time is $T = 1$, the HWM is equal to $A_t(1+y_{t+1})$ and regimes (2.4) and (2.5) are active depending if the portfolio management out- or underperforms the hurdle. We get $y_{0,t} = y_{h,t}$ and the HWM disappears in equation (2.6) that becomes:

$$\begin{cases} A_{t+1} = A_t(1+y_{t+1}) - \alpha A_t \left[y_{t+1} - y_{h,t}\right]^+, \\ B_{t+1} = B_t(1+y_{t+1}) + \alpha A_t \left[y_{t+1} - y_{h,t}\right]^+, \end{cases} \tag{8}$$

[11] Or of a European put option if we note that coefficient $-\alpha$ is negative and account for the put-call parity relationship.

[12] That is, this contractual value is not discounted.

[13] There exist funds with different reset times for the HWM and the B account.

and is reset at each date. In this setup a fixed proportion α of the return above the hurdle is allocated to class B at each period, which corresponds to a standard fulcrum scheme.

To summarize, the evolutions of account values depend on the portfolio management, that is, the sequence of portfolio returns (y_t), and on the allocation design characterized by hurdle rate $(y_{h,t})$, performance fee rate α, and reset time[14] T. The dynamic system is recursive, since value (A_t) has an autonomous dynamic, and value (B_t) is fixed later. Let us finally remark that the value of account A can decrease and even become smaller than the initial value A_0, or negative. Therefore, the HF can fail[15] before the contractual reset time T. We will consider in the theoretical analysis that the fund fails if A_t becomes negative before reset time. From equation (2.6), we see that the portfolio return is necessarily larger than -1 before the potential failure time, and account B is positive. From this theoretical point of view, fund failure arises as the consequence of an abnormal negative return. In practice, it is also possible that the fund manager decides to liquidate the fund if the losses on account A are too large, even if A_t is still positive, or if his/her fees B_t are too small.

2.2 Discussion of the High-Water Mark Scheme

Let us now discuss scheme $(2.1)-(2.5)$. Since $HWM_t \geq A_t$, regime (2.4) applies if the spread between the net portfolio return and the hurdle $y_{t+1} - y_{h,t}$ is negative. If the spread is negative, the total loss is allocated proportionally to each account. If the spread is positive and small, regime (2.4) still applies and the same return is applied to accounts A and B. If the spread is positive and large enough to hit the HWM, the allocation rule is no longer proportional. The gain is shared between accounts A and B, with an allocation more favorable for the managing firm [see (2.5)]. The values of accounts A and B can increase or decrease, but the effect of net portfolio return y_{t+1} is no longer symmetric.

If the reset time is $T = 1$, the dependence of $\Delta A_{t+1} = A_{t+1} - A_t$ (resp. $\Delta B_{t+1} = B_{t+1} - B_t$) with respect to net portfolio return y_{t+1} is described in Figure 1 (resp. Figure 2).

The value of the class A unit is a continuous increasing function of the net portfolio return with a change of slope at threshold $y_{0,t}$. The payoff on B account is a convex function of the return, which might be an incentive for the fund manager to

[14] It is important to distinguish the reset time and the termination date of an hedge fund. Whereas most hedge fund management contracts do not have a pre-specified termination date, a reset time is often indicated. The presence of a reset time has significant implications on fund management and returns, and has to be taken into account. By implicitely assuming an infinite reset time, a part of the literature considered rather unrealistic models [see e.g. Panageas, Westerfield (2009)]. Typically, in a continuous time framework, the reset time will imply jumps of an endogenous size at predetermined dates.

[15] A HF fails when the fund manager decides to liquidate the fund and gives back the remaining asset under management to investors. The decision for liquidation is not contractual, but is at the discretion of the fund manager.

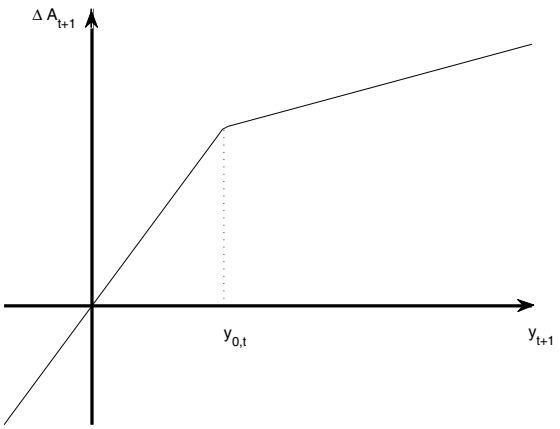


Fig. 1 ΔA_{t+1} as a function of y_{t+1} (unitary reset time)

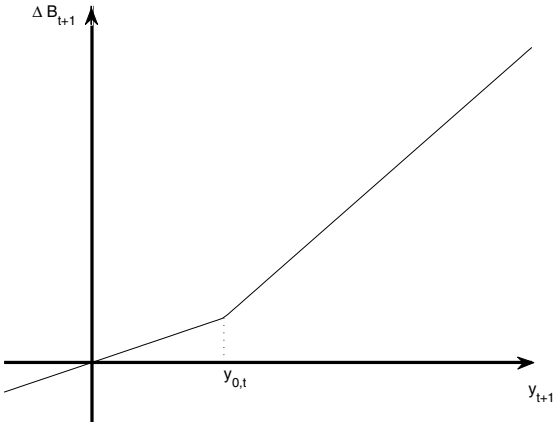


Fig. 2 ΔB_{t+1} as a function of y_{t+1} (unitary reset time)

take risk, i.e. to produce large positive returns at some date to feed account B. These extreme positive returns might have to increase in time due the increasingness of the high-water mark as function of past successes. This misleading intuition has been challenged by Carpenter(2000), Ross(2004), Panageas, Westerfield(2009) [see also the discussion in Section 4].

Let us also discuss this scheme if the fund manager invests only in a riskfree asset, $y_{t+1} = y_{f,t}$, with a riskfree return larger than the hurdle, $y_{f,t} \geq y_{h,t}$, say[16]. Since $A_t(1+y_{f,t}) \geq A_t(1+y_{h,t}) = HWM_t$, the fund manager would profit systematically of such a static riskfree investment. Surprisingly, this account allocation scheme is often used in the HF industry with a zero hurdle rate $y_{h,t} = 0$.

2.3 The Returns and Effective Performance Fees

A major point in the discussion of fund returns is the definition of returns in case of several accounts. Indeed, the following returns can be introduced:

i) the total portfolio net return y_{t+1},
ii) the return on B account[17]: $y_{B,t+1} = (B_{t+1} - B_t)/B_t$,
iii) the return on A account: $y_{A,t+1} = (A_{t+1} - A_t)/A_t$.

The fund returns available in the standard Hedge Funds Research (HFR) or Lipper-Tass databases are returns $(y_{A,t})$ corresponding to class A units. They can feature dynamics very different from the dynamics of (y_t) and $(y_{B,t})$. For instance, return $y_{A,t}$ is always smaller or equal to the total net porfolio return y_t. It coincides with it at some endogeneous periods, and is strictly below, otherwise. It can be important in the analysis to distinguish the reported HF return $y_{A,t}$ and the underlying total portfolio return y_t. As an illustration, the methodology proposed in Henriksson, Merton (1981) [see also Glosten, Jagannathan (1994), Agarwal, Naik (2004), Diez de los Rios, Garcia (2008)] to detect the market timing ability of a portfolio manager consists in running a regression of the HF return on a market return and on a put option payoff written on this market return, say, and to test if the optional component is significant[18]. Applied to reported HF return y_t, this optional effect will likely appear as a consequence of the HWM scheme, even if this effect is not present in the total portfolio return, that is, if the portfolio manager shows no market timing ability. This might explain why "this option like payoff (effect) is not restricted only to trend followers and risk arbitrageurs, but is a feature on a wide range of hedge funds strategies" [Agarwal, Naik (2004), p. 66]. Anyway, the first equation in (2.6) shows that the observed return $y_{A,t}$ is a complicated nonlinear function of $y_t, y_{t-1}, ..., y_{h,t}, y_{h,t-1}, ...$, function which is known from the prospectus[19].

[16] Under no arbitrage opportunity, this means that the contractual riskfree rate has been fixed at a level strictly smaller than the market riskfree rate.

[17] We have to choose a contractual positive B_0 initial value to give a meaning to this return.

[18] This methodology has to be applied on individual hedge funds, not on HF indices, to get this interpretation.

[19] In Getmanski, Lo, Makarov (2004), the observed return $y_{A,t}$ is written as a Moving Average $MA(2)$ process of the underlying porfolio return. This moving average representation is a linar stochastic approximation of the actual known nonlinear deterministic relation existing between the returns. Its interpretation, which neglects nonlinearity, can be misleading.

The ex-post performance allocation rate, i.e.:

$$\alpha_t = \frac{B_{t+1} - B_t}{A_{t+1} + B_{t+1} - (A_t + B_t)} = \frac{B_t y_{B,t}}{A_t y_{A,t} + B_t y_{B,t}}, \tag{9}$$

is not constant in time, can be erratic and rather different from the announced rate α. An effective performance allocation rate can be computed on a larger period to smooth the $\alpha_t' s$, for instance on the period $[0,T]$ corresponding to the time between resets. This effective performance allocation is:

$$\hat{\alpha}_T = \frac{B_T - B_0}{A_T + B_T - (A_0 + B_0)}, \tag{10}$$

and can also be different from α even for large T. Rate $\hat{\alpha}_T$ is likely strictly larger than α, since the total loss is assigned to account A, when the portfolio underperforms. It can even be larger due to the nonlinear allocation filtering which can create a convexity effect (see Appendix 1).

3 The Effects of the Scheme on i.i.d. Gaussian Portfolio Returns

In this section, we assume a zero riskfree rate, a zero hurdle rate $y_{h,t} = 0$, and *i.i.d.* Gaussian net portfolio returns $y_t \sim N(m, \sigma^2)$, where m (resp. σ^2) is the path-independent expected return (resp. volatility). Thus, we assume a constant hedge fund leverage ratio [see Getmanski, Lo, Makarov (2004), eq. 10] and do not consider the additional uncertainty associated with the hurdle. Except in the special case of unitary reset time in the standard allocation scheme (see Appendix 1), a theoretical analysis of the dynamics of bank accounts is difficult due to the nonlinear path dependent allocation schemes. The dynamic properties are discussed below by means of simulation studies.

In the standard High-Water Mark allocation scheme with zero hurdle, the joint dynamics of Class A value and high-water mark is characterized by the bivariate recursive system:

$$\begin{cases} A_{t+1} = A_t(1 + y_{t+1}) - \alpha[A_t(1 + y_{t+1}) - HWM_t]^+, \\ HWM_{t+1} = \max\left[HWM_t, A_t(1 + y_{t+1}) - \alpha(A_t(1 + y_{t+1}) - HWM_t)^+\right]. \end{cases} \tag{1}$$

with given initial condition (A_0, HWM_0). The bivariate process (A_t, HWM_t) is a Markov process. The joint transition distribution of (A_t, HWM_t) involves two partly degenerate distributions. Therefore, the joint bivariate transition is given by[20]:

$$\begin{aligned} & f_t(a_{t+1}, HWM_{t+1}) \\ & = \left\{ I_{a_{t+1} > HWM_t} \times \frac{1}{(1-\alpha)A_t\sigma\sqrt{2\pi}} \varphi\left[\frac{a_{t+1} - A_t(1+m) + \alpha[A_t(1+m) - HWM_t]}{(1-\alpha)A_t\sigma}\right] \right. \\ & \left. + I_{a_{t+1} < HWM_t} \times \frac{1}{A_t\sigma\sqrt{2\pi}} \varphi\left[\frac{a_{t+1} - A_t(1+m)}{A_t\sigma}\right] \right\} \otimes \varepsilon_{(HWM_{t+1} = Max(HWM_t, a_{t+1}))}, \end{aligned} \tag{2}$$

[20] Note that $y_{t+1} > y_{0,t}$, iff $A_{t+1} > HWM_t$.

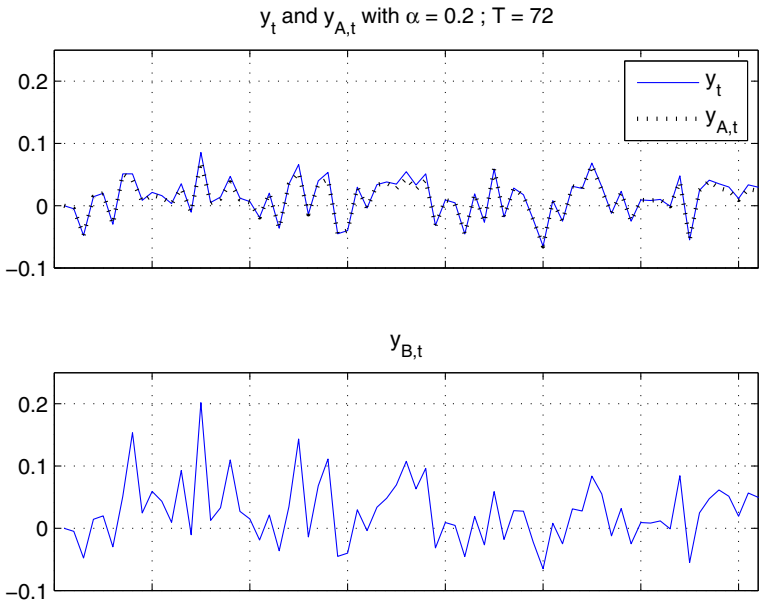


Fig. 3 Return Dynamics

where $\varepsilon_{(.)}$ denotes a point mass, φ the probability density function (pdf) of the standard normal distribution and $\otimes$ the tensor product.

To illustrate the consequences of the allocation scheme on accounts returns, let us consider risky returns following a Gaussian distribution with mean $m = 1\%$, and volatility $\sigma = 3.46\%$. We set the performance fee rate at $\alpha = 20\%$. The initial values of the accounts are $A_0 = 100$, $B_0 = 10$ and the reset time is set to $T = 72$ months$= 6$ years.

The return dynamics for y_t, $y_{A,t}$, $y_{B,t}$ are given in Figure 3. The return on management account is much more volatile than the underlying portfolio return and we observe the clustering for positive returns corresponding to the threshold effect of the HWM. The trajectories of y_t and $y_{A,t}$ are quite close[21]: the HWM effect is seen by the smoothing of peaks of y_t trajectories for the account A. These evolutions can be summarized in different ways. First, we compare the historical distribution of returns y_t and $y_{A,t}$. Second, we consider the associated autocorrelogram.

[21] It could be rather misleading to analyse the correlation between both returns in this dynamic framework. For instance, for a unitary reset time, we would have $y_{A,t} = y_t - \alpha y_t^+$. We see immediately that the conditional correlation between $y_{A,t}$ and y_t for "small" return $y_{A,t} < 0$ (resp. "large" return $y_{A,t} > 0$) is equal to 1 [resp. 1], whereas the unconditional correlation between the returns is positive, but significantly smaller than 1, with a value function of α.

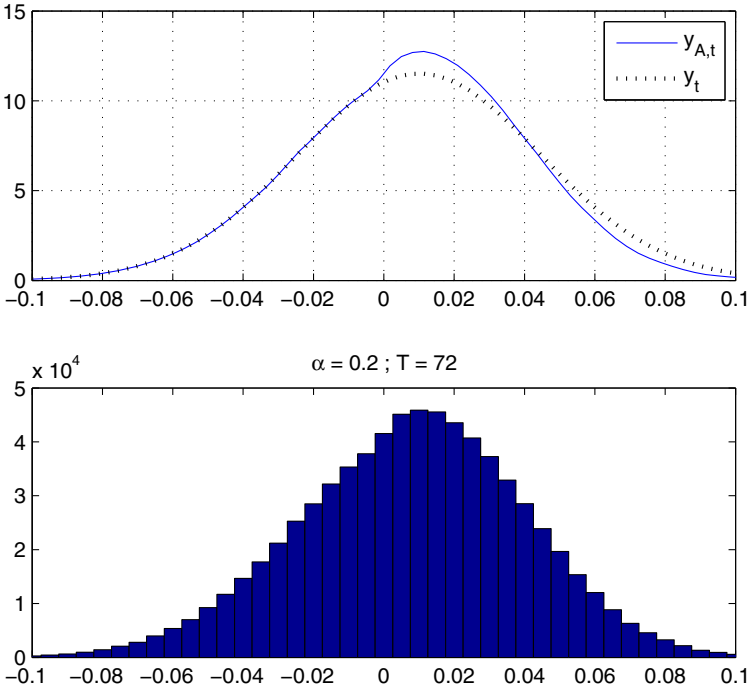


Fig. 4 Historical Distributions of Returns

The smoothed historical distributions of y_t and $y_{A,t}$ are given in the first panel of Figure 4 and the histogram of $y_{A,t}$ in the second panel. The presence of the management account explains the negative drift observed when passing from a positive portfolio return y_t to account A return. Indeed, the left part of the distribution is not impacted by the allocation scheme, whereas the right part is. The probability to observe high return is lower; the return distribution becomes more concentrated and skewed.

The nonlinear autoregressive effect due to the HWM barrier is difficult to detect from a standard linear analysis of serial dependence, but also from an analysis of the linear dependence between squared returns (see Figure 6). We observe a cycle effect in both autocorrelograms[22], which is just significant.

Let us now compare the characteristics of HF returns $y_{A,t+1}$, for different values of the performance fee rate α, $\alpha = 0\%, 10\%, 20\%, 50\%$, the limiting case $\alpha = 0\%$ corresponding to $y_{A,t+1} = y_{t+1}$. We fix the initial values to $A_0 = 100$, $B_0 = 10$. Fi-

[22] This is a consequence of the threshold autoregressive effects in the HWM dynamics [see Tong (1983)].

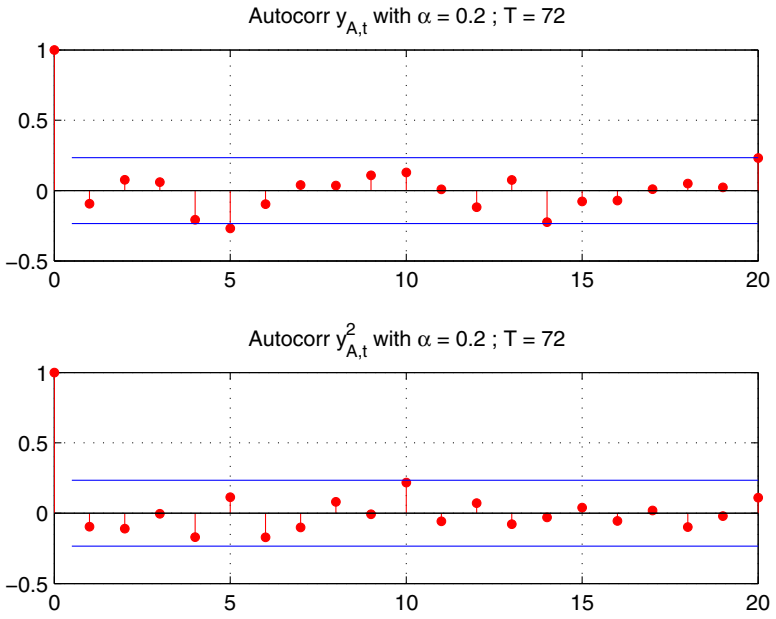


Fig. 5 ACF on Return and Squared Return

nally, we set m to 1%, consider different underlying annualized Sharpe performance ratio[23] for the portfolio return $P = \sqrt{12} \times m/\sigma = 0.5, 1, 1.5$ and different reset times for the fund, i.e. $T = 24$ (2 years), 48(4 years), 72(6 years).

Table 1 provides the mean, variance, annualized Sharpe performance, skewness, excess kurtosis and $5\% - 95\%$ quantiles of the average class A return on period $(0, T)$, that is $y_A(T) = (A_T - A_0)/(TA_0)$. These summary statistics are obtained with $S = 10000$ replications for each Monte-Carlo design.

For a zero performance fee, the return of class A unit is equal to the return of the underlying portfolio, i.e. $y_A(T) = \frac{1}{T}\left[\prod_{t=1}^{T}(1+y_t) - 1\right]$. For horizon $T \neq 1$, this return is no longer Gaussian and a convexity effect appears in the computation of the mean and the variance. For instance, we get:

$$E\left[y_A(T)\right] = \frac{1}{T}\left\{(1+m)^T - 1\right\} \simeq 1 + m + \frac{T-1}{2}m^2, \tag{3}$$

[23] The Sharpe performance ratio measures the annualized excess return per unit of annualized risk.

Table 1 Statistics on $y_A(T)$

Panel A: $T = 24$ **(2 years)**

Incentive fee α level	Mean	SD	Sharpe	Median	Skew	Exc. Kurt.	5%-Quant.	95%-Quant.
Sharpe ratio = 0.5								
0%	0.0116	0.0187	0.4375	0.0082	1.0815	2.0117	-0.0130	0.0472
10%	0.0095	0.0166	0.4028	0.0069	0.9282	1.4905	-0.0132	0.0406
20%	0.0074	0.0147	0.3585	0.0056	0.7649	1.0259	-0.0134	0.0346
50%	0.0020	0.0100	0.1408	0.0018	0.1794	0.0138	-0.0142	0.0188
Sharpe ratio = 1								
0%	0.0114	0.0091	0.8867	0.0104	0.5458	0.4780	-0.0021	0.0279
10%	0.0099	0.0081	0.8644	0.0092	0.4594	0.3658	-0.0023	0.0244
20%	0.0084	0.0072	0.8337	0.0079	0.3605	0.2722	-0.0026	0.0211
50%	0.0044	0.0047	0.6541	0.0044	-0.0748	0.2169	-0.0036	0.0121
Sharpe ratio = 1.5								
0%	0.0113	0.0060	1.3325	0.0109	0.3794	0.2175	0.0021	0.0220
10%	0.0100	0.0053	1.3234	0.0096	0.3242	0.1782	0.0017	0.0194
20%	0.0087	0.0047	1.3077	0.0084	0.2604	0.1514	0.0013	0.0168
50%	0.0049	0.0030	1.1832	0.0049	-0.0416	0.2505	0.0001	0.0098

Panel B: $T = 48$ **(4 years)**

Incentive fee α level	Mean	SD	Sharpe	Median	Skew	Exc. Kurt.	5%-Quant.	95%-Quant.
Sharpe ratio = 0.5								
0%	0.0129	0.0170	0.3809	0.0094	1.4832	3.5525	-0.0073	0.0462
10%	0.0106	0.0145	0.3675	0.0080	1.2944	2.7091	-0.0075	0.0389
20%	0.0085	0.0123	0.3468	0.0066	1.0988	1.9638	-0.0077	0.0322
50%	0.0031	0.0074	0.2141	0.0027	0.4239	0.2953	-0.0082	0.0161
Sharpe ratio = 1								
0%	0.0128	0.0081	0.7888	0.0120	0.6983	0.7160	0.0012	0.0280
10%	0.0111	0.0070	0.7901	0.0104	0.6100	0.5503	0.0008	0.0240
20%	0.0094	0.0060	0.7864	0.0090	0.5141	0.4055	0.0005	0.0203
50%	0.0050	0.0035	0.7168	0.0049	0.1242	0.1760	-0.0006	0.0110
Sharpe ratio = 1.5								
0%	0.0128	0.0054	1.1905	0.0124	0.4623	0.2722	0.0047	0.0225
10%	0.0112	0.0046	1.2056	0.0109	0.4086	0.2061	0.0041	0.0195
20%	0.0096	0.0039	1.2178	0.0094	0.3512	0.1498	0.0035	0.0166
50%	0.0054	0.0022	1.2134	0.0054	0.1246	0.0832	0.0019	0.0092

Panel C: $T = 72$ **(6 years)**

Incentive fee α level	Mean	SD	Sharpe	Median	Skew	Exc. Kurt.	5%-Quant.	95%-Quant.
Sharpe ratio = 0.5								
0%	0.0147	0.0181	0.3325	0.0104	2.1096	8.0992	-0.0046	0.0489
10%	0.0120	0.0148	0.3312	0.0087	1.8247	6.0745	-0.0048	0.0401
20%	0.0096	0.0121	0.3242	0.0073	1.5464	4.3958	-0.0050	0.0326
50%	0.0038	0.0065	0.2424	0.0032	0.6854	0.9820	-0.0056	0.0154
Sharpe ratio = 1								
0%	0.0147	0.0084	0.7090	0.0135	0.9212	1.5418	0.0031	0.0302
10%	0.0125	0.0071	0.7227	0.0117	0.8129	1.2204	0.0026	0.0254
20%	0.0105	0.0059	0.7334	0.0099	0.7005	0.9372	0.0021	0.0211
50%	0.0055	0.0031	0.7238	0.0054	0.2917	0.3452	0.0007	0.0109
Sharpe ratio = 1.5								
0%	0.0146	0.0056	1.0754	0.0141	0.6007	0.6736	0.0065	0.0245
10%	0.0126	0.0047	1.1035	0.0122	0.5348	0.5452	0.0057	0.0208
20%	0.0107	0.0039	1.1304	0.0104	0.4668	0.4310	0.0049	0.0175
50%	0.0058	0.0020	1.1872	0.0057	0.2266	0.1934	0.0027	0.0092

for small mean m, and:

$$
\begin{aligned}
V\left[y_A(T)\right] &= \frac{1}{T^2} V\left[\prod_{t=1}^{T}(1+y_t)\right] \\
&= \frac{1}{T^2}\left\{E\left[\prod_{t=1}^{T}(1+y_t)^2\right] - \left(E\left[\prod_{t=1}^{T}(1+y_t)\right]\right)^2\right\} \\
&= \frac{1}{T^2}\left\{\left[\sigma^2+(1+m)^2\right]^T - (1+m)^{2T}\right\}
\end{aligned}
$$

$$\simeq \frac{1}{T^2}\left[T(m^2+\sigma^2+2m)+\frac{T(T-1)}{2}(m^2+\sigma^2+2m)^2-T(m^2+2m)\right.$$
$$\left.-\ \frac{T(T-1)}{2}(m^2+2m)^2\right]$$
$$\simeq \frac{\sigma^2}{2}+(T-1)2m\sigma^2,$$

for small m, σ of a same magnitude. The convexity effects on these moments and the associated Sharpe ratio can be checked on all rows of Table 3 corresponding to $\alpha = 0$.

As expected from the design of management fees, the return distribution is shifted to the left. Thus, the mean, median and quantiles diminish when α increases. There is also a diminution of risk, since this distribution becomes more concentrated as observed on the values of the standard deviation and kurtosis. Finally, the distribution is right skewed for $\alpha = 0$, due to the convexity effect describe above, but the skewness diminishes when α increases due to the option interpretation of the *HWM*.

4 Endogeneous Portfolio Management

By considering i.i.d. Gaussian portfolio return in Section 3, we have implicitely assumed that the portfolio manager was investing in a kind of market portfolio, and in particular that his/her management strategy does not account for the existence of multiple accounts. The aim of this section is to discuss how the dynamics of account returns is modified with an endogenous investment strategy. In practice, the fund manager will account for an incentive mix such as reporting of good investor's performance, benefiting from the HWM on the management account, and controlling the risk of fund closure. In this section, we focus on mean-variance myopic strategies without taking into account the risk of fund closure. The strategies differ by the account value which is chosen as the main target. We consider the case of unitary reset times, where explicit strategies can be derived and analysed.

For illustration, let us assume that the fund manager invests only in a riskfree asset with zero riskfree rate and in a risky asset with i.i.d. Gaussian returns[24], denoted by y_t^*. With unitary reset time and the hurdle rate equal to the riskfree rate $y_{h,t} = y_{f,t} = 0$, the allocation between A and B accounts is given by (2.8):

$$\begin{cases} A_{t+1} = A_t(1+y_{t+1}) - \alpha A_t(y_{t+1})^+, \\ B_{t+1} = B_t(1+y_{t+1}) + \alpha A_t(y_{t+1})^+. \end{cases} \tag{1}$$

where y_{t+1} is the portfolio return. Let us now consider the portfolio allocation at date t. The total budget is allocated between the two assets: $W_t = A_t + B_t = a_{0,t} + a_t$, where $a_{0,t}$ (resp. a_t) is the value invested in the riskfree asset (resp. risky asset). At date $t+1$, the portfolio value becomes:

[24] This assumption is compatible with the standard Black-Scholes model.

$$W_{t+1} = a_{0,t} + a_t(1 + y^*_{t+1}) = W_t + a_t y^*_{t+1}.$$

We deduce the portfolio return as:

$$y_{t+1} = \frac{W_{t+1} - W_t}{W_t} = \delta_t y^*_{t+1}, \tag{2}$$

where $\delta_t = a_t/(A_t + B_t)$ denotes the fraction invested in risky asset. By substitution in (4.1), we get:

$$\begin{cases} A_{t+1} = A_t + \delta_t \left[A_t y^*_{t+1} - \alpha A_t (y^*_{t+1})^+\right], \\ B_{t+1} = B_t + \delta_t \left[B_t y^*_{t+1} + \alpha A_t (y^*_{t+1})^+\right], \end{cases} \tag{3}$$

and

$$A_{t+1} + B_{t+1} = (A_t + B_t)(1 + \delta_t y^*_{t+1}). \tag{4}$$

Let us now consider a myopic mean-variance investor[25], with absolute risk aversion[26] η. The optimal allocation depends on the account he/she is interested in.

i) If the account of interest is the total account $A + B$, the optimal allocation is the standard mean-variance efficient allocation [Markovitz (1952)] given by:

$$\delta^*_t = \frac{1}{A_t + B_t} \frac{1}{\eta} \frac{E(y^*_{t+1})}{V(y^*_{t+1})}. \tag{5}$$

Under the i.i.d. Gaussian assumption, the value invested in the risky asset is time dependent. As usual, the portfolio manager is proportionally investing less in risky asset, when $A_t + B_t$ increases. This total change in portfolio value, $(A_t + B_t)\delta^*_t y^*_{t+1} = \frac{1}{\eta}\frac{E(y^*_{t+1})}{V(y^*_{t+1})} y^*_{t+1}$, is i.i.d. Gaussian, whenever y^*_{t+1} is i.i.d. Gaussian.

ii) If the account of interest is account B, the efficient allocation becomes:

$$B_t \delta^*_{B,t} = \frac{1}{\eta} \frac{E[y^*_{t+1} + \alpha\gamma_t (y^*_{t+1})^+]}{V[y^*_{t+1} + \alpha\gamma_t (y^*_{t+1})^+]}, \tag{6}$$

where $\gamma_t = A_t/B_t$. As expected, the allocation is different from the standard allocation δ^*_t. It changes in time due to the evolution of both accounts (A_t, B_t). Moreover, the ratio between this allocation and the standard one shows a double effect: the effect of portfolio size, which diminishes from $A_t + B_t$ to B_t and implies an increase of the quantity invested in the risky asset; the effect of the optional component depends on time and tail distribution of the underlying return. The global effects is unclear.

[25] This corresponds to the two periods behavior analyzed in Christoffersen, Musto, Yilmaz (2013).

[26] We assume that the risk aversion is constant. Thus, the fund manager does not change his/her risk aversion as function of the size of the managed portfolio, or his/her past successes.

For instance, if γ_t is large, the investment in risky asset will become very small. Contrary to a usual belief, it is not guaranteed that giving an option to the fund manager makes him/her willing to take risk, even if he/she focus on the management account. This is compatible with the recent literature on incentives, in which several authors arrive to similar conclusions for instance by changing the utility function [Ross(2004)], introducing an infinite horizon [Panageas, Westerfield (2009)], or considering an option on the portfolio itself, not on the HWM [Carpenter(2000)]. As noted in this literature, if the value of account B becomes large, that is "if the HF manager has a substantial personal investment in the fund, this will inhibit excessive risk taking" [Fung, Hsieh (1999)]. This can lead to surprising consequences: for instance, at initial date 0, a small value of B_0 can be an incentive to take risk at the beginning; equivalently, introducing more frequent reset times with rather small B_0 can be an incentive to take risk regularly (ceteris paribus, i.e. for fixed gamma). In addition to this size effect, there is the optional feature since account B is a portfolio in the underlying asset and a call written on this asset. As noted in Hodder, Jackwerth (2007), this "generates risk-taking below the HWM, when the manager tries to assure that his/her incentive option will finish in the money". But "at performance levels modestly above the HWM, he/she reverses that strategy and opts for very low risk positions to lock in the option payoff".

iii) If the account of interest is account A, the efficient allocation is:

$$A_t \delta^*_{A,t} = \frac{1}{\eta} \frac{E[y^*_{t+1} - \alpha(y^*_{t+1})^+]}{V[y^*_{t+1} - \alpha(y^*_{t+1})^+]}. \tag{7}$$

This allocation depends on the evolution of account A only. The change in account value is:

$$\begin{aligned} A_{t+1} - A_t &= A_t \delta^*_{A,t} [y^*_{t+1} - \alpha(y^*_{t+1})^+] \\ &= cst\ [y^*_{t+1} - \alpha(y^*_{t+1})^+]. \end{aligned}$$

If the risky return is i.i.d. Gaussian, this change in value is still i.i.d., but no longer Gaussian.

iv) Finally, the fund manager can also own at date t a fraction ν_t of the fund, i.e. of account A [see e.g. the discussion in Fung, Hsieh (1999), or Kouwenberg, Ziemba (2007)]. Then his/her account of interest is $\nu_t A_{t+1} + B_{t+1}$, which leads to a mix of cases *ii)* and *iii)* above, if ν_t is taken exogenous.

In practice, it is difficult to know what is really the criterion selected by the fund manager. This is likely a mix, which takes into account his/her individual wealth, that is account B, and probably a fraction of account A. But he/she has also to account for the rankings of fund managers, which are regularly published in the press, and are a strong incentive for considering the preferences of fund investors[27]. To

[27] See Chevalier, Ellison (1997) for a deeper discussion of the agency conflict between fund investors and fund companies.

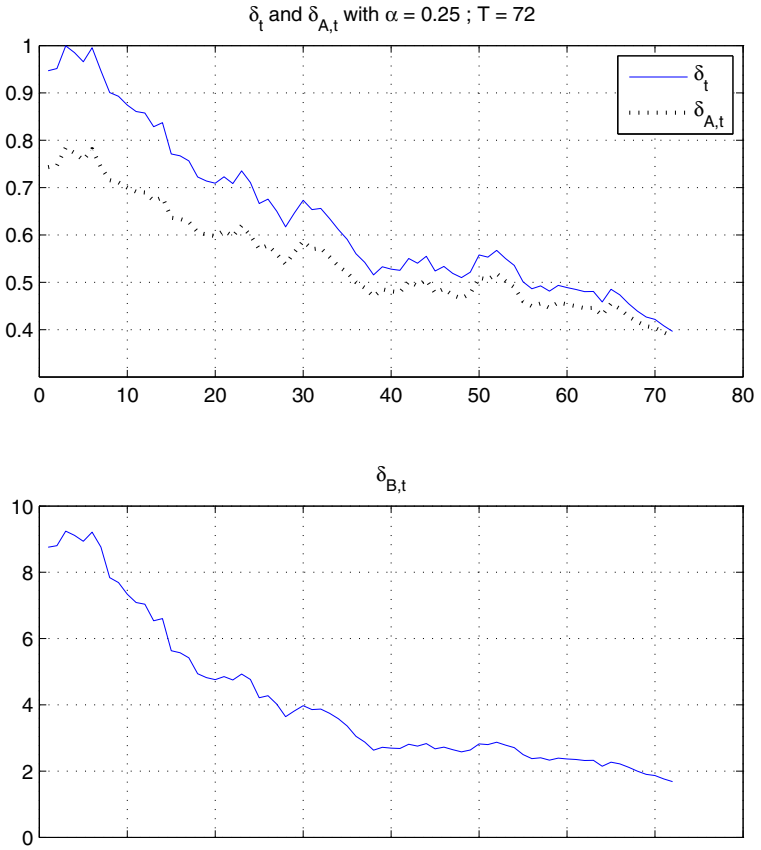


Fig. 6 Efficient Allocation in Risky Asset

illustrate the consequences of these portfolio managements on accounts returns, we consider risky returns following a Gaussian distribution with mean $m = 1\%$, and volatility $\sigma = 3.46\%$. We set the performance fee rate at $\alpha = 25\%$, with unitary reset time and the absolute risk aversion at $\eta = 0.08$. The initial values of the accounts are $A_0 = 100$, $B_0 = 10$. The length of the simulation period is $T = 72$. The explicit expressions of the mean and variance-covariance matrix of $[y^*_{t+1}, (y^*_{t+1})^+]$ are derived in Appendix 2. They are used to compute the optimal allocations.

Figure 6 displays the dynamics of efficient allocation in risky asset for the three strategies, that are δ^*_t, $\delta^*_{A,t}$, $\delta^*_{B,t}$.

The size effect is dominant in the three situations, where the allocation in risky asset diminishes in time. This shows the main role of the reset frequencies. If this frequency is the year, this might explain the empirical fact around Christmas discussed in Agarwal, Daniel, Naik (2011). We provide in Figure 7 the historical distributions

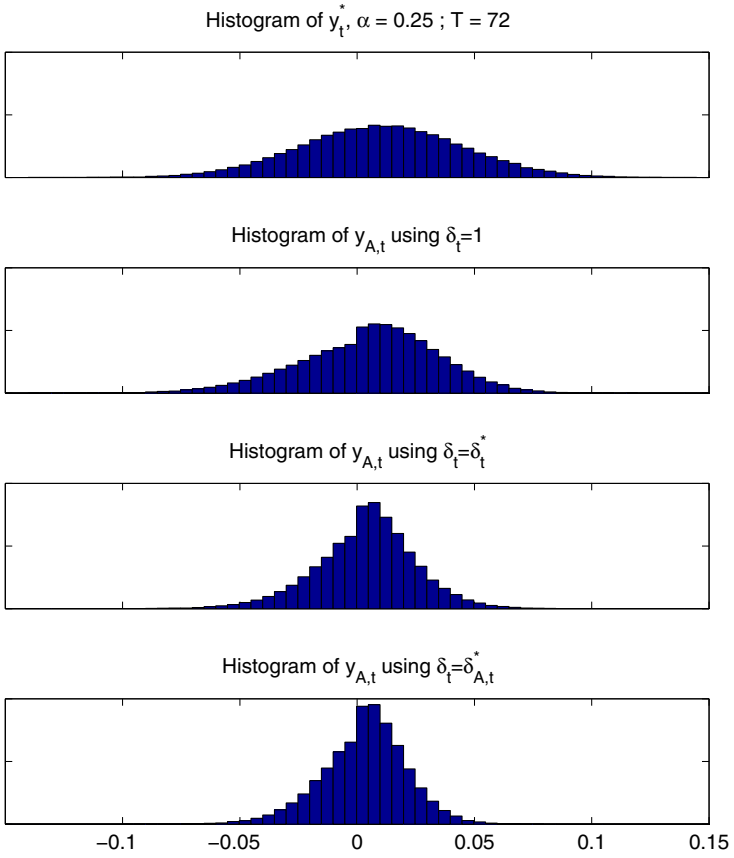


Fig. 7 Historical Distribution of Returns

of account A return when the managed portfolio is the market itself [$\delta_t = 1$], and for endogenous portfolio management with objectives $A + B$ and A, respectively. An endogenous portfolio management has clearly two effects: an increase of the discontinuity at zero and a more concentrated distribution.

However, the myopic mean-variance behaviour is not sufficient to create highly significant short term correlation on returns as shown on Figure 8. The serial correlation, which can be observed on real data, are more likely due to either the nonlinear dynamics of the basic assets introduced in the portfolio, or a non myopic, intertemporal portfolio management [see Darolles, Gourieroux (2014)]. In this respect, it could be interesting to reproduce the same simulation exercice with a market return

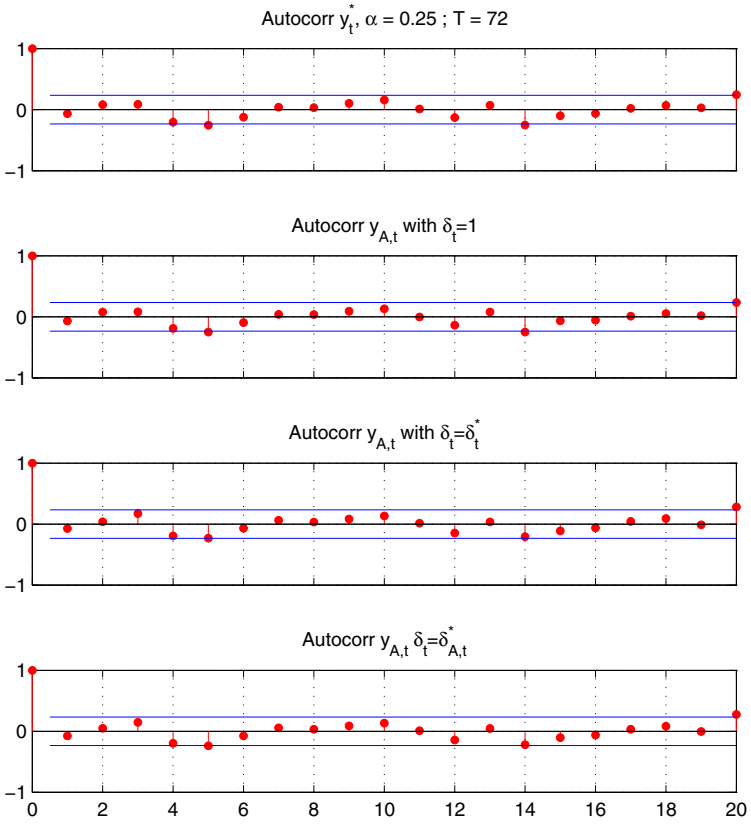


Fig. 8 ACF on Return

conditionally Gaussian, but including an ARCH effect. Indeed, this volatility effect could create linear serial correlation after passing by the nonlinear filter of HWM and provision account.

5 Conclusion

The selected HWM scheme for allocating gains and profits between the investor's account and management account has a significant impact on the performance of the investor's account. This effect is twofold. There is a direct effect on account A return due to the nonlinear scheme, especially the barrier effect included in the HWM. There is an additional indirect effect, when the fund manager ajusts his/her portfolio management to this scheme. These effects explain a part of the empirical

facts observed on hedge fund returns, such as the skewness of the return distribution, its discontinuity at zero, or some cyclical serial correlation. The special type of nonlinearity involved in this scheme can also lead to misleading interpretations for the analysis using thresholds effect, such as the study of market timing ability, or the comparison of unconditional correlations with correlations restricted to period of poor (or large) performances.

The hedge fund industry is known for its lack of transparency. Surprisingly, a lot of information is available in the prospectus of a fund, especially the scheme of allocation between the different accounts. A wise investor should analyse the consequences of these schemes on the performance of his own account before any investment in hedge funds. Similarly, it is important to take into account these schemes in the academic study of HF returns and of the behaviour of HF portfolio managers. In other terms, we have to correct the results for the management account bias and the provision account bias, and these corrections will differ due to the variability of schemes followed by individual hedge funds.

References

1. Abdulali, A.: The Bias Ratio: Measuring the Sharpe of Fraud. Protege Partners Quarterly Letter (2006)
2. Agarwal, V., Daniel, D., Naik, N.: Do Hedge Funds Manage their Reported Returns? Review of Financial Studies 24, 3281–3320 (2011)
3. Agarwal, V., Naik, N.: Multi-period Performance Persistence Analysis of Hedge Funds. Journal of Financial and Quantitative Analysis 30, 833–874 (2000)
4. Agarwal, V., Naik, N.: Risks and Portfolio Decision Involving Hedge Funds. Review of Financial Studies 17, 63–98 (2004)
5. Aragon, G., Nanda, V.: Tournament Behavior in Hedge Funds: High Water Marks, Fund Liquidation, and Managerial Stake. Review of Financial Studies 25, 937–974 (2012)
6. Bollen, N., Pool, V.: Do Hedge Fund Managers Misreport Returns? Evidence from the Pooled Distribution. Journal of Finance 64, 2257–2288 (2009)
7. Bougerol, P., Picard, D.: Strict Stationarity of Generalized Autoregressive Processes. Annals of Probability 20, 1714–1730 (1992)
8. Carpenter, J.: Does Option Conpensation Increase Managerial Risk Appetite. Journal of Finance 55, 2311–2331 (2000)
9. Chevalier, J., Ellison, G.: Risk-Taking by Mutual Funds as a Response to Incentives. Journal of Political Economy 105, 1167–1200 (1997)
10. Christoffersen, S., Musto, D., Yilmaz, B.: High Water Marks in Competitive Capital Markets (2013), available at SSRN: `http://ssrn.com/abstract=1314893`
11. Darolles, S., Gourieroux, C.: The Effects of Management and Provision Accounts on Hedge Fund Returns - Part II: The Loss Carry Forward Scheme. This issue, pp. 47–62 (2014)
12. Darolles, S., Gourieroux, C., Jasiak, J.: L-Performance with an Application to Hedge Funds. Journal of Empirical Finance 16, 671–685 (2009)
13. Diez de los Rios, A., Garcia, R.: Assessing and Valuing the Non-Linear Structure of Hedge Funds Returns (2008), available at SSRN: `http://ssrn.com/abstract=890739`
14. Elton, E., Gruber, M., Blake, C.: Incentive Fees and Mutual Funds. Journal of Finance 58, 779–804 (2003)

15. Fung, W., Hsieh, D.: A Primer on Hedge Funds. Journal of Empirical Finance 6, 309–331 (1999)
16. Getmanski, M., Lo, A., Makarov, I.: An Econometric Model of Serial Correlation and Illiquidity in Hedge Fund Returns. Journal of Financial Economics 74, 6–38 (2004)
17. Glosten, L., Jagannathan, R.: A Contingent Claim Approach to Performance Evaluation. Journal of Empirical Finance 1, 133–160 (1994)
18. Goetzmann, W., Ingersoll, J., Ross, S.: High-Water Marks and Hedge Fund Management Contracts. Journal of Finance 58, 1685–1717 (2003)
19. Henriksson, R., Merton, R.: On Market Timing and Investment Performance II: Statistical Procedures for Evaluating Forecasting Skills. Journal of Business 54, 513–533 (1981)
20. Hodder, J., Jackwerth, J.: Incentive Contracts and Hedge Fund Management. Journal of Financial and Quantitative Analysis 42, 811–826 (2007)
21. Kazemi, H., Li, Y.: Managerial Incentives and Shift of Risk-Taking in Hedge Funds (2009), available at SSRN: `http://ssrn.com/abstract=1364757`
22. Koh, F., Koh, W., Teo, M.: Asian Hedge Funds: Return Persistence Style and Fund Characteristics. Working Paper, Singapore Management University (2003)
23. Kouwenberg, R., Ziemba, W.: Incentives and Risk Taking in Hedge Funds. Journal of Banking and Finance 31, 3291–3310 (2007)
24. Lo, A.: Where Do Alphas Come From?: A New Measure of the Value of Active Investment Management. Journal of Investment Management 6, 1–29 (2008)
25. Markowitz, H.: Portfolio Selection. The Journal of Finance 7, 77–91 (1952)
26. Nelson, D.: Stationarity and Persistence in the GARCH(1,1) Model. Econometric Theory 6, 318–334 (1990)
27. Panageas, S., Westerfield, M.: High-Water Marks: High Risk Appetites? Convex Compensation, Long Horizons, and Portfolio Choice. Journal of Finance 64, 1–36 (2009)
28. Ross, S.: Compensation, Incentive and the Duality of Risk Aversion and Riskiness. Journal of Finance 59, 207–225 (2004)
29. Tong, H.: Threshold Models in Nonlinear Time Series Analysis. Springer, New York (1983)
30. Zuckerman, G.: Hedge Funds Grab More Fees as Their Popularity Increases. Wall Street Journal 244 (2004)

Appendix 1

Long Term Analysis of HWM Allocation Scheme

In HWM scheme (2.8), the dynamics of A account does not depend on the periodic reset of B account. The Net Asset Value (NAV) dynamics can be written as:

$$A_{t+1} = \left[1 + y_{t+1} - \alpha(y_{t+1} - y_{h,t})^+\right] A_t, \tag{1}$$

and (A_t) is an autoregressive process with stochastic autoregressive coefficient. Let us assume $y_{h,t} = 0$, and i.i.d. portfolio returns, with $y_t > -1/(1-\alpha)$. We can write:

$$A_{t+1} = \exp\left[\log(1 + y_{t+1} - \alpha y_{t+1}^+)\right] A_t, \tag{2}$$

and by recursive substitution:

$$A_t = A_0 \exp\left[\sum_{\tau=1}^{t} \log(1 + y_\tau - \alpha y_\tau^+)\right]. \tag{3}$$

Following the approach used in Nelson (1990), Bougerol, Picard (1992), we can determine the Lyapunov exponent of process (A_t) as follows. We have:

$$A_t = A_0 \exp\left[t \frac{1}{t} \sum_{\tau=1}^{t} \log(1 + y_\tau - \alpha y_\tau^+)\right] \tag{4}$$

$$\simeq A_0 \exp\left[tE \log(1 + y_t - \alpha y_t^+)\right], \tag{5}$$

for large t, by the Law of Large Number. Thus, the long term return on class A account is:

$$r_{\infty,A} = \lim_{t\to\infty} \frac{1}{t} \log(A_t/A_0) = E \log(1 + y_t - \alpha y_t^+). \tag{6}$$

Since $log(1+x) \leq x$, we note that:

$$r_{\infty,A} \leq E(y_t - \alpha y_t^+) = Ey_t - \alpha E y_t^+ \leq (1-\alpha) Ey_t. \tag{7}$$

As expected, this rate is strictly smaller than the long term rate on the portfolio crudely adjusted for performance rate α, i.e. $(1-\alpha)Ey_t$. It can also be significantly smaller than $Ey_t - \alpha E(y_t^+)$, with a difference which increases with the variability on (y_t).

Appendix 2

First- and Second-Order Moments of the Truncated Normal

Let us consider a Gaussian variable with mean m and unitary variance 1. The variable can be written as: $Y = m + U$, $U \sim N(0,1)$.

i) First-Order Moments

We have:

$$\begin{aligned} E[Y^+] &= E[(m+U)^+] \\ &= \int_{-m}^{\infty} (m+u)\varphi(u)du \\ &= m\int_{-m}^{\infty} \varphi(u)du + \int_{-m}^{\infty} u\varphi(u)du \\ &= m[1 - \Phi(-m)] - \int_{-m}^{\infty} \frac{d\varphi(u)}{du} du \\ &= m\Phi(m) + \varphi(m), \end{aligned}$$

where φ [resp. Φ] is the pdf [resp. cdf] of the standard normal, by using the symmetry of the standard normal. Therefore: $[EY, EY^+] = [m, m\Phi(m) + \varphi(m)]$.

ii) Second-Order Moments

Let us consider the expected squared variables, that are: $E[Y^2]$, $E[YY^+]$, $E[(Y^+)^2]$, and introduce $Y^- = Max(-Y, 0)$. We have: $Y = Y^+ - Y^-$ and $E[Y^-Y^+] = 0$. Thus:

$$\begin{aligned} E[Y^2] &= 1 + m^2 \\ E[YY^+] &= E[(Y^+)^2]. \end{aligned}$$

Therefore, all second-order moments are directly deduced from the quantity $E[(Y^+)^2]$. We get:

$$\begin{aligned} E[(Y^+)^2] &= E[((m+U)^+)^2] \\ &= \int_{-m}^{\infty} (m+u)^2 \varphi(u) du \\ &= m^2 \int_{-m}^{\infty} \varphi(u) du + 2m \int_{-m}^{\infty} u\varphi(u) du + \int_{-m}^{\infty} u^2 \varphi(u) du \\ &= m^2 \Phi(m) + 2m\varphi(m) - \int_{-m}^{\infty} u d\varphi(u) \\ &= m^2 \Phi(m) + 2m\varphi(m) - u\varphi(u)]_{-m}^{\infty} + \int_{-m}^{\infty} \varphi(u) du, \\ &= m^2 \Phi(m) + m\varphi(m) + \Phi(m). \end{aligned}$$

We deduce:

$$\begin{aligned} V \begin{bmatrix} Y \\ Y^+ \end{bmatrix} &= E \left[\begin{pmatrix} Y \\ Y^+ \end{pmatrix} (Y, Y^+) \right] - E \begin{pmatrix} Y \\ Y^+ \end{pmatrix} E (Y, Y^+) \\ &= \begin{pmatrix} 1 & \Phi(m) \\ \Phi(m) & m^2 \Phi(m) + m\varphi(m) + \Phi(m) - [m\Phi(m) + \varphi(m)]^2 \end{pmatrix}. \end{aligned}$$

The Effects of Management and Provision Accounts on Hedge Fund Returns – Part II: The Loss Carry Forward Scheme

Serge Darolles and Christian Gourieroux

Abstract. In addition to active portfolio management, hedge funds are characterized by the allocation of portfolio performance between the external investors and the management firm accounts. This allocation can take different forms, such as the Loss Carry Forward scheme, and some of them can be coupled with performance smoothing techniques. This paper shows that this additional smoothing component might explain some empirical facts observed on the distribution and the dynamics of hedge fund returns.

1 Introduction

In addition to an active portfolio management[1], hedge funds (HF) are characterized by the allocation of portfolio performance between the external investors and the management firm accounts. There exist almost as many account allocation schemes as hedge funds shares. This explains why any precautionary investor, regulator, or researcher should study in details the prospectus of the funds, and in particular the fee structure. This paper completes the discussion of the effect of the High Water Mark (HWM) allocation scheme in Darolles, Gourieroux (2014). The HWM scheme basically describes the allocation between the account invested by external clients, called class A units, and the account invested by the management firm, called class B units. The Loss Carry Forward Scheme introduced in this paper can in addition

Serge Darolles
Universite Paris-Dauphine and CREST
e-mail: serge.darolles@dauphine.fr

Christian Gourieroux
CREST and University of Toronto
e-mail: christian.gourieroux@ensae.fr

[1] The active management includes the possibility for the hedge fund manager to invest in illiquid assets, in derivatives, in junk assets, and last but not least to borrow in such assets to increase his leverage.

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_3,

include a provision account used to smooth the performance of the class A account. We describe the Loss Carry Forward (LCF) allocation scheme in Section 2 and the dynamics of the allocation between the A, B accounts and the reserve account C. We also characterize the returns of the different accounts for a given trajectory of the total portfolio return. This additional smoothing component might increase the impact of the fee structure on the hedge fund return characteristics. Section 3 compares the portfolio and fund returns for the LCF allocation schemes, when the portfolio returns are independent and identically Gaussian distributed. The i.i.d. Gaussian assumption on portfolio returns corresponds to a rather exogenous portfolio management. This assumption allows us to focus on the way the hedge fund manager will account for the existence of multiple accounts in his/her management strategy. We emphasize the special role of the provision account in this scheme. Section 4 contains conclusions. Proofs are gathered in Appendices.

2 The Loss Carry Forward Scheme

We first introduced the basic LCF scheme without provision account. We then consider a more complex scheme including a provision account.

2.1 The Basic Scheme

The basic Loss Carry Forward (LCF) allocation scheme is parametrized by a performance fee rate α, a hurdle rate $y_{h,t}$, and a reset time T. The difference with the HWM scheme described in Darolles, Gourieroux (2014) is the definition of the predeterminated path dependent scheme.

(*a*) Allocation between A and B Accounts

Let us first consider two accounts, with respective values A_t, B_t at date $t, t = 0, ..., T$. The A account is invested by external clients while the B account is invested by the management firm. The contractual hurdle rate is denoted by $y_{h,t}, y_{h,t} \geq 0$, and is assumed to be predetermined and observable at date t [see Darolles, Gourieroux (2014)]. The global portfolio value $A_t + B_t$ is invested and provides at the end of the period a return y_{t+1}. The change in portfolio value $(A_t + B_t)y_{t+1}$ can be positive, or negative. The possibility of negative return has to be considered seriously for HF, especially when they use a high leverage ratio, i.e. borrow a lot on financial markets. This change in total portfolio value has to be allocated between the two accounts. As in the HWM framework [see Darolles, Gourieroux (2014)], the performance fee is not charged if the fund is globally in a deficit of performance, called loss carry forward[2] (LCF). This measure of deficit is recursively defined by $LCF_0 = 0$ and:

[2] The loss carry forward is an accounting technique that applies the current year's losses to future years gains in order to reduce tax liability.

$$LCF_t = -\left[LCF_{t-1} + A_{t-1}(y_t - y_{h,t-1})\right]^-, \tag{1}$$

$$= -A_{t-1}\left[y_t - (A_{t-1}y_{h,t-1} - LCF_{t-1})/A_{t-1}\right]^-, \tag{2}$$

where $X^- = \max(-X, 0)$. The LCF is always nonpositive and corresponds to the cumulated negative performance. The hurdle rate $y_{h,t-1}$ is fixing an objective for the portfolio return. If this objective is not reached, that is if $y_t < y_{h,t-1}$, this is considered as a loss and the measure of deficit increases. The LCF_t becomes negative if y_t is not large enough to cover (potential) previous losses.

Then, the allocation depends on LCF and is driven by the following updating equations:

$$\begin{cases} A_{t+1} = A_t(1+y_{t+1}) - \alpha A_t\left[y_{t+1} - (A_t y_{h,t} - LCF_t)/A_t\right]^+, \\ B_{t+1} = B_t(1+y_{t+1}) + \alpha A_t\left[y_{t+1} - (A_t y_{h,t} - LCF_t)/A_t\right]^+, \end{cases} \tag{3}$$

where α, $\alpha > 0$, is the performance rate. Thus the management firm (B account) receives a bonus, if the portfolio return is sufficiently large, i.e. if $y_{t+1} > A_t y_{h,t} - LCF_t$, receive nothing otherwise.

The fee rate α, $\alpha = 20\%$, say, is often presented at a first place when promoting a fund, whereas the complicated formulas (2.2), (2.3) can only be revealed by the careful reading of the prospectus. Therefore a naive investor may have the impression that the management firm receives at date $t+1$ the quantity $(A_t + B_t)y_{t+1}(1+\alpha)$. This is clearly not the case. The payment to the management firm includes some incentives to get extreme positive performance in order to increase the bonus and to optimize the reduction of tax liability. As important as the fee rate is of course the choice of the hurdle rate and its dynamics.

At short term horizon equal to 1, the future account values involve the payoff of a European call written on y_{t+1}, with predetermined path dependent strike equal to $y_{0,t} = (A_t y_{h,t} - LCF_t)/A_t$.

The recursive equations (2.3) are valid on period $\{0, T-1\}$. At reset time T, the management account is reset to the contractual initial value B_0 and the LCF reset to zero.

If the reset time is $T = 1$, the LCF is always set to zero, $y_{0,t} = y_{h,t}$, and the recursive equation (2.3) can be simplified and becomes:

$$\begin{cases} A_{t+1} = A_t(1+y_{t+1}) - \alpha A_t\left[y_{t+1} - y_{h,t}\right]^+, \\ B_{t+1} = B_t(1+y_{t+1}) + \alpha A_t\left[y_{t+1} - y_{h,t}\right]^+. \end{cases} \tag{4}$$

that corresponds to the HWM scheme [see Darolles, Gourieroux (2014)]. Therefore, the HWM and LCF schemes are equivalent for a unitary reset time. The dependence of the change of account value $\Delta A_{t+1} = A_{t+1} - A_t$ (resp. $\Delta B_{t+1} = B_{t+1} - B_t$) with respect to net portfolio return y_{t+1} is described in Figure 1 (resp. Figure 2).

When $T = 1$, the value of the class A unit is a continuous increasing function of the net portfolio return with a change of slope at threshold $y_{0,t}$. The payoff on B account is a convex function of the return. This convexity property shows the incentive mechanism.

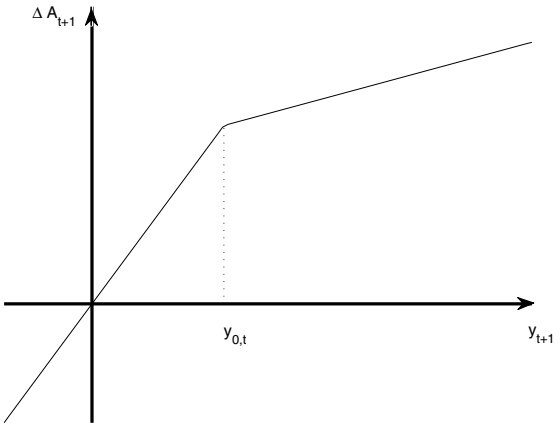


Fig. 1 ΔA_{t+1} as a function of y_{t+1} (unitary reset time)

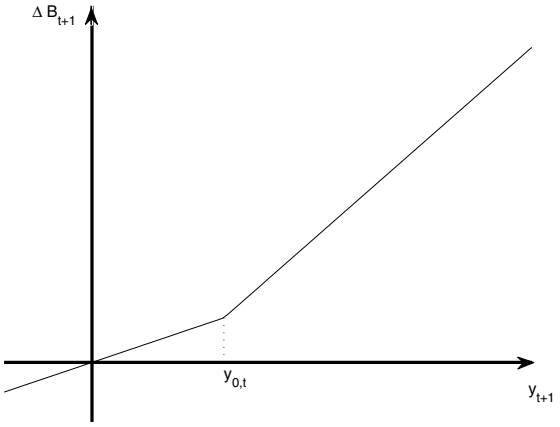


Fig. 2 ΔB_{t+1} as a function of y_{t+1} (unitary reset time)

(*b*) The Case of a Zero Hurdle Rate

Larger the hurdle rate, greater is the incentive for the fund manager to take risk and to increase the leverage in order to get a high bonus. In the HF industry the hurdle rate is generally positive and indexed on some basic rate such as the LIBOR. However, a significant number of HF set the hurdle rate to zero, that is do not adjust for a riskfree rate. We consider this special LCF scheme in this section to better highlight the link with the HWM framework.

Proposition 1: The HWM and LCF schemes are identical for zero hurdle rate, with $LCF_t = A_t - HWM_t$.

Proof: see Appendix 1.

For a nonzero hurdle rate, the HWM and LCF approaches differ by their discounting scheme and the dependence of ΔA_{t+1} (resp. ΔB_{t+1}) with respect to net portfolio return y_{t+1} is more complex.

Proposition 2: For zero hurdle rate, there exists a one-to-one relationship between the trajectories of the portfolio return y_t and the return of the investors' account $y_{A,t}$. More precisely, we can deduce the underlying portfolio return as a deterministic function:

$$y_t = g(y_{A,t}, y_{A,t-1}, ..., y_{A,1}, A_0), \ say.$$

Proof: By the transformation in Figure 1, we have:

$$A_t = A_{t-1} + b(y_t, y_{0,t-1}),$$

and by recursive substitution:

$$A_t = b^*(y_t, A_{t-1}, A_{t-2}, ..., A_0), \ say,$$

where b^* is one-to-one in the first argument y_t. Thus, by introducing return $y_{A,t}$, we deduce the formula of Proposition 2.

□

The return $y_{A,t}$ on the investors' account is regularly reported by the HF manager and use to promote the fund. They do not report the underlying portfolio return y_t in order not to reveal clearly their portfolio management, but also the actual level of fees. As a consequence, the academic literature is often using the return $y_{A,t}$ as a proxy of y_t, that is neglects the effect of the management fee. Proposition 2 shows that we are able to derive the underlying portfolio return from the return of account A by simply inverting the filter, which defines the accounts allocation. Even if the data on portfolio return are not made directly observable by the fund manager, we can recursively reconstruct them. Of course the relation between y_t and $y_{A,t}$ is not static, and no deterministic link of the type $y_t = g^*(y_{A,t})$, say, will be detected by a joint plot of $(y_t, y_{A,t})$. When the hurdle rate is nonzero, we still have a one-to-one relationship conditional on the knowledge of the hurdle rate history, that is:

$$y_t = g(y_{A,t}, y_{A,t-1}, ..., y_{A,1}, y_{h,t-1}, y_{h,t-2}, ..., A_0), \ say.$$

2.2 An Allocation Scheme with Provision Account

More sophisticated allocation scheme can include a third account, called provision account[3]. This scheme involves additional allocation parameters characterizing the allocation between the external investors' A account and the provision C account.

(*a*) Allocation between A, B and C Accounts

Let us now consider three accounts, with respective values A_t, B_t and C_t at date t, $t = 0, ..., T$. The global portfolio value $A_t + B_t + C_t$ is invested and provides at the end of the period a return denoted by y_{t+1}. Then, the change in total portfolio value is $(A_t + B_t + C_t)y_{t+1}$. As in Section 2.1., we first assume that the return on B account is always allocated to the corresponding class. We only consider how $(A_t + C_t)y_{t+1}$ has to be allocated between the three accounts depending on some predetermined regimes.

We consider below an allocation process based on a modified LCF measure of performance deficit. In this case, the LCF can be interpreted as the negative part of a virtual provision account (whereas the value of the actual provision account has to be always positive). Hence, at any date t, the sum $LCF_t + C_t$ is only impacted by one of its two components, the other one being zero. The LCF starts to be negative when the provision account is empty and C_t starts to be positive when the LCF is null . For expository purpose, the allocation scheme is described below in two steps to highlight the smoothing technique.

i) Three accounts - no smoothing

A proportion β of the change in the A+C accounts value up to the hurdle rate, that is $(A_t + C_t)(y_{t+1} - y_{h,t})$, is allocated to the provision account, under the positivity constraint on this account. The loss carry forward is defined by:

$$LCF_{t+1} = -\left[LCF_t + C_t + \beta(A_t + C_t)(y_{t+1} - y_{h,t})\right]^-, \tag{5}$$

and the corresponding provision account value is:

$$C_{t+1} = \left[LCF_t + C_t + \beta(A_t + C_t)(y_{t+1} - y_{h,t})\right]^+, \tag{6}$$

with initial conditions $LCF_0 = C_0 = 0$. Then, the values of accounts A and B are deduced from the dynamics of the provision account by the following equations:

$$\begin{cases} A_{t+1} = (A_t + C_t)(1 + y_{t+1}) - C_{t+1}, \\ B_{t+1} = B_t(1 + y_{t+1}). \end{cases} \tag{7}$$

[3] In HF literature, this account is called reserve account. It seems preferable to avoid this terminology, which will become misleading if some Basel type of regulation is applied to HF in a near future.

By construction, the provision account value (resp. the LCF) is always nonnegative (resp. nonpositive). Moreover, only one of the LCF and C value can be different from zero at any given date.

When $C_t = 0$, equation (2.5) reduces to the standard LCF recursive equation (2.1). When $C_t > 0$ (and $LCF_t = 0$), a capital appreciation $(A_t + C_t)(y_{t+1} - y_{h,t}) > 0$ will increase the value of the provision account, whereas the LCF will stay equal to zero. Finally, if $C_t > 0$ and there is a large capital depreciation up to the hurdle rate, the provision account is set to zero and the complete return allocated to the A account.

ii) Three accounts with smoothing

We now add to the previous allocation scheme the smoothing component. This effect is obtained through a change in the recursive equation (2.6) giving the C account dynamics. We assume that a proportion of the provision account is allocated to the external investors' and management firm accounts in case of bad portfolio performance. The recursive system becomes:

$$LCF_{t+1} = -\left[LCF_t + C_t + \beta(A_t + C_t)(y_{t+1} - y_{h,t})\right]^-, \tag{8}$$

for the LCF,

$$C_{t+1} = [1 - \varphi_A(y_{t+1}) - \varphi_B(y_{t+1})]\left[LCF_t + C_t + \beta(A_t + C_t)(y_{t+1} - y_{h,t})\right]^+, \tag{9}$$

for the provision account, and:

$$\begin{cases} A_{t+1} = (A_t + C_t)(1 + y_{t+1}) + [\varphi_A(y_{t+1}) - 1]\left[LCF_t + C_t + \beta(A_t + C_t)(y_{t+1} - y_{h,t})\right]^+, \\ B_{t+1} = B_t(1 + y_{t+1}) + \varphi_B(y_{t+1})\left[LCF_t + C_t + \beta(A_t + C_t)(y_{t+1} - y_{h,t})\right]^+, \end{cases} \tag{10}$$

for A and B accounts, where the smoothing functions φ_A, φ_B are positive and such that $\varphi_A + \varphi_B \leq 1$.

A simple scheme assumes constant smoothing functions $\varphi_A(y) = \varphi_A$, $\varphi_B(y) = \varphi_B$, say. For instance, if φ_A and φ_B are such that $\varphi_A + \varphi_B = 1$, and if moreover $\beta = 1$, the provision account is always empty, and the scheme reduces to the standard *LCF* scheme with two accounts described in Section 2.2.

However, more sophisticated smoothing functions are introduced in the hedge fund industry. For instance, we can fix a predetermined level[4] $y_{0,t} < 0$, different from the hurdle rate, and define the smoothing functions as:

$$\varphi_A(y_{t+1}) = \varphi_B(y_{t+1}) = \frac{1}{2}\min\left[1, \left(\frac{y_{t+1}}{y_{0,t}}\right)^+\right]. \tag{11}$$

[4] The level $y_{0,t}$ can be constant and set for example to -1% to smooth small negative returns.

Thus, if $y_{t+1} < y_{0,t}$, we get $\varphi_A(y_{t+1}) = \varphi_B(y_{t+1}) = \frac{1}{2}$, and a full use of the provision account to smooth A (and B) return. If $y_{0,t} < y_{t+1} < 0$, we have a partial smoothing. Finally, if $y_{t+1} > 0$, we get $\varphi_A(y_{t+1}) = \varphi B(y_{t+1}) = 0$ and the previous account is feeded to insure the fund against future potential losses.

(*b*) Returns and Asset Values

By analogy with the standard scheme, we can consider different returns. The most important ones are:

i) The total portfolio return: y_{t+1};
ii) The return for class A account: $y_{A,t+1} = (A_{t+1} - A_t)/A_t$;
iii) The return associated with both A and C accounts: $y_{A,C,t+1} = (A_{t+1} + C_{t+1} - (A_t + C_t))/(A_t + C_t)$.

Indeed, it is important to distinguish the net asset value (NAV) for class A, i.e. A_t, and the value including also the provision account, i.e. $A_t + C_t$. The net asset value A_t is provided for at least two purposes. This is the accounting value which has to be introduced by the investors in their balance sheet. This is also the benchmark for the selling price proposed by the fund management to an investor who wants to redeem its investment. This NAV A_t is a kind of bid price (i.e. selling price), which is smaller or equal to the "fair value" of the fund equal to $A_t + C_t$.

Clearly, the provision account creates a "conditional return smoothing" when passing from y_t to y_t^A, to follow the terminology of Bollen, Pool (2008). However, this (known) smoothing is much more complicated than usually described in the academic literature [see e.g. Bollen, Pool (2008), eq 7].

3 The Effects of the Scheme on i.i.d. Gaussian Portfolio Returns

In this section, we assume a zero riskfree rate, a zero hurdle rate $y_{h,t} = 0$, and $i.i.d.$ Gaussian net portfolio returns $y_t \sim N(m, \sigma^2)$, where m (resp. σ^2) is the path-independent expected return (resp. volatility). Thus, we assume a constant hedge fund leverage ratio [see Getmanski, Lo, Makarov (2004), eq. 10] and do not consider the additional uncertainty associated with the hurdle. Except in the special case of unitary reset time in the standard allocation scheme for which the LCF and HWM coincide [see Darolles, Gourieroux (2014)], a theoretical analysis of the dynamics of bank accounts is difficult due to the nonlinear path dependent allocation schemes. The dynamic properties are discussed below by means of simulation studies.

3.1 The Loss Carry Forward Allocation Scheme (Without Provision Account)

From Proposition 1, we know that the LCF scheme is identical to the HWM scheme for a zero hurdle rate. The associated $LCF^* = LCF$ trajectory is given in the fourth panel of Figure 3.

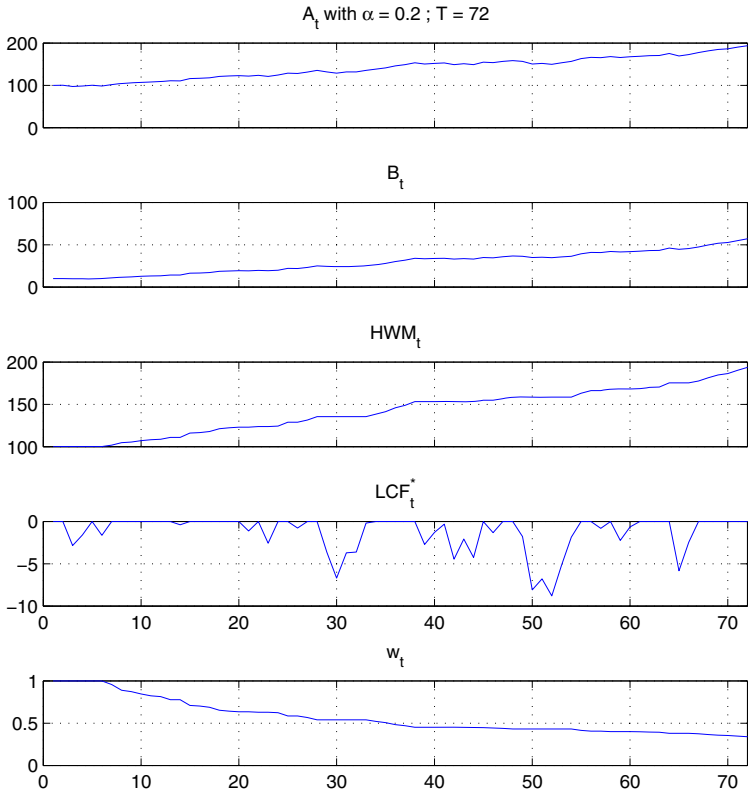


Fig. 3 Trajectories of Account Values, HWM, LCF^* (without provision account)

We display in Figure 3 the trajectories of the two account values, the HWM, the implied $LCF_t^* = A_t - HWM_t$ (see Proposition 1) and the relative weights of both accounts, i.e. the ratio $w_t = \frac{A_t}{B_t}\frac{B_0}{A_0}$.

Due to the selected performance fee rate of the portfolio management, the two account values are increasing, but this increase is larger for the management account than for the investor's account. We also observe that the ratio w_t is decreasing in time and clearly different from the announced $1 - \alpha = 80\%$.

3.2 The Allocation Scheme with Provision Account

We display in Figure 4 the trajectories of the three accounts A, B, C, and the LCF. We consider independent risky returns following a Gaussian distribution with mean $m = 1\%$, and variance $\sigma^2 = 1\%$, set the provision rate at $\beta = 25\%$, and use the smoothing functions (2.11) with level $y_0 = -1\%$.

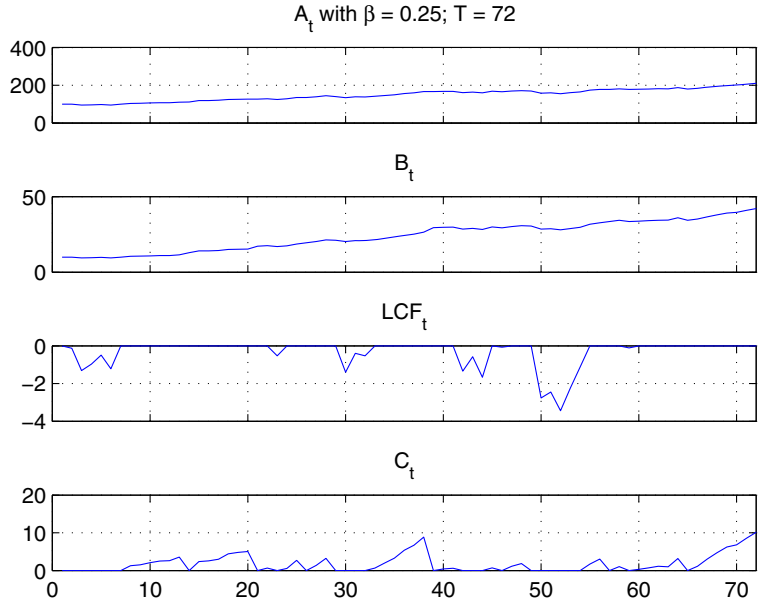


Fig. 4 Trajectories of Account Values and LCF (with provision account)

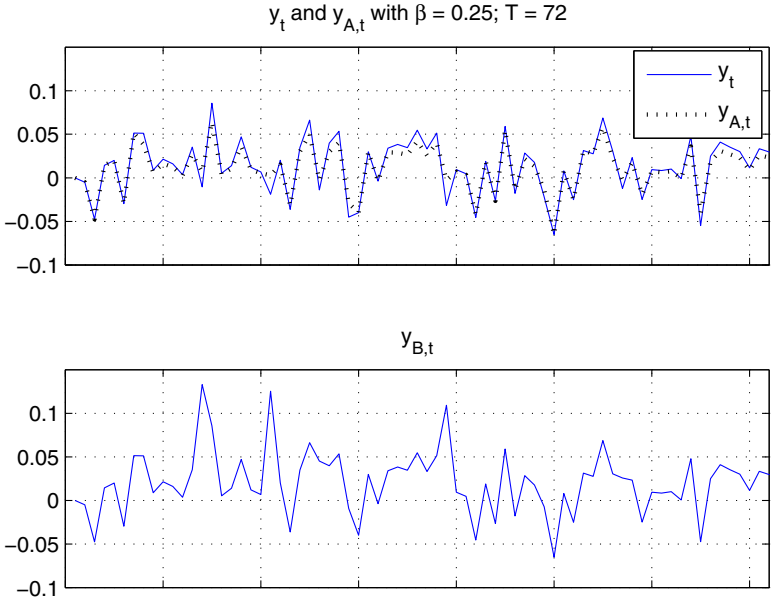


Fig. 5 Return Dynamics (with provision account)

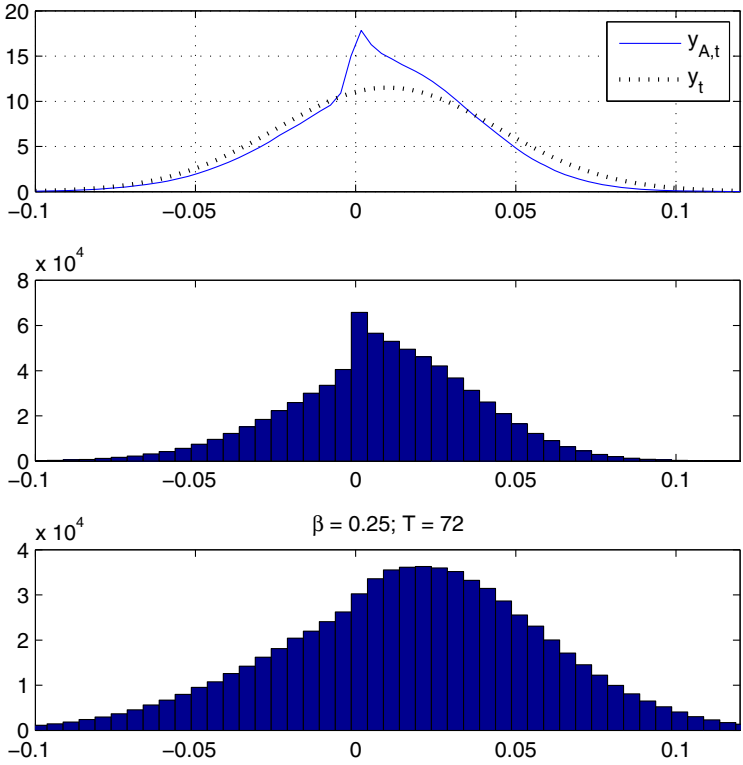


Fig. 6 Historical Distributions of Returns (with provision account)

The return dynamics for $y_t, y_{A,t}, y_{B,t}$ are provided in Figure 5. We observe that the presence of a provision account smooths the investor's account return. This makes more marketable the published HF returns $y_{A,t}$ by reducing the value of the usual fund risk indicators such as the return volatility.

i) Historical distribution of returns

As in the HWM allocation scheme [see Darolles, Gourieroux (2014)], the return dynamics can be summarized in different ways. First, we compare the historical distributions of returns y_t and $y_{A,t}$ in Figure 6. In presence of a provision account, the two sides of the distribution are modified. The left side (corresponding to negative return) is moved to the right, that is, we get less negative returns, especially around zero. Moreover, the right part is also impacted, due to the smoothing rule used in this simulation. The high positive returns are less frequent, but the probability to observe small positive returns increases. Thus, the provision account implies right skewness and discontinuity on the return distribution, which is clearly seen on the histogram of $y_{A,t}$ provided in the second panel of Figure 6. The discontinuity is less pronounced with return computed on two consecutive periods (3^d panel of Figure 6), which

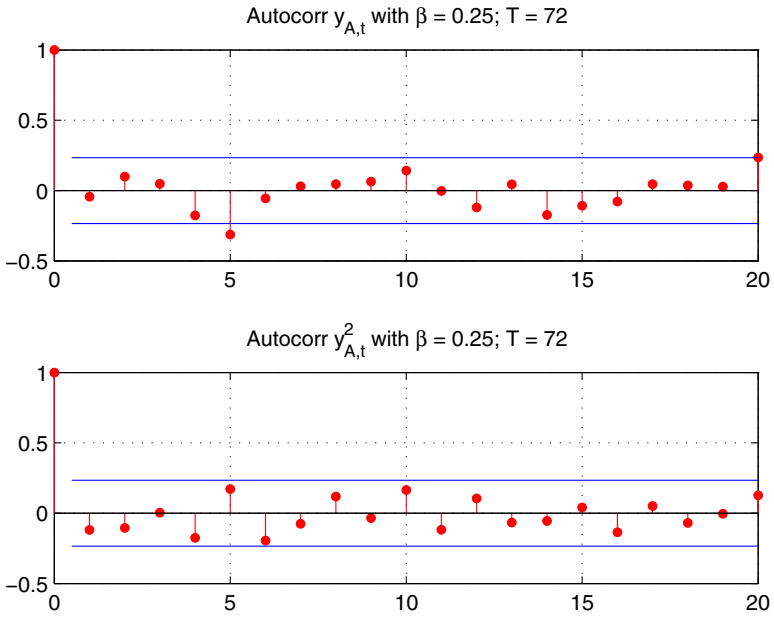


Fig. 7 ACF on Return and Squared Return (with provision account)

is compatible with the observation by Bollen, Pool (2009) that the discontinuity can disappear when the horizon increases. These empirical facts have already been documented in the literature. However, they have been explained by either fraud [Abdulali (2006)], misreporting of returns, if the manager fully report gains, but delays reporting losses [see e.g. Bollen, Pool (2009)], survivorship bias [Brown, Goetzmann, Ibbotson (1999)], or backfilling bias, when both superior and inferior performers stop reporting [Ackermann, McEnally, Ravenscraft (1999)]. In fact, the bias ratio is likely a consequence of the (transparent) design of the allocation scheme between the three accounts.

ii) Return dynamics

The nonlinear autoregressive effect due to the provision account is still difficult to detect from a simple linear analysis of serial dependence (see Figure 7), even if the cycle effect due to the threshold autoregressive dynamics (2.10) [see Tong (1983)] becomes more significant. This cycle effect implies in particular negative autocorrelations at periodic lags. This dependence created by the account allocation scheme is not able to explain the positive short term persistence emphasized in the HF literature [see e.g. Agarwal, Naik (2000), Getmanski, Lo, Makarov (2004)], but is

Table 1 Statistics on $y_A(T)$ (with provision account)

Panel A: $T = 24$ **(2 years)**

Provision β level	Mean	SD	Sharpe	Median	Skew	Exc. Kurt.	5%-Quant.	95%-Quant.
Sharpe ratio = 0.5								
0%	0.0116	0.0187	0.4375	0.0082	1.0815	2.0117	-0.0130	0.0472
5%	0.0110	0.0181	0.4310	0.0079	1.0431	1.8856	-0.0130	0.0456
10%	0.0105	0.0175	0.4241	0.0076	1.0037	1.7617	-0.0130	0.0439
20%	0.0095	0.0165	0.4086	0.0070	0.9222	1.5221	-0.0131	0.0404
Sharpe ratio = 1								
0%	0.0114	0.0091	0.8867	0.0104	0.5458	0.4780	-0.0021	0.0279
5%	0.0110	0.0088	0.8840	0.0101	0.5216	0.4443	-0.0021	0.0269
10%	0.0106	0.0085	0.8810	0.0098	0.4967	0.4129	-0.0021	0.0259
20%	0.0098	0.0079	0.8736	0.0091	0.4449	0.3583	-0.0022	0.0240
Sharpe ratio = 1.5								
0%	0.0113	0.0060	1.3325	0.0109	0.3794	0.2175	0.0021	0.0220
5%	0.0110	0.0058	1.3344	0.0105	0.3615	0.1990	0.0020	0.0212
10%	0.0106	0.0056	1.3361	0.0102	0.3435	0.1827	0.0019	0.0204
20%	0.0098	0.0052	1.3385	0.0095	0.3073	0.1591	0.0018	0.0189

Panel B: $T = 48$ **(4 years)**

Provision β level	Mean	SD	Sharpe	Median	Skew	Exc. Kurt.	5%-Quant.	95%-Quant.
Sharpe ratio = 0.5								
0%	0.0129	0.0170	0.3809	0.0094	1.4832	3.5525	-0.0073	0.0462
5%	0.0124	0.0163	0.3787	0.0091	1.4406	3.3529	-0.0073	0.0444
10%	0.0118	0.0157	0.3763	0.0088	1.3976	3.1588	-0.0074	0.0425
20%	0.0108	0.0146	0.3703	0.0081	1.3101	2.7869	-0.0074	0.0389
Sharpe ratio = 1								
0%	0.0128	0.0081	0.7888	0.0120	0.6983	0.7160	0.0012	0.0280
5%	0.0124	0.0078	0.7901	0.0116	0.6794	0.6806	0.0011	0.0270
10%	0.0120	0.0076	0.7911	0.0112	0.6603	0.6468	0.0010	0.0259
20%	0.0111	0.0070	0.7925	0.0104	0.6213	0.5842	0.0009	0.0239
Sharpe ratio = 1.5								
0%	0.0128	0.0054	1.1905	0.0124	0.4623	0.2722	0.0047	0.0225
5%	0.0124	0.0052	1.1950	0.0120	0.4510	0.2590	0.0045	0.0217
10%	0.0120	0.0050	1.1993	0.0116	0.4398	0.2468	0.0044	0.0209
20%	0.0111	0.0046	1.2076	0.0108	0.4177	0.2260	0.0041	0.0193

Panel C: $T = 72$ **(6 years)**

Provision β level	Mean	SD	Sharpe	Median	Skew	Exc. Kurt.	5%-Quant.	95%-Quant.
Sharpe ratio = 0.5								
0%	0.0147	0.0181	0.3325	0.0104	2.1096	8.0992	-0.0046	0.0489
5%	0.0141	0.0173	0.3327	0.0100	2.0462	7.6148	-0.0047	0.0468
10%	0.0134	0.0165	0.3328	0.0096	1.9832	7.1516	-0.0047	0.0447
20%	0.0122	0.0150	0.3322	0.0088	1.8577	6.2843	-0.0047	0.0405
Sharpe ratio = 1								
0%	0.0147	0.0084	0.7090	0.0135	0.9212	1.5418	0.0031	0.0302
5%	0.0141	0.0081	0.7126	0.0130	0.8981	1.4693	0.0030	0.0289
10%	0.0136	0.0077	0.7162	0.0126	0.8749	1.3992	0.0029	0.0277
20%	0.0125	0.0071	0.7229	0.0117	0.8283	1.2663	0.0026	0.0254
Sharpe ratio = 1.5								
0%	0.0146	0.0056	1.0754	0.0141	0.6007	0.6736	0.0065	0.0245
5%	0.0141	0.0053	1.0822	0.0136	0.5871	0.6452	0.0063	0.0235
10%	0.0136	0.0051	1.0891	0.0131	0.5736	0.6180	0.0061	0.0226
20%	0.0126	0.0047	1.1026	0.0122	0.5469	0.5674	0.0056	0.0208

compatible with the negative autocorrelation detected in Bollen, Pool (2009), when lagged returns are just above zero[5].

[5] A linear analysis of serial correlation can also be rather misleading. Indeed conditional serial correlations can be very different. For instance, it is equal to zero when $y_{A,t}$ is sufficiently large, since $y_{A,t} = y_t$, but will become significant when $y_{A,t}$ is small, due to the effect of the optional component which depends on the past. These different levels of conditional serial correlations are just consequences of the HWM schemes. We cannot necessarily conclude that a "manager smooths more likely losses than gains" [Bollen, Pool (2008), (2009)].

iii) Summary statistics on return

Let us now compare the characteristic of HF returns $y_{A,t+1}$, for different values of the provision rate β assigned to account C, $\beta = 0\%$, 5%, 10%, 20%; the limiting case $\beta = 0\%$ corresponds to $y_{A,t+1} = y_{t+1}$. All other parameters are set to the values used to compute Table 1.

We observe that the distribution is shifted to the left when the β parameter increases, but this shift is less pronounced than in the scheme without provision account. Moreover, the risk parameters also diminish when the β parameter increases. In consequence, the Sharpe ratio is stable, and then is less sensitive to the management fee politics. The skewness and kurtosis parameters also decrease with β.

4 Conclusion

The LCF scheme used for allocating gains and profits between the investor's account, management account and provision account has a significant impact on the performance of the investors' account. The first effect is related to the nonlinearity of the scheme, especially the barrier effects, An additional effect is introduced by the smoothing component associated with the provision account. These two effects explain a part of the empirical facts observed on hedge fund returns, such as the skewness of the return distribution, it discontinuity at zero, or some cyclical serial correlation.

We see that the complexity of the formulas defining the allocation schemes and also the diversity of these schemes, which depend on the choice of the free rate, sequence of hurdle rate, the rate of the capital appreciation/depreciation and the smoothing functions. This diversity makes difficult the comparison of what is proposed by different funds. From a regulatory point of view, there is a need for a standardization of these allocation schemes, that is of the way the "bonuses" of the HF management firms are computed.

References

1. Abdulali, A.: The Bias Ratio: Measuring the Sharpe of Fraud. Protege Partners Quarterly Letter (2006)
2. Ackermann, C., McEnally, R., Ravenscraft, D.: The Performance of Hedge Funds: Risk, Return and Incentives. The Journal of Finance 54, 833–874 (1999)
3. Agarwal, V., Naik, N.: Multi-period Performance Persistence Analysis of Hedge Funds. Journal of Financial and Quantitative Analysis 30, 833–874 (2000)
4. Bollen, N., Pool, V.: Conditional Return Smoothing in the Hedge Fund Industry. Journal of Financial and Quantitative Analysis 43, 267–298 (2008)
5. Bollen, N., Pool, V.: Do Hedge Fund Managers Misreport Returns? Evidence from the Pooled Distribution. The Journal of Finance (2009) (forthcoming)
6. Brown, S., Goetzmann, W., Ibbotson, R.: Offshore Hedge Funds: Survival and Performance 1989-1995. Journal of Business 72, 91–117 (1999)
7. Darolles, S., Gourieroux, C.: The Effects of Management and Provision Accounts on Hedge Fund Returns - Part I: The High Water Mark Scheme. This Issue, pp. 23–45 (2014)

8. Getmanski, M., Lo, A., Makarov, I.: An Econometric Model of Serial Correlation and Illiquidity in Hedge Fund Returns. Journal of Financial Economics 74, 6–38 (2004)
9. Tong, H.: Threshold Models in Nonlinear Time Series Analysis. Springer, New York (1983)

Appendix 1

Proof of Proposition 1

i) Let us first consider the HWM scheme and denote by $LCF_t^* = A_t - HWM_t$ the implied LCF associated with this scheme. The recursion for the HWM scheme is:

$$\begin{cases} A_{t+1} = A_t(1+y_{t+1}) - \alpha A_t \left[y_{t+1} - (HWM_t - A_t)/A_t\right]^+, \\ HWM_{t+1} = \max\left(HWM_t, A_{t+1}\right), \end{cases}$$

or equivalently,

$$\begin{cases} A_{t+1} = A_t(1+y_{t+1}) - \alpha A_t \left(y_{t+1} + \frac{LCF_t^*}{A_t}\right)^+, \\ LCF_{t+1}^* = -\left(LCF_t^* + A_{t+1} - A_t\right)^-. \end{cases}$$

We get the two following regimes:

- Regime 1: $LCF_t^* + A_t y_{t+1} > 0$,
  with:
  $$A_{t+1} = A_t(1+y_{t+1}) - \alpha A_t \left(y_{t+1} + \frac{LCF_t^*}{A_t}\right). \tag{1}$$
  Then:
  $$\begin{aligned} LCF_t^* + A_{t+1} - A_t &= LCF_t^* + A_{t+1} - \alpha A_t \left(y_{t+1} + \frac{LCF_t^*}{A_t}\right) \\ &= (1-\alpha)(LCF_t^* + A_t y_{t+1}) > 0. \end{aligned}$$
  We deduce that:
  $$LCF_{t+1}^* = 0. \tag{2}$$
- Regime 2: $LCF_t^* + A_1 y_{t+1} < 0$.
  We get:
  $$A_{t+1} = A_t(1+y_{t+1}). \tag{3}$$
  Thus, $LCF_t^* + A_{t+1} - A_t = LCF_t^* + A_t y_{t+1} < 0$, and we deduce that:
  $$LCF_{t+1}^* = LCF_t^* + A_t y_{t+1}. \tag{4}$$

ii) Let us now consider the recursion for the LCF scheme:

$$\begin{cases} A_{t+1} = A_t(1+y_{t+1}) - \alpha A_t \left[y_{t+1} + LCF_t/A_t\right]^+ \\ LCF_{t+1} = -\left(LCF_t + A_t y_{t+1}\right)^-. \end{cases}$$

We get the two following regimes:

- Regime 1: $LCF_t + A_t y_{t+1} > 0$,

$$\begin{cases} A_{t+1} = A_t(1 + y_{t+1}) - \alpha A_t \left[y_{t+1} + LCF_t / A_t \right], \\ LCF_{t+1} = 0. \end{cases} \tag{5}$$

- Regime 2: $LCF_t + A_t y_{t+1} < 0$,

$$\begin{cases} A_{t+1} = A_t(1 + y_{t+1}), \\ LCF_{t+1} = LCF_t + A_t y_{t+1}. \end{cases} \tag{6}$$

The recursive equations $(1.1) - (1.4)$ are identical to the equations $(1.5) - (1.6)$. Proposition 1 follows by noting that the initial values of the LCF and implied LCF are the same: $LCF_0^* = A_0 - HWM_0 = 0$, $LCF_0 = 0$.

□

How to Detect Linear Dependence on the Copula Level?

Vladik Kreinovich, Hung T. Nguyen, and Songsak Sriboonchitta

Abstract. In many practical situations, the dependence between the quantities is linear or approximately linear. Knowing that the dependence is linear simplifies computations; so, is is desirable to detect linear dependencies. If we know the joint probability distribution, we can detect linear dependence by computing Pearson's correlation coefficient. In practice, we often have a copula instead of a full distribution; in this case, we face a problem of detecting linear dependence based on the copula. Also, distributions are often heavy-tailed, with infinite variances, in which case Pearson's formulas cannot be applied. In this paper, we show how to modify Pearson's formula so that it can be applied to copulas and to heavy-tailed distributions.

1 Introduction: Traditional Approach to Detecting Linear Dependence

Locally, linear dependencies are ubiquitous.

Dependencies between quantities are often described by smooth (even analytical) functions $y = f(x_1, \ldots, x_n)$. An analytical function can be expanded in Taylor series

Vladik Kreinovich
Department of Computer Science,
University of Texas at El Paso, 500 W. University, El Paso, TX 79968, USA
e-mail: vladik@utep.edu

Hung T. Nguyen
Department of Mathematical Sciences,
New Mexico State University, Las Cruces, New Mexico 88003, USA
e-mail: hunguyen@nmsu.edu

Songsak Sriboonchitta
Department of Economics, Chiang Mai University, Chiang Mai, Thailand
e-mail: songsak@econ.chiangmai.ac.th

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_4,

around each point $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)})$:

$$y = f(x^{(0)}) + \sum_{i=1}^{n} c_i \cdot (x_i - x_i^{(0)}) + \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} \cdot (x_i - x_i^{(0)}) \cdot (x_j - x_j^{(0)}) + \ldots \qquad (1)$$

For values x_i close to $x_i^{(0)}$, we can safely ignore terms which are quadratic in $x_i - x_i^{(0)}$ (or of higher order), and thus, approximate the dependence by a linear function $y \approx f(x^{(0)}) + \sum\limits_{i=1}^{n} c_i \cdot (x_i - x_i^{(0)})$.

Linear dependencies are often global.

In many practical situations, linear dependencies extend beyond local, they hold even for situations in which differences $x_i - x_i^{(0)}$ are reasonably large.

It is important to know if we have a linear dependence.

Linear dependencies make computations easier. For example, there are efficient algorithms for solving systems of linear equations, while a solution to the system of non-linear equations is, in general, NP-hard; see, e.g., [10].

An exact linear dependence is easy to detect.

Let us first consider the ideal case, when estimation and measurement errors can be safely ignored, and the dependence is exactly linear. In this case, if we have K situations in which we measured all the values x_i and y, then, based on the corresponding values $(x_1^{(k)}, \ldots, x_n^{(k)}, y^{(k)})$, $k = 1, 2, \ldots, K$, we can check the dependence is linear by checking whether the corresponding system of linear equations with unknowns c_i has a solution:

$$y^{(k)} = f(x^{(0)}) + \sum_{i=1}^{n} c_i \cdot \left(x_i^{(k)} - x_i^{(0)}\right), \quad k = 1, \ldots, K. \qquad (2)$$

As we have mentioned, there exist efficient algorithms for checking solvability of such a linear system.

How the presence of an approximate linear dependence is detected now.

Since linear dependencies make computations easier, it is desirable to detect them even when we only have an approximate linear dependence: e.g., due to measurement or approximation errors, or due to actual non-linear terms in the dependence, or due to the fact that the value of the quantity y is only approximately determined by the values $x_1, \ldots, x_n$.

In the case of the exact linear dependence, possible values of the tuple $(x_1, \ldots, x_n, y)$ form a linear surface $y = f(x^{(0)}) + \sum\limits_{i=1}^{n} c_i \cdot (x_i - x_i^{(0)})$. When we observe

the frequency with which different tuples occur, we get a probability distribution on this surface.

In the case of an approximate linear dependence, tuples can deviate from the surface corresponding to the exact linear equation. In this case, the probability distribution is no longer limited to this surface. Instead, we have a probability distribution on the $(n+1)$-dimensional space. Let $\rho(x_1,\ldots,x_n,y)$ denote the probability density of this probability distribution.

In traditional statistics, in the simplest case $n=1$, the linearity of the corresponding dependence can be gauged by computing the Pearson's correlation coefficient (see, e.g., [21]). For a 2-D distribution with a cumulative distribution function $F(x,y)=\text{Prob}(X\le x\,\&\,Y\le y)$ corresponding to probability density $\rho(x,y)$, Pearson's correlation coefficient is defined as

$$r(F)=\frac{C_{XY}}{\sigma_X\cdot\sigma_Y}, \tag{3}$$

where

$$C_{XY}\stackrel{\text{def}}{=}E[(X-E[X])\cdot(Y-E[Y])]=E[X\cdot Y]-E[X]\cdot E[Y]=$$
$$\int x\cdot y\cdot\rho(x,y)\,dxdy-E[X]\cdot E[Y], \tag{4}$$

$$E[X]\stackrel{\text{def}}{=}\int x\cdot\rho(x,y)\,dxdy,\ \ E[Y]\stackrel{\text{def}}{=}\int y\cdot\rho(x,y)\,dx, \tag{5}$$

$$\sigma_X\stackrel{\text{def}}{=}\sqrt{V_X},\ \ \sigma_Y\stackrel{\text{def}}{=}\sqrt{V_Y}, \tag{6}$$

$$V_X\stackrel{\text{def}}{=}E[(X-E[X])^2]=E[X^2]-(E[X])^2=$$
$$\int x^2\cdot\rho(x,y)\,dxdy-\left(\int x\cdot\rho(x,y)\,dxdy\right)^2, \tag{7}$$

$$V_Y\stackrel{\text{def}}{=}E[(Y-E[Y])^2]=E[Y^2]-(E[Y])^2=$$
$$\int y^2\cdot\rho(x,y)\,dxdy-\left(\int y\cdot\rho(x,y)\,dxdy\right)^2. \tag{8}$$

In the case of an exact linear dependence $Y=c_0+c_1\cdot X$, this coefficient $r(F)$ is equal to 1 if $c_1>0$ and to -1 if $c_1<0$. Vice versa, if $r(F)=\pm1$, this means that with probability 1, we have $Y=c_0+c_1\cdot X$ for appropriate coefficients c_0 and c_1.

In general, values $r(F)\ne0$ indicate that there is an approximate linear dependence – and the closer $|r(F)|$ to 1, the closer is the the actual dependence to a linear one.

Validating a linear model.

The square $R^2=(r(F))^2$ is used, in statistics, as a "measure of fit" which is used to validate the linear model: the closer this square to 1, the better the fit.

2 Detecting Linear Dependence Based on a Copula: Formulation of the First Problem

Need for copulas.

In the general case, a distribution of a random variable X can be described by the cumulative distribution function $F_X(x) \stackrel{\text{def}}{=} \text{Prob}(X \le x)$, and a joint distribution of two variables X and Y can be described by the cumulative distribution function $F(x,y) \stackrel{\text{def}}{=} \text{Prob}(X \le x \& Y \le y)$.

A problem with this description is that it depends on the units in which we describe x and y. For example, if we use meters instead of feet to describe x, or if we use a logarithmic scale of decibels instead of a linear scale of energy to describe noise, we get different cumulative distribution functions $F(x,y)$.

It is desirable to describe the dependence between x and y in a way which is independent on the units for measuring x and y. Such a description is known as a *copula*. The main idea behind a copula is that, once we know the probability distribution, we no longer need to use any artificial units to describe each of the quantities x and y:

- to describe the value of x, we can use the probability $F_X(x) = \text{Prob}(X \le x)$; and
- to describe the value of y, we can use the probability $F_Y(y) = \text{Prob}(Y \le y)$.

Thus, instead of asking for a value $F(x,y) = \text{Prob}(X \le x \& Y \le y)$ corresponding to given real numbers x and y, we can ask for a value $C(a,b)$ of this probability corresponding to given probabilities $a = F_X(x)$ and $b = F_Y(y)$.

Formally, the copula is defined as a function $C(a,b)$ for which $a = F_X(x)$ and $b = F_Y(y)$ imply that $F(x,y) = c(a,b)$, i.e., equivalently, as a function for which $F(x,y) = C(F_X(x), F_Y(y))$ for all x and y.

Copulas are useful.

Copulas have been successfully used to describe dependencies in many application areas, including econometrics; see, e.g., [9, 17, 19].

Formulation of the problem.

We need to be able to detect linear dependence between the quantities x and y based only on the copula $C(a,b)$ that describes their dependence.

3 Detecting Linear Dependence Based on a Copula: Main Idea and the Resulting Definition

Main idea behind the new definition.

We consider a situation in which we know the copula $C(a,b)$ but we do not know the marginal distributions $F_X(x)$ and $F_Y(y)$. We would like to know whether there exist

some marginal distributions for which the dependence between the corresponding random variables x and y is linear, i.e., for which, for which, for the corresponding probability distribution $F(x,y) = C(F_X(x), F_Y(y))$, the Pearson's coefficient is equal either to 1 or to -1.

For different marginal distributions, we have different values of the Pearson's correlation coefficient. The possibility to have $r(F) = 1$ for *at least one* pair of the marginal distributions means that the *maximum* L^+ of $r(F)$ over all pairs of possible marginal distributions is equal to 1. Thus, we can use this maximum to gauge to what extent a given copula represents an increasing linear dependence.

Similarly, the possibility to have $r(F) = -1$ for *at least one* pair of marginal distributions means that the *minimum* L^- of $r(F)$ over all such pairs is equal to -1. Thus, we can use this minimum to gauge to what extent a given copula represents a decreasing linear dependence. So, we arrive at the following definition.

Definition

Let a copula $C(a,b)$ be given. By *measures of linearity* corresponding to this copula, we mean the values

$$L^- \stackrel{\text{def}}{=} \min_{F_X(x), F_Y(y)} r(C(F_X(x), F_Y(y))); \tag{9a}$$

$$L^+ \stackrel{\text{def}}{=} \max_{F_X(x), F_Y(y)} r(C(F_X(x), F_Y(y))), \tag{9b}$$

where $r(F)$ denote Pearson's correlation coefficient (3) corresponding to $F(x,y) = C(F_X(x), F_Y(y))$, and the minimum and maximum are taken over all possible marginal probability distributions $F_X(x)$ and $F_Y(y)$.

Thus defined values L^- and L^+ depend only on the copula.

In the above definition, we fix a copula $C(a,b)$, and we consider all possible 2-D probability distributions $F(x,y)$ corresponding to this copula. Therefore, the above-defined values L^- and L^+ depend only on the copula.

The values L^- and L^+ describe the possibility of a linear dependence.

If $L^+ = 1$, this means that there exist marginal distributions $F_X(x)$ and $F_Y(y)$ for which $r(F) = 1$, i.e., for which the corresponding random variables X and Y are linearly related by an increasing linear dependence $Y = c_0 + c_1 \cdot X$, with $c_1 > 0$. Similarly, if $L^- = -1$, this means that the exist marginal distributions $F_X(x)$ and $F_Y(y)$ for which $r(F) = -1$, i.e., for which the corresponding random variables X and Y are linearly related by a decreasing linear dependence $Y = c_0 + c_1 \cdot X$, with $c_1 < 0$.

In general, values $L^+ > 0$ or $L^- < 0$ indicate that there is an approximate linear dependence – and the closer $|L^+|$ or $|L^-|$ to 1, the closer is the approximate dependence to a linear one.

How to define the corresponding measure of fit.

For validating a linear model, as a measure of fit M, it is reasonable to take the largest possible value of the traditional measure of fit $R^2 = (r(F))^2$ over all possible probability distributions corresponding to the given copula.

If the largest value of $(r(F))^2$ is attained when $r(F) > 0$, then $L^+ \geq |L^-|$, and the above-defined measure of fit is equal to $(L^+)^2$. If the largest value of $(r(F))^2$ is attained when $r(F) < 0$, then $|L^-| \geq L^+$, and the above-defined measure of fit is equal to $(L^-)^2$. These two cases can be combined into a single formula

$$M = \max((L^-)^2, (L^+)^2).$$

How to actually compute L^- and L^+ based on $F(x,y)$: an idea.

A direct application of the above definition based on the known probability distribution $F(x,y)$ seems computationally expensive: first, we need to compute the copula, and then, based on this copula, we need to solve two optimization problems. It turns out that it is possible to compute L^- and L^+ more efficiently.

This possibility is related to the fact that, once we know a joint distribution $F(x,y)$ for non-discrete random variables X and Y (i.e., for random variables for which the corresponding marginal distributions $F_X(x)$ and $F_Y(y)$ are continuous functions), we can explicitly describe all other random variables (X',Y') with the same copula as (X,Y).

Indeed, by definition of the copula, for the original random pair (X,Y), we have $F(x,y) = C(F_X(x), F_Y(y))$. Thus, we have $C(a,b) = F(F_X^{-1}(a), F^{-1}(b))$, where $F^{-1}(x)$ denotes an inverse function. Since the pair (X',Y') is described by the same copula $C(a,b)$ as the pair (X,Y), the distribution function $F'(x',y')$ for this pair has the form $F'(x',y') = C(F_{X'}(x'), F_{Y'}(y'))$, where $F_{X'}(x')$ and $F_{Y'}(y')$ are the corresponding marginal distributions. Substituting the above expression for the copula $C(a,b)$ into this formula, we conclude that $F'(x',y') = F(a(x'), b(y'))$, where we denoted $a(x') \stackrel{\text{def}}{=} F_X^{-1}(F_{X'}(x'))$ and $b'(x') \stackrel{\text{def}}{=} F_X^{-1}(F_{X'}(x'))$.

By definition of a cumulative distribution function $F(x,y) =$
$\text{Prob}(X \leq x \& Y \leq y)$, the formula $F'(x',y') = F(a(x'), b(y'))$ means that $\text{Prob}(X' \leq x' \& Y' \leq y') = \text{Prob}(X \leq a(x') \& Y \leq b(y'))$.

Since the cumulative distribution functions are non-decreasing, the inverses $F_X^{-1}(a)$ and $F^{-1}(b)$ are also non-decreasing and thus, the compositions $a(x')$ and $b(y')$ are also non-decreasing. So, the condition $X \leq a(x')$ is equivalent to $A(X) \leq x'$, where $A(x)$ denotes an inverse function to $a(x)$, and similarly the condition $Y \leq a(y')$ is equivalent to $B(Y) \leq y'$, where $B(y)$ denotes an inverse function to $b(y)$. Thus, we conclude that $\text{Prob}(X' \leq x' \& Y' \leq y') = \text{Prob}(A(X) \leq x' \& B(Y) \leq y')$. In other words, the probability distribution of the pair (X',Y') is exactly the same as the probability distribution of the pair $(A(X), B(Y))$.

Vice versa, one can easily check that if we take any two strictly increasing functions $A(x)$ and $B(y)$, then for the pair (X',Y') with $X' = A(X)$ and $Y' = B(Y)$, we get the exact same copula as for the original pair (X,Y).

In other words, all possible probability distributions (X',Y') corresponding to the same copula $C(a,b)$ as the pair of random variables (X,Y) can be obtained by considering appropriate non-decreasing transformations $X' = A(X)$ and $Y' = B(Y)$. For the variables, mean, variance, covariance, and correlation can be explicitly determined in terms of the functions $A(x)$ and $B(y)$. Thus, we arrive at the following easier-to-compute equivalent formulas for describing the desired measures of linearity L^- and L^+.

Towards an easier-to-compute equivalent definition of L^- and L^+.

Let (X,Y) be random variables corresponding to a copula $C(a,b)$. Then, the measures of linearity L^- and L^+ can be computed as

$$L^- = \min_{A(x),B(y)} r(A(X),B(Y)),\ \ L^+ = \max_{A(x),B(y)} r(A(X),B(Y)), \tag{9c}$$

where maximum and minimum are taken over all possible non-decreasing functions $A(x)$ and $B(y)$, and $r(A(X),B(Y))$ is the Pearson's correlation coefficient relating the random variables $A(X)$ and $B(Y)$.

By definition of Pearson's correlation coefficient $r(F)$, we conclude that

$$L^- = \min_{A(x),B(y)} L(A,B);\ \ L^+ = \max_{A(x),B(y)} L(A,B), \tag{10}$$

where

$$L(A,B) \stackrel{\text{def}}{=} \frac{C(A,B)}{\sigma(A)\cdot\sigma(B)}, \tag{11}$$

$$C(A,B) = E[(A(X)\cdot B(Y))] - E[A(X)]\cdot E[B(Y)] =$$

$$\int A(x)\cdot b(y)\cdot\rho(x,y)\,dx\,dy -$$

$$\left(\int A(x)\cdot\rho(x,y)\,dx\,dy\right)\cdot\left(\int B(y)\cdot\rho(x,y)\,dx\,dy\right), \tag{12}$$

$$\sigma(A) \stackrel{\text{def}}{=} \sqrt{V(A)},\ \ \sigma(B) \stackrel{\text{def}}{=} \sqrt{V(B)}, \tag{13}$$

$$V(A) \stackrel{\text{def}}{=} E[A^2(X)] - (E[A(X)])^2 =$$

$$\int A^2(x)\cdot\rho(x,y)\,dx\,dy - \left(\int A(x)\cdot\rho(x,y)\,dx\,dy\right)^2, \tag{14}$$

$$V(B) \stackrel{\text{def}}{=} E[B^2(X)] - (E[B(X)])^2 =$$

$$\int B^2(y)\cdot\rho(x,y)\,dx\,dy - \left(\int B(y)\cdot\rho(x,y)\,dx\,dy\right)^2. \tag{15}$$

Comment.

Strictly speaking, the above equivalence between copulas and non-linear re-scalings requires that we consider only strictly increasing functions $a(x)$ and $b(y)$, for which the inverses $A(x)$ and $B(y)$ are also strictly increasing. However, one can easily show that any non-decreasing function $A(x)$ can be approximated, with any given accuracy, by a strictly increasing one: e.g., we can approximate $A(x)$ by $A(x)+\varepsilon\cdot x$ for a sufficiently small $\varepsilon>0$. Thus, in (10), it does not matter whether we take only strictly increasing functions or all non-decreasing ones.

Explicit expressions for L^- and L^+ in terms of the copula.

The above equivalent reformulation was intended for the case when we still need to compute the copula. However, even when we already know the copula $C(a,b)$, the above reformulation can still simplify computations.

Indeed, the formula (9c) can be applied to any probability distribution corresponding to a given copula. In particular, it is well known that the copula itself is a probability distribution on the box $[0,1]\times[0,1]$, corresponding to uniform marginal distributions $F_X(x)=\text{Prob}(X\le x)=x$ and $F_Y(y)=\text{Prob}(Y\le y)=y$. For this probability distribution, $F(x,y)=C(x,y)$ and thus, $\rho(x,y)=\dfrac{\partial^2 C(x,y)}{\partial x\partial y}$. For this probability density, we can apply the above formulas (10)–(15), and compute the desired values L^- and L^+.

4 How to Actually Compute L^- and L^+

Analysis of the problem.

In accordance with the above idea, for computing L^- and L^+, we will use the easier-to-compute equivalent reformulation (10) of the original definition of these two measures of linearity.

According to calculus, one way to find minimum and maximum of an expression is to equate the derivative to 0. In our case, we need to situations when the unknowns are two functions $A(x)$ and $B(y)$, the rules for corresponding differentiation are described in variational calculus; see, e.g., [7].

Here, $\sigma(B)$ does not depend on $A(x)$, so, by using the usual rules of differentiating the ratio, we get:

$$\frac{\delta}{\delta A(x)}L(A,B)=\frac{1}{\sigma(B)}\cdot\frac{\delta}{\delta A(x)}\left(\frac{C(A,B)}{\sigma(A)}\right)=$$

$$\frac{1}{\sigma(B)}\cdot\frac{\delta}{\delta A(x)}\cdot\frac{\dfrac{\delta C(A,B)}{\delta A(x)}\cdot\sigma(A)-C(A,B)\cdot\dfrac{\delta\sigma(A)}{\delta A(x)}}{\sigma^2(A)}. \tag{16}$$

Thus, the derivative is equal to 0 if

$$\frac{\delta C(A,B)}{\delta A(x)} \cdot \sigma(A) - C(A,B) \cdot \frac{\delta \sigma(A)}{\delta A(x)} = 0. \tag{17}$$

Since $\sigma(A) = \sqrt{V(A)}$, the chain rule for differentiation implies that

$$\frac{\delta \sigma(A)}{\delta A(x)} = \frac{1}{2\sigma(A)} \cdot \frac{\delta V(A)}{\delta A(x)}. \tag{18}$$

For $V(A) = \int A^2(x) \cdot \rho(x,y)\,dx\,dy - \left(\int A(x) \cdot \rho(x,y)\,dx\,dy\right)^2$, we get

$$\frac{\delta V(A)}{\delta A(x)} = 2A(x) \cdot \int \rho(x,y)\,dy - 2E[A(X)] \cdot \int \rho(x,y)\,dy. \tag{19}$$

Similarly, for

$$C(A,B) = \int A(x) \cdot B(y) \cdot \rho(x,y)\,dx\,dy - \left(\int A(x) \cdot \rho(x,y)\,dx\,dy\right) \cdot \left(\int B(y)) \cdot \rho(x,y)\,dx\,dy\right), \tag{20}$$

we get

$$\frac{\delta C(A,B)}{\delta A(x)} = \int B(y) \cdot \rho(x,y)\,dx\,dy - E[B(Y)] \cdot \int \rho(x,y)\,dy. \tag{21}$$

Thus, the above equation (17) takes the form

$$C_1 \cdot \int B(y) \cdot \rho(x,y)\,dx\,dy + C_2 \cdot A(x) \cdot \int \rho(x,y)\,dy + C_3 \cdot \int \rho(x,y)\,dy = 0 \tag{22}$$

for some constants C_i. From this equation, we can determine $A(x)$ as

$$A(x) = a_1 + a_2 \cdot E[B(Y)\,|\,X = x], \tag{23}$$

where a_i are appropriate constants, and the conditional expected value

$$E[B(Y)\,|\,X = x] \tag{24}$$

has the form

$$E[B(Y)\,|\,X = x] = \frac{\int B(y) \cdot \rho(x,y)\,dx\,dy}{\int \rho(x,y)\,dx\,dy}. \tag{25}$$

By differentiating with respect to $B(y)$, we get a similar equation

$$B(y) = b_1 + b_2 \cdot E[A(X)\,|\,Y = y], \tag{26}$$

for appropriate constants b_1 and b_2.

These expressions depend on constants a_i and b_j which need to be determined. To make the expressions easier, we can take into account that the correlation coefficient does not change if we apply a linear transformation to the variables. Thus, instead of the functions $A(x)$ and $B(y)$, we can use arbitrary linear re-scalings $a + a' \cdot A(x)$ and $b + b' \cdot B(y)$. We can use this ambiguity to normalize the functions $A(x)$ and $B(y)$, e.g., by setting $A(0) = B(0) = 0$ and $A(1) = B(1) = 1$. By applying these conditions to the above formula for $B(y)$, we conclude that

$$B(0) = 0 = b_1 + b_2 \cdot E[A(X) \,|\, Y = 0], \tag{27}$$

$$B(1) = 1 = b_1 + b_2 \cdot E[A(X) \,|\, Y = 1]. \tag{28}$$

Subtracting the first equation from the second one, we get

$$1 = b_2 \cdot (E[A(X) \,|\, Y = 1] - E[A(X) \,|\, Y = 0]), \tag{29}$$

hence

$$b_2 = \frac{1}{E[A(X) \,|\, Y = 1] - E[A(X) \,|\, Y = 0]}. \tag{30}$$

From the equation (27) for $B(0)$, we can now conclude that

$$b_1 = -\frac{E[A(X) \,|\, Y = 0]}{E[A(X) \,|\, Y = 1] - E[A(X) \,|\, Y = 0]}. \tag{31}$$

Substituting the expressions for b_1 and b_2 into the formula (26) for $B(y)$, we thus conclude that

$$B(y) = \frac{E[A(X) \,|\, Y = y] - E[A(X) \,|\, Y = 0]}{E[A(X) \,|\, Y = 1] - E[A(X) \,|\, Y = 0]}. \tag{32}$$

Similarly, we get

$$A(x) = \frac{E[B(Y)) \,|\, X = x] - E[B(Y) \,|\, X = 0]}{E[B(Y) \,|\, X = 1] - E[B(Y) \,|\, X = 0]}. \tag{33}$$

Resulting algorithm.

Formulas (32) and (33) prompts the following natural iterative algorithm. We start with arbitrary initial functions $A^{0)}(x)$ and $B^{0)}(y)$, e.g., with functions $A^{(0)}(x) = x$ and $B^{(0)}(y) = y$. Then, on each iteration, once we know the values $A^{(k)}(x)$ and $B^{(k)}(y)$, we compute the values corresponding to the next iteration as follows:

$$A^{(k+1)}(x) = \frac{E[B^{(k)}(Y)) \,|\, X = x] - E[B^{(k)}(Y) \,|\, X = 0]}{E[B^{(k)}(Y) \,|\, X = 1] - E[B^{(k)}(Y) \,|\, X = 0]}, \tag{34}$$

$$B^{(k+1)}(y) = \frac{E[A^{(k)}(X) \,|\, Y = y] - E[A^{(k)}(X) \,|\, Y = 0]}{E[A^{(k)}(X) \,|\, Y = 1] - E[A^{(k)}(X) \,|\, Y = 0]}. \tag{35}$$

We stop when the new functions $A^{(k+1)}(x)$ and $B^{(k+1)}(y)$ are close to functions $A^{(k)}(x)$ and $B^{(k)}(y)$ from the previous iteration: e.g., when the differences do not exceed some threshold ε:

$$|A^{(k+1)}(x) - A^{(k+1)}(x)| \le \varepsilon; \;\; |B^{(k+1)}(y) - B^{(k+1)}(y)| \le \varepsilon. \tag{36}$$

We then take $A^{(k+1)}(x)$ and $B^{(k+1)}(y)$ as the desired functions $A(x)$ and $B(y)$. Based on these functions, we use the formula (11) to compute the desired value L^+.

Comment.

As a result of this algorithm, we get functions A and B which minimize and maximize the expression (9c), and we have already shown that the resulting minimum L^- and maximum L^+ depend only on the copula. Thus, the *result* of applying this algorithm depends only on the copula – and do not depend on the marginal distributions.

However, since we start with *some* distribution $\rho(x,y)$ corresponding to the given copula, the conditional expectations computed on each iteration will be, in general, *different*. In other words, if we start with the distributions $F(x,y)$ corresponding to different marginal distributions $F_X(x)$ and $F_Y(y)$, then:

- on each iteration, we get *different* functions, but
- for all starting distributions (X,Y) corresponding to the same copula, in the limit (after all the iterations) we get functions $A(x)$ and $B(y)$ for which the distribution of the pair $(X',Y') = (A(X),B(Y))$ is *the same* – namely, the distribution which, among all distributions corresponding to the given copula, maximizes (or minimizes) the Pearson correlation coefficient $r(F)$.

Example.

To make sure that this algorithm makes sense, let us analyze what happens when we apply this algorithm to the standard case of two jointly distributed correlated Gaussian variables.

Let us start with the simplest initial functions $A^{(0)}(x) = x$ and $B^{(0)}(y) = y$. For these functions, the formulas (34) and (35) for computing the next iteration $A^{(1)}(x)$ and $B^{(1)}(y)$ take the form

$$A^{(1)}(x) = \frac{E[Y \,|\, X = x] - E[Y \,|\, X = 0]}{E[Y \,|\, X = 1] - E[Y \,|\, X = 0]}, \tag{37}$$

$$B^{(1)}(y) = \frac{E[X \,|\, Y = y] - E[X \,|\, Y = 0]}{E[X \,|\, Y = 1] - E[X \,|\, Y = 0]}. \tag{38}$$

It is know that when variables X and Y have a Gaussian joint distribution, then $E[Y \,|\, X = x]$ is a linear function of x, i.e.,

$$E[Y \,|\, X = x] = c_0 + c_1 \cdot x \tag{39}$$

for some constant c_0 and c_1. Substituting this expression (30) into the formula (37), we get

$$A^{(1)}(x) = \frac{(c_0 + c_1 \cdot x) - (c_0 + c_1 \cdot 0)}{(c_0 + c_1 \cdot 1) - (c_0 + c_1 \cdot 0)} = \frac{c_1 \cdot x}{c_1} = x. \tag{40}$$

Similarly, we get $B^{(1)}(y) = y$.

Here, we have $A^{(1)}(x) = A^{(0)}(x)$ and $B^{(1)}(y) = B^{(0)}(y)$ for all x ad y, so we stop iterations, and take $A(x) = A^{(1)}(x) = x$ and $B(y) = B^{(1)}(y) = y$. For these functions $A(x) = x$ and $B(y) = y$, the expression (11) becomes the usual expression (3) for the Pearson's correlation coefficient $r(F)$. So, for the usual Gaussian case, the above algorithm converges and leads to the desired result.

Important mathematical subtleties.

1°. There are cases when the above algorithm – and even the definition (9) – do not lead to the desired result.

For example, if $Y = X$ when $X \geq 0$ and $Y = X - Z^2$ for $X < 0$, where Z is a random variable which is independent of X, then the maximum in (9) is attained when we take $A(x) = x$ for $x \geq 0$, $A(x) = 0$ for $x \leq 0$, and similarly, $B(y) = y$ for $y \geq 0$ and $B(y) = 0$.

For these functions $A(x)$ and $B(y)$, we have $A(X) = B(Y)$ and thus, $L(A,B) = 1$. This value seems to indicate that X and Y are perfectly correlated, but in reality, they are only correlated when $X \geq 0$ and $Y \geq 0$ and they are definitely not well correlated when $X < 0$ and $Y < 0$.

This counterintuitive feature of the definition (9) appeared because we allowed functions $A(x)$ and $B(y)$ which are constant on some intervals. To avoid this counterintuitive feature, it is therefore reasonable to make sure that functions $A(x)$ and $B(y)$ are never constant. The functions $A(x)$ and $B(y)$ are supposed to be non-decreasing. Non-decreasing means that the derivative is non-negative, while constant means derivative is 0. Thus, it makes sense to select a small positive number $\delta > 0$ and, in the definition (9), only consider functions for which $A'(x) \geq \delta$ and $B'(y) \geq \delta$ for all x and y.

2°. Another important issue is the existence of the functions $A(x)$ and $B(y)$ which maximize $L(A,B)$. In general, a continuous function is guaranteed to attain its maximum value on a given domain D only if this domain is *compact*. A known Ascoli-Arzela theorem states that a compact class of functions should be uniformly continuous; for smooth functions, this means that there should be an upper bound M on the derivatives, such that $A'(x) \leq M$ and $B'(y) \leq M$ for all x and y.

3°. Because of Comments 1 and 2, it makes sense to fix two positive real numbers $\delta < M$ and to restrict ourselves only to functions $A(x)$ and $B(y)$ for which $\delta \leq A'(x) \leq M$ and $\delta \leq B'(y) \leq M$.

5 Case of Heavy-Tailed Distribution: Second Related Problem

Need to go beyond Pearson's correlation coefficient.

Pearson's correlation coefficient $r(F)$, as defined by the formula (3), implicitly assumes that the marginal distributions for X and Y have finite variance. In reality, however, many econometric-related distributions are *heavy-tailed*, with infinite variance. Let us show how we can extend the above definitions to the heavy-tailed case. For that, we first need to briefly recall the need for heavy-tailed distributions.

Heavy-tailed distributions are ubiquitous.

In many practical situations, e.g., in economics and finance, we encounter heavy-tailed probability distributions, i.e., distributions for which the variance is infinite. These distributions surfaced in the 1960s, when Benoit Mandelbrot, the author of fractal theory, empirically studied the fluctuations and showed [12] that larger-scale fluctuations follow the power-law distribution, with the probability density function $\rho(y) = A \cdot y^{-\alpha}$, for some constant $\alpha \approx 2.7$. For this distribution, variance is infinite.

The above empirical result, together with similar empirical discovery of heavy-tailed laws in other application areas, has led to the formulation of *fractal theory*; see, e.g., [13, 14].

Since then, similar heavy-tailed distributions have been empirically found in other financial situations [2, 3, 4, 16, 22, 23], and in many other application areas [1, 8, 13, 15, 20].

Utility: reminder.

People's economic behavior is determined by their preferences. A standard way to describe preferences of a decision maker is to use the notion of *utility* u; see, e.g., [5, 11, 19]. According to decision theory, a user prefers an alternative for which the expected value $\sum_{i=1}^{n} p_i \cdot u_i$ of the utility is the largest possible. Alternative, we can say that the expected value $\sum_{i=1}^{n} p_i \cdot U_i$ of the *disutility* $U \stackrel{\text{def}}{=} -u$ is the smallest possible.

Disutility caused by probabilistic uncertainty.

If we know the exact value of a quantity, then we can make an optimal decision based on this value. If we do not know the exact value – e.g., if we only know the probability distribution $\rho(y)$ on the set of all possible values – then we have to make a decision based on *some* value m. Since the actual value y is, in general, different from m, this decision is not as perfect as the decision based on the exact knowledge y.

For example, if we knew exactly what will be the future price y of a certain financial instrument (e.g., stock), then (after applying an appropriate future-related discount), we will be able to find the exact price that we are willing to pay for this instrument. In practice, we do not know this future price; at best, we know the

probability of future value. As a result, we set up a price corresponding to some "expected" value m.

- If the actual value y is smaller than our prediction m, then we overpay and thus, lose money on this transaction.
- If the actual value y is larger than m, this means that we may have missed an opportunity to invest in this instrument.

In both cases, the difference between the actual value x and the selected value m leads to disutility.

Let $U(d)$ denote the disutility caused by the difference $d = y - m$. When the value m has been selected, the average disutility is equal to $\int U(y-m)\cdot\rho(y)\,dy$. We select the value m for which this disutility is the smallest possible. The resulting minimal disutility is the disutility caused by the probabilistic uncertainty:

$$d_U(X) \stackrel{\text{def}}{=} \min_m E[U(Y-m)] = \min_m \int U(y-m)\cdot\rho(y)\,dy. \tag{41}$$

What if y partly depends on a known quantity x?

If the desired quantity y is somewhat dependent on another (known) quantity x, then, once we know x, we thus have more knowledge about y and hence, our uncertainty-caused disutility will decrease.

It is reasonable to take the percentage of this decrease as a measure of dependence between x and y.

Case of linear dependence.

In this paper, we are interested in the case of linear dependence $y = c_0 + c_1 \cdot x$. A linear dependence is either increasing or decreasing.

If we expect the dependence to be increasing, then it makes sense to consider dependencies with $c_1 \geq 0$. Among all such dependencies, we should select the values c_0 and $c_1 \geq 0$ for which the expected disuility $E[U(Y-(c_0+c_1\cdot X)]$ is the smallest possible. The resulting remaining disutility is equal to

$$d_U^+(Y\,|\,X) = \min_{c_0;c_1\geq 0} E[U(Y-(c_0+c_1\cdot X)] =$$

$$\min_{c_0;c_1\geq 0} \int U(y-(c_0+c_1\cdot x))\cdot\rho(x,y)\,dx\,dy. \tag{42}$$

The corresponding decrease $D_U^+(Y\,|\,X)$ in disutility can be thus estimated as

$$D_U^+(Y\,|\,X) \stackrel{\text{def}}{=} \frac{d_U(Y) - d_U^+(Y\,|\,X)}{d(Y)}. \tag{43}$$

Similarly, if we expect the dependence of y on x to be decreasing, we should consider dependencies with $c_1 \leq 0$. Among all such dependencies, we should also

select the values c_0 and $c_1 \le 0$ for which the expected disuility $E[U(Y-(c_0+c_1\cdot X)]$ is the smallest possible. The resulting remaining disutility is equal to

$$d_U^-(Y\,|\,X) = \min_{c_0;c_1\le 0} E[U(Y-(c_0+c_1\cdot X)] =$$

$$\min_{c_0;c_1\le 0} \int U(y-(c_0+c_1\cdot x))\cdot \rho(x,y)\,dx\,dy. \tag{44}$$

The corresponding decrease $D_U^-(Y\,|\,X)$ in disutility can be thus estimated as

$$D_U^-(Y\,|\,X) \stackrel{\text{def}}{=} \frac{d_U(Y)-d_U^-(Y\,|\,X)}{d_U(Y)}. \tag{45}$$

How is this idea related to Pearson's correlation coefficient?

It turns out that the Pearson's correlation coefficient $r(F)$ corresponds to the quadratic disutility function $U(d)=d^2$. Specifically, for the case when $U(d)=d^2$, as one can easily check:

- the optimal value m is the mean of the random variable Y: $m=E[Y]$;
- the corresponding value $d_U(Y)$ is equal to the variance $V(Y)$;
- for $r(F)\ge 0$, the decrease $D_U^+(Y\,|\,X)$ is equal to $R^2=(r(F))^2$; and
- for $r(F)\le 0$, the decrease $D_U^-(Y\,|\,X)$ is equal to $R^2=(r(F))^2$.

How to modify the above definition so that it depends only on the copula.

Let us assume that we have a copula $C(a,b)$ and a disutility function $U(d)$. We can then define the corresponding measures of linearity L^- and L^+ as the maximum, correspondingly, of the expression $D_U^-(Y\,|\,X)$ or of the expression $D_U^+(Y\,|\,X)$ over all possible probability distributions $F(x,y)=C(F_X(x),F_Y(y))$ corresponding to the given copula $C(a,b)$.

This definition clearly depends only on the copula (and not on the marginal distributions).

An easier-to-compute equivalent reformulation.

Similarly to the case of the Pearson's correlation coefficient, we can show that the above definitions can be reformulated in an easier-to-compute equivalent form. Namely, for a joint distribution of two random variables X and Y, the above measures of linearity L_U^- and L_U^+ can be equivalently defined as

$$L^- = \max_{A(x),B(y)} D_U^-(B(Y)\,|\,A(X)),\ \ L^+ = \max_{A(x),B(y)} D_U^+(B(Y)\,|\,A(X)), \tag{46}$$

where maximum is taken over all possible non-decreasing functions $A(x)$ an $B(y)$, and the values $D_U^{\pm}$ are defined by the formulas (41)–(45).

Acknowledgments. This work was supported in part by the National Science Foundation grants HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and DUE-0926721, by Grants 1 T36 GM078000-01 and 1R43TR000173-01 from the National Institutes of Health, and by a grant N62909-12-1-7039 from the Office of Naval Research.

References

1. Beirlant, J., Goegevuer, Y., Teugels, J., Segers, J.: Statistics of Extremes: Theory and Applications. Wiley, Chichester (2004)
2. Chakrabarti, B.K., Chakraborti, A., Chatterjee, A.: Econophysics and Sociophysics: Trends and Perspectives. Wiley-VCH, Berlin (2006)
3. Chatterjee, A., Yarlagadda, S., Chakrabarti, B.K.: Econophysics of Wealth Distributions. Springer, Italia (2005)
4. Farmer, J.D., Lux, T. (eds.): Applications of statistical physics in economics and finance. A Special Issue of the Journal of Economic Dynamics and Control 32(1), 1–320 (2008)
5. Fishburn, P.C.: Utility Theory for Decision Making. John Wiley & Sons Inc., New York (1969)
6. Gabaix, X., Parameswaran, G., Vasiliki, P., Stanley, H.E.: Understanding the cubic and half-cubic laws of financial fluctuations. Physica A 324, 1–5 (2003)
7. Gelfand, I.M., Fomin, S.V.: Calculus of Variations. Dover, New York (2000)
8. Gomez, C.P., Shmoys, D.B.: Approximations and Randomization to Boost CSP Techniques. Annals of Operations Research 130, 117–141 (2004)
9. Jaworski, P., Durante, F., Härdle, W.K., Ruchlik, T. (eds.): Copula Theory and Its Applications. Springer, Heidelberg (2010)
10. Kreinovich, V., Lakeyev, A., Rohn, J., Kahl, P.: Computational Complexity and Feasibility of Data Processing and Interval Computations. Kluwer, Dordrecht (1997)
11. Luce, R.D., Raiffa, R.: Games and Decisions: Introduction and Critical Survey. Dover, New York (1989)
12. Mandelbrot, B.: The variation of certain speculative prices. J. Business 36, 394–419 (1963)
13. Mandelbrot, B.: The Fractal Geometry of Nature. Freeman, San Francisco (1983)
14. Mandelbrot, B., Hudson, R.L.: The (Mis)behavior of Markets: A Fractal View of Financial Turbulence. Basic Books (2006)
15. Markovich, N. (ed.): Nonparametric Analysis of Univariate Heavy-Tailed Data: Research and Practice. Wiley, Chichester (2007)
16. McCauley, J.: Dynamics of Markets, Econophysics and Finance. Cambridge University Press, Cambridge (2004)
17. McNeil, A.J., Frey, R., Embrechts, P.: Quantitative Risk Management: Concepts, Techniques, Tools. Princeton University Press, Princeton (2005)
18. Nelsen, R.B.: An Introduction to Copulas. Springer, Heidelberg (1999)
19. Raiffa, H.: Decision Analysis. Addison-Wesley, Reading (1970)
20. Resnick, S.I.: Heavy-Tail Phenomena: Probabilistic and Statistical Modeling. Springer, New York (2007)
21. Sheskin, D.J.: Handbook of Parametric and Nonparametric Statistical Procedures. Chapman and Hall/CRC, Boca Raton, Florida (2011)

22. Stoyanov, S.V., Racheva-Iotova, B., Rachev, S.T., Fabozzi, F.J.: Stochastic models for risk estimation in volatile markets: a survey. Annals of Operations Research 176, 293–309 (2010)
23. Vasiliki, P., Stanley, H.E.: Stock return distributions: tests of scaling and universality from three distinct stock markets. Physical Review E: Statistical, Nonlinear, and Soft Matter Physics 77(3, pt. 2). Publ. 037101 (2008)

An Innovative Financial Time Series Model: The Geometric Process Model

Jennifer S.K. Chan, Connie P.Y. Lam, and S.T. Boris Choy

Abstract. Geometric Process (GP) model is proposed as an alternative model for financial time series. The model contains two components: the mean of an underlying renewal process and the ratio which measures the direction and strength of the dynamic trend pattern over time. They simultaneously account for the uncertainty on the mean and the autoregressive and time-varying nature of the volatility. Compare to the popular GARCH and SV models, this model is simple and easy to implement using the least squares (LS) method. We extend the GP model to analyze the daily asset price range which exhibit threshold and asymmetric effects for some exogenous variables. Models are selected according to mean square error (MSE). Finally forecasting are performed for the best model that allows for both threshold and asymmetric effects.

1 Introduction

In asset price series, an important feature is the time-varying variance called heteroskedasticity. A related and important measure of heteroskedasticity is volatility which refers to the variation of price over time. Volatility has become a standard risk measure in financial markets. Two main classes of models have been proposed to capture the dynamics of the volatility precisely, and they are the generalised autoregressive conditional heteroskedastic (GARCH) models (Bollerslev, 1986) and the stochastic volatility (SV) models (Hull and White, 1987).

These models usually employ squared close to close asset returns to model return volatility. However many papers have shown the intra-day range to be a far more efficient measure of return volatility, e.g. see Parkinson [20], Garman and Klass [12] and recently Anderson and Bollerslev [1] and Chen *et al.* [8]. The daily range has

Jennifer S.K. Chan · Connie P.Y. Lam · S.T. Boris Choy
School of Mathematics and Statistics,
The University of Sydney, NSW 2006, Australia
e-mail: jchan@maths.usyd.edu.au

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_5,

been used in volatility models: Alizadeh et al. (2002) incorporated the range with the stochastic volatility model and Chou [9], Brandt and Jones [3], Chen *et al.* [8] proposed range-based heteroskedastic models, using observed range data and a link between the range and volatility. To capture the trend dynamic in range data, Chan *et al.* [7] extended the conditional autoregressive range (CARR) model of Chou [9] to incorporate the Geometric Process (GP) model of Lam [16].

Lam [16] first proposed to model a monotone trend directly by a monotone process called the GP. A GP is related to a renewal process (RP) (Feller [11]) which is an arrival process with event count $N(t)$ and independent and identically distributed interarrival intervals $Y_i > 0$ such that $N(t) = n$ if $\sum_{t=1}^{n} Y_i \leq t < \sum_{i=1}^{n+1} Y_i$. A sequence of positive random variables $X_1, X_2, \ldots$ forms a GP if there exists a positive real number a such that $\{Y_t = a^{t-1}X_t,\ t = 1,2,\ldots\}$ is a RP. The real number a is called the ratio of the GP. Note that the latent RP $\{Y_t\}$ is stationary and the observed GP $\{X_t\}$ is increasing if $a < 1$, decreasing if $a > 1$ and stationary if $a = 1$. If we denote the mean and variance for the latent RP $\{Y_t\}$ to be μ and σ^2 respectively, the mean and variance for the observed data $\{X_t\}$ are given by

$$E(X_t) = \mu/a^{t-1} \quad \text{and} \quad Var(X_t) = \sigma^2/a^{2(t-1)} \tag{1}$$

respectively. The GP model identifies effects on the trend movement by two components: the mean μ of the underlying RP and the ratio a which measures the direction and strength of the trend dynamic.

The original GP model mainly focused on modeling the interarrival time of a series of events in reliability and maintenance problem in system engineering (Lam [19]). Chan et al. [5] first extended the application of GP model to health science by applying threshold GP models to describe multiple trends at different stages of development for the Severe Acute Respiratory Syndrome (SARS) epidemic in 2003. More applications of GP model on Poisson count times series in clinical trials can be found in Wan and Chan ([24], [25]) in the analyses of bladder cancer data. The GP model was also extended to binary data in Chan and Leung [6] with an application to methadone clinic data. This paper advances the application of GP model to finance.

The GP model has several advantages over the popular financial time series models such as GARCH, defined as

$$X_t = \mu_t + \varepsilon_t$$

where $\varepsilon_t = \sigma_t z_t$,

$$\sigma_t^2 = \alpha + \sum_{i=1}^{p} \beta_i \varepsilon_{t-i}^2 + \sum_{i=1}^{q} \gamma_i \sigma_{t-i}^2 \tag{2}$$

and z_t is a white noise process. Firstly, a GARCH model applies to return data and treats volatility as an unobserved process. Hence the modelling of volatility is less efficient due to the lack of volatility information. On the other hand, a GP model describes directly the intra-day range which has been shown in many researches to be a far more efficient measure of return volatility as it utilizes two pieces of

information (the high and low prices) rather than just the closing price. Secondly, a GARCH model needs a separate volatility equation in (2) to allow for heteroskedasticity whereas equation (1) for a GP model shows that both of its mean and variance change with a and hence it allows for heteroskedasticity without a separate model.

Thirdly, financial time series often exhibit volatility clustering. Volatility clustering has periods of elevated volatility interspersed among more tranquil periods. This effect often produces distinct trend patterns that should be accommodated in a model. The ratio parameter a of a GP model can model the trend movement directly and gives a straightforward interpretation for the progression of trends. Moreover the two components, the mean μ and ratio a, can be modeled separately and hence the model distinguishes effects on the underlying stationary process Y_t from effects on the strength and direction of trend movement. Lastly with the inherent geometric structure, forecast of volatility using $E(X_t)$ in equation (1) is simple and straightforward whereas it relies on the more complicated volatility equation in (2) condition on previous unobserved ε_{t-i} and volatility σ_{t-i} in the GARCH model.

The objective of this paper is to extend the GP model to capture many important features in the intra-day price range of some stock markets. Firstly, we include covariate effects in the mean μ and ratio a parameters to describe the dynamics of mean and variance in equation (1). Moreover as the price range may be subject to abrupt and unanticipated asymmetric effects from certain variables, the model is further extended to incorporate regime switching to allow model change after some threshold times as well as when the outcomes exceed certain threshold levels. We show that the extended GP model provides a simple analytical tool for analyzing daily price range series.

For model implementation, Lam [17] first proposed the non-parametric (NP) least squares (LS) approach. Chan et al. [5] applied the LS and log-LS approaches to analyze the SARS epidemic data in 2003. By adopting a lifetime distribution to the underlying RP Y_i, the model can be implemented by a parametric approach. Lam and Chan [18] investigated the statistical inference and properties of the maximum likelihood (ML) estimators with lognormal distribution and Chan et al. [4] considered gamma distribution. Wan and Chan ([24], [25]) adopted the Bayesian approach and Chan and Leung [6] compared all three LSE, ML and Bayesian approaches for the binary GP model. Although parametric inference has been a popular choice of inference, it sometimes fails for complicated model especially when data deviate considerably from the distributional assumptions. Nonparametric inference is released from distribution assumption and hence offers an attractive choice of model implementation. Despite nonparametric inference may be less efficient than parametric inference, such disadvantage will be lessen if the data size is large enough, usually the case for financial time series. We adopt the LS approach and show how the model can be easily implemented under this approach.

The paper is presented as follows. Section 2 introduces the GP model and its extension to capture covariate, multiple trends and asymmetric effects respectively. Section 3 describes the method of inference. Then the models are fitted to four intra-day asset price range data in Section 4 and finally, a conclusion is given in Section 5.

2 The GP Model and Its Extensions

2.1 Extension to Covariate Effect

Original GP model adopts a constant mean and a constant ratio over time. The adoption of a homogeneous mean over time is over-simplified in many cases. For example, business cycle and seasonal effect commonly appear in financial time series. As the outcome measures may evolve over time subject to different internal and external effects, covariates should be incorporated into the mean μ and ratio a of the GP model to allow for these effects. The resulting mean and ratio become time-varying mean μ_t and ratio a_t functions respectively.

Adopting the framework of linear models (LM), the mean function is linked to a linear function $\eta_{\mu t}$ of p covariates $z_{tk},\ k = 1,\dots p$ using a log link function. In subsequent analysis, the model adopts one covariate z_t ($p = 1$) and hence the time-varying mean function as log-linked to this covariate is defined as

$$\mu_t = \exp(\eta_{\mu t}) = \exp(\beta_{\mu 0} + \beta_{\mu 1} z_t). \tag{3}$$

Moreover the ratio a which models the direction and strength of the movement can change gradually over time t. The time effects can be modeled by different functional forms, for example, t or $\ln t$. Other covariates can also be included to allow for their effects on the trend movement. The extended model is called the adaptive GP (AGP) model because the mean μ_t and ratio a_t functions adapt to changes in covariates and hence the model can adaptively model the progression of trend movement over time. In subsequent analysis, we consider either

$$\begin{array}{ll} \text{a constant ratio (C):} & a_t = \exp(\eta_{at}) = \exp(\beta_{a0}), \text{ or} \\ \text{a time-evolving ratio (R):} & a_t = \exp(\eta_{at}) = \exp(\beta_{a0} + \beta_{a1} \ln t). \end{array} \tag{4}$$

Since a_t affects both the mean and variance of X_t, the variance of the resulting GP model will change over time with different volatility levels. The vector of parameters is $\beta = (\beta_\mu^T, \beta_a^T)$ where $\beta_\mu = (\beta_{\mu 0}, \beta_{\mu 1})^T$ and $\beta_a = (\beta_{a0})$ or $(\beta_{a0}, \beta_{a1})^T$ for a constant or time-evolving ratio function respectively. Note that the latent process $\{Y_t\}$ is now a stochastic process (SP) in general.

2.2 Extension to Threshold Time

Persistent changes may occur in a financial time series when some external or structural factors take place from certain time points T called the threshold times. Chan et al. [5] extended the GP model to the threshold GP (TGP) model by fitting a separate GP to each stage, growing, stabilizing and declining, of an epidemic as identified by turning points using the LS method of inference.

Let $\mathscr{T}_m,\ m = 1,\dots,M$ be the thresholds time for the m-th GP. When $\mathscr{T}_m \le t < \mathscr{T}_{m+1}$, the mean and variance for X_t are

$$E(X_t) = \mu_{tm}/a_t m^{t-\mathscr{T}_m} \quad \text{and} \quad Var(X_t) = \sigma_m^2/a_t m^{2(t-\mathscr{T}_m)}, \tag{5}$$

where

$$\mu_{tm} = \exp(\eta_{\mu tm}) = \exp(\beta_{\mu 0m} + \beta_{\mu 1m} z_t), \tag{6}$$

$$a_{tm} = \exp(\eta_{atm}) = \exp(\beta_{a0m}), \tag{7}$$

$\mathscr{T}_1 = 1$, $\mathscr{T}_m = 1 + \sum_{j=1}^{m-1} n_j$, $m = 2, \ldots, M$, n_m is the number of observations for the m-th GP, and $\sum_{m=1}^{M} n_m = n$. The extended model is called the threshold time GP (TTGP) model.

To estimate the threshold times when trends change their movements, Chan et al. [5] proposed a moving window technique in which separate GP model is applied to each subset of data of fixed length L starting from time $i = 1$, $i = 2$ and so on up to $i = n - L + 1$. Since the ratio a of a GP changes according to the moving windows, threshold times $\mathscr{T}_m$ can be located when a changes from "less than 1" to "greater than 1" or vice versa. As different window widths L produce different sets of parameter estimates and threshold times, an optimal L is selected from a range which gives the least Adjusted Mean Square Error ($AMSE$). The penalized term in $AMSE$ is a scalar multiple of the number of parameters. Although this method may detect several threshold times simultaneously, it is computationally intensive as both the window width L and the window $(i, i+L-1)$ have to vary in detecting the threshold times $\mathscr{T}_m$.

A more direct way is to estimate the threshold times condition on M. We first set $M = 2$ and search $\mathscr{T}_2$ over certain interval not too close to the end points 1 and n. Condition on each threshold time $\mathscr{T}_2$, two GP models are fitted to data with time $t < \mathscr{T}_2$ and $t \geq \mathscr{T}_2$ respectively. Optimal $\mathscr{T}_2$ is chosen to minimize the MSE in (15). Then M is set to 3 and the search for $\mathscr{T}_3$ given $\mathscr{T}_2$ is similarly repeated from the remaining time points not too close to $\mathscr{T}_2$, 1 and n. This method is essentially a partial LS method where β is obtained by LS method but $\{\mathscr{T}_m\}$ by searching. The number of threshold times M can be chosen by some model selection criteria, say the cross validation (CV).

2.3 Extension to Threshold Outcome

Financial time series are sometimes subject to abrupt and unanticipated asymmetric effects due to a certain risk variable. We assume that a temporary model shift occur when an observable lag-d risk variable, W_{t-d}, $t = d+1, \ldots, n$, exceeds certain threshold levels. Let $\mathscr{W}_h$, $h = 1, \ldots, H$ be the latent threshold levels for W_{t-d}. When $\mathscr{W}_h \leq w_{t-d} < \mathscr{W}_{h+1}$, the mean and variance for X_t are

$$E(X_t) = \mu_{th}/a_{th}^{t-1} \quad \text{and} \quad Var(X_t) = \sigma_h^2/a_{th}^{2(t-1)}, \tag{8}$$

where

$$\mu_{th} = \exp(\eta_{\mu th}) = \exp(\beta_{\mu 0h} + \beta_{\mu 1h} z_t), \tag{9}$$

$$a_{th} = \exp(\eta_{ath}) = \exp(\beta_{a0h} + \beta_{a1h} \ln t), \tag{10}$$

$\mathscr{W}_1 = 0$ or $-\infty$ depending on whether W_t is positive continuous or continuous and $\mathscr{W}_{H+1} = \infty$. The extended model is called the threshold level GP (TLGP) model. The risk variable W_t is preferable to be positively related to X_t and contains some market information. Examples include lagged values of X_t, or lagged values of an exogenous factor, such as international market movement, a financial index or interest rates.

To estimate the threshold levels $\mathscr{W}_h$, we first set $H = 2$. The search for threshold level $\mathscr{W}_2$ for the risk variable W_t is similar to that of the threshold time model. Conditioning on $d = 1$, we search $\mathscr{W}_2$ over certain interval which is approximately the median to the 95 percentile of W_t because $\mathscr{W}_2$ which indicates a model shift should be large in general. For each threshold value $\mathscr{W}_2$ in the interval, two GP models are fitted to data x_t with $w_{t-d} < \mathscr{W}_2$ and $w_{t-d} \geq \mathscr{W}_2$ respectively. Then we set $d = 2$ and fit two GP models for each $\mathscr{W}_2$ similarly. Optimal $\mathscr{W}_2$ is chosen to minimize the *MSE* in (15) over a range of d. Then we may set $H = 3$ and repeat the search for $\mathscr{W}_3$ again. This method is as well a partial LSE method where β is obtained by LS method but $\{\mathscr{W}_h\}$ by searching.

3 Methodology of Inference

The least squares (LS) method is perhaps the simplest method in parameter estimation. Lam [17] considered the LS method on $\ln X_t$ and Chan et al. [5] adopted the LS method on both $\ln X_t$ and X_t. In this paper, we adopt the LS method on X_t by minimizing the sum of squared errors *SSE* given by

$$SSE = \sum_{t=1}^{n} [X_t - E(X_t)]^2 \tag{11}$$

where $E(X_t)$, for the AGP model with constant (C) and time-varying (R) ratio, the TTGP model with threshold times (T) and the TLGP model with threshold levels (L) are given by (1), (5) and (8) respectively, the mean functions μ_t by (3), (6) and (9) respectively and the ratio functions by (4), (7) and (10) respectively.

To solve for the parameter estimates β that minimize the *SSE* in the score equation $SSE'(\beta) = 0$, we use the Newton Raphson (NR) iterative procedure. In each NR iteration, current parameter estimates $\beta^{(v)}$ in the v-th iteration are updated to $\beta^{(v+1)}$ in the $(v+1)$-th iteration by

$$\beta^{(v+1)} = \beta^{(v)} - [SSE''(\beta^{(v)})]^{-1} SSE'(\beta^{(v)}) \tag{12}$$

and the procedure continues until $\| \beta^{(v+1)} - \beta^{(v)} \|$ is sufficiently small. Then the LS estimates are given by $\widehat{\beta}_{LSE} = \beta^{(v+1)}$. The first and second order derivatives as

required in the NR procedure are given by vector $SSE'(\beta^{(v)})$ and matrix $SSE''(\beta^{(v)})$ respectively with elements

$$\frac{\partial SSE}{\partial \beta_{jkm}} = -2\sum_{t=1}^{n}(x_t - \hat{x}_t)\,\hat{x}_t\,z'_{jkt} \tag{13}$$

$$\frac{\partial^2 SSE}{\partial \beta_{j_1k_1m_1}\partial \beta_{j_2k_2m_2}} = -2\sum_{t=1}^{n} z'_{j_1k_1t}\,z'_{j_2k_2t}\,\hat{x}_t\,(x_t - 2\hat{x}_t) \tag{14}$$

where $\hat{x}_t = \widehat{E}(X_t|X_t = x_t, \beta = \beta^{(v)})$, $j, j_1, j_2 = \mu, a$; $k, k_1, k_2 = 0, 1$; $m, m_1, m_2 = 1, \ldots, G$ ($G = 1$ for AGP model, $G = M$ for TTGP model and $G = H$ for TLGP model); $z'_{\mu 0t} = 1$, $z'_{\mu 1t} = z_t$, $z'_{a0t} = -(t-1)$ and $z'_{a1t} = -(t-1)\ln t$. Standard error estimates are given by the square root of the diagonal elements in the inverse of the second order derivative matrix $SSE''(\beta^{(v+1)})$.

Estimates for σ^2, σ_m^2 and σ_h^2 in (1), (5) and (8) respectively are given by the mean sum of squared residuals $\frac{1}{n_l}\sum_{t=T_l}^{T_l+n_l-1}(x_t - \hat{x}_t)^2$ where n_l denote the data size in the l-th GP when $\mathscr{T}_l \le t < \mathscr{T}_{l+1}$ for the TTGP model or $\mathscr{W}_l \le w_{t-d} < \mathscr{W}_{l+1}$ for the TLGP model. The method and standard error calculation can be easily implemented in R.

4 Empirical Study

4.1 The Intra-day Range Data

We analyze the intra-day high-low prices from four stock markets, obtained from the website "finance.yahoo.com". The data is collected from January 1, 2000 to December 31, 2006 and it includes four Asia-Pacific Economic Cooperation (APEC) financial markets which are in order Nikkei 225 Index (N225, Japan), Hang Seng Index (HSI, Hong Kong), All Ordinaries Index (AORD, Australia) and Taiwan weighted index (TWII, Taiwan). The variable of interest is the daily range which is the differences between the log of the daily maximum $x_{max,t}$ and minimum $x_{min,t}$ indice defined as

$$x_t = [\ln(x_{max,t}) - \ln(x_{min,t})] \times 100.$$

Figures 1 to 4 show that the data exhibit different trend patterns. Volatility clustering is also one of salient features about daily range data. Hence each daily range data is fitted to four proposed GP models, namely, the AGP model with constant (C) and time-varying (R) ratio, the TTGP (T) model and the TLGP (L) model that allow for these characteristics in the data.

4.2 Numerical Results

Experience show that the daily range z'_t of the Standard & Poors 500 (US) with delay $d = 1$ ($z_t = z'_{t-1}$) is a significant covariate (correlation coefficient r ranges from 0.33 to 0.36 for the four regions). Hence it is included in the mean functions (3), (6) and

(9) for the AGP, TTGP and TLGP models respectively. For the ratio function, we adopt both constant (C) and time-varying (R) function in (4) for the AGP model, constant in (7) for the TTGP model and time-varying in (10) for the TLGP model.

For the threshold models, we set the number of threshold times and threshold levels to be both $M = H = 2$. For the TLGP model, we choose the threshold variable to be X_t and search the lag d from 1 to 5 because autocorrelation often drops with increasing lag so that the optimal lag will not be very large. The lag 1-5 autocorrelations r_j for X_t are (0.38,0.36,0.38,0.31,0.33), (0.46,0.47,0.48,0.45,0.45), (0.37,0.35,0.31,0.33,0.25) and (0.48,0.44,0.46,0.44,0.42) respectively for the four regions which are higher than the correlation between X_t and Z_{t-1}. Optimal lags are often related to the lag with high autocorrelation.

Tables 1 to 4 report parameter estimates, standard errors in parenthesis, and *MSEs* for the four daily range data fitted to the four proposed GP models. Significant parameters are indicated by '$*$'. Parameters $\beta_{\mu 1l}, l = 1,2$ indicate the effect of US daily range while $\beta_{akl}, k = 0,1, l = 1,2$ reveal distinct trend movements after allowing for the US daily range effect. These effects can be viewed in Figures 1(a),(b) to 4(a),(b) which plot the fitted daily range $\hat{x}_t$ over time for the TTGP and TLGP models and their trends are described as below.

The trend of the daily range of N225 shows a drop ($\beta_{a0} > 0$ in model C) over time in general. Specifically, it drops slowly before June 22, 2004 ($\beta_{a01} > 0,\ \mathscr{T}_2 = 1059$ in T) and increases faster from a lower level afterwards ($\beta_{\mu 02} < \beta_{\mu 01},\ \beta_{a02} < 0$ in T). With the lag d in TLGP model estimated to be 3, 7.6% of daily range ($x_{t-3} \geq 2.69$) are modeled by the upper-level model and they remain constant over time ($\beta_{a02},\ \beta_{a12}$ insignificant, $\mathscr{W}_2 = 2.69$ in L) while the rest of daily range increase at a decreasing rate and then decrease over time ($\beta_{a01} < 0,\ \beta_{a11} > 0$ in L). The US daily range has a positive and significant effect ($\beta_{\mu 1l} > 0$) for all the four models.

The general trend of the daily range of HSI is again a drop over time but it drops faster before Mar 18, 2003 and slower afterwards. The lag d is estimated to be 5 and 29% of daily range ($x_{t-5} \geq 1.6$) are modeled by the upper-level model. They drop at a decreasing rate over time while the rest also drop over time. The upper-level model starts at a higher level but approaches the lower-level model. The US daily range has a positive and significant effect for all the four models except after Mar 18, 2003 and when the lag-5 daily range is below 1.60.

The daily range of AORD also trend downward in general. In particular, they drop before Feb 14, 2005 but increase afterwards. The estimated lag d is 3 and 32% of daily range ($x_{t-3} \geq 1.33$) are modeled by the upper-level model. They remain constant over time while the rest drop at a decreasing rate. The US daily range has positive and significant effects for all the four models.

Lastly, the daily range of TWII also trend downward over time. They increase sharply till Nov 27, 2000 and drop afterwards. With a lag of one, about 18% of daily range lie above 2.30 and they remain constant over time while the rest increase at a decreasing rate and then drop. The US daily range has positive and significant effects for all the four models except after Nov 27, 2000 and when the lag one daily range is above 2.30.

4.3 Model Selection

A natural way to compare models is to use a criterion that is based on the trade-off between model fit and model complexity. For nonparametric inference, Chan et al. [5] adopted the Penalized Mean Squares of Error (*PMSE*) measure that adds to the *MSE* a term which penalizes the number of parameters as follow

$$PMSE = 2pk \ln\left(\sum_{t=1}^{n} x_t/n\right) + MSE$$

where

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(x_t - \hat{x}_t)^2, \tag{15}$$

the penalized term in *PMSE* accounts for the scale of measurement X_t, the size of data n and the number of parameters p estimated using the LSE method. The size of constant k adjusts for the level of penalty. However the value of k can be quite arbitrary and we set $k = 0$ to indicate that no penalty is applied.

Result show that *MSE*s drop consistently across the four models for all data. The TLGP model shows the best *MSE* and hence is chosen to be the best model. Result shows that a small portion of daily range should adopt a separate model with larger mean and variance in response to previous changes. Trends in mean and variance over time for the TLGP model can be viewed in Figures 1(b),(c) to 4(b)(c). Specifically, Figures 1(c) to 4(c) give the bounds for the lower and upper fitted daily range which are two standard deviation from the mean when the variance is calculated using (8). These intervals called the predictive intervals show a noncoverage of 4.9%, 18.9%, 12.3% and 7.4% respectively for the four regions. Their σ_l^2 estimates are (0.430,0.943), (0.232,0.588), (0.112,0.236) and (0.502,1.214) respectively. Note that the proportions need not be close to 5% because we do not have the normality assumption. Higher proportions indicate that the data are more volatile.

Obviously the upper-level model has a wider predictive interval than that of the lower-level model. However such difference often lessens over time as the daily range become less volatile over the measurement period. Hence the predictive intervals for both lower- and upper-level models become shorter over time. This is particularly the case for the daily range of N225 and TWII.

4.4 Forecasting

Forecasting is one of the main objectives for model fitting. Twenty day forecasts from Jan 2, 2007 are performed for the four regions using the chosen TLGP model. The *MSE*s during the forecasting period are 0.280, 0.453, 0.070 and 0.134 respectively and they are all less than the fitted *MSE*s in Tables 1-4 except the daily range of HSI which exhibit high volatility. The lower *MSE*s for regions other than HK are probably due to the low level of daily range during the forecasting period so

Table 1 Parameter estimates with standard errors in parenthesis and model assessment measures for the N225 data

	Threshold	Mean parameter		Ratio parameter		
	thres.	intercept	US.	intercept	time	
Model	$\mathscr{T}_2$ or $\mathscr{W}_2$	$\beta_{\mu 0l}$	$\beta_{\mu 1l}$	$\beta_{a0l} \times 10^3$	$\beta_{a1l} \times 10^3$	*MSE*
C		0.4432*	0.1297*	0.2654*		0.5003
		(0.0315)	(0.0103)	(0.0261)		
R		0.3066*	0.1205*	-2.2730*	0.3420*	0.4928
		(0.0434)	(0.0107)	(0.5252)	(0.0707)	
T	$t < 1059$	0.4218*	0.1137*	0.0947*		0.4751
		(0.0337)	(0.0109)	(0.0405)		
	$t \geq 1059$	-0.2977	0.2165*	-0.6694*		
		(0.2254)	(0.0653)	(0.1492)		
L	$x_{t-3} < 2.69$	0.2336*	0.1263*	-2.4985*	0.3680*	0.4688
		(0.0499)	(0.0132)	(0.5721)	(0.0768)	
	$x_{t-3} \geq 2.69$	0.8208*	0.0674*	0.7512	-0.0774	
		(0.0916)	(0.0202)	(1.1761)	(0.1598)	

* Parameter is significant at 5% significant level.

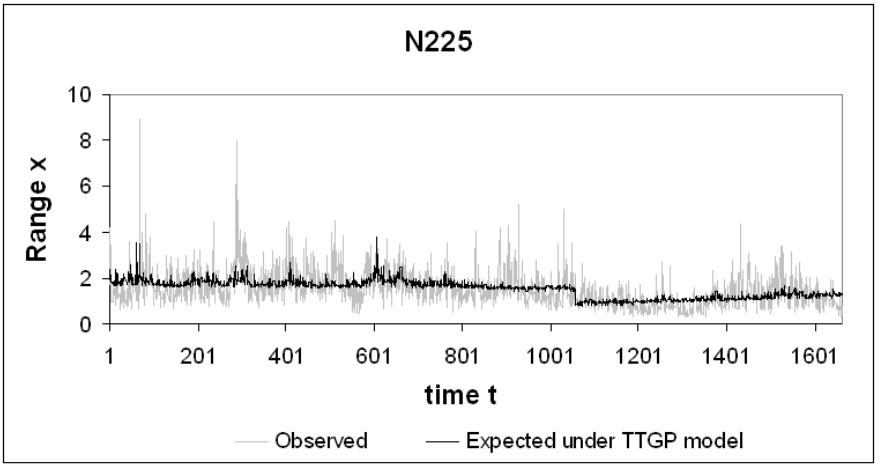


Fig. 1(a) Observed and fitted N225 daily range using TTGP model

that all data adopt the lower-level model. For the daily range of HSI, about 65% of data adopt the upper-level model. Noncoverage of the predictive intervals are 0.10, 0.50, 0.00 and 0.10 respectively for the four regions, showing satisfactory forecasting performance except the HSI daily range. Indeed, forecasting performance can be improved using a more predictive risk variable.

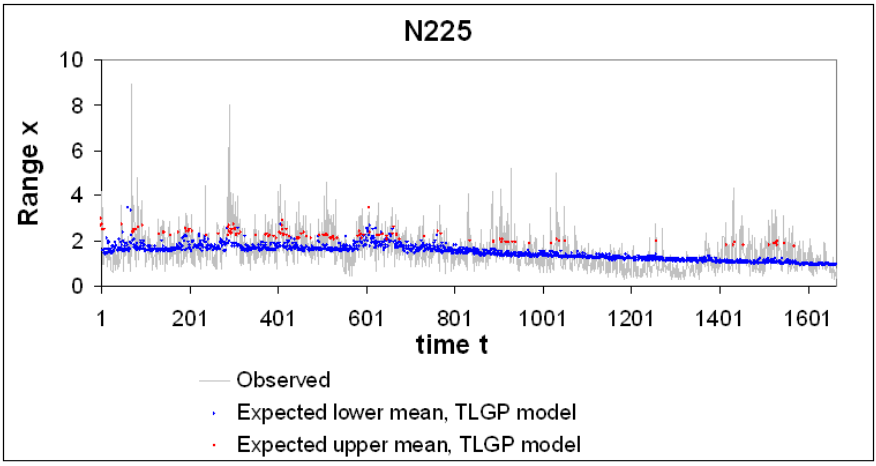


Fig. 1(b) Observed and fitted N225 daily range using TLGP model

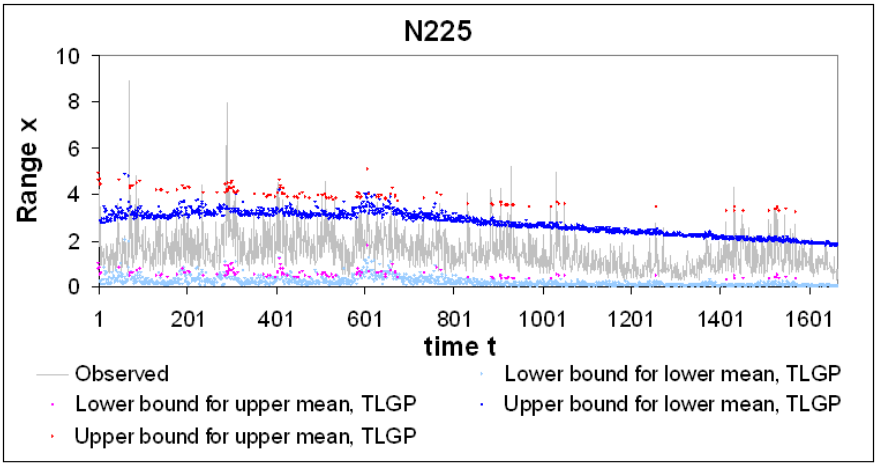


Fig. 1(c) Observed, fitted lower and upper N225 daily range within 2 sd using TLGP model

Table 2 Parameter estimates with standard errors in parenthesis and model assessment measures for the HSI data

	Threshold	Mean parameter		Ratio parameter		
	thres.	intercept	US.	intercept	time	
Model	$\mathscr{T}_2$ or $\mathscr{W}_2$	$\beta_{\mu 0l}$	$\beta_{\mu 1l}$	$\beta_{a0l} \times 10^3$	$\beta_{a1l} \times 10^3$	*MSE*
C		0.7278*	0.0572*	0.6684*		0.3619
		(0.0347)	(0.0133)	(0.0313)		
R		0.9063*	0.0696*	4.3482*	-0.4997*	0.3465
		(0.0409)	(0.0132)	(0.4947)	(0.0670)	
T	$t < 764$	0.7772*	0.0929*	1.1301*		0.3424
		(0.0358)	(0.0135)	(0.0686)		
	$t \geq 764$	0.3354*	-0.0192	0.6111*		
		(0.1383)	(0.0478)	(0.0957)		
L	$x_{t-5} < 1.60$	0.7421*	0.0132	1.9865*	-0.1933	0.3355
		(0.0841)	(0.0235)	(0.9027)	(0.1192)	
	$x_{t-5} \geq 1.60$	0.8950*	0.0967*	4.5496*	-0.5372*	
		(0.0513)	(0.0132)	(0.7749)	(0.1092)	

* Parameter is significant at 5% significant level.

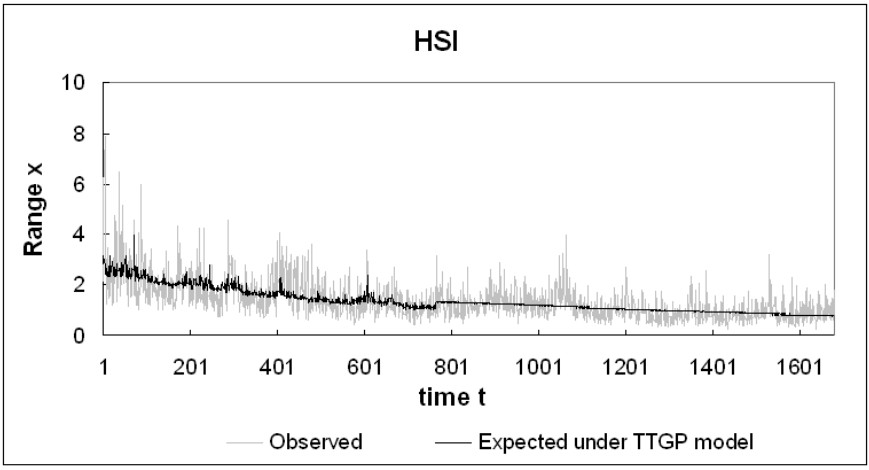


Fig. 2(a) Observed and fitted HSI daily range using TTGP model

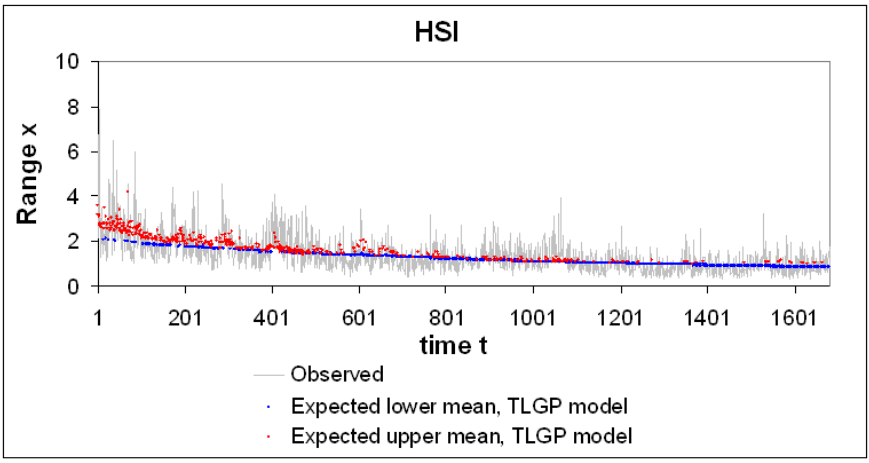


Fig. 2(b) Observed and fitted HSI daily range using TLGP model

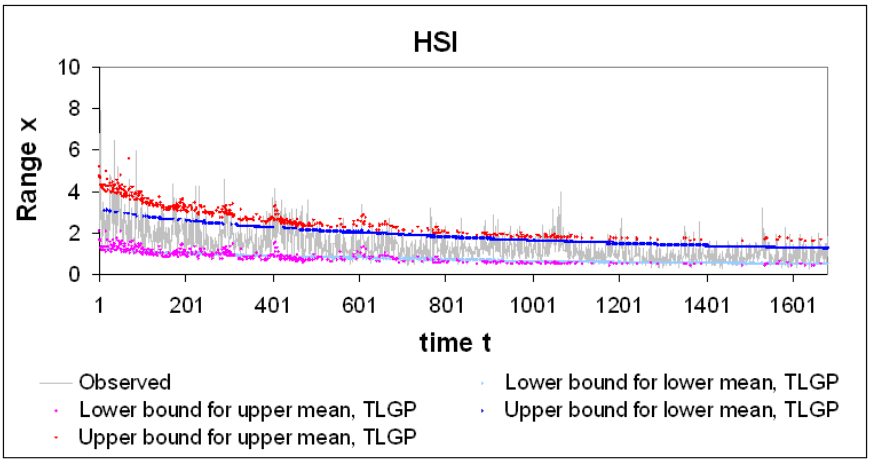


Fig. 2(c) Observed, fitted lower and upper HSI daily range within 2 sd using TTGP model

Table 3 Parameter estimates with standard errors in parenthesis and model assessment measures for the AORD data

	Threshold	Mean parameter		Ratio parameter		
	thres.	intercept	US.	intercept	time	
Model	$\mathscr{T}_2$ or $\mathscr{W}_2$	$\beta_{\mu 0l}$	$\beta_{\mu 1l}$	$\beta_{a0l} \times 10^3$	$\beta_{a1l} \times 10^3$	*MSE*
C		-0.5295*	0.2273*	0.0687		0.1463
		(0.0600)	(0.0148)	(0.0518)		
R		-0.2122*	0.2350*	5.2465*	-0.6899*	0.1357
		(0.0728)	(0.0144)	(0.8182)	(0.1089)	
T	$t < 1260$	-0.3708*	0.2247*	0.4329*		0.1315
		(0.0599)	(0.0151)	(0.0748)		
	$t \geq 1260$	-0.8151	0.4474*	-0.6219*		
		(0.4974)	(0.0964)	(0.3168)		
L	$x_{t-3} < 1.33$	-0.3155*	0.3181*	6.5713*	-0.8733*	0.1232
		(0.0895)	(0.0184)	(0.9201)	(0.1219)	
	$x_{t-3} \geq 1.33$	-0.1290	0.1302*	1.0083	-0.1312	
		(0.1435)	(0.0288)	(1.8059)	(0.2419)	

* Parameter is significant at 5% significant level.

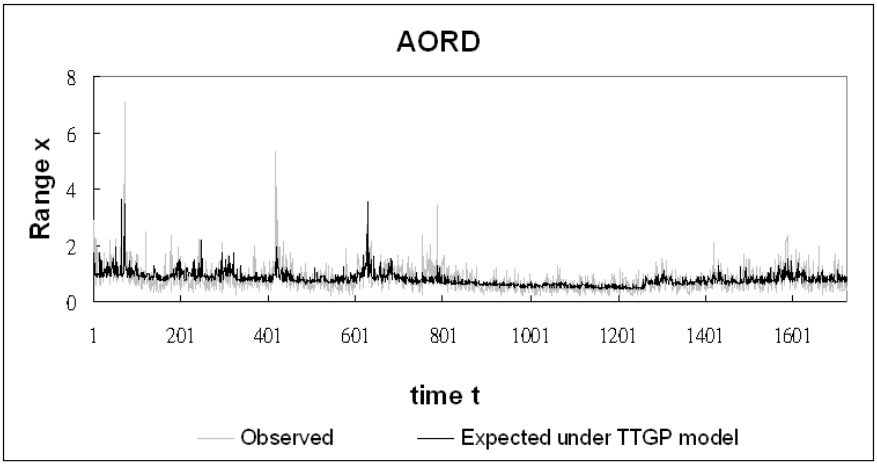


Fig. 3(a) Observed and fitted AORD daily range using TTGP model

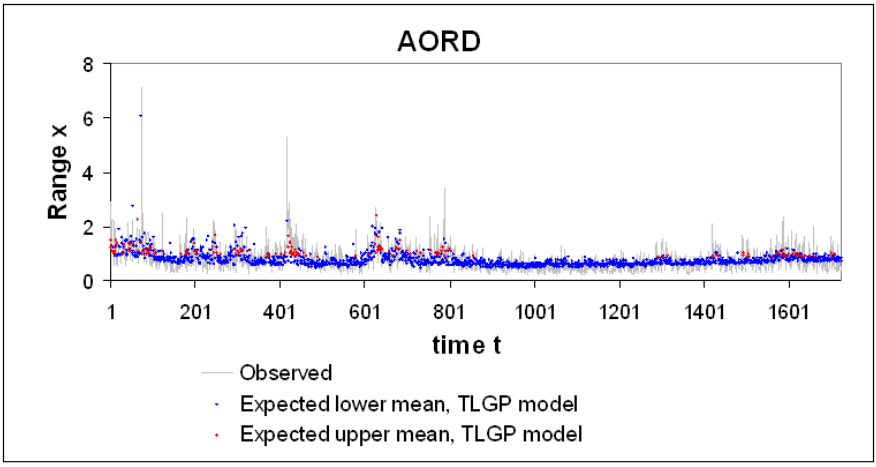


Fig. 3(b) Observed and fitted AORD daily range using TLGP model

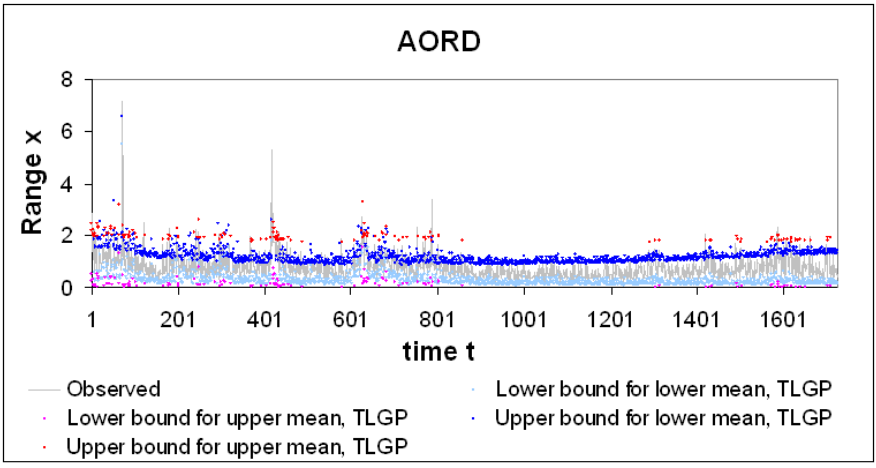


Fig. 3(c) Observed, fitted lower and upper AORD daily range within 2 sd using TLGP model

Table 4 Parameter estimates with standard errors (*in italics*) and model assessment measures for the TWII data

	Threshold	Mean parameter		Ratio parameter		
	thres.	intercept	US.	intercept	time	
Model	$\mathscr{T}_2$ or $\mathscr{W}_2$	$\beta_{\mu 0l}$	$\beta_{\mu 1l}$	$\beta_{a0l} \times 10^3$	$\beta_{a1l} \times 10^3$	*MSE*
C		0.9190*	0.0382*	0.5927*		0.7170
		(0.0285)	(0.0110)	(0.0241)		
R		0.7933*	0.0299*	-1.9257*	0.3423*	0.7073
		(0.0372)	(0.0112)	(0.4582)	(0.0622)	
T	$t < 211$	0.4316*	0.1005*	-2.6725*		0.6863
		(0.0634)	(0.0187)	(0.3378)		
	$t \geq 211$	0.8727*	0.0198	0.6715*		
		(0.0400)	(0.0137)	(0.0329)		
L	$x_{t-1} < 2.30$	0.6374*	0.0397*	-1.8549*	0.3176*	0.7060
		(0.0452)	(0.0150)	(0.4987)	(0.0677)	
	$x_{t-1} \geq 2.30$	1.0889*	-0.0010	-0.5150	0.1433	
		(0.0558)	(0.0156)	(0.8797)	(0.1238)	

* Parameter is significant at 5% significant level.

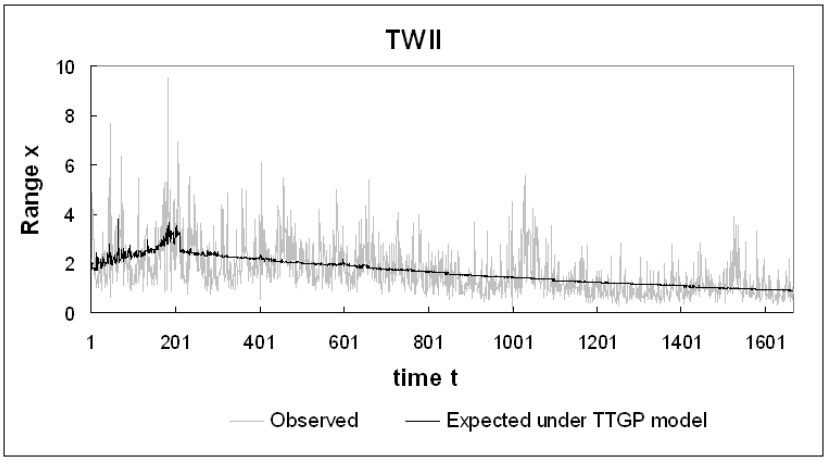


Fig. 4(a) Observed and fitted TWII daily range using TTGP model

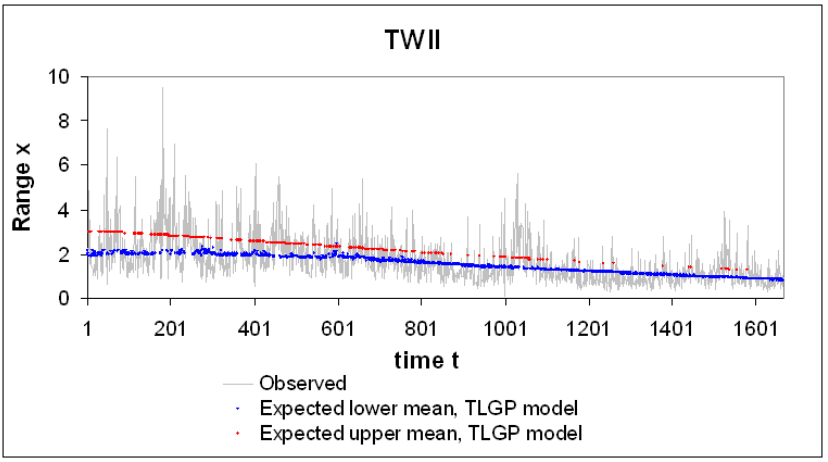


Fig. 4(b) Observed and fitted TWII daily range using TLGP model

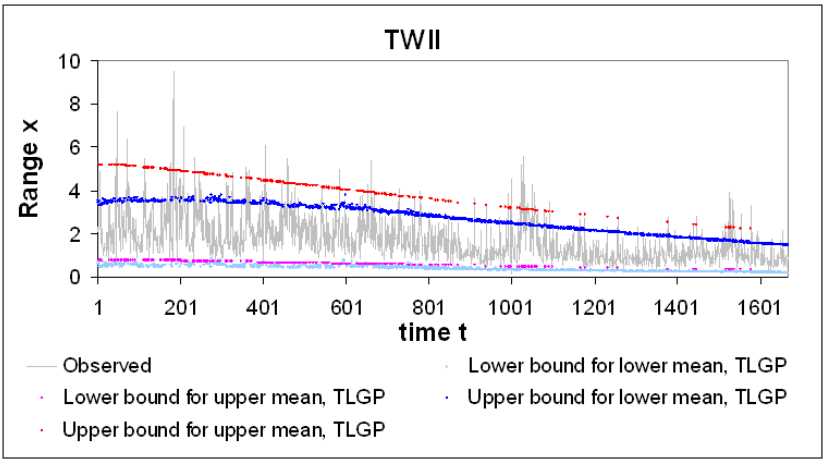


Fig. 4(c) Observed, fitted lower and upper TWII daily range within 2 sd using TLGP model

5 Conclusion

In this paper, we generalize the GP models to financial time series with different features and propose four extended GP models to allow for these features. The extended GP models show distinct trend movements, identify significant covariate effects and detect threshold times and threshold levels that indicate shift of models. Moreover the models allow variance to change over time and forecasts are simple and straight forward. Adopting the LS method, model implementation is also greatly simplified.

Four extended GP models including the AGP model with separate linear function of covariates for the mean μ_t and ratio a_t, the TTGP models with threshold times

and the TLGP model with threshold levels are proposed and applied to analyze the intra-day price range from four stock markets of four Asian cities and countries. Result shows that the TLGP model is the best model. The model identifies significant trend movements and covariate effects. Noncoverages of predictive intervals, which are two standard deviations from either sides of the mean, range from 5% to 19%. Twenty day forecasts show reasonable MSEs and noncoverages of predictive intervals. The only exception is the daily range of HSI which are much more volatile during the forecasting period. In summary, the proposed GP models are simple, easy to implement and give reliable estimates of the mean and volatility.

References

1. Andersen, T., Bollerslev, T.: Answering the skeptics: yes, standard volatility models do provide accurate forecasts. International Economic Review 39, 885–905 (1998)
2. Bollerslev, T.: Generalized autoregressive conidtional heteroskedasticity. Journal of Econometrics 31, 307–327 (1986)
3. Brandt, M.W., Jones, C.S.: Volatility forecasting with range-based EGARCH models. Journal of Business and Economic Statistics 24, 470–486 (2006)
4. Chan, J.S.K., Lam, Y., Leung, D.Y.P.: Statistical inference for geometric processes with gamma distributions. Computional Statistics and Data Analysis 47, 565–581 (2004)
5. Chan, J.S.K., Yu, P.L.H., Lam, Y., Ho, A.P.K.: Modeling SARS data using threshold geometric process. Statistics in Medicine 25, 1826–1839 (2006)
6. Chan, J.S.K., Leung, D.Y.P.: A new approach to the modelling longitudinal binary data with trend: the binary geometric process model. Computational Statistics 25, 505–536 (2010)
7. Chan, J.S.K., Lam, C.P.Y., Yu, P.L.H., Choy, S.T.B., Chen, C.W.S.: A Bayesian conditional autoregressive geometric process model for range data. Computational Statistics and Data Analysis 56, 3006–3019 (2012)
8. Chen, C.W.S., Gerlach, R.H., Lin, E.M.H.: Forecast volatility from threshold heteroskedastic range models. Computational Statistics and Data Analysis, on Statistical & Computational Methods in Finance 52, 2990–3010 (2008)
9. Chou, R.: Forecasting Financial Volatilities With Extreme Values: The Conditional Autoregressive Range (CARR) Model. Journal of Money Credit and Banking 37, 561–582 (2005)
10. Engle, R.F.: Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50, 987–1008 (1982)
11. Feller, W.: Fluctuation theory of recurrent events. Transactions of the American Mathematical Society 67, 98–119 (1949)
12. Garman, M.B., Klass, M.J.: On the estimation of price volatility from historical data. Journal of Business 53, 67–78 (1980)
13. Gilks, W.R., Richardson, S., Spiegelhalter, D.J.: Markov Chain Monte Carlo in Practice. Chapman and Hall, UK (1996)
14. Heston, S.L.: A closed-form solution for options with stochastic volatility with applications to bond and currency options. Review of Financial Studies 6, 327–343 (1993)
15. Hull, J.C., White, A.: An analysis of the bias in option pricing caused by by a stochastic volatility. Advances in Futures and Options Research 3, 29–61 (1988)
16. Lam, Y.: Geometric process and replacement problem. Acta Mathematicae Applicatae Sinica 4, 366–377 (1988)

17. Lam, Y.: Nonparametric inference for geometric processes. Commun. Statist. Theory Meth. 21, 2083–2105 (1992)
18. Lam, Y., Chan, J.S.K.: Statistical inference for geometric processes with lognormal distribution. Computional Statistics and Data Analysis 27, 99–112 (1998)
19. Lam, Y.: The Geometric Process and it's applications. World Scientific Publishing Co. Pte. Ltd. (2007)
20. Parkinson, M.: The extreme value method for estimating the variance of the rate of return. Journal of Business 53, 61–65 (1980)
21. Smith, A.F.M., Roberts, G.O.: Bayesian Computation via the Gibbs Sampler and Related Markov Chain Monte Carlo Methods. Journal of the Royal Statistical Society B 55, 3–23 (1993)
22. Spiegelhalter, D., Thomas, A., Best, N.: Bayesian inference using Gibbs sampling for Window version (2000), The website for WinBUGS is `http://www.mrc-bsu.cam.ac/bugs`
23. Spiegelhalter, D., Best, N.G., Carlin, B.P., Van der Linde, A.: Bayesian Measures of Model Complexity and Fit (with Discussion). Journal of the Royal Statistical Society B 64, 583–616 (2002)
24. Wan, W.Y., Chan, J.S.K.: A new approach for handling longitudinal count data with zero inflation and overdispersion: Poisson Geometric Process model. Biometrical Journal 51, 556–570 (2009)
25. Wan, W.Y., Chan, J.S.K.: Bayesian analysis of robust Poisson geometric process model using heavy-tailed distributions. Computational Statistics and Data Analysis 55, 687–702 (2011)

Residual Based Cusum Test for Parameter Change in AR-GARCH Models

Sangyeol Lee and Jiyeon Lee

Abstract. In this paper we consider the problem of testing for a parameter change in AR(1)-GARCH(1,1) models based on the residual cusum test. It is shown that the limiting distribution of the residual cusum test statistic is the sup of a Brownian bridge. Through a simulation study, it is demonstrated that the proposed test performs adequately. A real data analysis is provided for illustration.

1 Introduction

Since Page (1955), the problem of testing for a parameter change has been an important issue in economics, engineerin and medicine, and a vast number of articles have been published in various research areas. For earlier work, we refer to Csörgő and Horváth (1997). The change point problem has drawn much attention from many researchers in time series analysis since time series often suffer from structural changes owing to changes of policy and critical social events. It is well known that detecting a change point is a crucial task and ignoring it can lead to a false conclusion: see, for example, Hamilton (1994), page 450.

The GARCH model has long been popular in financial time series analysis. Inclán and Tiao's (1994) cumulative sum (cusum) test was originally designed for testing for variance changes and allocating their locations in iid samples. Later, it is demonstrated that the same idea can be extended to various time series models such as ARMA-GARCH, ARCH regression, Poisson GARCH, multivariate GARCH models, tail indices, and diffusion processes: see Lee and Park (2001), Lee, Ha, Na and Na (2003), Lee, Tokutsu and Maekawa (2004), Lee, Nishiyama and Yosida (2006), Lee and Song (2008), Kang and Lee (2009), Kim and Lee (2009), and Na, Lee and Lee (2012, 2013). See also the papers cited therein for a general review.

The cusum test is designed based on checking the discrepancy between the sequentially obtained estimators and the one obtained from the whole observations.

Sangyeol Lee · Jiyeon Lee
Department of Statistics, Seoul National University, Seoul, 151-747, Korea
e-mail: sylee@stats.snu.ac.kr

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_6,

This estimates-based approach performs well in many situations but has been proven to suffer from severe size distortions and low powers in GARCH type models. Particularly, Song and Lee (2008) studied the change point problem in ARMA-GARCH models without conducting a simulation study. Lee, Tokutsu and Maekawa (2004) made an effort to overcome the defect by considering the cusum test based on residuals and demonstrated its validity. A simulation study therein illustrated that the residual based test discards correlation effects and much improves the performance of the test. However, their simulation study was restricted to the pure GARCH model case and there has been a demand to investigate the performance of the cusum test in more general GARCH type models. Motivated by this, we study the residual based cusum test in ARMA-GARCH models since they are widely used in practice. Special attention is paid to AR(1)-GARCH(1,1) models for simplicity although our method could be extended to general ARMA-GARCH models.

The organization of this paper is as follows. In Section 2, we introduce the residual cusum test and show that its limiting distribution is the sup of a Brownian bridge. In Section 3, we perform a simulation study and conduct a real data analysis. Finally, in Section 4, we provide concluding remarks.

2 Residual Based Cusum Test

Suppose that $\{y_t\}$ satisfies the following equation:

$$\begin{aligned} y_t &= \phi y_{t-1} + \varepsilon_t \\ \varepsilon_t &= h_t \cdot \xi_t \\ h_t^2 &= \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}^2, \end{aligned} \tag{1}$$

where ξ_t are iid random variables with zero mean and unit variance, ϕ is a real number in $(-1, 1)$, and ω, α, β are nonnegative real numbers with $\alpha + \beta < 1$. We assume that $E|\varepsilon_t|^{4+\delta} < \infty$ and $E|\xi_t|^{4+\delta} < \infty$ for some $\delta > 0$.

Given the observations $y_1, \ldots, y_n$, our objective is to test the following hypotheses:

$H_0 : \theta = (\omega, \alpha, \beta)$ remains the same for the whole series. vs.

H_1 : not H_0.

For a test, as in Lee et al. (2004), we consider the cusum test based on residuals $\left\{\widehat{\xi}_t^2\right\}$ obtained from equation (1) such as

$$\tilde{T}_n := \frac{1}{\sqrt{n}\widehat{\tau}} \max_{q+1 \le k \le n} \left| \sum_{t=q+1}^{k} \widehat{\xi}_t^2 - \frac{k-q}{n-q} \sum_{t=q+1}^{n} \widehat{\xi}_t^2 \right|, \tag{2}$$

where $\widehat{\tau}^2$ is an estimator of $\tau^2 = Var\left(\xi_t^2\right)$, q is a positive integer, and $\widehat{\xi_t^2} = (y_t - \hat{\phi} y_{t-1})^2 / \widehat{h_t^2}$, which are obtained by estimating the unknown parameters $\phi, \omega, \alpha, \beta$. These estimators play an important role to detect changes in the GARCH parameters in the presence of changes, while the iid property of the true errors remains when

no change occurs. From this reasoning, one can anticipate that the residual cusum test should be more stable and produce better powers.

Now, we construct the residual cusum test. Similarly to

$$h_t^2 = a + \alpha \sum_{j=0}^{\infty} \beta^j \varepsilon_{t-1-j}^2,$$

we define

$$\widehat{h_t^2} = \widehat{a} + \widehat{\alpha} \sum_{j=0}^{q_n-1} \widehat{\beta}^j \widehat{\varepsilon}_{t-1-j}^2,$$

where $a = \dfrac{\omega}{1-\beta}$, $\widehat{\varepsilon}_t = y_t - \widehat{\phi} y_{t-1}$, and $\widehat{\phi}$, $\widehat{a}$, $\widehat{\alpha}$, $\widehat{\beta}$ are the estimators based on $y_1, \ldots, y_n$ for ϕ, a, α, β with

$$\begin{aligned} \sqrt{n}\left(\widehat{\phi} - \phi\right) = O_P(1), \sqrt{n}(\widehat{a} - a) = O_P(1), \\ \sqrt{n}(\widehat{\alpha} - \alpha) = O_P(1), \sqrt{n}\left(\widehat{\beta} - \beta\right) = O_P(1) \end{aligned} \tag{3}$$

under the null hypothesis, and q_n is a sequence of positive integers such that $q_n \to \infty$ and $q_n/n \to 0$ as $n \to \infty$. In practice, one can either employ the quasi maximum likelihood estimators (QMLEs) as in Lee and Song (2008) or two-step estimators, namely, one first estimates ϕ by the least squares method and then estimate the GARCH parameters based on the least squares residuals. It is noteworthy that our cusum test is not suitable to detect the change of the autoregressive parameter ϕ, and further, a change of ϕ does not much affect the performance of the cusum test. The following is a main result of this section.

Theorem 1. *Let*

$$\widehat{T}_n := \frac{1}{\sqrt{n}\widehat{\tau}} \max_{q_n+1 \le k \le n} \left| \sum_{t=q_n+1}^{k} \widehat{\xi}_t^2 - \frac{k-q_n}{n-q_n} \sum_{t=q_n+1}^{n} \widehat{\xi}_t^2 \right|, \tag{4}$$

where $\widehat{\tau}^2 = \dfrac{1}{n-q_n} \sum_{t=q_n+1}^{n} \widehat{\xi}_t^4 - \left(\dfrac{1}{n-q_n} \sum_{t=q_n+1}^{n} \widehat{\xi}_t^2 \right)^2$. *Then if* $n\rho^{2q_n} \to 0$ *for all* $\rho \in (0,1)$, *under* H_0,

$$\widehat{T}_n \xrightarrow{d} \sup_{0 \le u \le 1} \left| \overset{0}{\mathscr{B}}(u) \right|, \ n \to \infty,$$

where $\overset{0}{\mathscr{B}}$ *denotes a Brownian bridge, namely,* $\overset{0}{\mathscr{B}}$ *is a Gaussian process on [0,1] with mean zero and* $Cov(\overset{0}{\mathscr{B}}(s), \overset{0}{\mathscr{B}}(t)) = s \wedge t - st$ *for all* $s,t \in [0,1]$.

Remark. A typical example of q_n is $[(\log n)^{\zeta}]$ with $\zeta > 1$. The proof below is similar to that of Theorem 1 of Lee et al. (2004) and is presented without detailing all

algebras. However, in our proof, we do not need their (A4), since this condition may be violated in our case, so that our proof is a lot simpler than the original one. The above theorem may be generalized to ARMA-GARCH models without serious troubles. We leave this as a task of our future study.

Proof. We decompose $\widehat{\xi_t^2}$ into $A_{1t}+A_{2t}+A_{3t}$, where

$$A_{1t}=\xi_t^2\cdot\frac{h_t^2}{\widehat{h_t^2}},\ A_{2t}=\frac{2\varepsilon_t y_{t-1}(\phi-\hat{\phi})}{\widehat{h_t^2}},\ A_{3t}=\frac{\varepsilon_t^2 y_{t-1}^2(\phi-\hat{\phi})^2}{\widehat{h_t^2}}.$$

We first deal with A_{1t}. We express

$$A_{1t}=\xi_t^2+\left(\frac{h_t^2}{\widehat{h_t^2}}-1\right)\cdot\xi_t^2=\xi_t^2+B_t+C_t,$$

where

$$B_t=\frac{h_t^2-\widehat{h_t^2}}{h_t^2}\cdot\xi_t^2\ \text{ and }\ C_t=\frac{\left(h_t^2-\widehat{h_t^2}\right)^2}{h_t^2\cdot\widehat{h_t^2}}\cdot\xi_t^2.$$

We first show that

$$\frac{1}{\sqrt{n}}\max_{q_n+1\le k\le n}\left|\sum_{t=q_n+1}^{k}B_t-\frac{k-q_n}{n-q_n}\sum_{t=q_n+1}^{n}B_t\right|=o_P(1). \tag{5}$$

Note that

$$\begin{aligned}h_t^2-\widehat{h_t^2}&=(a-\widehat{a})+(\alpha-\widehat{\alpha})\sum_{j=0}^{q_n-1}\beta^j\varepsilon_{t-1-j}^2\\&\quad+\hat{\alpha}\sum_{j=0}^{q_n-1}\widehat{\beta^j}(\varepsilon_{t-1-j}^2-\hat{\varepsilon}_{t-1-j}^2)\\&\quad+\widehat{\alpha}\sum_{j=0}^{q_n-1}\left(\beta^j-\widehat{\beta^j}\right)\varepsilon_{t-1-j}^2+\alpha\sum_{j=q_n}^{\infty}\beta^j\varepsilon_{t-1-j}^2\\&:=\sum_{i=1}^{5}B_{it}/\left(\xi_t^2/h_t^2\right).\end{aligned} \tag{6}$$

To show (5) with B_t replaced by B_{4t}, we use the invariance principle for the strong mixing process: see Carrasco and Chen (2002) and Theorem 1.7 of Peligrad (1986) and follow the arguments in the proof of argument (6) of Lee et al. (2004). In a similar fashion, one can readily show that (5) with B_t replaced by $\sum_{i=1}^{3}B_{it}$ is $o_P(1)$. Then, (5) follows from the fact $E\left[\frac{1}{\sqrt{n}}\sum_{t=q_n+1}^{n}|B_{5t}|\right]=O(\sqrt{n}\beta^{q_n})=o(1)$.

Now, we prove

$$\frac{1}{\sqrt{n}} \max_{q_n+1\le k\le n} \left| \sum_{t=q_n+1}^{k} C_t - \left(\frac{k-q_n}{n-q_n}\right) \sum_{t=q_n+1}^{n} C_t \right| = o_P(1). \tag{8}$$

In view of (6), we have that

$$C_t \leqslant d \sum_{i=1}^{5} B_{it}^2/\widehat{h}_t^2 \qquad \text{for some } d>0.$$

It is easy to see that $\frac{1}{\sqrt{n}} \sum_{t=q_n+1}^{n} \sum_{i=1}^{4} B_{it}^2/\widehat{h}_t^2 = o_P(1)$. Since $\frac{1}{\sqrt{n}} \sum_{t=q_n+1}^{n} B_{5t}^2 = O_P\left(\sqrt{n}\beta^{2q_n}\right)$, (8) follows. Combining (5) and (8), we have

$$\frac{1}{\sqrt{n}} \max_{q_n+1\le k\le n} \left| \sum_{t=q_n+1}^{k} A_{1t} - \left(\frac{k-q_n}{n-q_n}\right) \sum_{t=q_n+1}^{n} A_{1t} \right|$$

$$-\frac{1}{\sqrt{n}} \max_{q+1\le k\le n} \left| \sum_{t=q_n+1}^{k} \xi_t^2 - \left(\frac{k-q_n}{n-q_n}\right) \sum_{t=q_n+1}^{n} \xi_t^2 \right| = o_P(1). \tag{9}$$

Meanwhile, in a similar fashion to the above, one can show that

$$\frac{1}{\sqrt{n}} \max_{q_n+1\le k\le n} \left| \sum_{t=q_n+1}^{k} A_{it} - \left(\frac{k-q_n}{n-q_n}\right) \sum_{t=q_n+1}^{n} A_{it} \right| = o_P(1),\ i=2,3, \tag{10}$$

where we have used Donsker's invariance principle for the partial sum process of martingale differences $y_{t-1}\xi_t$ in dealing with A_{2t}, while we only used the moment condition and stationarity in dealing with A_{3t}. Further, we can easily verify that $\widehat{\tau}^2 \xrightarrow{P} \tau^2$. Hence, it follows from (9) and (10) that $\widehat{T}_n - T_n = o_P(1)$, where

$$T_n = \frac{1}{\sqrt{n}\tau} \max_{q_n+1\le k\le n} \left| \sum_{t=q_n+1}^{k} \xi_t^2 - \frac{k-q_n}{n-q_n} \sum_{t=q_n+1}^{n} \xi_t^2 \right|,$$

and the theorem follows from Donsker's invariance principle (cf. Billingsley (1999)). This completes the proof. □

3 Empirical Study

3.1 Simulation Study

In this section, we evaluate the performance of the test statistic $\hat{T}_n$ through a simulation study. In this simulation we perform a test at nominal level 0.05. The empirical sizes and power are calculated as the rejection number of the null hypothesis out of 1000 repetitions.

In order to see the performance of $\hat{T}_n$, we consider the model

$$y_t = \phi y_{t-1} + h_t \cdot \xi_t$$
$$h_t^2 = \omega + \alpha y_{t-1}^2 + \beta h_{t-1}^2,$$

where y_0 is assumed to be 0 and $\{\xi_t\}$ are iid standard nomal random variables. Now we consider the problem of test the following hypotheses:

$H_0 : \theta = (\omega, \alpha, \beta)$ are constant during the time $t = 1, \cdots, n$. vs.

$H_1 : \theta$ changes to $\theta' = (\omega', \alpha', \beta')$ at $n/2$.

Here we evaluate $\widehat{T}_n$ with sample sizes $n = 500, 800, 1000, 2000$ and use $q_n = \left[(\log n)^{3/2}\right]$. The empirical sizes and powers are summarized in Tables 1-3.

Tables 1-3 show that $\widehat{T}_n$ has no severe size distortions in most cases. It can be seen that when $\alpha + \beta$ is close to 1 (see Tables 2 and 3), $\widehat{T}_n$ exhibits some size distortions for small sample sizes. In fact, the empirical size gets very close to the nominal level 0.05 as n increases in all the cases. As mentioned earlier, this is because $\widehat{\xi_t^2}$ behaves asymptotically like iid ξ_t^2, unaffected by the GARCH parameters. Meanwhile, we can see that the powers are more than 0.9 at the sample size 2000. Generally, the cusum test in GARCH models requires a much larger sample size to make an accurate inference compared to the iid sample case. It seems that the GARCH data with volatility makes it harder to identify small changes. All these results indicate that $\widehat{T}_n$ performs adequately for the GARCH parameter change test.

Table 1 $(\phi, \omega, \alpha, \beta) = (0.3, 0.5, 0.2, 0.2)$

$(\phi', \omega', \alpha', \beta')$	$n = 500$	$n = 800$	$n = 1000$	$n = 2000$
size	0.059	0.058	0.057	0.055
$(0.3, 3.0, 0.2, 0.2)$	0.564	0.743	0.790	0.978
$(0.3, 0.5, 0.6, 0.2)$	0.768	0.912	0.962	0.999
$(0.3, 0.5, 0.2, 0.6)$	0.938	0.984	0.997	1.000
$(0.3, 3.0, 0.6, 0.2)$	0.386	0.604	0.713	0.948

Table 2 $(\phi, \omega, \alpha, \beta) = (0.3, 0.1, 0.4, 0.4)$

$(\phi', \omega', \alpha', \beta')$	$n = 500$	$n = 800$	$n = 1000$	$n = 2000$
size	0.067	0.065	0.059	0.056
$(0.3, 0.4, 0.4, 0.4)$	0.989	1.000	1.000	1.000
$(0.3, 0.1, 0.1, 0.4)$	0.350	0.663	0.859	1.000
$(0.3, 0.1, 0.4, 0.1)$	0.504	0.873	0.965	1.000
$(0.3, 0.4, 0.1, 0.1)$	0.443	0.604	0.653	0.902

Table 3 $(\phi, \omega, \alpha, \beta) = (0.3, 0.1, 0.2, 0.7)$

$(\phi', \omega', \alpha', \beta')$	$n = 500$	$n = 800$	$n = 1000$	$n = 2000$
size	0.071	0.057	0.054	0.057
$(0.3, 0.4, 0.2, 0.7)$	0.797	0.973	0.993	1.000
$(0.3, 0.1, 0.2, 0.2)$	0.242	0.523	0.763	1.000
$(0.3, 0.1, 0.1, 0.7)$	0.138	0.246	0.431	0.941
$(0.3, 0.4, 0.7, 0.2)$	0.474	0.616	0.685	0.907

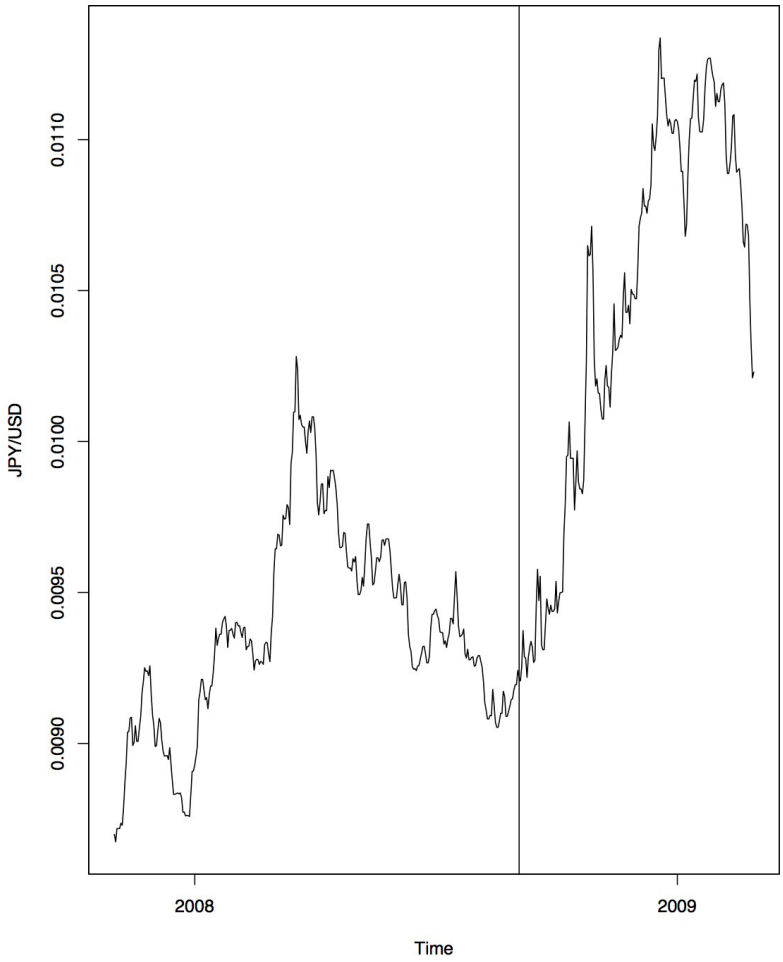


Fig. 1 Plot of daily JPY/USD data

3.2 Real Data Analysis

In this section, we apply our test to the foreign exchange rate(JPY/USD) data. The data is a daily log return of the JPY/USD exchange rate from Nov 1, 2007 to Feb 29, 2009 with 486 observations. Based on the SACF, SPACF, AIC and BIC results, we fit the AR(1) model to the data. Moreover, Figures 1 and 2 show that the returns have some volatility clustering phenomenon. By examining the Ljung-Box and LM-ARCH tests, it is revealed that the GARCH(1,1) model is reasonable to this series. From this, the AR(1)-GARCH(1,1) model in (1) is fitted to the data and the QMLE for the parameter is obtained as $(\hat{\phi}, \hat{\omega}, \hat{\alpha}, \hat{\beta})$ = (0.293,0.011,0.085,0.888).

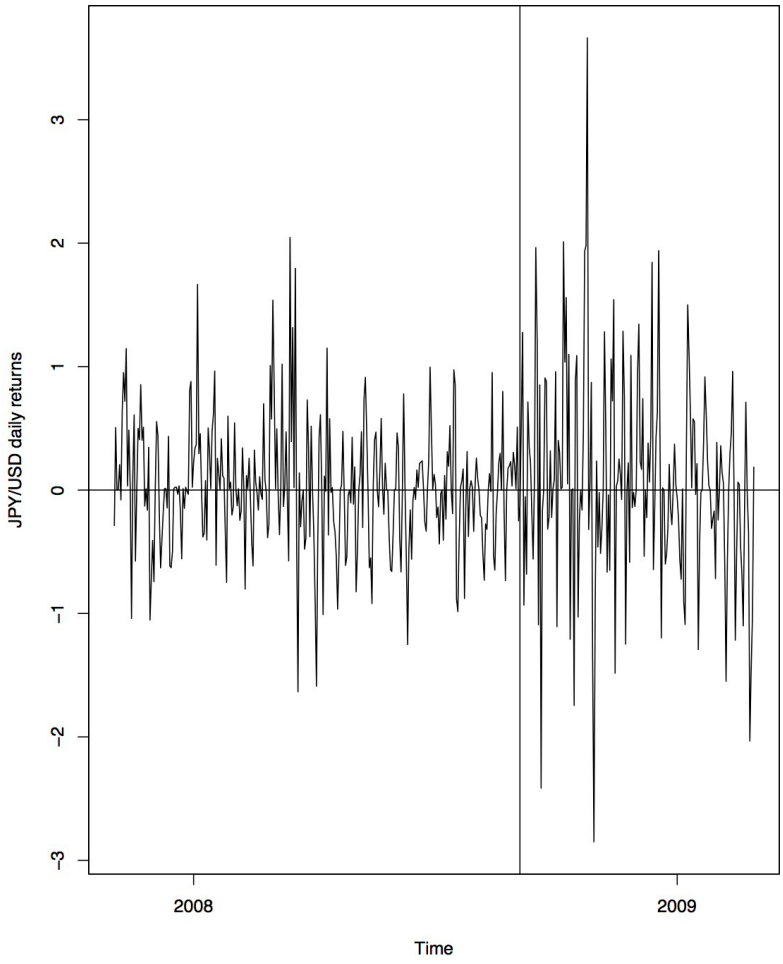


Fig. 2 Plot of daily log return JPY/USD data

To test H_0, we use $q_n = \left[(\log n)^{3/2}\right]$ and the critical value 1.342 at the nominal level 0.05 (see the horizon line in Figure 3). As seen in Figure 3, the $\widehat{T}_{n,k}$ plot with $\widehat{T}_{n,k} = \frac{1}{\sqrt{n}\hat{\tau}} \left| \sum_{t=q_n+1}^{k} \hat{\xi}_t^2 - \frac{k-q_n}{n-q_n} \sum_{t=q_n+1}^{n} \hat{\xi}_t^2 \right|$ shows that the maximum value of $\widehat{T}_{nk}$ is 1.43 at Sep, 4, 2008 (see the vertical line in Figures 1-3). The QMLE for the data in the first period from Nov 1, 2007 to Sep, 4, 2008 is obtained as $(\hat{\phi}_1, \hat{\omega}_1, \hat{\alpha}_1, \hat{\beta}_1)$ = (0.291,0.022,0.071,0.830) and the QMLE for data in the second period from Sep, 5, 2008 to Feb 29, 2009 is obtained as $(\hat{\phi}_2, \hat{\omega}_2, \hat{\alpha}_2, \hat{\beta}_2)$ = (0.278,0.017,0.051,0.922). The above results indicate that the GARCH parameters experience a significant change.

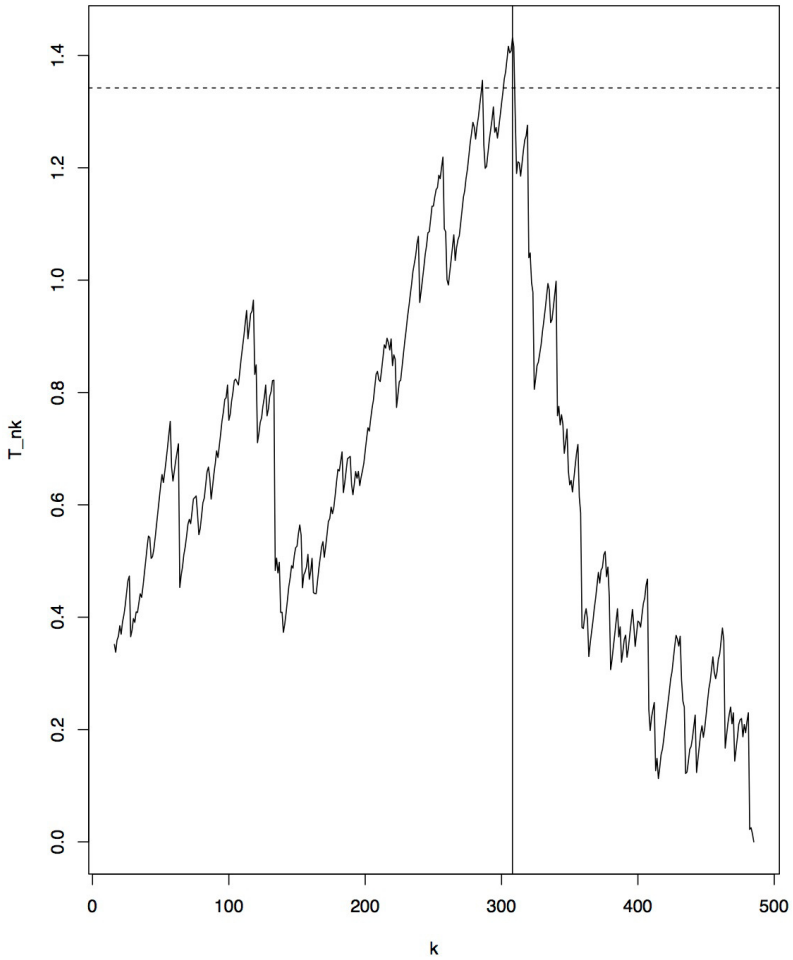


Fig. 3 Plot of $\widehat{T}_{nk}$

4 Concluding Remarks

In this paper, we proposed a residual based cusum test and derived that the test statistic is asymptotically distributed as the sup of a Brownian bridge under regularity conditions. This paper was motivated to overcome the drawbacks of the estimates based cusum test. The simulation result showed that our test performs adequately and a real data analysis was illustrated. Overall, it is believed that our test is a functional tool for testing for a paramter change in AR(1)-GARCH(1,1) models. We anticipate that the residual cusum test can be extended to other type of GARCH models. We leave the task of extension as our future study.

Acknowledgements. We would like to thank the referee for his careful reading. This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MEST) (No. 2012R1A2A2A01046092).

References

1. Billingsley, P.: Convergence of Probability Measures, 2nd edn. Wiley, New York (1999)
2. Carrasco, M., Chen, X.: Mixing and moment properties of various GARCH and stochastic volatility models. Econometric Theory 18, 17–39 (2002)
3. Csörgő, M., Horváth, L.: Limit Theorems in Change-Point Analysis. Jhon Wiley & Sons Ltd., West Sussex (1997)
4. Hamilton, J.D.: Time Series Analysis. Princeton University Press, New Jersey (1994)
5. Inclán, C., Tiao, G.C.: Use of cumulative sums of squares for retrospective detection of changes of variances. J. Amer. Statist. Assoc. 89, 913–923 (1994)
6. Kang, J., Lee, S.: Parameter change test for random coefficient integer-valued autoregressive processes with application to polio data analysis. J. Time Series Anal. 30, 239–258 (2009)
7. Kim, M., Lee, S.: Test for tail index change in stationary time series with Pareto type marginal distribution. Bernoulli 15, 325–356 (2009)
8. Lee, S., Ha, J., Na, O., Na, S.: The Cusum Test for Parameter Change in Time Series Models. Scand. J. Statist. 30, 781–796 (2003)
9. Lee, S., Nishiyama, Y., Yosida, N.: Test for parameter change in diffusion processes by cusum statistics based on one-step estimators. Ann. Inst. Statist. Math. 58, 211–222 (2006)
10. Lee, S., Park, S.: The cusum of squares test for scale changes in infinite order moving average processes. Scand. J. Statist. 28, 625–644 (2001)
11. Lee, S., Song, J.: Test for parameter change in ARMA models with GARCH errors. Statist. Probab. Letters 78, 1990–1998 (2008)
12. Lee, S., Tokutsu, Y., Maekawa, K.: The residual cusum test for parameter change in regression models with ARCH errors. J. Japan Statist. Soc. 34, 173–188 (2004)
13. Na, S., Lee, J., Lee, S.: Change point test for copula ARMA-GARCH models. J. Time Series Anal. 33, 554–569 (2012)
14. Na, O., Lee, J., Lee, S.: Change point detection in SCOMDY models. AStA Advancs in Statistical Analysis 97, 215–238 (2013)
15. Page, E.S.: A test for change in a parameter occurring at an unknown point. Biometrika 42, 523–527 (1955)

16. Picard, D.: Testing and estimating change-points in time series. Adv. Appl. Prob. 17, 841–867 (1985)
17. Peligrad, M.: Recent advances in the central limit theorem and its weak invariance principle for mixing sequences of random varaibles. In: Eberlein, E., Taqqu, M.S. (eds.) Dependence in Probability and Statistics, pp. 193–223. Birkhäuser, Boston (1986)

Dependence and Association Concepts through Copulas

Zheng Wei, Tonghui Wang*, and Wararit Panichkitkosolkul

Abstract. In this paper, dependence concepts such as affiliation, left-tail decreasing, right-tail increasing, positively regression dependent, and positively quadrant dependent are studied in terms of copulas. Relationships among these dependent concepts are obtained. An affiliation is a notion of dependence between two positively dependent random variables and some measures of it are provided. It has been shown that the affiliation property is preserved using bilinear extensions of subcopula. As an application, the affiliation property of skew-normal copula is investigated. For illustration of dependent concepts and their relationships, several examples are given.

1 Introduction

With the rapid development of mathematical finance and risk management in the last two decades, more and more attention has been paid to creating some practical statistical models beyond normal settings to improve competitive performance in finance and insurance fields. The copula is one of the most important models used in mathematical finance. Specifically, copulas, introduced in [25], are used to model multivariate data as they account for the dependence structure and provide a flexible representation of the multivariate distribution. Copulas are multivariate

Zheng Wei
Department of Mathematical Sciences, New Mexico State University, USA
e-mail: weizheng@nmsu.edu

Tonghui Wang
Department of Mathematical Sciences, New Mexico State University, USA,
and College of Science, Northwest A & F University, China
e-mail: twang@nmsu.edu

Wararit Panichkitkosolkul
Department of Mathematics and Statistics, Thammasat University, Thailand
e-mail: wararit@mathstat.sci.tu.ac.th

* Corresponding author.

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_7,

distributions with $[0,1]$-uniform marginal, which contain the most multivariate dependence structure properties and do not depend on the marginals. For references, see [10], [7], [20], and [22].

In analysis of auction theory, valuations of different bidders (modeled as random variables) could be affiliated. In similar situations in econometrics, when dependence of random variables is a concern, the theory of affiliated copulas, which will be defined in next section, offers an appropriate approach. Recently, Rinotta and Scarsini studied the total positivity order for multivariate normal distributions in [20]. The importance of the affiliation properties in application of auction theory can be found in [14], [4], [19], [24] and [21].

As an extension of normal settings, multivariate skew normal distributions are widely used in almost all fields for almost three decades. For references on skew normal distributions, see [1], [2], [8], and many other papers listed in the website of Azzalini [3]. The concept of affiliation on the class of multivariate skew normal family has not been investigated in the literature.

This paper is organized as follows. Dependence and association concepts as well as their relationships are obtained in Section 2. Bilinear extension method of a two dimensional subcopula together with their affiliation property is studied in Section 3. Average and local measures of affiliation are provided with several examples in Section 4. Conditions under which the bivariate skew-normal copulas are affiliated are discussed in Section 5.

2 Basic Concepts

Following the notions of [25], we have definition of affiliation.

Definition 1. The random variables X and Y are said to be **affiliated** (or **positively likelihood ratio dependent**(PLRD)) if

$$h(x,y^*)h(x^*,y) \leq h(x,y)h(x^*,y^*) \tag{1}$$

holds for all $x^* \leq x$ and $y^* \leq y$, where $h(\cdot,\cdot)$ is the joint density function of (X,Y).

Recall that a copula C is a function $C(\cdot,\cdot) : [0,1]^2 \to [0,1]$ satisfying
(i) $C(u,0) = C(0,v) = 0$, for $u,v \in [0,1]$,
(ii) $C(u,1) = u, C(1,v) = v$, for $u,v \in [0,1]$, and
(iii) For any $(u,v) \leq (u',v')$, $C(u',v') - C(u,v') - C(u',v) + C(u,v) \geq 0$.

Sklar's theorem states that if H is the joint distribution of (X,Y), then there is a copula C such that $H(x,y) = C(F(x),G(y))$ for $(x,y) \in R^2$. Copula characterizes dependence structures and dependence measures which is also independent of marginal distributions. It can be viewed as a joint distribution of two random variables U and V located on $[0,1]$. Motivated by this, we give the corresponding affiliation definition for a copula as follows.

Definition 2. A copula $C(u,v)$ is said to be **affiliated** if

$$c(u,v^*)c(u^*,v) \leq c(u,v)c(u^*,v^*) \tag{2}$$

holds for all $u^* \leq u$ and $v^* \leq v$, where $c(\cdot,\cdot)$ is the joint density corresponding to copula $C(u,v)$ with $c(u,v) = \frac{\partial^2 C(u,v)}{\partial u \partial v}$.

Remark. It is true that the random variables X and Y are affiliated if and only if their corresponding copula is affiliated. Indeed, suppose X and Y are affiliated. Let $h(x,y)$ and $c(u,v)$ be the corresponding density function and copula density, respectively. Then $h(x,y) = c(F(x),G(y))f(x)g(y)$, where $F(x)$ and $G(y)$ are cumulative distribution functions (CDF) of X and Y, respectively. Since X and Y are affiliated, by definition,

$$h(x,y^*)h(x^*,y) \leq h(x^*,y^*)h(x,y), \quad x^* \leq x, y^* \leq y,$$

which is equivalent to

$$\begin{aligned} &c(F(x),G(y^*))f(x)g(y^*)c(F(x^*),G(y))f(x^*)g(y) \\ &\quad \leq c(F(x^*),G(y^*))f(x^*)g(y^*)c(F(x),G(y))f(x)g(y), \end{aligned}$$

which is reduced to

$$c(F(x),G(y^*))c(F(x^*),G(y)) \leq c(F(x^*),G(y^*))c(F(x),G(y)).$$

Since both F and G are distribution functions and therefore non-decreasing, for any $u^* \leq u, v^* \leq v$, let $x^* = F^{-1}(u^*), x = F^{-1}(u)$ and $y^* = G^{-1}(v^*), y = G^{-1}(v)$, where $F^{-1}(u) = \inf\{x \in \mathbb{R} | F(x) \geq u\}$. Therefore

$$c(u,v^*)c(u^*,v) \leq c(u^*,v^*)c(u,v).$$

The converse relation can be proved similarly. □

For the connection between affiliation property and positively quadrant dependence (PQD), let us recall the definition of PQD given below.

Definition 3. A copula $C : [0,1] \times [0,1] \mapsto [0,1]$ is said to be **positively quadrant dependent** if $C(u,v) \geq uv$ holds for all u, v.

The following Lemma of [10] will be used in the proof of our next result.

Lemma 1. *Let $H(x,y)$, $F(x)$, and $G(y)$ be the joint, and marginal CDFs of X and Y, respectively. If X and Y are positively quadrant dependent, then*

$$H(x,y) = F(x)G(y) + w(x,y) \quad x,y \in \mathbb{R} \tag{3}$$

with $w(x,y)$ satisfying the following conditions:
(i) $w(x,y) \geq 0$ for all x and y,
(ii) $w(x,\infty) = w(\infty,y) = w(x,-\infty) = w(-\infty,y) = 0$, for all x and y, and
(iii) $\frac{\partial^2 w(x,y)}{\partial x \partial y} \geq 0$.

Recall that a function $w(u,v)$ is **totally positive of order-2** (TP2) if

$$w(u',v)w(u,v') \leq w(u',v')w(u,v) \qquad \text{for all } u' \leq u, v' \leq v.$$

Also $w(u,v)$ is said to be 2**-increasing** if

$$w(u',v') + w(u,v) - w(u',v) - w(u,v') \geq 0 \qquad \text{for all } u' \leq u, v' \leq v.$$

Using above lemma, it is easy to prove the similar result given below.

Proposition 1. *If a copula $C(u,v)$ can be written as*

$$C(u,v) = uv + W(u,v) \qquad \textit{for all } u \textit{ and } u,$$

where $W(u,v)$ satisfying the following conditions:
(i) $W(u,v) \geq 0$,
(ii) $W(u,1) = W(1,v) = W(u,0) = W(0,v) = 0$,
(iii) $\frac{\partial^2 W(u,v)}{\partial v \partial u} \geq -1$,
(iv) $\frac{\partial^2 W(u,v)}{\partial v \partial u}$ *is a function with TP2 property and is 2-increasing, then the copula C is affiliated.*

Note that if we let $W(u,v) = C(u,v) - uv$, then by Theorem 1 below, we know that $C(u,v)$ is affiliated implies it is PQD, then by Lemma 1, conditions (i), (ii), and (iii) hold for W, but (iv) does not necessarily hold.

Example 2.1. For the CDF of Farlie-Gumbel-Morgenstern bivariate distribution [8]:

$$F(x,y) = F_X(x)F_Y(y)[1 + \rho(1 - F_X(x))(1 - F_Y(y))], \qquad -1 \leq \rho \leq 1,$$

the corresponding copula is $C(u,v) = uv[1 + \rho(1-u)(1-v)]$, $-1 \leq \rho \leq 1$. By the remark after Definition 2, it is easy to see Farlie-Gumbel-Morgenstern family is affiliated for $0 \leq \rho \leq 1$. □

In order to show that affiliation implies PQD, we need the following definition.

Definition 4. The random variable Y is said to be **positively regression dependent** in X, denoted by $PRD(Y|X)$, if $P(Y \leq y|X = x)$ is non-increasing in x for all y. Y is said to be **left-tail decreasing** in X, denoted by $LTD(Y|X)$, if $P(Y \leq y|X \leq x)$ is non-increasing in x for all y.

Corresponding to copulas, we have the following definition.

Definition 5. The random variable V in C is said to be **positively regression dependent** in U, denoted by $PRD(V|U)$, if $p_u(u,v)$ is non-increasing in u for all v, where $p_u(u,v) = \frac{\partial C(u,v)}{\partial u}$. V in C is said to be **left-tail decreasing** in U, denoted by $LTD(V|U)$, if $\frac{C(u,v)}{u}$ is non-increasing in u for all v.

Proposition 2. *The following result gives the relationship between X and Y and their corresponding U and V.*

(i) The random variables X and Y are positively regression dependent in X if and only if V in the corresponding copula $C(u,v)$ is positively regression dependent in U.

(ii) The random variables X and Y are left-tail decreasing in X if and only if V in the corresponding copula $C(u,v)$ is left-tail decreasing in U.

(iii) The random variables X and Y are PQD if and only if U and V in the corresponding copula $C(u,v)$ are PQD.

Proof. We prove (i) and (ii) only, and the proof of (iii) is trivial. For (i), suppose Y is Positively regression dependent in X. Let $H_{y|x}(y|x)$ be the conditional CDF of Y given $X = x$. By definition, it is non-increasing in x. Also

$$\begin{aligned} H_{y|x}(y|x) &= \int_{-\infty}^{y} h_{y|x}(t|x)dt = \int_{-\infty}^{y} \frac{h(x,t)}{f(x)}dt = \int_{-\infty}^{y} \frac{\partial}{\partial t}\left(\frac{\partial C(F(x),G(t))}{\partial x}\right)\frac{1}{f(x)}dt \\ &= \int_{-\infty}^{y} \frac{\partial}{\partial t} p_u(F(x),G(t))dt = p_u(F(x),F(y)) - p_u(F(x),0) \\ &= p_u(F(x),G(y)). \end{aligned}$$

Since both F,G are distribution functions and therefore non-decreasing, so that, $H_{y|x}(y|x)$ is non-increasing in x if and only if $p_u(u,v)$ is non-increasing in u.

(ii) Let Y is Left-tail decreasing in X, by definition, $P(Y \le y|X \le x) = \frac{H(x,y)}{F(x)}$ is non-increasing in x. Since both F,G are distribution functions and therefore non-decreasing, so that, $LTD(Y|X)$ if and only if $LTD(V|U)$. □

Theorem 1. *Let $C : [0,1] \times [0,1] \mapsto [0,1]$ be a copula, then the following implications are true.*

$$\textit{Affiliation} \quad \Rightarrow \quad PRD(V|U) \quad \Rightarrow \quad LTD(V|U) \quad \Rightarrow \quad PQD.$$

Proof. Suppose that U and V are affiliated.

To show it implies $PRD(V|U)$, for any $u^* < u, v^* < v$, we have, by definition,

$$c(u,v^*)c(u^*,v) \le c(u,v)c(u^*,v^*) \quad \Rightarrow \quad \frac{C(v|u)}{c(u|v)} \le \frac{C(v|u^*)}{c(u|v^*)}.$$

Let $G(v|u) = \frac{c(v|u)}{C(v|u)}$, then we have $G(v|u) \ge G(v|u^*)$ for all $u^* < u, v^* < v$. Note that $G(u|v) = \frac{\partial \ln(C(v|u))}{\partial v}$. We obtain

$$1 - \ln(C(v|u)) = \int_{v}^{1} G(t|u)dt \ge \int_{v}^{1} G(t|u^*)dt = 1 - \ln(C(v|u^*)).$$

Thus, $C(v|u^*) \ge C(v|u)$ for $u^* < u$, which implies $PRD(V|U)$.

For $PRD(V|U) \Rightarrow LTD(V|U)$, we need the fact that for any interval $I \subseteq [0,1]$,

$$P(V > v|U \in I) = \frac{\int_I P(V > v|U = u)dP(U \leq u)}{P(U \in I)}.$$

$LTD(V|U)$ is equivalent to the $Pr(V > v|U \leq u)$ is non-decreasing in u for all v, which, in turn, is equivalent to

$$P(V > v|U \leq u) \geq P(V > v|U \leq u^*)$$

for $u^* < u$ and all v. This is also equivalent to $P(V > v|u^* < U \leq u) \geq P(V > v|U \leq u^*)$, for all $u > u^*$. Note that,

$$\begin{aligned} P(V > v|u^* < U \leq u) &= \frac{\int_{u^*}^{u} P(V > v|U = u)dP(U \leq u)}{P(u^* < U \leq u)} \\ &\geq \frac{P(V > v|U = u^*)\int_{u^*}^{u} dP(U \leq u)}{P(u^* < U \leq u)} = P(V > v|U = u^*) \\ &\geq \frac{\int_{-\infty}^{u^*} P(V > v|U = u)dPr(U \leq u)}{P(-\infty < U \leq u^*)} = P(V > v|U \leq u^*), \end{aligned}$$

which implies $LTD(V|U)$.

For $LTD(V|U) \Rightarrow PQD$, note that

$$P(V \leq v|U \leq u) \geq P(V \leq v|U \leq 1) = P(V \leq v) = v, \tag{4}$$

which is equivalent to $C(u,v) \geq uv$.

Note that the FGM-copula has properties PRD, LTD, PQD for $0 \leq \rho \leq 1$. Several counterexamples of the converse relations of Theorem 2.1 can be found in [25] and [15].

3 Invariance of Affiliation of Subcopula through Bilinear Interpolation

In this section, we are going to discuss the affiliation property of copula, which is obtained from a subcopula through the method of bilinear interpolation.

Definition 6. A *two-dimensional* **subcopula** is a function C' with the following properties:

(a) Domain of C' is $S_1 \times S_2$, where S_1 and S_2 are subsets of $[0,1]$ containing 0 and 1,

(b) C' is 2-increasing and $C'(u,0) = C'(0,v) = 0$,

(c) For every u in S_1 and every v in S_2, and

$$C'(u,1) = u \quad \text{and} \quad C'(1,v) = v.$$

Also, any sub-copula can be extended to a copula, but the extension is generally non-unique. Here we introduce one popular method called **bilinear interpolation** [20]:

Definition 7. Let C' be a sub-copula with domain $S_1 \times S_2$, now for any $(a,b) \in [0,1]^2$, let a_1 and a_2 be, respectively, the greatest and least elements of $\overline{S_1}$ that satisfy $a_1 \le a \le a_2$; and let b_1 and b_2 be, respectively, the greatest and least elements of $\overline{S_2}$ that satisfy $b_1 \le b \le b_2$, where $\overline{S}$ is the closure of set S . Note that if a is in $\overline{S_1}$, then $a_1 = a = a_2$; and if b is in $\overline{S_2}$, then $\overline{b_1} = b = \overline{b_2}$. Now let

$$\lambda = \begin{cases} \frac{a-a_1}{a_2-a_1} & \text{if } a_1 < a_2 \\ 1 & \text{if } a_1 = a_2 \end{cases}$$

and

$$\mu = \begin{cases} \frac{b-b_1}{b_2-b_1} & \text{if } b_1 < b_2 \\ 1 & \text{if } b_1 = b_2. \end{cases}$$

The copula C given by

$$C(a,b) = (1-\lambda)(1-\mu)C'(a_1,b_1) + (1-\lambda)\mu C'(a_1,b_2) + \lambda(1-\mu)C'(a_2,b_1) + \lambda\mu C'(a_2,b_2),$$

is a well defined copula.

The following result shows that the invariance between a subcopula and its bilinear interpolation of affiliation property.

Theorem 2. *Let C' be a sub-copula over $S_1 \times S_2$, and $C : [0,1]^2 \to [0,1]$ be the copula, which is constructed by bilinear interpolation from C'.*

(i) If C' is affiliated, then C is also affiliated. Furthermore, if C' is not affiliated, then C is also not affiliated.

(ii) If C' is PQD, then C is also PQD. Furthermore, if C' is not PQD, then C is also not PQD.

Proof. Let $a < c, b < d$. Suppose $a_1, a_2, b_1, b_2, c_1, c_2, d_1, d_2$ and $\lambda_1, \mu_1, \lambda_2$ and μ_2 are defined according to the method of bilinear interpolation.

Then

$$C(a,b) = \frac{a_2-a}{a_2-a_1}\frac{b_2-b}{b_2-b_1}C'(a_1,b_1) + \frac{a_2-a}{a_2-a_1}\frac{b-b_2}{b_2-b_1}C'(a_1,b_2) + \frac{a-a_2}{a_2-a_1}\frac{b_2-b}{b_2-b_1}C'(a_2,b_1) + \frac{a-a_2}{a_2-a_1}\frac{b-b_2}{b_2-b_1}C'(a_1,b_1).$$

Then

$$c(a,b) = \frac{\partial^2 C(a,b)}{\partial a \partial b} = \frac{C'(a_1,b_1) - C'(a_1,b_2) - C'(a_2,b_1) + C'(a_2,b_2)}{(a_2-a_1)(b_2-b_1)}.$$

To show (i), we shall consider the following cases,

Case 1. suppose $a_1 = c_1$, $a_2 = c_2$, $b_1 = d_1$ and $b_2 = d_2$, in this case, $c(a,b) = c(c,d) = c(a,d) = c(c,b)$. Thus $c(a,b)c(c,d) \geq c(a,d)c(c,b)$ holds.

Case 2. suppose $a_1 < c_1, a_2 < c_2, b_1 = d_1$ and $b_2 = d_2$. then we have $c(a,b) = c(c,b)$ and $c(a,d) = c(c,d)$. Thus $c(a,b)c(c,d) \geq c(a,d)c(c,b)$ holds.

Case 3. suppose $a_1 = c_1$, $a_2 = c_2$, $b_1 < d_1$ and $b_2 < d_2$, the proof follows from Case 1 and Case 2.

Case 4. suppose $a_1 < c_1$, $a_2 < c_2$, $b_1 < d_1$ and $b_2 < d_2$, then,

$$\begin{aligned}
c(a,b)c(c,d) &= \frac{C'(a_1,b_1) - C'(a_1,b_2) - C'(a_2,b_1) + C'(a_2,b_2)}{(a_2-a_1)(b_2-b_1)} \\
&\quad \times \frac{C'(c_1,d_1) - C'(c_1,d_2) - C'(c_2,d_1) + C'(c_2,d_2)}{(c_2-c_1)(d_2-d_1)} \\
&= \frac{c'(a_2,b_2)}{(a_2-a_1)(b_2-b_1)} \frac{c'(c_2,d_2)}{(c_2-c_1)(d_2-d_1)} \\
&\geq \frac{c'(a_2,d_2)}{(a_2-a_1)(b_2-b_1)} \frac{c'(c_2,b_2)}{(c_2-c_1)(d_2-d_1)} \\
&= \frac{C'(a_1,d_1) - C'(a_1,d_2) - C'(a_2,d_1) + C'(a_2,d_2)}{(a_2-a_1)(b_2-b_1)} \\
&\quad \times \frac{C'(c_1,b_1) - C'(c_1,b_2) - C'(c_2,b_1) + C'(c_2,b_2)}{(c_2-c_1)(d_2-d_1)} = c(a,d)c(c,b),
\end{aligned}$$

Note that the inequality above holds because affiliation property of $c'(u,v)$. Therefore, $c(a,b)c(c,d) \geq c(a,d)c(c,b)$ holds. This completes the proof of (i).

For (ii), assume that C' is PQD, and for any $a,b \in [0,1]$,

$$\begin{aligned}
C(a,b) &= (1-\lambda)(1-\mu)C'(a_1,b_1) + (1-\lambda)\mu C'(a_1,b_2) + \lambda(1-\mu)C'(a_2,b_1) + \lambda\mu C'(a_2,b_2) \\
&\geq (1-\lambda)(1-\mu)a_1b_1 + (1-\lambda)\mu a_1b_2 + \lambda(1-\mu)a_2b_1 + \lambda\mu a_2b_2 \\
&= ab,
\end{aligned}$$

The last equality hold since $(1-\lambda)a_1 + \lambda a_2 = a$ and $(1-\mu)b_1 + \mu b_2 = b$, therefore, C is PQD as desired. □

Example 1. For the subcopula C and its mass function c given below:

U \ V	1/3	2/3	1
1/3	1/3	1/3	1/3
2/3	1/3	2/3	2/3
1	1/3	2/3	1

U \ V	1/3	2/3	1
1/3	1/3	0	0
2/3	0	1/3	0
1	0	0	1/3

it is easy to see that C is affiliated so that the corresponding copula constructed by the bilinear extension is also affiliated.

4 Average and Local Measures of Affiliations

Copula characterizes dependence structures and dependence measures. For example, random variables X and Y are independent if and only if their corresponding copula $C(u,v) = uv$. A measure of dependence indicates in some particular manner how closely the random variables X and Y are related; Hence a variety of measures are needed to reveal the nature of affiliation dependence. We review measures of an affiliation discussed in [11]. Let T denote the average measure of the affiliation for $-\infty < x_1 < x_2 < \infty$ and $-\infty < y_1 < y_2 < \infty$, that is,

$$T = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{y_2}\int_{-\infty}^{x_2}[h(x_2,y_2)h(x_1,y_1) - h(x_1,y_2)h(x_2,y_1)]dx_1dy_1dx_2dy_2.$$

Also, it could be defined as average measure for affiliation of copula,

$$T_C = \int_0^1\int_0^1\int_0^{v_2}\int_0^{u_2}[c(u_2,v_2)c(u_1,v_1) - c(u_1,v_2)c(u_2,v_1)]du_1dv_1du_2dv_2.$$

After some calculation, we can get $\frac{1}{2}\tau = T$, where τ is Kendall's τ(See [17]).

For discrete copula, we give the following discrete average measure,

$$T_C = \sum_{i=0}^{n}\sum_{j=0}^{m}\sum_{k=0}^{i}\sum_{l=0}^{j}[c(u_k,v_l)c(u_i,v_j) - c(u_i,v_l)c(u_k,v_j)]. \tag{5}$$

Holland and Wang[5, 6] defined the local dependence index for affiliation as

$$\gamma(x,y) = \frac{\partial^2 \log h(x,y)}{\partial x \partial y}.$$

Also, it can be defined for copula

$$\gamma(u,v) = \frac{\partial^2 \log c(u,v)}{\partial u \partial v}.$$

We list several properties of this local measure of affiliation:

(i) $-\infty < \gamma(u,v) < \infty$.

(ii) $\gamma(u,v) = 0$ for all u,v if and only if U and V are independent.

(iii) If X and Y have a bivariate normal distribution with correlation coefficient ρ, then $\gamma(x,y) = \frac{\rho}{1-\rho^2}$, a constant.

Example 2. Consider the experiment of tossing an unbalanced coin 3 times with success rate p. Let X be the total number of heads observed and Y be the number of heads on the second toss. Then the joint density of X and Y is given by

Y \ X	0	1	2	3
0	$(1-p)^3$	$2(1-p)^2p$	$(1-p)p^2$	0
1	0	$(1-p)^2p$	$2p^2(1-p)$	p^3

The corresponding copula density of U and V is

V \ U	$(1-p)^3$	$(1-p)^2(1+2p)$	$1-p^3$	1
$1-p$	$(1-p)^3$	$2(1-p)^2p$	$(1-p)p^2$	0
1	0	$(1-p)^2p$	$2p^2(1-p)$	p^3

Note that U and V are affiliated, its bilinear interpolation is also affiliation. By (5), the average measure for this discrete copula is $T=(1-p)^5p+8p^3(1-p)^3+2(1-p)^2p^4+(1-p)p^5$. Note that if the coin is balanced then $T=3/16$.

Example 3. Consider the experiment of tossing a unbalanced coin with success rate p. Let X be value 2^K, where K is number of tosses until the first head occurs, and Y be the number of heads in the first toss. Note that $E(X)$ does not exists for $p<1/2$. For $p\in[0,1]$, the joint distribution of X and Y is

Y \ X	2^0	2^1	...	2^n	...
0	0	$p(1-p)$	...	$p(1-p)^{n-1}$	...
1	p	0	...	0	...

and its corresponding copula is,

V \ U	p,	$p(1-p)+p$,	...,	$(1-p)(1-(1-p)^{n-1})+p$,	...,	1
p	0,	$p(1-p)$,	...,	$p(1-p)^{n-1}$,	...,	p
1	p,	$p(1-p)+p$,	...,	$p(1-p)^{n-1}+p$,	...,	1

This is a discrete copula which is not PQD. The average measure for this discrete copula $T=-p^2(1-p)-p^2(1-p)^2-\cdots-p^2(1-p)^n-\cdots=-p(1-p)<0$.

5 Conditions on Affiliation in the Bivariate Skew Normal Family

We first recall the definition of the multivariate skew-normal distribution which are given in [2]. A k-dimensional random variable Z is said to have a multivariate skew-normal distribution if it is continuous with density function

$$2\phi_k(\mathbf{z};\Sigma)\Phi(\alpha^T\mathbf{z}), \qquad \mathbf{z}\in\mathbb{R}^k,$$

where $\phi_k(\mathbf{z};\Sigma)$ is the k-dimensional normal density with zero mean and correlation matrix Σ, $\Phi(\cdot)$ is the CDF of $N(0,1)$, and α is a k-dimensional vector. Here we only consider the case where $k=2$. The density of (X,Y) is given by

$$h(x,y) = 2\phi_\rho(x,y)\Phi(\alpha_1 x+\alpha_2 y), \tag{6}$$

where

$$\phi_\rho(x,y) = (2\pi)^{-1}(1-\rho^2)^{-1/2}\exp\left\{\frac{1}{2(1-\rho^2)}(x^2-2\rho xy+y^2)\right\},$$

and α_1 and $\alpha_2\in\mathbb{R}$ are skewness parameters.

Theorem 3. *Consider the bivariate skew normal random vector (X,Y) with density given by (6). Then X and Y are affiliated if and only if $\rho\geq 0$ and $\alpha_1\alpha_2\leq 0$.*

Proof. For the "*if*" part, assume that $\rho\geq 0$ and $\alpha_1\alpha_2\leq 0$. By Lemma 3.5 of Rinott and Scarsini [20], we know that $\rho\geq 0$ implies that bivariate normal density is affiliated. That is

$$\phi_\rho(x',y)\phi_\rho(x,y')\leq\phi_\rho(x',y')\phi_\rho(x,y) \qquad \text{for all } x'<x \text{ and } y'<y. \tag{7}$$

Now it is sufficient to show that X and Y in $\Phi(\alpha_1 x+\alpha_2 y)$ are affiliated. Without loss of generality, we assume that $\alpha_1<0,\alpha_2>0$. For any $x'<x, y'<y$, we have

$$\alpha_1 x+\alpha_2 y'\leq\alpha_1 x'+\alpha_2 y'\leq\alpha_1 x'+\alpha_2 y$$

and

$$\alpha_1 x+\alpha_2 y'\leq\alpha_1 x+\alpha_2 y\leq\alpha_1 x'+\alpha_2 y.$$

Since Φ is log concave, we have

$$\frac{\log\Phi(\alpha_1 x'+\alpha_2 y)-\log\Phi(\alpha_1 x'+\alpha_2 y')}{\alpha_2(y-y')}\leq\frac{\log\Phi(\alpha_1 x+\alpha_2 y)-\log\Phi(\alpha_1 x+\alpha_2 y')}{\alpha_2(y-y')},$$

which implies

$$\log\Phi(\alpha_1 x'+\alpha_2 y)-\log\Phi(\alpha_1 x'+\alpha_2 y')\leq\log\Phi(\alpha_1 x+\alpha_2 y)-\log\Phi(\alpha_1 x+\alpha_2 y').$$

Thus

$$\log\left[\Phi(\alpha_1 x' + \alpha_2 y)\Phi(\alpha_1 x + \alpha_2 y')\right] \leq \log\left[\Phi(\alpha_1 x + \alpha_2 y)\Phi(\alpha_1 x' + \alpha_2 y')\right],$$

which is reduced to

$$\Phi(\alpha_1 x' + \alpha_2 y)\Phi(\alpha_1 x + \alpha_2 y') \leq \Phi(\alpha_1 x + \alpha_2 y)\Phi(\alpha_1 x' + \alpha_2 y'). \tag{8}$$

Combining (7) and (8), we obtain

$$\phi_\rho(x',y)\Phi(\alpha_1 x' + \alpha_2 y)\phi_\rho(x,y')\Phi(\alpha_1 x + \alpha_2 y') \leq \phi_\rho(x,y)\Phi(\alpha_1 x + \alpha_2 y)\phi_\rho(x',y')\Phi(\alpha_1 x' + \alpha_2 y'),$$

so that X and Y are affiliated.

For the "$only\,if$" part, assume that X and Y are affiliated. It is suffices to show that if conditions $\rho \geq 0$ and $\alpha_1\alpha_2 < 0$ are not satisfied, then there exist $x' < x$, $y' < y$ such that

$$\phi_\rho(x,y)\Phi(\alpha_1 x + \alpha_2 y)\phi_\rho(x',y')\Phi(\alpha_1 x' + \alpha_2 y') < \phi_\rho(x',y)\Phi(\alpha_1 x' + \alpha_2 y)\phi_\rho(x,y')\Phi(\alpha_1 x + \alpha_2 y')$$

which is equivalent to

$$\frac{\rho}{1-\rho^2}(x-x')(y-y') + \log\left[\frac{\Phi(\alpha_1 x + \alpha_2 y)\Phi(\alpha_1 x' + \alpha_2 y')}{\Phi(\alpha_1 x' + \alpha_2 y)\Phi(\alpha_1 x + \alpha_2 y')}\right] < 0. \tag{9}$$

Now consider the following cases.

Case 1. For $\rho < 0$ and $\alpha_1\alpha_2 \leq 0$, without loss of generality, we assume that $\alpha_1 \leq 0, \alpha_2 \geq 0$, if we pick $x' = y' = 0$, and $y = \exp(x)$, then (9) is reduced to

$$\begin{aligned}&\frac{\rho}{1-\rho^2}x\exp(x) + \log\left[\frac{\frac{1}{2}\Phi(\alpha_1 x + \alpha_2\exp(x))}{\Phi(\alpha_2\exp(x))\Phi(\alpha_1 x)}\right]\\ &\quad= \frac{\rho}{1-\rho^2}x\exp(x) - \log\left(\Phi(\alpha_1 x)\right) + \log\left[\frac{\frac{1}{2}\Phi(\alpha_1 x + \alpha_2\exp(x))}{\Phi(\alpha_2\exp(x))}\right],\end{aligned}$$

which goes to $-\infty$ as x tends to ∞.

Case 2. For $\rho > 0$ and $\alpha_1\alpha_2 > 0$, without loss of generality, we assume that $\alpha_1 > 0, \alpha_2 > 0$, if we pick $x' = y' = 0$, and $y = 1$, then (9) is reduced to

$$\frac{\rho}{1-\rho^2}x + \log\left[\frac{\frac{1}{2}\Phi(\alpha_1 x + \alpha_2)}{\Phi(\alpha_2)\Phi(\alpha_1 x)}\right],$$

which goes to $-\infty$ as x tends to ∞.

Case 3. For $\rho < 0$ and $\alpha_1\alpha_2 > 0$, then the first part and second part of (9) are all strictly negative, therefore, the desired result follows. □

Acknowledgments. The authors would like to thank Professor Hung T. Nguyen for introducing this interesting topic to us and anonymous referees for their helpful comments which led to improvement of this paper.

References

1. Azzalini, A., Dalla Valle, A.: The multivariate skew-normal distribution. Biometrika 83, 715–726 (1996)
2. Azzalini, A., Capitanio, A.: Statistical applications of the multivariate skew normal distribution. Journal of the Royal Statistical Society: Series B 61, 579–602 (1999)
3. Azzalini, A.: The Skew-Normal Probability Distribution and related distributions, such as the skew-*t*, `http://azzalini.stat.unipd.it/SN/`
4. De Castro, L.I.: Equilibrium existence and Revenue ranking of auctions. Mimeo, University Carlos III (2007)
5. Holland, P.W., Wang, Y.J.: Regional dependence for continuous bivariate densities. Communications in Statistics-Theory and Methods 16, 193–206 (1987a)
6. Holland, P.W., Wang, Y.J.: Dependence function for continuous bivariate densities. Communications in Statistics-Theory and Methods 16, 863–876 (1987b)
7. Joe, H.: Multivariate Models and Dependence Concepts. Chapman & Hall, London (1997)
8. Johnson, N.L., Kott, S.: On some generalized Farlie-Gumbel-Morgenstern distributions. Communications in Statistics 4(5), 415–427 (1975)
9. Karlin, S., Rinott, Y.: Classes of orderings of measures and related correlation inequalities, I. Multivariate totally positive distributions. Journal of Multivariate Analysis 10, 467–498 (1980)
10. Lai, C.D., Xie, M.: A new family of positive quadrant dependent bivariate distributions. Statistics and Probability Letters 46, 359–364 (2000)
11. Balakrishnan, N., Lai, C.D.: Continuous Bivariate Distributions, 2nd edn. Rumsby Scientific Publishing, Adelaide (2009)
12. Mayor, G., Suner, J., Torrens, J.: Sklar's theorem in finite settings. IEEE transactions on Fuzzy Systems 15(3), 410–416 (2007)
13. Milgrom, P.R., Weber, R.J.: A theory of auctions and competitive bidding. Econometrica 50(5), 1089–1122 (1982)
14. Monteiro, P.K., Moreira, H.: First-price auctions without affiliation. Economics Letters 91(1), 1–7 (2006)
15. Muller, A., Scarsini, M.: Archimedean copulae and positive dependence. Applied Mathematics Working Paper Series 25 (2003)
16. Nelsen, R.B.: An Introduction to Copulas. Springer, New York (2006)
17. Nelsen, R.B.: On measures of association as measures of positive dependence. Statistics and Probability Letters 14, 269–274 (1992)
18. Nguyen, H.T.: Statistics with Copulas: An Inviation to Modern Statistics (class notes) (2012)
19. Pinkse, J., Tan, G.: The affiliation effect in first-price auctions. Econometrica 73, 263–277 (2005)
20. Rinott, Y., Scarsini, M.: Total positivity order and the normal distribution. Journal of Multivariate Analysis 97, 1251–1261 (2006)
21. Rodriguez, G.E.: First price auctions: Monotonicity and uniqueness. International Journal of Game Theory 29(3), 413–432 (2000)

22. Roncalli, T.: Gestion des risques multiples, Technical report, Groupe de Recherche Opréationelle, Crédit Lyonnais
23. Sklar, A.: Fonctions de r*é*partition *à n* dimensions et leurs marges. Publ. Inst. Statist. Univ., Paris 8, 8229–8231 (1959)
24. Li, T., Zhang, B.: Testing for affiliation in first-price auctions using entry behavior. International Economic Review 51(3), 837–849 (2010)
25. Tong, Y.L.: Probability Inequalities in Multivariate Distributions. Academic Press, New York (1980)
26. Wang, T., Li, B., Gupta, A.K.: Distribution of quadratic forms under skew normal settings. Journal of Multivariate Analysis 100, 533–545 (2009)

Pairs Trading via Three-Regime Threshold Autoregressive GARCH Models

Cathy W.S. Chen*, Max Chen, and Shu-Yu Chen

Abstract. Pairs trading is a popular strategy on Wall Street. Most pairs trading strategies are based on a minimum distance approach or cointegration method. In this paper, we propose an alternative model to the process of pair return spread. Specifically, we model the return spread of potential stock pairs as a three-regime threshold autoregressive model with GARCH effects (TAR-GARCH), and the upper and lower regimes in the model are used as trading entry and exit signals. An application to the Dow Jones Industrial Average Index stocks is presented.

1 Introduction

Pairs trading is a popular strategy on Wall Street. It became well known by the Quant team led by Nunzio Tartagli from Morgan Stanley in the mid-1980s. The main principle underlying pairs trading is the simple idea of reversion, which is the process of identifying two stocks whose prices move together closely. When the spread between them widens, short the high price one and long the low price one. If the past is a good mirror of the future, then prices will converge and the pairs trading will result in profit. There are many ways to find stocks which are moving together. The simplest one is the Minimum Squared Distance method (MSD), which involves calculating the sum of squared deviations between two normalized stock prices and choosing the one which has the minimum value. Gatev, Goetzmann and Rouwenhorst (2006) give detailed results using US CRSP stock prices. Do and Faff (2010) examine the validity of MSD in more recent datasets. Another way to find stocks which are moving together utilizes the equilibrium relation among stocks,

Cathy W.S. Chen · Shu-Yu Chen
Department of Statistics, Feng Chia University, Taiwan
e-mail: chenws@mail.fcu.edu.tw

Max Chen
Department of Finance, Ming Chuan University, Taiwan

* Corresponding author.

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_8,

referred to as the cointegration approach. A group of nonstationary stock prices can have a common stochastic trend (cf. Engel and Granger, 1987). Vidyamurthy (2004) describes how to apply this method to pairs trading. Kawasaki, Tachiki, Udaka, and Hirano (2003) use this cointegration method to find stock pairs in the Tokyo Stock Exchange. Perlin (2009) applies it to the Brazilian stock market. The third way models the spread as a mean-reverting a Gaussian Markov chain and trading is triggered when the forecasting spread is different from the subsequent spread in a significant level. Elliott, van der Hoek, and Malcolm (2005) show how to estimate this model in detail. To obtain appropriate investment decisions, observations of the spread are compared with predictions from calibrated model.

In previous literature, the pair spread is assumed to be a single regime process. However, practitioners often find that the pair spread seems to switch between different regimes, and the usual pairs trading methods fail to identify potential arbitrage opportunities. A recent empirical study by Bock and Mestel (2009) proposes a two-state, first-order Markov-switching process to model the spread and apply it to their trading rules. Since financial time series often exhibit some stylized facts such as volatility clustering, asymmetry in conditional mean and variance, mean reversion, and fat-tailed distributions, it is important to develop an appropriate model which can capture these stylized facts.

To capture the dynamic features of volatility, the popular choices are the autoregressive conditional heteroscedastic (ARCH) and generalized ARCH (GARCH) models of Engle (1982) and Bollerslev (1986), which allow the conditional volatility to be predicted from its lagged terms and past news. Both ARCH and GARCH models are widely employed for describing dynamic volatility in financial time series. Bollerslev, Chou, and Kroner (1992) advocate that a GARCH(1,1) model would usually be sufficient for most financial time series.

In this paper, we propose an alternative: a three-regime threshold nonlinear GARCH model, with a fat-tailed error distribution (TAR-GARCH), to capture mean and volatility asymmetries in financial markets. The salient feature of this model is that it can capture asymmetries in the average return, volatility level, mean reversion, and volatility persistence. For a brief review of the TAR model in finance, refer to Chen, So, and Liu (2011). We employ a Bayesian method, based on Markov chain Monte Carlo (MCMC) methods, allowing simultaneous inference for all unknown parameters in a TAR-GARCH model.

The remainder of this study proceeds as follows. Section 2 introduces the three-regime TAR-GARCH model with a fat-tailed error distribution which is applied to identify pairs trading signals. Bayesian estimation is also briefly discussed in this section. Section 3 presents some results for stocks from the Dow Jones 30 index. These stocks are the most liquid stocks in the US market that traders can buy and sell at any time. Conclusions are presented in Section 4.

2 Methodology

We would like to model the return spread of potential stock pairs as a three-regime threshold autoregressive model with GARCH effects (TAR-GARCH), and the upper and lower regimes in the model are used as trading entry and exit signals.

2.1 Threshold AR Model with GARCH Effect

Li and Li (1996) model both mean and volatility asymmetry in a double threshold (DT-)ARCH model; Brooks (2001) further generalizes this to a double threshold GARCH model. Chen, Chiang, and So (2003) further allow an exogenous threshold variable (U.S. market news) and nonlinear mean spill-over effects. Chen and So (2006) propose a threshold heteroskedastic model to capture the mean and variance asymmetries which allows the threshold variable to be formulated with auxiliary variables. This avoids subjectively choosing the threshold variable and enables the relative importance of the auxiliary variables to be examined after model fitting. Most of these studies focus only on two-regime models. The model used here is a three-regime threshold nonlinear GARCH model, with a fat-tailed error distribution, to capture mean and volatility asymmetries in financial markets, which has been studied by Chen, Gerlach, and Lin (2010). This model is characterized by several non-linear factors commonly observed in practice, such as asymmetry in declining and rising patterns of a process. In fact, all mean and volatility parameters are allowed to change between regimes. Due to its complexity, Bayesian estimation and inference for this class of model is considered.

The three-regime model is specified as:

$$
\begin{aligned}
y_t &= \begin{cases} \phi_0^{(1)} + \phi_1^{(1)} y_{t-1} + a_t, & y_{t-d} < c_1 \\ \phi_0^{(2)} + \phi_1^{(2)} y_{t-1} + a_t, & c_1 \le y_{t-d} < c_2 \\ \phi_0^{(3)} + \phi_1^{(3)} y_{t-1} + a_t, & y_{t-d} \ge c_2 \end{cases} \\
a_t &= \sqrt{h_t}\varepsilon_t, \qquad \varepsilon_t \overset{\text{i.i.d.}}{\sim} t_\nu^*, \\
h_t &= \begin{cases} \alpha_0^{(1)} + \alpha_1^{(1)} a_{t-1}^2 + \beta_1^{(1)} h_{t-1}, & y_{t-d} < c_1 \\ \alpha_0^{(2)} + \alpha_1^{(2)} a_{t-1}^2 + \beta_1^{(2)} h_{t-1}, & c_1 \le y_{t-d} < c_2 \\ \alpha_0^{(3)} + \alpha_1^{(3)} a_{t-1}^2 + \beta_1^{(3)} h_{t-1}, & y_{t-d} \ge c_2, \end{cases}
\end{aligned} \tag{1}
$$

where c_1 and c_2 are the threshold values that satisfy $-\infty = c_0 < c_1 < c_2 < c_3 = \infty$; h_t is $\mathrm{Var}(y_t|y_1,\ldots,y_{t-1})$; the integer d is the threshold lag; t_ν^* is a standardized Student-t error distribution with a mean of zero and a variance of one. Some standard restrictions on the variance parameters are given.

$$
\alpha_0^{(j)} > 0,\ \alpha_1^{(j)}, \beta_1^{(j)} \ge 0 \qquad \text{and} \qquad \alpha_1^{(j)} + \beta_1^{(j)} < 1, \tag{2}
$$

The lagged return y_{t-1} is included in the model (1) in order to test zero serial correlations. We would like to know whether the series has a statistically significant

lag-1 autocorrelation which indicates the lagged returns might be useful in predicting y_t. When the lag-one autocorrelation is not statistically significant, it indicates that potential pair arbitrage opportunities may not exist.

Bayesian estimation requires the specification of a likelihood and prior distributions on the model parameters. We select prior distributions that are mostly uninformative, so that the data dominates inference via the likelihood.

2.2 Priors and Likelihood

Let the full parameter vector be denoted as $\theta = (\phi_1, \phi_2, \alpha_1, \alpha_2, c, \nu, d)'$, where $\phi_j = (\phi_0^{(j)}, \phi_1^{(j)})$, $\alpha_j = (\alpha_0^{(j)}, \alpha_1^{(j)}, \beta_1^{(j)})'$, and $c = (c_1, c_2)'$ and d_0 denotes the maximum delay lag. The conditional likelihood function of the model is:

$$L(\theta \mid y) = \prod_{t=2}^{n} \left\{ \sum_{j=1}^{2} \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{(\nu-2)\pi}} \frac{1}{\sqrt{h_t}} \left[1 + \frac{(y_t - \mu_t)^2}{(\nu-2)h_t}\right]^{-\frac{\nu+1}{2}} I_{jt} \right\}, \quad (3)$$

where $\mu_t = \phi_0^{(j)} + \phi_1^{(j)} y_{t-1}$ and I_{jt} is an indicator variable of $I(c_{j-1} \le y_{t-d} < c_j)$.

Our prior settings are similar to those used by Chen, Gerlach, and Lin (2010). A Gaussian prior distribution is assumed for $\phi_j \sim N(\phi_{j0}, V_j)$, constrained for mean stationarity, where $\phi_{j0} = 0$ and V_j^{-1} is a matrix with 'large' numbers on the diagonal. With the maximum delay d_0, we assume a discrete uniform prior $p(d) = \frac{1}{d_0}$ for d. To ensure the required constraint equation (2) on $p(\alpha_j)$, we adopt a uniform prior $p(\alpha_j)$ over the region which is the indicator $I(S_j)$, $j = 1, 2$, where S_j is the set of α_j that satisfies the restriction in (2). The prior for the threshold parameters c, as proposed by Chen, Gerlach, and Lin (2010), for three regimes is:

$$c_1 \sim \text{Unif}(lb_1, ub_1) \ ; \ c_2 | c_1 \sim \text{Unif}(lb_2, ub_2),$$

where we can set $lb_1 = h$ and $ub_1 = (1 - 2h)$. If we choose $h = 0.1$, then $c_1 \in (0.1, 0.8)$. Further, set $ub_2 = (1 - h)$ and $lb_2 = c_1 + h$. The prior for (c_1, c_2) is flat over the region ensuring $c_1 + h \le c_2$ and at least $100h\%$ of observations are contained in each regime. Finally, the degrees of freedom ν is re-parametrised to $\nu^* = \nu^{-1}$ with uniform prior $I(\nu^* \in [0, 0.25])$, this restricts $\nu > 4$, so that the first four moments of the error distribution are finite.

We assume a prior independence among the groupings $\phi_1, \phi_2, \alpha_1, \alpha_2, c$, ν, and d. Multiplication of each prior followed by the conditional likelihood function in (3) leads to the conditional posterior density for each parameter group. Detailed conditional posteriors can be found in Chen, Gerlach, and Lin (2010). Except for parameter d, the conditional posterior distributions for each remaining parameter group are non-standard. We thus incorporate the Metropolis and Metropolis-Hastings (MH) methods to draw the MCMC iterates for the other parameter groups, see Chen and So (2006) for the discussion on the MH method. To speed up convergence and allow optimal mixing, we employ an adaptive MH-MCMC algorithm that combines a random walk Metropolis and an independent kernel MH algorithm. We extensively

examine trace plots and autocorrelation function (ACF) plots from multiple runs of the MCMC sampler for each parameter to confirm convergence and to infer adequate coverage. We set the MCMC sample size N sufficiently large, discarding the burn-in iterates, and keep the last $N - M$ iterates for inference.

2.3 Pairs Selection

The series of returns are calculated by taking differences of the logarithms of the daily closing price, $r_t^j = \ln(P_t^j / P_{t-1}^j)$, where P_t^j is the closing price index of asset j on day t. Our implementation of pairs trading has two stages. The procedure is given as follows:

Stage 1: We calculate the MSD between the two normalized price series among the pairs. The formula of MSD is given as follows.

$$MSD = \sum_{t=1}^{n} (P_t^A - P_t^B)^2, \tag{4}$$

where P_t^i is normalized price of asset i at time t. Five pairs are selected with the smallest MSD.
Stage 2: We calculate the return spread between the selected pairs, $y_t = r_t^A - r_t^B$, and fit a three-regime TAR model with GARCH effect to the return spread. Once the model is fitted, the upper and lower threshold values in the model are used as trading entry and exit signals.

We open a position in a pair when the pair return spread (y_t) is larger (smaller) than the high (low) threshold value, as estimated by the TAR-GARCH model, that is, we short (long) A and long (short) B. We unwind the position when the pair return spread crosses over the same threshold value again. If the spread doesn't cross before the end of the last trading day of the trading period, gains or losses are calculated at the end of the last trade of the trading period.

The average trading return on the short stock A and long stock B position is calculated as follows:

$$r_1 = \frac{1}{D}\left[-\ln\frac{P^A_{sold}}{P^A_{bought}} + \ln\frac{P^B_{sold}}{P^B_{bought}}\right], \tag{5}$$

where D stands for the number of holding days. On the other hand, the average trading return on the long stock A and short stock B position is given as follows.

$$r_2 = \frac{1}{D}\left[\ln\frac{P^A_{sold}}{P^A_{bought}} - \ln\frac{P^B_{sold}}{P^B_{bought}}\right], \tag{6}$$

where D stands for the number of holding days.

3 Empirical Results

The daily close prices (adjusted for dividends and splits) of constituents of the Dow Jones Industrial Average Index (DJIA) are used as an illustration. The data are obtained from Yahoo Finance US over a 7-year time period, from January 2, 2006 to May 31, 2013. The in-sample period of this study is from January 2, 2006 to February 28, 2013 and the out-of-sample period is from March 1, 2013 to May 31, 2013.

The companies that comprise the DJIA are 3M (MMM), Alcoa (AA), American Express (AXP), AT&T (T), Bank of America (BAC), Boeing (BA), Caterpillar (CAT), Chevron Corporation (CVX), Cisco Systems (CSCO), Coca-Cola (KO), Dupont (DD), ExxonMobil (XOM), General Electric (GE), Hewlett-Packard (HPQ), The Home Depot (HD), Intel (INTC), IBM (IBM), Johnson & Johnson (JNJ), JPMorgan Chase (JPM), McDonald's (MCD), Merck (MRK), Microsoft (MSFT), Pfizer (PFE), Procter & Gamble (PG), Travelers (TRV), UnitedHealth Group Incorporated (UNH), United Technologies Corporation (UTX), Verizon Communications (VZ), Wal-Mart (WMT), and Walt Disney (DIS).

Table 1 gives the descriptive statistics of the DJIA stock prices. Table 1 shows that 13 companies have a standard deviation of greater than 10, and IBM has the highest standard deviation, 41.67. A volatile stock will have a high standard deviation. Figure 1 shows the time series plots of 30 DJIA company's daily closing prices.

We calculate the MSD between the two normalized price series, and the number of possible pairs is 435 (i.e. C_2^{30}). The five pair trading candidates are

Pair 1: Caterpillar (CAT) vs Chevron Corporation (CVX)
Pair 2: IBM (IBM) vs Johnson & Johnson (JNJ)
Pair 3: Merck (MRK) vs Microsoft (MSFT)
Pair 4: Verizon (VZ) vs Wal-Mart Stores Inc.(WMT)
Pair 5: Wal-Mart Stores Inc. (WMT) vs The Walt Disney Company (WAL).

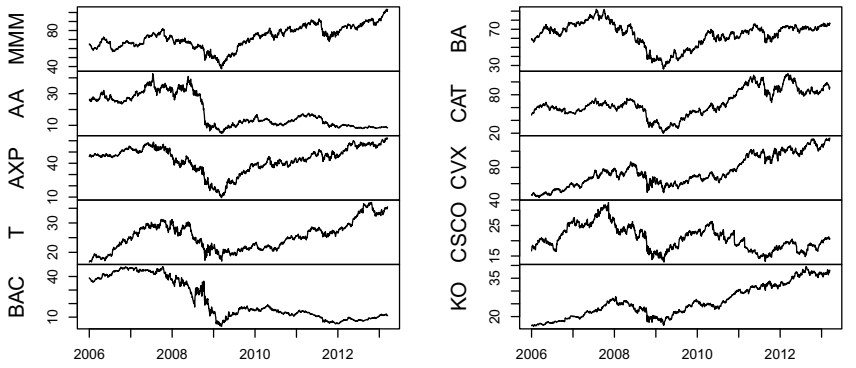


Fig. 1 The times series plots of 30 DJ company's prices

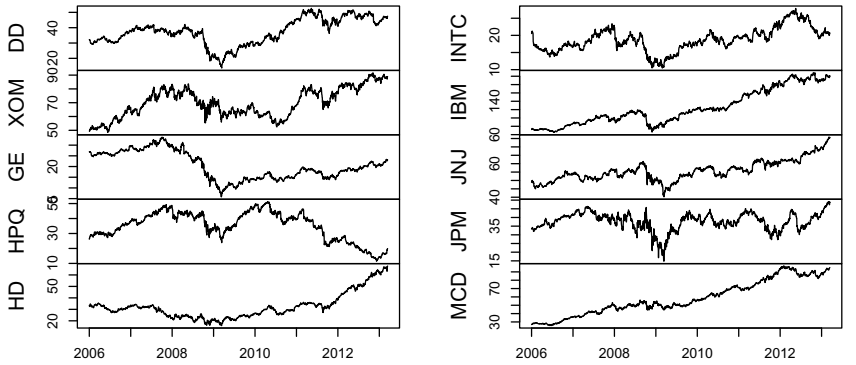


Fig. 1 (*continued*)

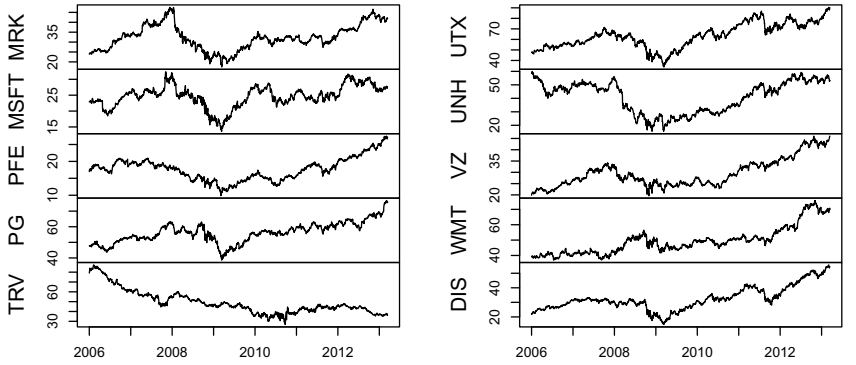


Fig. 1 (*continued*)

The MSD values of the five pairs are shown in Table 2. All five return spread series exhibit the standard property: they have fat-tailed distributions, as indicated by the highly significant Jarque-Bera normality test statistics, a joint test for the absence of skewness and kurtosis which are given in Table 2. Descriptive statistics of the DJIA stock returns are given in Table 3. Table 3 shows that the returns of all companies are close to zero, and this result coincides with the mean reversion theory. Furthermore, there are 6 companies' stock index returns that are negative.

We then fit a three-regime TAR model with GARCH effect to these five selected pairs' return spreads. Allowing y_{t-1} in the conditional mean helps account for possible autocorrelations in the pairs' return spreads. Once the model is fitted, the upper and lower threshold values in the model are used as trading entry and exit signals. When the return spread is above (below) the upper (lower) threshold value, we then short (long) one share A stock and long (short) one share B stock. Once the position is open and the spread falls back to the threshold, the position is closed.

Table 1 The descriptive statistics of the DJIA stock prices from January 2, 2006 to February 28, 2013

Company	Symbol	Mean	Maximun	Minum	Std
3M	MMM	72.30	103.59	37.40	12.41
Alcoa	AA	18.50	42.63	4.99	9.84
American Express	AXP	43.59	62.38	9.47	11.35
AT&T	T	25.37	36.98	16.69	4.75
Bank of America	BAC	22.51	47.13	3.09	14.58
Boeing	BA	64.12	92.06	26.24	13.33
Caterpillar	CAT	67.05	112.91	19.87	20.77
Chevron Corporation	CVX	75.19	116.21	42.71	18.95
Cisco Systems	CSCO	20.78	32.55	13.01	3.96
Coca -Cola	KO	25.76	39.45	16.13	6.27
DuPont	DD	37.33	52.57	13.68	8.51
ExxonMobil	XOM	69.90	91.67	48.45	10.49
General Electric	GE	20.38	33.71	5.83	6.42
Technology	HPQ	35.06	51.30	11.46	9.51
The Home Depot	HD	32.36	67.79	15.80	11.10
Intel	INTC	18.73	27.86	10.47	3.47
IBM	IBM	123.84	208.22	65.28	41.68
Johnson&Johnson	JNJ	55.72	75.78	40.23	6.19
JPMorgan Chase	JPM	37.59	49.14	14.78	5.15
MeDonald's	MCD	58.55	97.03	25.57	20.45
Merck	MRK	32.28	47.56	17.45	6.30
Microsoft	MSFT	24.62	32.39	13.62	3.45
Pfizer	PFE	17.66	27.48	9.85	3.35
Procter & Gamble	PG	56.10	76.82	38.60	6.51
Travlers	TRV	49.24	87.37	26.74	12.17
United Technologies Corp.	UTX	63.43	90.51	33.84	11.76
UnitedHealth Group Incorporated	UNH	40.95	59.75	15.51	11.55
Verizon Communications Inc.	VZ	29.54	46.05	19.63	6.26
Wal-Mart Stores Inc.	WMT	49.34	75.78	36.96	8.91
The Walt Disney Company	DIS	32.52	55.73	14.77	8.19

Table 2 Stock pairs with the smallest MSD and Jarque-Bera test for pair return spreads

Pairs	Company A	Company B	MSD	Jarque-Bera test	
				Statistic	p-value
1	Caterpillar(CAT)	Chevron Corporation(CVX)	454.94	2363.27	0.00
2	IBM (IBM)	Johnson & Johnson(JNJ)	649.18	5515.45	0.00
3	Merck (MRK)	Microsoft(MSFT)	661.11	7845.45	0.00
4	Verizon(VZ)	Wal-Mart Stores Inc.(WMT)	681.79	4031.88	0.00
5	Wal-Mart Stores Inc.(WMT)	The Walt Disney Company(WAL)	701.61	1755.15	0.00

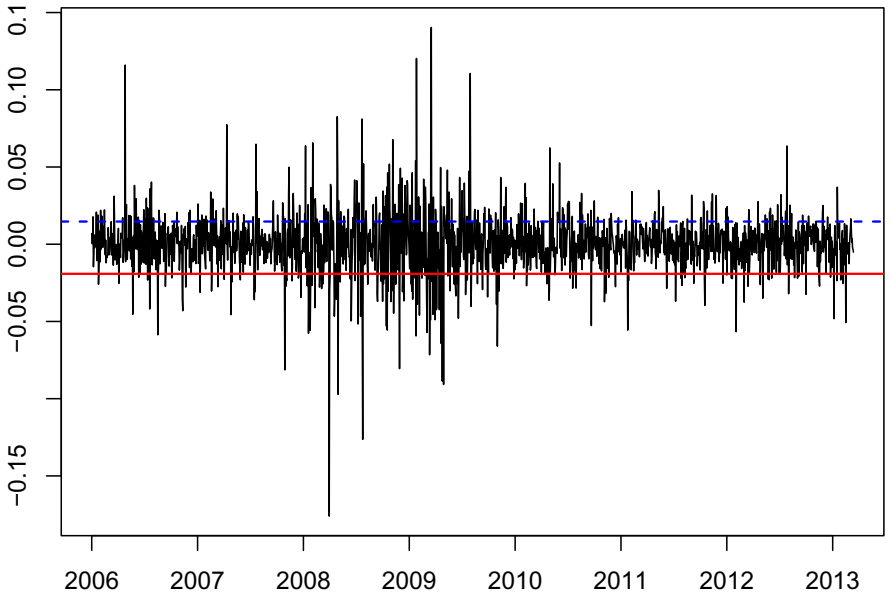


Fig. 2 MRK-MSFT pair return spread from January 2, 2006 to February 28, 2013

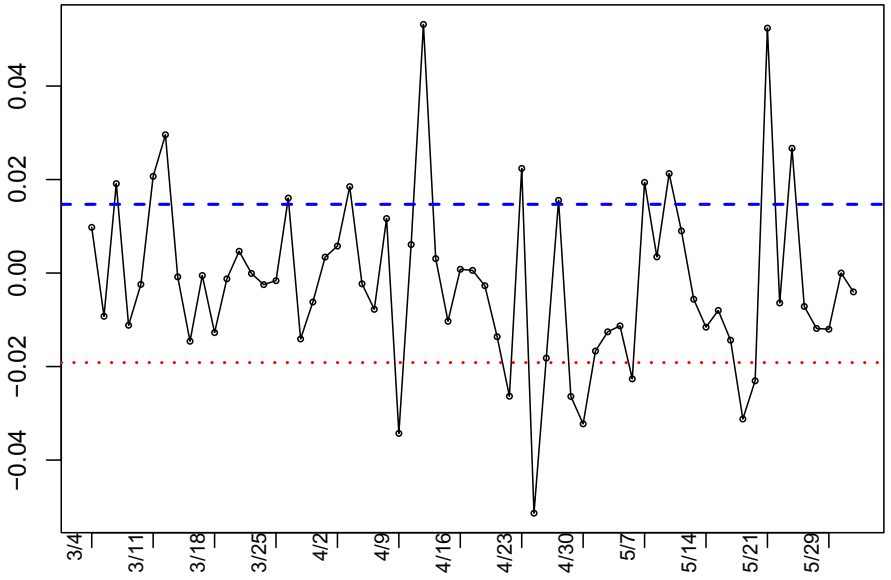


Fig. 3 MRK-MSFT pair return spread from March 1, 2013 to May 31, 2013

Table 3 The descriptive statistics of the DJIA stock log returns from January 2, 2006 to February 28, 2013

Company	Symbol	Mean	Maximun	Minum	Std
3M	MMM	0.0003	0.0941	−0.0938	0.0156
Alcoa	AA	−0.0006	0.2077	−0.1742	0.0315
American Express	AXP	0.0002	0.1877	−0.1933	0.0283
AT&T	T	0.0004	0.1506	−0.0799	0.0154
Bank of America	BAC	−0.0007	0.3020	−0.3422	0.0422
Boeing	BA	0.0001	0.1437	−0.0805	0.0198
Caterpillar	CAT	0.0004	0.1375	−0.1569	0.0233
Chevron Corporation	CVX	0.0005	0.1895	−0.1333	0.0188
Cisco Systems	CSCO	0.0001	0.1482	−0.1768	0.0212
Coca-Cola	KO	0.0005	0.1298	−0.0907	0.0125
DuPont	DD	0.0002	0.1088	−0.1205	0.0197
ExxonMobil	XOM	0.0003	0.1587	−0.1504	0.0174
General Electric	GE	−0.0001	0.1800	−0.1365	0.0221
Technology	HPQ	−0.0002	0.1353	−0.2238	0.0217
The Home Depot	HD	0.0004	0.1316	−0.0858	0.0193
Intel	INTC	0.0000	0.1119	−0.1318	0.0204
IBM	IBM	0.0006	0.1089	−0.0611	0.0145
Johnson&Johnson	JNJ	0.0002	0.1154	−0.0798	0.0106
JPMorgan Chase	JPM	0.0002	0.2240	−0.2325	0.0315
MeDonald's	MCD	0.0007	0.0897	−0.0830	0.0130
Merck	MRK	0.0003	0.1192	−0.1595	0.0180
Microsoft	MSFT	0.0001	0.1707	−0.1247	0.0185
Pfizer	PFE	0.0003	0.0975	−0.1121	0.0157
Procter & Gamble	PG	0.0003	0.0972	−0.0823	0.0118
Travlers	TRV	−0.0004	0.2005	−0.2274	0.0213
United Technologies Corp.	UTX	0.0004	0.1281	−0.0917	0.0169
UnitedHealth Group Incorporated	UNH	−0.0001	0.2984	−0.2059	0.0244
Verizon Communications Inc.	VZ	0.0005	0.1369	−0.0842	0.0150
Wal-Mart Stores Inc.	WMT	0.0003	0.1051	−0.0840	0.0130
The Walt Disney Company	DIS	0.0005	0.1484	−0.1026	0.0192

We performed 20,000 MCMC iterations and discarded the first 8,000 iterates as a burn-in sample for each data series. The parameter estimates for the model in each selected pairs' return spreads are summarized in Table 4, which include posterior medians and standard deviations (Std.) for each parameter. Note that the c_1 and c_2 are the threshold values. All five pairs return spreads clearly display no series correlation across the three regimes, in response to the lag-one spread return. Since in-sample fitting may contain information for out-of-sample trading, AR lag-one coefficient remains in our proposed model. Regarding volatility persistence ($\alpha_1^{(j)} + \beta_1^{(j)}$), the second regime displays the lowest level of persistence across spreads. We can see from Table 4 that most of the high and low threshold values are opposite signs, indicating possible different regime processes.

Table 5 shows the mean returns of companies in five pairs from March 1, 2013 to March 31, 2013, and the mean return of five pairs. It is a one-month out-of-sample result. In a similar way, we also calculate a three-month out-of-sample result. Table 6 shows the mean returns of companies in five pairs from March 1, 2013 to May 28,

Table 4 Bayesian inference of three-regime TAR model with GARCH effect for the five pair return spreads

Pairs	CAT-CVX		IBM-JNJ		MRK-MSFT		VZ-WMT		WMT-WAL	
Par.	Med	Std	Med	Std	Med	Std	Med	Std	Med	Std
$\phi_0^{(1)}$	0.0000	0.0031	−0.0001	0.0012	−0.0012	0.0028	0.0017	0.0011	−0.0041	0.0028
$\phi_1^{(1)}$	0.0591	0.0971	−0.0297	0.0710	0.0855	0.0811	0.1242	0.0644	−0.0942	0.0921
$\phi_0^{(2)}$	0.0000	0.0005	0.0002	0.0003	0.0008	0.0004	0.0003	0.0004	−0.0002	0.0004
$\phi_1^{(2)}$	0.0558	0.0685	0.0068	0.0680	0.0229	0.0469	0.0029	0.0712	−0.0633	0.0595
$\phi_0^{(3)}$	−0.0006	0.0020	−0.0007	0.0017	0.0023	0.0024	0.0005	0.0022	−0.0014	0.0021
$\phi_1^{(3)}$	0.0200	0.0789	0.0855	0.0790	−0.0470	0.0772	−0.0088	0.0916	0.0943	0.0863
$\alpha_0^{(1)}$	0.0001	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000
$\alpha_1^{(1)}$	0.0281	0.0223	0.0308	0.0169	0.0182	0.0154	0.0575	0.0192	0.0379	0.0178
$\beta_1^{(1)}$	0.9289	0.0459	0.9474	0.0283	0.9674	0.0232	0.9124	0.0276	0.9279	0.0323
$\alpha_0^{(2)}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$\alpha_1^{(2)}$	0.0393	0.0194	0.0698	0.0219	0.0758	0.0228	0.0542	0.0191	0.0365	0.0174
$\beta_1^{(2)}$	0.9159	0.0229	0.8890	0.0190	0.8696	0.0222	0.8993	0.0245	0.9104	0.0156
$\alpha_0^{(3)}$	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000
$\alpha_1^{(3)}$	0.0870	0.0260	0.0287	0.0156	0.0291	0.0211	0.0359	0.0170	0.0870	0.0255
$\beta_1^{(3)}$	0.7847	0.0675	0.9313	0.0353	0.9400	0.0387	0.9039	0.0460	0.8880	0.0360
c_1	−0.0184	0.0015	−0.0062	0.0018	−0.0192	0.0020	−0.0052	0.0020	−0.0172	0.0027
c_2	0.0105	0.0046	0.0103	0.0018	0.0147	0.0032	0.0139	0.0021	0.0113	0.0032
ν	5.5336	0.7293	4.2734	0.2758	4.2981	0.2910	5.1025	0.5995	5.1695	0.6468

Table 5 Company returns in five pairs and pairs returns from March 1, 2013 to March 28, 2013

Pairs	Company A	Mean Return	Company B	Mean Return	No. of Trading	Pairs Return
1	CAT	-0.260 %	CVX	0.086 %	4	-1.173 %
2	IBM	0.263 %	JNJ	0.321 %	4	3.494 %
3	MRK	0.241 %	MSFT	0.122 %	3	3.849 %
4	VZ	0.267 %	WMT	0.255 %	7	4.758 %
5	WMT	0.255 %	WAL	0.138 %	5	1.019 %

Table 6 Company returns in five pairs and pairs returns from March 1, 2013 to May 31, 2013

Pairs	Company A	Mean Return	Company B	Mean Return	No. of Trading	Pairs Return
1	CAT	-0.090 %	CVX	0.090 %	13	10.448 %
2	IBM	0.047 %	JNJ	0.160 %	17	3.225 %
3	MRK	0.160 %	MSFT	0.363 %	17	15.780 %
4	VZ	0.075 %	WMT	0.087 %	24	8.082 %
5	WMT	0.087 %	WAL	0.208 %	11	1.600 %

2013, and the mean return of five pairs. Figure 2 is a time series plot of the third pair's return spread (MRK-MSFT) during in-sample period. The red and blue lines are estimated threshold values, c_1 and c_2, respectively. Figure 3 shows the pairs return spread of asset MRK and its pair, MSFT, during the out-of-sample period. Again, red and blue lines locate at threshold values which are employed as trading entry and exit signals.

From Tables 5 and 6, we find that there are 4.6 round trips trading on average in the one-month period and 16.4 round trips trading on average in the three-month period. The average 5 pairs profits are 2.389% and 7.827%, respectively. The pair returns increase with the trading horizons in most pairs. This indicates that the longer the trading horizon, the better the mean reversion process works.

For a comparison, we would like to consider the cointegration approach. In the cointegration pairs trading literature, there is a potential problem, that is, we can't find enough cointegration pairs in the sample. Hakkio and Rush (1991) had observed that "cointegration is a long-run concept and hence requires long spans of data to give tests for cointegration much power rather than merely large numbers of observations." Indeed when we apply the cointegration test to the five selected pairs, only the third pair (MRK/ MSFT) is found to be cointegrated in the in-sample period.

$$P_{MRK} = -4.13 + 1.47P_{MSFT} + \varepsilon_{IN}.$$

We then define the residual out-of-sample as

$$\varepsilon_{OUT} = P_{MRK,OUT} + 4.13 - 1.47P_{MSFT,OUT}.$$

When the indicator (defined as $\varepsilon_{OUT}/\varepsilon_{IN}$) is larger (smaller) than one (negative one), then we short (long) one share of the first stock (MRK) and long (short) 1.47 shares of the second stock (MSFT). In this case, we short 1 share MRK, and long 1.47 shares of MSFT. The final pairs average return is 0.85% $(\frac{1}{41} \times (-\ln(\frac{41.84}{47.38}) + 1.47 \times \ln(\frac{32.38}{27.76}))$ in the three-month out of sample period. The profit is significantly less than that of the proposed method which yields an average return 15.780%.

4 Conclusions

In this study, we model the daily return spread of stock pairs as a three-regime TAR-GARCH process, and the upper and lower regimes in the model are used as trading entry and exit signals. We apply the trading rules to the Dow Jones Industrial Average Index stocks. The in-sample period of this study is from January 2, 2006 to February 28, 2013 and the out-of-sample period is from March 1, 2013 to May 31, 2013.

The empirical results suggest that the combination of MSD and TAR-GARCH trading rules generate positive excess returns, relative to the underlying stocks. The average pairs trading profits are 2.389% and 7.827% in the one-month and

three-month trading periods, respectively. With the proposed three-regime TAR-GARCH pairs trading strategy, traders can reap adequate profits from the Dow Jones 30 stocks. Transaction costs and rolling in-sample fitting and out-of-sample trading will be analyzed in future studies.

Acknowledgements. We thank the editor and anonymous referee for their insightful and helpful comments, which have improved this paper. Cathy W.S. Chen is supported by the grant (NSC 101-2118-M-035-006-MY2) from the National Science Council (NSC) of Taiwan.

References

1. Bock, M., Mestel, R.: A regime-switching relative value arbitrage rule. In: Operations Research Proceedings 2008. Springer, Heidelberg (2009)
2. Bollerslev, T.: Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics 31, 307–327 (1986)
3. Bollerslev, T., Chou, R.Y., Kroner, K.F.: ARCH modeling in finance: a review of the theory and empirical evidence. Journal of Econometrics 52, 5–59 (1992)
4. Brooks, C.: A double-threshold GARCH model for the French Franc/Deutschmark exchange rate. Journal of Forecasting 20, 135–143 (2001)
5. Chen, C.W.S., Chiang, T.C., So, M.K.P.: Asymmetrical reaction to US stock-return news: evidence from major stock markets based on a double-threshold model. The Journal of Economics and Business 55, 487–502 (2003)
6. Chen, C.W.S., So, M.K.P.: On a threshold heteroscedastic model. International Journal of Forecasting 22, 73–89 (2006)
7. Chen, C.W.S., Gerlach, R., Lin, A.M.H.: Falling and explosive dormant and rising markets via multiple-regime financial time series models. Applied Stochastic Models in Business and Industry 26, 28–49 (2010)
8. Chen, C.W.S., So, M.K.P., Liu, F.C.: A review of threshold time series models in finance. Statistics and Its Interface 4, 167–182 (2011)
9. Do, B., Faff, R.: Does simply pairs trading still work? Financial Analysts Journal 66, 83–95 (2010)
10. Engle, R.F.: Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50, 987–1008 (1982)
11. Engle, R.F., Granger, C.W.: Co-integration and error correction: representation, estimation, and testing. Econometrica 55, 251–276 (1987)
12. Elliott, R.J., Van Der Hoek, J., Malcolm, W.P.: Pairs trading. Quantitative Finance 5, 271–276 (2005)
13. Gatev, E., Goetzmann, W.N., Rouwenhorst, K.G.: Pairs trading: Performance of a relative value trading arbitrage rule. The Review of Financial Studies 19, 797–827 (2006)
14. Hakkio, C.S., Rush, M.: Cointegration: how short is the long run? Journal of International Money and Finance 10, 571–581 (1991)
15. Kawasaki, Y., Tachiki, S., Udaka, H., Hirano, T.: A characterization of long-short trading strategies based on cointegration. In: Proceedings of IEEE International Conference on Computational Intelligence for Financial Engineering, pp. 411–416 (2003)
16. Li, C.W., Li, W.K.: On a double-threshold autoregressive heteroscedastic time series model. Journal of Applied Econometrics 11, 253–274 (1996)

17. Perlin, M.: Evaluation of pairs-trading strategy at the Brazilian financial market. Journal of Derivatives and Hedge Funds 15, 122–136 (2009)
18. Vidyamurthy, G.: Pairs Trading: Quantitative Methods and Analysis. John Wiley \& Son, Inc., New Jersey (2004)

Testing Dependencies in Term Structure of Interest Rates

Kian-Guan Lim

Abstract. In this paper we study the term structure of interest rates and test the rational expectations hypothesis using single regression equations and then multivariate regression equations. Single regression equations are found to produce results that are sensitive to outliers due to finite sample. Multivariate regression equations produce results that are less sensitive to outliers due to a larger sample size, and in our sample, yield a borderline rejection of the rational expectatons hypothesis. We apply a distance covariance test statistic measuring the deviation from independence between the forward forecast errors and present information variables. This measure is asymptotically distributed to be bounded below by χ_1^2 for usual ranges of critical region, and does not require any distributional assumption. The rational expectation hypothesis is more clearly rejected using the distance covariance metric. There is thus preliminary evidence that distributional and linearity mis-specification of the rationality hypothesis in the term structure could potentially biased toward non-rejection of an otherwise generally unsustainable hypothesis.

1 Introduction

There is a copious amount of work on testing term structure theories up to the early 2000s. Most of the studies either test the rational expectations hypothesis within the term structure context, or else test the predictability of future spot interest rates. However, it is one of a few areas of financial economics where empirical results have been rather diverse and different with not much of a conclusion. [9] found that forward interest rates on a contract due over a long maturity do not predict the corresponding future spot rates well. [7] had firmly rejected the rational expectations hypothesis of the interest rate term structure. But [2] concluded that correlation

Kian-Guan Lim
Lee Kong Chian Business School, Singapore Management University, 50 Stamford Road, Singapore 178899
e-mail: kgl@smu.edu.sg

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_9,

between the forward and spot rates seemed positively high. [6] also found forward rates to be good predictors when the forward horizon grew longer. [4] cited a number of studies that showed opposite results regarding the rational expectations hypothesis on interest rates in UK, in Germany, and in other European countries. [5] remarked that, "Tests generally reject the hypothesis that expectations are rational, which is then a rejection of the expectations theory of the term structure..."

Though there have been enough studies and replications to show that the rational expectations hypothesis on the term structure of interest rates (in short, "REH") does not work in many situations, it remains a controversy as to when it would work and when it would not. Some new methods have surfaced now and then, such as the vector autoregression approach by [3]. In this paper, we contribute to the literature by employing a new distance covariance test statistic to show how distributional and linearity mis-specification of the rationality hypothesis in the term structure could potentially biased toward non-rejection of an otherwise generally unsustainable hypothesis.

2 Deriving the Spot and Forward Rates

Daily annualized U.S. Treasury bill investment yields and Treasury bond market bond equivalent yields (based on semi-annual compounding) with constant maturities of 1-month, 3-month, 6-month, 1-year, 2-year, 3-year, 5-year, 7-year, 10-year, and 20-year, are available from the U.S. Federal Reserve System database.[1] The daily time series is from 31 July 2001 till June 2013. We use the end of month calendar data on the above dataset, so that at each end of month starting from July 2001 till June 2013, we have time series with a sample length of 144 months. We use t to denote the time corresponding to the number of months starting from July 2001. Thus, data on August 2001 corresponds to $t = 2$, September 2001 corresponds to $t = 3$, and so on. The Fed data reported are in terms of investment yields. These investment yields are first converted to annualized spot rates $S_t(0,T)$ where subscript t denotes spot rate available as at time t indexed in months from July 2001, and T refers to the term in years of the spot rate.

A spot rate with a term T implies that a \$1 put in a deposit at time t will yield $\$(1+S_t(0,T))^T$ at time $T > t$. Note that the rates in the Fed database are all annualized rates. For Treasury bills with maturities less than a year, the investment yields are also the spot rates. For Treasury notes and bonds with maturities greater than or equal to a year, say $T \geq 1$ year maturity, the T-year annualized spot rate is $S_t(0,T) = (1+y_t(0,T)/2)^2 - 1$ where $y_t(0,T)$ is the corresponding annualized investment yield. For each $t = 1,2,\ldots,144$, a cubic spline curve is fitted to pass through all the discrete annualized spot rates $S_t(0,n)$, $n = \frac{1}{12}, \frac{3}{12}, \frac{6}{12}, 1,2,3,5,7,10,20$ in terms of n years; each pair of adjacent spot rates is connected with a cubic polynomial. Moreover, the ends of the curves at both the right side and the left side of a yield

[1] Source is webpage
`http://www.federalreserve.gov/releases/h15/data.htm`

rate are connected such that their first and second derivatives equal. The latter are sometimes called smooth pasting conditions. The additional two conditions used to identify all the cubic polynomials of a cubic spline are the slope conditions at the start and end points of the spline. Once these are defined, then the method of cubic spline is termed the "clamped" method. We set the starting and end slopes to zeros.

Thus for each day at end of month, indexed by t, the following annualized spot rates can be obtained from the fitted spline: $S_t(0,\frac{1}{12})$, $S_t(0,\frac{2}{12})$, $S_t(0,\frac{3}{12})$, $\ldots, S_t(0,T), \ldots, S_t(0,20)$, where T denotes the number of years into the future from t. For example, $T=\frac{x}{12}$ denotes x months after t, while $T=3\,\frac{2}{12}$ denotes 3 years and 2 months, or 38 months, after t. The annualized continuously compounded spot rate over time interval $(0,T)$ starting at T is given by $R_t^{(T)}=\frac{1}{T}\ln(1+S_t(0,T))^T$, or simply $R_t^{(T)}=\ln(1+S_t(0,T))$. The superscript $''(T)''$ denotes the spot term of T years.

We can also obtain the time t annualized forward rate $f_t(k,k+\frac{1}{12})$ denoting a contract at t that \$1 can be deposited at time k years forward of t, and yielding $\$\big(1+f_t(k,k+\frac{1}{12})\big)^{\frac{1}{12}}$ at time $t+k+\frac{1}{12}$ or a month later, for $k\geq 0$. Thus at each t, there is a corresponding series of 239 forward rates as follows: $f_t(0,\frac{1}{12}), f_t(\frac{1}{12},\frac{2}{12}), f_t(\frac{2}{12},\frac{3}{12}), \ldots, f_t(19\frac{11}{12},20)$. By no arbitrage argument,

$$\left(1+f_t(k,k+\frac{1}{12})\right)^{\frac{1}{12}}=\frac{\big(1+S_t(0,k+\frac{1}{12})\big)^{k+\frac{1}{12}}}{(1+S_t(0,k))^k},$$

so $\ln\big(1+f_t(k,k+\frac{1}{12})\big)=12\left[(k+\frac{1}{12})\ln\big(1+S_t(0,k+\frac{1}{12})\big)-k\ln(1+S_t(0,k))\right]$.

We define annualized continuously compounded forward rate $F_t^{(k,\frac{1}{12})}$ to be $\ln\big(1+f_t(k,k+\frac{1}{12})\big)$. The superscripts $''(k,\frac{1}{12})''$ denotes deposit at k year after t, and collection at $\frac{1}{12}$ year or a month after deposit. Then,

$$\begin{aligned}F_t^{(k,\frac{1}{12})} &= 12\left[(k+\frac{1}{12})R_t^{(k+\frac{1}{12})}-kR_t^{(k)}\right]\\ &= (12k+1)R_t^{(k+\frac{1}{12})}-12kR_t^{(k)}.\end{aligned}\tag{1}$$

It is instructive to note that for $k=0$, $F_t^{(0,\frac{1}{12})}=R_t^{(\frac{1}{12})}$, for $k=\frac{1}{12}$, $F_t^{(\frac{1}{12},\frac{1}{12})}=2R_t^{(\frac{2}{12})}-R_t^{(\frac{1}{12})}$, for $k=\frac{2}{12}$, $F_t^{(\frac{2}{12},\frac{1}{12})}=3R_t^{(\frac{3}{12})}-2R_t^{(\frac{2}{12})}$, for $k=\frac{3}{12}$, $F_t^{(\frac{3}{12},\frac{1}{12})}=4R_t^{(\frac{4}{12})}-3R_t^{(\frac{3}{12})}$, and so on.

If the market is risk-neutral, then for $N=1,2,3,\ldots,240$,

$$\exp\left(\frac{N}{12}R_t^{(\frac{N}{12})}\right)=E_t\left[\prod_{i=0}^{N-1}\exp\left(\frac{1}{12}R_{t+\frac{i}{12}}^{(\frac{1}{12})}\right)\right],$$

where $E_t[\cdot]$ denotes expectation conditional on all information available at time t. When $N=1$, $\exp\left(\frac{1}{12}R_t^{(\frac{1}{12})}\right)=E_t\left[\exp\left(\frac{1}{12}R_t^{(\frac{1}{12})}\right)\right]$. When $N=2$,

$$\exp\left(\frac{2}{12}R_t^{(\frac{2}{12})}\right) = E_t\left[\exp\left(\frac{1}{12}R_t^{(\frac{1}{12})} + \frac{1}{12}R_{t+\frac{1}{12}}^{(\frac{1}{12})}\right)\right].$$

When $N = 3$,

$$\exp\left(\frac{3}{12}R_t^{(\frac{3}{12})}\right) = E_t\left[\exp\left(\frac{1}{12}R_t^{(\frac{1}{12})} + \frac{1}{12}R_{t+\frac{1}{12}}^{(\frac{1}{12})} + \frac{1}{12}R_{t+\frac{2}{12}}^{(\frac{1}{12})}\right)\right],$$

and so on. Even if the market is risk-averse, the equation is correct provided the conditional probability distribution is a risk-neutral distribution, which is an empirical distribution subjected to a Girsanov transformation under condition of no arbitrage. The latter is sometimes termed as the fundamental theorem of asset pricing.

Taking natural logarithms on both sides, and dividing by $\frac{N}{12}$, we have

$$R_t^{(\frac{N}{12})} = \frac{12}{N}\ln E_t\left[\prod_{i=0}^{N-1}\exp\left(\frac{1}{12}R_{t+\frac{i}{12}}^{(\frac{1}{12})}\right)\right].$$

In the academic literature (see [3] for example), this is usually approximated by the relationship

$$\begin{aligned} R_t^{(\frac{N}{12})} &= \frac{12}{N}\sum_{i=0}^{N-1}\frac{1}{12}E_t\left[R_{t+\frac{i}{12}}^{(\frac{1}{12})}\right] \\ &= \frac{1}{N}\sum_{i=0}^{N-1}E_t\left[R_{t+\frac{i}{12}}^{(\frac{1}{12})}\right]. \end{aligned} \tag{2}$$

The approximation allows for tractable analytical formulation, and is acceptable when the spot rates are small. The approximation can be explained by the following example: if $S_t(0,\frac{1}{12})$ and $S_{t+\frac{1}{12}}(0,\frac{1}{12})$ are conditionally independent, then

$$\begin{aligned} &\ln E_t\left[(1+S_t(0,\frac{1}{12}))^{\frac{1}{12}}(1+S_{t+\frac{1}{12}}(0,\frac{1}{12}))^{\frac{1}{12}}\right] \\ &= \ln E_t(1+S_t(0,\frac{1}{12}))^{\frac{1}{12}} + \ln E_t(1+S_{t+\frac{1}{12}}(0,\frac{1}{12}))^{\frac{1}{12}} \\ &\approx E_t\ln(1+S_t(0,\frac{1}{12}))^{\frac{1}{12}} + E_t\ln(1+S_{t+\frac{1}{12}}(0,\frac{1}{12}))^{\frac{1}{12}} \\ &= \frac{1}{12}E_t\left[(R_t^{(\frac{1}{12})} + R_{t+\frac{1}{12}}^{(\frac{1}{12})}\right]. \end{aligned}$$

We shall work with empirical distribution, so all random variables $\{X_t\}$ in our study belong to a product sample space $\Omega = \mathscr{R}\times\mathscr{R}\times\ldots\times\mathscr{R}$, where $\mathscr{R}$ is the real line. The filtered probability space is $(\Omega,\mathscr{F},\{\mathscr{F}\}_t,\mathscr{P})$ where $\{\mathscr{F}\}_t$ is a natural filtration, and $\mathscr{P}$ is the joint probability measure on the sigma algebra $\mathscr{F}$.

When we include the existence of risk aversion (or risk preference) in the market, then Eq. (2) can be generalized to the following

$$R_t^{(\frac{N}{12})} = \frac{1}{N}\sum_{i=0}^{N-1} E_t\left[R_{t+\frac{i}{12}}^{(\frac{1}{12})}\right] + C_{N-1}, \tag{3}$$

where C_{N-1} is a risk premium per unit of time, or per year in this case, that may vary with term represented by N, but not with t. It may be thought of as a time-homogeneous variable risk premium of the term structure of interest rates. The use of subscript $N-1$ for C here instead of N is arbitrary, and does not have affect the theoretical construction in any way.

3 Term Structure of Interest Rates

We shall derive the rational expectation theory of term structure in this section and formulate several testable hypotheses. We consider spot rate spaced at monthly intervals (dropping the elaboration of "annualized continuously compounded"): $R_t^{(\frac{N}{12})}$, where $N \in \mathscr{I}^+$ is a positive integer, and is the number of months of the term structure.

Applying Eq. (3) recursively for $N = 1,2,3,\ldots$ and so on, we can obtain the following.

$$R_{t+}^{(\frac{1}{12})} = E_t\left[R_{t+}^{(\frac{1}{12})}\right] + C_0, \tag{4}$$

where $t+$ refers to a very small positive interval close to zero.

$$\begin{aligned} R_t^{(\frac{2}{12})} &= \frac{1}{2}\left\{E_t\left[R_t^{(\frac{1}{12})}\right] + E_t\left[R_{t+\frac{1}{12}}^{(\frac{1}{12})}\right]\right\} + C_1 \\ &= \frac{1}{2}\left\{R_t^{(\frac{1}{12})} + E_t\left[R_{t+\frac{1}{12}}^{(\frac{1}{12})}\right]\right\} + C_1. \end{aligned} \tag{5}$$

$$R_t^{(\frac{3}{12})} = \frac{1}{3}\left\{R_t^{(\frac{1}{12})} + E_t\left[R_{t+\frac{1}{12}}^{(\frac{1}{12})}\right] + E_t\left[R_{t+\frac{2}{12}}^{(\frac{1}{12})}\right]\right\} + C_2. \tag{6}$$

$$R_t^{(\frac{4}{12})} = \frac{1}{4}\left\{R_t^{(\frac{1}{12})} + E_t\left[R_{t+\frac{1}{12}}^{(\frac{1}{12})}\right] + E_t\left[R_{t+\frac{2}{12}}^{(\frac{1}{12})}\right] + E_t\left[R_{t+\frac{3}{12}}^{(\frac{1}{12})}\right]\right\} + C_3. \tag{7}$$

If we take two times the LHS and RHS of Eq. (5), then subtract the LHS and RHS of Eq. (4) accordingly, we obtain:

$$2R_t^{(\frac{2}{12})} - R_{t+}^{(\frac{1}{12})} = E_t\left[R_{t+\frac{1}{12}}^{(\frac{1}{12})}\right] + (2C_1 - C_0), \tag{8}$$

where we take the limit that $E_t[R_{t+\delta}^{(\frac{1}{12})}] = R_t^{(\frac{1}{12})}$ as $\delta \downarrow 0$. We shall not put any further assumption on the limit of C_0 as $\delta \downarrow 0$. Since the LHS of Eq. (8) is $F_t^{(\frac{1}{12},\frac{1}{12})}$, as $\delta \downarrow 0$, from Eq. (1), we have

$$E_t\left[R_{t+\frac{1}{12}}^{(\frac{1}{12})}\right] = F_t^{(\frac{1}{12},\frac{1}{12})} + (C_0 - 2C_1). \tag{9}$$

Similarly, subtracting Eq. (5) from Eq. (6), we obtain:

$$E_t\left[R_{t+\frac{2}{12}}^{(\frac{1}{12})}\right] = F_t^{(\frac{2}{12},\frac{1}{12})} + (2C_1 - 3C_2). \tag{10}$$

Subtracting Eq. (6) from Eq. (7), we obtain:

$$E_t\left[R_{t+\frac{3}{12}}^{(\frac{1}{12})}\right] = F_t^{(\frac{3}{12},\frac{1}{12})} + (3C_2 - 4C_3). \tag{11}$$

In general, we obtain:

$$E_t\left[R_{t+\frac{k}{12}}^{(\frac{1}{12})}\right] = F_t^{(\frac{k}{12},\frac{1}{12})} + (kC_{k-1} - [k+1]C_k). \tag{12}$$

We can always write, from Eq. (9),

$$R_{t+\frac{1}{12}}^{(\frac{1}{12})} = a_0 + F_t^{(\frac{1}{12},\frac{1}{12})} + \varepsilon_{t+\frac{1}{12}}, \tag{13}$$

where $a_0 = C_0 - 2C_1$, and $E_t\left[\varepsilon_{t+\frac{1}{12}}\right] = 0$. The rational expectations hypothesis would add a substantive content to the above regression Eq. (13) by the condition that

$$E_t\left[F_t^{(\frac{1}{12},\frac{1}{12})}\varepsilon_{t+\frac{1}{12}}\right] = 0.$$

In other words, $F_t^{(\frac{1}{12},\frac{1}{12})} \in \phi_t$ (part of the information set ϕ_t available to the market at time t) is independent (implying also zero covariance with) of $\varepsilon_{t+\frac{1}{12}}$. Economically, this means that any information contained in $F_t^{(\frac{1}{12},\frac{1}{12})}$ which is available at t, cannot be used to make any non-zero forecast of the next innovation $\varepsilon_{t+\frac{1}{12}}$ occuring from t to $t+\frac{1}{12}$. More generally, any information available at t is independent of innovation or surprise $\varepsilon_{t+\frac{1}{12}}$. We shall adhere to this construction of independence rather than mere contemporaneous zero correlation which is a weaker and linear concept.

From Eq. (13), we can express

$$R_{t+\frac{1}{12}}^{(\frac{1}{12})} - F_t^{(\frac{1}{12},\frac{1}{12})} = a_0 + a_1\left(F_t^{(\frac{1}{12},\frac{1}{12})} - R_t^{(\frac{2}{12})}\right) + \varepsilon_{t+\frac{1}{12}}, \tag{14}$$

where

$$E_t\left[\left(F_t^{(\frac{1}{12},\frac{1}{12})}-R_t^{(\frac{2}{12})}\right)\varepsilon_{t+\frac{1}{12}}\right]=0.$$

Under the rational expectations hypothesis, $H_0: a_1 = 0$ in regression Eq. (14).

Nominal interest rates are often found to be non-stationary in the literature (see [1] and [8]), but are usually cointegrated with other interest rates. Therefore, by constructing the dependent and explanatory variables as differences in interest rates, regression Eq. (14) is well specified.

In the same way, we can construct regression equations from Eq. (10) and Eq. (11) as follows to test for the rational expectations hypothesis in the term structure of interest rates.

From Eq. (10),

$$R_{t+\frac{2}{12}}^{(\frac{1}{12})}-F_t^{(\frac{2}{12},\frac{1}{12})}=b_0+b_1\left(F_t^{(\frac{2}{12},\frac{1}{12})}-R_t^{(\frac{3}{12})}\right)+\eta_{t+\frac{2}{12}}, \tag{15}$$

where

$$E_t\left[\left(F_t^{(\frac{2}{12},\frac{1}{12})}-R_t^{(\frac{3}{12})}\right)\eta_{t+\frac{2}{12}}\right]=0,$$

and $H_0: b_1 = 0$.

From Eq. (11),

$$R_{t+\frac{3}{12}}^{(\frac{1}{12})}-F_t^{(\frac{3}{12},\frac{1}{12})}=c_0+c_1\left(F_t^{(\frac{3}{12},\frac{1}{12})}-R_t^{(\frac{4}{12})}\right)+\varepsilon_{t+\frac{3}{12}}, \tag{16}$$

where

$$E_t\left[\left(F_t^{(\frac{3}{12},\frac{1}{12})}-R_t^{(\frac{4}{12})}\right)\varepsilon_{t+\frac{3}{12}}\right]=0,$$

and $H_0: c_1 = 0$.

4 Tests of Dependencies

We first test Eqs. (14), (15), and (16) individually. Next we combine them into a stacked regression to test under a joint hypothesis that all their slopes are zeros. Finally we consider the vector of the 3 different dependent variables, and the vector of the 3 different explanatory variables, and test for independence between the vectors based on the distance covariance metrics.

Single Regression Test

For the 144 months (using end-of-month data) from July 2001 till June 2013, let t denote end of July 2001, $t+\frac{1}{12}$ denote end of August 2001, and so on. Let 143×1 vector $Y=(y_1,y_2,\ldots,y_j,,\ldots,y_{143})'$ where $y_j=R_{t+\frac{j}{12}}^{(\frac{1}{12})}-F_{t+\frac{j-1}{12}}^{(\frac{1}{12},\frac{1}{12})}$. Let 143×2 matrix X of explanatory variables be

$$\begin{pmatrix} 1 & F_t^{(\frac{1}{12},\frac{1}{12})} - R_t^{(\frac{2}{12})} \\ 1 & F_{t+\frac{1}{12}}^{(\frac{1}{12},\frac{1}{12})} - R_{t+\frac{1}{12}}^{(\frac{2}{12})} \\ \vdots & \vdots \\ 1 & F_{t+\frac{i-1}{12}}^{(\frac{1}{12},\frac{1}{12})} - R_{t+\frac{i-1}{12}}^{(\frac{2}{12})} \\ \vdots & \vdots \\ 1 & F_{t+\frac{142}{12}}^{(\frac{1}{12},\frac{1}{12})} - R_{t+\frac{142}{12}}^{(\frac{2}{12})} \end{pmatrix}$$

Let 143×1 vector of disturbances be $E = \left(\varepsilon_{t+\frac{1}{12}}, \varepsilon_{t+\frac{2}{12}}, \ldots, \varepsilon_{t+\frac{143}{12}}\right)'$. Let $A = (a_0, a_1)'$. Then Eq. (14) can be represented by $Y = XA + E$, and the least squares estimates $\hat{A}$ can be obtained as $(X'X)^{-1}(X'Y)$. The null of $H_0 : a_1 = 0$ can thus be tested.

To test Eq. (15), we avoid overlapping data problem, and use the subsample of time series $t, t+\frac{2}{12}, t+\frac{4}{12}, t+\frac{6}{12}, \ldots, t+\frac{142}{12}$ (72 sample points). Let 71×1 vector $Y = (y_1, y_2, \ldots, y_j, , \ldots, y_{71})'$ where $y_j = R_{t+\frac{2j}{12}}^{(\frac{1}{12})} - F_{t+\frac{2(j-1)}{12}}^{(\frac{2}{12},\frac{1}{12})}$. Let 71×2 matrix X of explanatory variables be

$$\begin{pmatrix} 1 & F_t^{(\frac{2}{12},\frac{1}{12})} - R_t^{(\frac{3}{12})} \\ 1 & F_{t+\frac{2}{12}}^{(\frac{2}{12},\frac{1}{12})} - R_{t+\frac{2}{12}}^{(\frac{3}{12})} \\ \vdots & \vdots \\ 1 & F_{t+\frac{2(j-1)}{12}}^{(\frac{2}{12},\frac{1}{12})} - R_{t+\frac{2(j-1)}{12}}^{(\frac{3}{12})} \\ \vdots & \vdots \\ 1 & F_{t+\frac{142}{12}}^{(\frac{2}{12},\frac{1}{12})} - R_{t+\frac{142}{12}}^{(\frac{3}{12})} \end{pmatrix}$$

The 71×1 vector of disturbances is $E = \left(\eta_{t+\frac{2}{12}}, \eta_{t+\frac{4}{12}}, \ldots, \eta_{t+\frac{142}{12}}\right)'$. Let $B = (b_0, b_1)'$. Then $Y = XB + E$ and the null of $H_0 : b_1 = 0$ is tested.

To test Eq. (16), again we avoid overlapping data problem, and use the subsample of time series $t, t+\frac{3}{12}, t+\frac{6}{12}, t+\frac{9}{12}, \ldots, t+\frac{141}{12}$ (48 sample points). Let 47×1 vector $Y = (y_1, y_2, \ldots, y_j, , \ldots, y_{47})'$ where $y_j = R_{t+\frac{3j}{12}}^{(\frac{1}{12})} - F_{t+\frac{3(j-1)}{12}}^{(\frac{3}{12},\frac{1}{12})}$. Let 47×2 matrix X of explanatory variables be

$$\begin{pmatrix} 1 & F_t^{(\frac{3}{12},\frac{1}{12})} - R_t^{(\frac{4}{12})} \\ 1 & F_{t+\frac{3}{12}}^{(\frac{3}{12},\frac{1}{12})} - R_{t+\frac{3}{12}}^{(\frac{4}{12})} \\ \vdots & \vdots \\ 1 & F_{t+\frac{3(j-1)}{12}}^{(\frac{3}{12},\frac{1}{12})} - R_{t+\frac{3(j-1)}{12}}^{(\frac{4}{12})} \\ \vdots & \vdots \\ 1 & F_{t+\frac{138}{12}}^{(\frac{3}{12},\frac{1}{12})} - R_{t+\frac{138}{12}}^{(\frac{4}{12})} \end{pmatrix}$$

The 47×1 vector of disturbances is $E = \left(\varepsilon_{t+\frac{3}{12}}, \varepsilon_{t+\frac{6}{12}}, \ldots, \varepsilon_{t+\frac{141}{12}}\right)'$. Let $C = (c_0, c_1)'$. Then $Y = XC + E$ and the null of $H_0 : c_1 = 0$ is tested.

The above single regressions can be construed as testing if present information variables $F_{t+\frac{k(j-1)}{12}}^{(\frac{k}{12},\frac{1}{12})} - R_{t+\frac{k(j-1)}{12}}^{\frac{k+1}{12}}$ can predict future surprises in spot rates or the forward forecast errors. The results for the above single regression tests are reported in Table 1 as follows.

Table 1 Single Regression Tests of Term Structure Dependency

Statistic	Eq. (14)	Eq. (15)	Eq. (16)
Sample Size	143	71	47
Constant Estimate	-0.085^{***}	-0.154^{***}	-0.249^{***}
(t-Statistic)	(-3.943)	(-3.714)	(-2.814)
Slope Estimate	1.217^{***}	-0.021	-0.468
(t-Statistic)	(2.985)	(-0.087)	(-1.050)
F-Statistic	8.911^{***}	0.008	1.103
(p-Value)	(0.003)	(0.931)	(0.299)
R^2	0.059	0.000	0.024

*** indicates rejection of null of zero at significance level of 1%.

The negative constant shows that forward rate carries a positive risk premium. In other words, for a borrower to lock in a forward deposit rate in the future, he or she has to pay a positive risk premium over and above the expected future spot rate. This implies a risk or volatility of future increases in spot interest rates, or else more risk aversion or excess demand on the part of borrowers than lenders. Table 1 also shows that out of the 3 cases of single regressions, one rejects the REH as the estimated slope coefficient is significantly different from zero, while the other two did not reject the null of zero slope. The results for the single regressions are similar across different k for the explanation variables, and so we do not report all the details.

As a robust observation, single regression equations such as Eqs. (14), (15), (16), are found to produce results that are sensitive to outliers due to finite sample. For

example, if we remove two to three large forecast errors, the slope estimate could change drastically from being significantly different from zero to being insignificantly different from zero. Employing generalized least squares did not alter the substantive results.

Joint Test

Now to perform a joint multivariate regression test, we define random variable $\tilde{Y}_1$ to have the 47×1 random sample

$$\left(\begin{array}{ccccc} R_{t+\frac{1}{12}}^{(\frac{1}{12})} - F_t^{(\frac{1}{12},\frac{1}{12})} & R_{t+\frac{4}{12}}^{(\frac{1}{12})} - F_{t+\frac{3}{12}}^{(\frac{1}{12},\frac{1}{12})} & R_{t+\frac{7}{12}}^{(\frac{1}{12})} - F_{t+\frac{6}{12}}^{(\frac{1}{12},\frac{1}{12})} & \vdots & R_{t+\frac{139}{12}}^{(\frac{1}{12})} - F_{t+\frac{138}{12}}^{(\frac{1}{12},\frac{1}{12})} \end{array} \right).$$

Define random variable $\tilde{Y}_2$ to have the 47×1 random sample

$$\left(\begin{array}{c} R_{t+\frac{2}{12}}^{(\frac{1}{12})} - F_t^{(\frac{2}{12},\frac{1}{12})} \\ R_{t+\frac{5}{12}}^{(\frac{1}{12})} - F_{t+\frac{3}{12}}^{(\frac{2}{12},\frac{1}{12})} \\ R_{t+\frac{8}{12}}^{(\frac{1}{12})} - F_{t+\frac{6}{12}}^{(\frac{2}{12},\frac{1}{12})} \\ \vdots \\ R_{t+\frac{140}{12}}^{(\frac{1}{12})} - F_{t+\frac{138}{12}}^{(\frac{2}{12},\frac{1}{12})} \end{array} \right).$$

Define random variable $\tilde{Y}_3$ to have the 47×1 random sample

$$\left(\begin{array}{c} R_{t+\frac{3}{12}}^{(\frac{1}{12})} - F_t^{(\frac{3}{12},\frac{1}{12})} \\ R_{t+\frac{6}{12}}^{(\frac{1}{12})} - F_{t+\frac{3}{12}}^{(\frac{3}{12},\frac{1}{12})} \\ R_{t+\frac{9}{12}}^{(\frac{1}{12})} - F_{t+\frac{6}{12}}^{(\frac{3}{12},\frac{1}{12})} \\ \vdots \\ R_{t+\frac{141}{12}}^{(\frac{1}{12})} - F_{t+\frac{138}{12}}^{(\frac{3}{12},\frac{1}{12})} \end{array} \right).$$

Stack the above vectors to become a 141×1 $Y^* = (Y_1', Y_2', Y_3')'$. Similarly define 47×1 X_1 to be

$$\left(\begin{array}{c} F_t^{(\frac{1}{12},\frac{1}{12})} - R_t^{(\frac{2}{12})} \\ F_{t+\frac{3}{12}}^{(\frac{1}{12},\frac{1}{12})} - R_{t+\frac{3}{12}}^{(\frac{2}{12})} \\ F_{t+\frac{6}{12}}^{(\frac{1}{12},\frac{1}{12})} - R_{t+\frac{6}{12}}^{(\frac{2}{12})} \\ \vdots \\ F_{t+\frac{138}{12}}^{(\frac{1}{12},\frac{1}{12})} - R_{t+\frac{138}{12}}^{(\frac{2}{12})} \end{array} \right).$$

Define 47×1 X_2 to be

$$\begin{pmatrix} F_t^{(\frac{2}{12},\frac{1}{12})} - R_t^{(\frac{3}{12})} \\ F_{t+\frac{3}{12}}^{(\frac{2}{12},\frac{1}{12})} - R_{t+\frac{3}{12}}^{(\frac{3}{12})} \\ F_{t+\frac{6}{12}}^{(\frac{2}{12},\frac{1}{12})} - R_{t+\frac{6}{12}}^{(\frac{3}{12})} \\ \vdots \\ F_{t+\frac{138}{12}}^{(\frac{2}{12},\frac{1}{12})} - R_{t+\frac{138}{12}}^{(\frac{3}{12})} \end{pmatrix}.$$

Define 47×1 X_3 to be

$$\begin{pmatrix} F_t^{(\frac{3}{12},\frac{1}{12})} - R_t^{(\frac{4}{12})} \\ F_{t+\frac{3}{12}}^{(\frac{3}{12},\frac{1}{12})} - R_{t+\frac{3}{12}}^{(\frac{4}{12})} \\ F_{t+\frac{6}{12}}^{(\frac{3}{12},\frac{1}{12})} - R_{t+\frac{6}{12}}^{(\frac{4}{12})} \\ \vdots \\ F_{t+\frac{138}{12}}^{(\frac{3}{12},\frac{1}{12})} - R_{t+\frac{138}{12}}^{(\frac{4}{12})} \end{pmatrix}.$$

Similarly, stack the above vectors to become a 141×1 $X^* = (X_1', X_2', X_3')'$.

Let 47×3 matrix D_1 be $\begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ \vdots & & \\ 1 & 0 & 0 \end{pmatrix}$.

Let 47×3 matrix D_2 be $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ \vdots & & \\ 0 & 1 & 0 \end{pmatrix}$.

Let 47×3 matrix D_3 be $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ \vdots & & \\ 0 & 0 & 1 \end{pmatrix}$.

Stack D_1, D_2, D_3 to become a 141×3 D^*. Concatenate $Z = (D^*, X^*)$ to be a 141×4 matrix. Let $\Theta = (\theta_1, \theta_2, \theta_3, \theta_4)'$. Let 141×1 disturbance term be U. Then perform regression on $Y^* = Z\Theta + U$. The covariance of U in this case is heteroskedastic since $\text{var}(\varepsilon_t) = \sigma_1^2$, $\text{var}(\eta_t) = \sigma_2^2$, $\text{var}(\varepsilon_t) = \sigma_3^2$, are different from each other. Moreover, $\text{cov}(\varepsilon_t, \eta_t) = \sigma_{1,2} \neq 0$, $\text{cov}(\varepsilon_t, \varepsilon_t) = \sigma_{1,3} \neq 0$, and $\text{cov}(\eta_t, \varepsilon_t) = \sigma_{2,3} \neq 0$. Let 3×3

matrix V be $\begin{pmatrix} \sigma_1^2 & \sigma_{1,2} & \sigma_{1,3} \\ \sigma_{1,2} & \sigma_2^2 & \sigma_{2,3} \\ \sigma_{1,3} & \sigma_{2,3} & \sigma_3^2 \end{pmatrix}$. Then $\text{cov}(U) = V \otimes I_{47\times 47}$ and we test for the $H_0 : \theta_4 = 0$ using generalized least squares. The results are reported in Table 2.

Distance Covariance Test

The above stacked multivariate regression or sometimes known as seemingly unrelated regressions, assumes linearity in the relationship between forecast errors and present information variables, as well as an assumption of the distribution of the disturbances, usually Gaussian. The linearity and Gaussian assumptions may lead to mis-specification biases, particularly in finite samples. To avoid these biases, we employ the distance covariance metric of [10] as a test of the independence of the forward forecast errors and the present information variables.

We consider Y^* to be realizations from a random 3×1 vector $(\widetilde{Y}_1, \widetilde{Y}_2, \widetilde{Y}_3)'$. Also consider X^* to be realizations from a random 3×1 vector $(\widetilde{X}_1, \widetilde{X}_2, \widetilde{X}_3)'$. We then employ the distance covariance method to test the independence of these two random vectors.

[11] used the idea that two random functions, in our case vectors $\widetilde{Y}^*$ and $\widetilde{X}^*$, are independent if and only if the characteristic function of the joint distribution, $f_{\widetilde{Y}^*,\widetilde{X}^*}(t,s)$ equals the product of their characteristic functions, $f_{\widetilde{Y}^*}(t) f_{\widetilde{X}^*}(s)$. The theoretical distance covariance metric is formulated as

$$V^2(\widetilde{Y}^*, \widetilde{X}^*) = \int_{\mathscr{R}^6} |f_{\widetilde{Y}^*,\widetilde{X}^*}(t,s) - f_{\widetilde{Y}^*}(t) f_{\widetilde{X}^*}(s)|^2 \times (c_p c_q |t|_p^{1+p} |s|_q^{1+q})^{-1} dt\, ds,$$

where $|t|_p$ is the usual Euclidean norm in $\mathscr{R}^p$, and $c_p = \dfrac{\pi^{(1+p)/2}}{\Gamma((1+p)/2)}$. The normalized sample equivalent of this metric is $T\hat{V}^2/S \xrightarrow{d} Q$ as sample size $T \to \infty$, where

$$S = T^{-2} \sum_{n=1}^{T} \sum_{m=1}^{T} |X_n - X_m||Y_n - Y_m|.$$

$\hat{V}^2 = T^{-2} \sum_{n=1}^{T} \sum_{m=1}^{T} A_{nm} B_{nm}$ where $A_{nm} = a_{nm} - \overline{a_{n\cdot}} - \overline{a_{\cdot m}} + \overline{a_{\cdot\cdot}}$, $a_{nm} = |X_n - X_m|$, $\overline{a_{n\cdot}} = T^{-1} \sum_{m=1}^{T} a_{nm}$, $\overline{a_{\cdot m}} = T^{-1} \sum_{n=1}^{T} a_{nm}$, and $\overline{a_{\cdot\cdot}} = T^{-2} \sum_{n=1}^{T} \sum_{m=1}^{T} a_{nm}$. According to [11] Theorem 6, Q is asymptotically distributed as This measure is asymptotically distributed to be bounded below by χ_1^2 for usual ranges of the critical region, such as 10%, 5%, or 1%, and does not require any distributional assumption.

We report the results of the joint tests and the distance covariance test statistic in Table 2.

Table 2 shows that in the joint multivariate test, the common slope coefficient is estimated at close to one, and is significantly different from zero at a p-value of 9.8%. It is a borderline case whereby REH may appear to hold. However, the distance covariance test yields a Q-statistic of 3.4607 with a p-value smaller than

Table 2 Joint Tests of Term Structure Dependency

Statistic	Estimate	z-Statistic	p-Value
$\hat{\theta}_1$	-0.0346	(-0.232)	0.816
$\hat{\theta}_2$	-0.369^{**}	(-2.310)	0.021
$\hat{\theta}_3$	1.708^{***}	(10.052)	0.000
$\hat{\theta}_4$	1.108^{*}	(1.654)	0.098
Statistic	Estimate		Upper Bound for p-Value
Q-Test	3.4607^{*}		0.0628

Sample size of joint test is 141. *** indicates rejection of null of zero at significance level of 1%, ** indicates rejection of null of zero at significance level of 5%, and * indicates rejection of null of zero at significance level of 10%.

0.0628. This is clearly a stronger rejection of the REH than in the case of the linear multivariate joint test.

5 Conclusions

Rational expectations constraints are set up in the term structure relationship between forward interest rate forecasts and present information variables containing differences of forward and spot rates. REH would suggest that present information values did not have any impact on the future forecast surprises. Using linear single regressions, we show results that are sensitive to outliers such as unusually high forecast errors in some months.

In joint test involving more than one regression equation, the rational expectation hypothesis is more clearly rejected using the distance covariance metric. There is thus preliminary evidence that distributional and linearity mis-specification of the rationality hypothesis in the term structure could potentially biased toward non-rejection of an otherwise generally unsustainable hypothesis.

Acknowledgments. This work was supported by a research grant from the Sim Kee Boon Institute for Financial Economics at the Singapore Management University. The excellent research assistance of Jessica Zhang is gratefully acknowledged.

References

1. Campbell, J.Y., Clarida, R.H.: The Term Structure of Euromarket Interest Rates. Journal of Monetary Economics 19, 25–44 (1987)
2. Campbell, J.Y., Shiller, R.J.: Cointegration and Tests of Present Value Models. Journal of Political Economy 95, 1062–1088 (1987)
3. Campbell, J.Y., Shiller, R.J.: Yield Spreads and Interest Rate Movements: A Bird's Eye View. Review of Economic Studies 58, 495–514 (1991)
4. Cuthbertson, K., Bredin, D.: The Expectations Hypothesis of the Term Structure: The Case of Ireland. The Economic and Social Review 31, 267–281 (2000)

5. Fair, R.C.: Estimating Term Structure Equations using Macroeconomic Variables. Cowles Foundation Discussion Paper No. 1634 (2008)
6. Fama, E.F., Bliss, R.R.: The Information in Long-Maturity Forward Rates. American Economic Review 77, 680–692 (1987)
7. Hansen, L.P., Hodrick, R.J.: Forward Rates as Optimal Predictors of Future Spot Rates. Journal of Political Economy 88, 829–853 (1980)
8. Newbold, P., Leybourne, S., Sollis, R., Wohar, M.E.: US and UK Interest Rates 1890-1934: New Evidence on Structural Breaks. Journal of Money Credit and Banking 33, 235–250 (2001)
9. Shiller, R.J.: The Volatility of Long-Term Interest Rates and Expectations Theories of the Term Structure. Journal of Political Economy 87, 1190–1219 (1979)
10. Szekely, G.J., Rizzo, M.L.: Brownian Distance Covariance. The Annals of Applied Statistics 3, 1236–1265 (2009)
11. Szekely, G.J., Rizzo, M.L., Bakirov, N.K.: Measuring and Testing Dependence by Correlation of Distances. The Annals of Statistics 35, 2769–2794 (2007)

Joint Distributions of Random Sets and Their Relation to Copulas

Bernhard Schmelzer

Abstract. Random sets are set-valued random variables. They have been applied in various fields like stochastic geometry, statistics, economics, engineering or computer science, and are often used for modeling uncertainty. This paper is concerned with joint distributions of random sets. Generalizations of the Choquet theorem are presented which state that the joint distribution of random sets can be characterized by multivariate analogues of capacity functionals. Furthermore, it is shown how copulas can be used to describe the relation between a joint distribution of random sets and their marginal distribution.

1 Introduction

Roughly speaking, random sets are random variables whose values are sets. They have been applied in various fields like stochastic geometry, statistics, economics, engineering or computer science. Random sets are frequently interpreted as imprecise observations of random variables [4]. Thus, they can be used to model uncertainty when there is only vague knowledge about a random variable or when only imprecise or incomplete observations are available. In this interpretation, it is assumed that the values of a random set contain the values of the true (but unavailable) random variable.

Just like distributions of random variables can be uniquely described by distribution functions, distributions of random sets can be characterized by set functions with special properties, so-called *capacity* or *containment functionals* [5]. This is stated by the well-known Choquet theorem (or Matheron-Kendall-Choquet theorem) [5] which is of central importance in random set theory. Capacity and containment functionals have been interpreted as *plausibility* and *belief functions* in evidence theory [15] and *upper* and *lower probabilities* [3] in imprecise probability

Bernhard Schmelzer
Salzburger Strasse 14, 6300 Woergl, Austria
e-mail: bernhard.schmelzer@uibk.ac.at

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_10,

theories [16]. For further information on random sets the reader is referred to the textbooks [5, 7, 10].

This paper will present generalizations of the Choquet theorem to the multidimensional case; some of the presented ideas first appeared in [14]. More precisely, multivariate analogues of capacity functionals will be introduced and it will be demonstrated that the latter fully characterize joint distributions of random sets. Furthermore, the paper is addressed to the question if or how copulas can be used to describe the relation of joint distributions of random sets and their marginal distributions. Copulas are a popular tool for dependence modeling in statistics (see [20] for an introduction). A proposition will be presented which shows that copulas can be used to relate joint containment functionals to its marginal containment functionals. This proposition suggests that, in general, a single copula is not enough to completely describe dependence of two random sets.

The plan of the paper is as follows. In Section 2 basic facts about random sets are reviewed. Section 3 summarizes the most important results from [14] concerning joint distributions of random sets and characterization by set functions. Section 4 is addressed to the question how copulas can be used to describe the dependence of random sets.

2 Random Sets

Random sets can be seen as random variables whose values are subsets of some given set $\mathbb{E}$. These values are called *focal sets* of the random set. The simplest case arises when a random set consists of finitely many focal sets. In this case, one speaks of finite random sets or *Dempster-Shafer structures* [3, 15]. Each of the focal sets X_i, $i = 1, \dots, \ell$ comes with a probability weight p_i such that $\sum p_i = 1$.

A finite random set X with focal sets $X_1, \dots, X_\ell$ can be seen as a set-valued random variable by defining an ℓ-point probability space $\Omega = \{1, \dots, \ell\}$ with probability weights $\{p_1, \dots, p_\ell\}$. The assignment $X : i \mapsto X_i$ is the defining set-valued random variable.

Following Dempster and Shafer [3, 15] one can consider two set functions associated with a random set. For each event $B \subseteq \mathbb{E}$ these set functions are defined by

$$\widetilde{\varphi}(B) = \sum_{X_i \subseteq B} p_i, \quad \varphi(B) = \sum_{X_i \cap B \neq \emptyset} p_i$$

Clearly, $\widetilde{\varphi}(B) \leq \varphi(B)$ for all $B \subseteq \mathbb{E}$, $\widetilde{\varphi}(\emptyset) = \varphi(\emptyset) = 0$ and $\widetilde{\varphi}(\mathbb{E}) = \varphi(\mathbb{E}) = 1$ if all focal elements are non-empty. Note that $\widetilde{\varphi}$ and φ are dual set functions which means that

$$\widetilde{\varphi}(B) = 1 - \varphi(B^{\mathsf{c}})$$

where B^{c} denotes the set-theoretic complement of B. Furthermore, $\widetilde{\varphi}$ and φ have special properties: The set function $\widetilde{\varphi}$ is *completely monotone* (or monotone of infinite order) which means that for any $k \geq 2$, and $B_1, \dots, B_k \subseteq \mathbb{E}$ it holds that

$$\widetilde{\varphi}\left(\bigcup_{j=1}^{k} B_j\right) \geq \sum_{\emptyset \neq J \subseteq \{1,\ldots,k\}} (-1)^{|J|+1}\, \widetilde{\varphi}\left(\bigcap_{j\in J} B_j\right) \tag{1}$$

By duality, one has that φ is *completely alternating* (or alternating of infinite order), i.e.,

$$\varphi\left(\bigcap_{j=1}^{k} B_j\right) \leq \sum_{\emptyset \neq J \subseteq \{1,\ldots,k\}} (-1)^{|J|+1}\, \varphi\left(\bigcup_{j\in J} B_j\right) \tag{2}$$

Note that for probability measures, equality holds in equations (1) and (2). A completely monotone set function yielding 0 for the empty set and 1 for $\mathbb{E}$ is called a *belief function* whereas a completely alternating set function yielding 0 for the empty set and 1 for $\mathbb{E}$ is called a *plausibility function* [10, 15]. In the framework of imprecise probabilities [16] $\widetilde{\varphi}$ and φ can be interpreted as *lower* and *upper probabilities*. This is due to the fact that the two set functions bound the values of all probability distributions on $\mathbb{E}$ induced by random variables whose values are contained in the focal sets of the random set (see [6], for example).

Example 1. Let $\mathbb{E} = \{x_1, x_2, x_3\}$ and let X be the (finite) random set with focal sets $X_1 = \{x_1, x_2\}$, $X_2 = \{x_3\}$, $X_3 = \{x_2, x_3\}$ and probability weights p_1, p_2, p_3. Then one obtains the following values for the associated belief and plausibility function.

B	$\emptyset$	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1,x_2\}$	$\{x_1,x_3\}$	$\{x_2,x_3\}$	$\mathbb{E}$
$\widetilde{\varphi}(B)$	0	0	0	p_2	p_1	p_2	p_2+p_3	1
$\varphi(B)$	0	p_1	p_1+p_3	p_2+p_3	p_1+p_3	1	1	1

When the basic set $\mathbb{E}$ is finite then also its power set $2^{\mathbb{E}}$, i.e., the set of all subsets of $\mathbb{E}$, is finite: it has $2^{|\mathbb{E}|}$ elements. In this case, a finite random set X with focal elements $X_1, \ldots, X_\ell$ induces a *basic probability assignment* m on $2^{\mathbb{E}}$ [15]:

$$\mathrm{m}(A) = \begin{cases} p_i & \text{if } A = X_i \\ 0 & \text{else} \end{cases}$$

Note that $\mathrm{m}(\emptyset) = 0$ if all focal sets are non-empty and that $\sum_{A\subseteq\mathbb{E}} \mathrm{m}(A) = 1$. By using the basic probability assignment m the belief and the plausibility function associated with X can be written as

$$\widetilde{\varphi}(B) = \sum_{A\subseteq B} \mathrm{m}(A), \quad \varphi(B) = \sum_{A\cap B\neq\emptyset} \mathrm{m}(A)$$

On the other hand, given a belief function $\widetilde{\varphi}$ on $2^{\mathbb{E}}$ there exists a unique basic probability assignment m on $2^{\mathbb{E}}$ such that $\widetilde{\varphi}(B) = \sum_{A\subseteq B} \mathrm{m}(A)$. An explicit formula to compute m from $\widetilde{\varphi}$ is given by the so-called Moebius inversion formula [15]:

$$\mathrm{m}(A) = \sum_{B\subseteq A} (-1)^{|A\setminus B|} \widetilde{\varphi}(B) \tag{3}$$

where $A \setminus B$ denotes the set-theoretic difference.

In general, random sets are random elements defined on an arbitrary probability space Ω equipped with a σ-algebra Σ and a probability measure P. Furthermore, it is usual to assume that $\mathbb{E}$ is an infinite set satisfying certain topological properties (more precisely, $\mathbb{E}$ is a locally compact Hausdorff second countable space) and that the focal sets are closed subsets of $\mathbb{E}$ because this implies favorable properties. A *random closed set* is then a map from Ω into the closed subsets of $\mathbb{E}$, denoted by $\mathscr{F}$, such that for each compact subset K of $\mathbb{E}$ its *upper inverse* [10] $X^-(K)$ of K under X defined by

$$X^-(K) = \{\omega \in \Omega : X(\omega) \cap K \neq \emptyset\}$$

is measurable, i.e., $X^-(K)$ is an element of Σ.

Given a random closed set X one can define a set function on the family of compact subsets of $\mathbb{E}$, denoted by $\mathscr{K}$, which assigns to each $K \in \mathscr{K}$ the probability of the upper inverse, i.e.,

$$\varphi(K) = \mathrm{P}(X^-(K)) = \mathrm{P}(\{\omega : X(\omega) \cap K \neq \emptyset\})$$

It can be shown that the set function φ has the following three properties:

(1) $\varphi(\emptyset) = 0$
(2) φ is completely alternating, i.e., for every finite family of compact sets Equation (2) holds.
(3) φ is continuous from above, i.e., for each decreasing sequence $\{K_n\}_{n=1}^{\infty}$ (i.e., $K_{n+1} \subseteq K_n$ for all $n \geq 1$) it holds that

$$\lim_{n\to\infty} \varphi(K_n) = \varphi\left(\bigcap_{n=1}^{\infty} K_n\right).$$

A set function satisfying these three properties is called a *capacity functional* [5]. Note that the terms capacity functional and plausibility function essentially mean the same but stem from different theories. A minor difference is that a capacity functional does not generally assign the value 1 to $\mathbb{E}$ since the latter is, in general, not compact. But if one assumes that the values of the random set are non-empty with probability 1, then it holds that

$$\sup\{\varphi(K) : K \in \mathscr{K}\} = 1$$

The dual set function $\widetilde{\varphi}$ of the capacity functional φ is called a *containment functional* [5] and assigns to every complement K^{c} of a compact set K the probability that the random set X is contained in K^{c}, i.e.,

$$\widetilde{\varphi}(K^{\mathsf{c}}) = 1 - \varphi(K) = \mathrm{P}\big(\{\omega : X(\omega) \subseteq K^{\mathsf{c}}\}\big)$$

Obviously, containment functionals are essentially the same as belief functions.

Example 2. Random sets with basic probability space $(0,1]$ equipped with the Lebesgue measure λ (uniform probability distribution) are frequently used for uncertainty modeling. They are sometimes called *random sets of indexable type* [1]. As

an example a *Chebychev random set* [12] is considered. Such a random set is useful for modeling uncertainty about a variable where the only information available are values for the mean μ and the standard deviation σ. The random set assigns to each $\omega \in \Omega = (0,1]$ the interval in which the variable lies with a probability of at least $1-\omega$ irrespective of the distribution of the variable. The intervals are constructed from the well-known Chebychev inequality and are given by

$$X(\omega) = \left[\mu - \frac{\sigma}{\sqrt{\omega}}, \mu + \frac{\sigma}{\sqrt{\omega}}\right]$$

In addition to the fact that the random set X is of indexable type, it is special inasmuch, that the focal sets are linearly ordered by set inclusion, i.e., for $\omega_2 \geq \omega_1$ it holds that $X(\omega_2) \subseteq X(\omega_1)$. Thus, upper inverses $X^-(K)$ are always intervals and the values of the capacity functional associated with X are the lengths of these intervals, i.e., for a compact set K one has

$$\varphi(K) \;=\; \lambda\Big(\big(0, \max\{\omega \;:\; X(\omega) \cap K \;\neq\; \emptyset\}\big]\Big) \;=\; \max\{\omega \;:\; X(\omega) \cap K \;\neq\; \emptyset\}$$

This further implies that φ is *maxitive*, i.e., for $K, L \in \mathscr{K}$ one has

$$\varphi(K \cup L) = \max\{\varphi(K), \varphi(L)\}$$

Example 3. Every finite random set can be represented as a random set of indexable type. For the random set from Example 1 this can be done by dividing $(0,1]$ into the subintervals $I_1 = (0, p_1]$, $I_2 = (p_1, p_1 + p_2]$ and $I_3 = (p_1 + p_2, 1]$. Note that the lengths of the intervals correspond to the probability weights. Then X can be defined as an infinite random set by assigning to $\omega \in (0,1]$ the focal set X_i with i such that $\omega \in I_i$.

Similar to the case where $\mathbb{E}$ is finite, a capacity functional uniquely determines a probability distribution on the family $\mathscr{F}$ of closed subsets of $\mathbb{E}$. The following classes of subsets of $\mathscr{F}$ constitute important classes of events ($K \in \mathscr{K}$):

$$\mathscr{F}_K = \{F \in \mathscr{F}(\mathbb{E}) : F \cap K \neq \emptyset\}, \quad \mathscr{F}^K = \{F \in \mathscr{F}(\mathbb{E}) : F \cap K = \emptyset\}$$

Note that these events are related to the containment and the capacity functional of a random closed set X in the following way:

$$\widetilde{\varphi}(K^{\mathsf{c}}) = \mathrm{P}\big(X \in \mathscr{F}^K\big), \quad \varphi(K) = \mathrm{P}\big(X \in \mathscr{F}_K\big)$$

The smallest family of events, i.e., the smallest σ-algebra containing all events of the form $\mathscr{F}_K$, K ranging through $\mathscr{K}$, is called *Effros-σ-algebra* and denoted by $\mathscr{B}(\mathscr{F})$. The Choquet theorem (or Choquet-Kendall-Matheron theorem [5, 7, 10]) states a one to one correspondence between capacity functionals and probability distributions on $\mathscr{F}$. More precisely, the theorem reads as follows.

Theorem 1 ([5, 7, 10]). *Let $\varphi : \mathscr{K} \to [0,1]$ be a capacity functional. Then there exists a unique probability measure Π on $\mathscr{B}(\mathscr{F})$ such that $\varphi(K) = \Pi\big(\mathscr{F}_K\big)$ for all $K \in \mathscr{K}$.*

By duality, a probability measure Π on $\mathscr{F}$ is also uniquely determined by a containment functional $\widetilde{\varphi}$ via the relation

$$\widetilde{\varphi}\big(K^{\mathsf{c}}\big) = \Pi\big(\mathscr{F}^K\big)$$

Note that in the case where $\mathbb{E}$ is finite the Moebius inversion formula (3) admits the explicit computation of the density of the distribution Π.

3 Joint Distributions of Random Sets

In the last section it has been mentioned that the distribution of a random (closed) set can be fully characterized by its capacity (or containment) functional. In [14] it has been shown that similar results hold concerning the joint distribution of finitely many random sets. The aim of this section is to summarize the most important results from [14] without giving the proofs (for proofs the reader is referred to [14]). For the sake of simplicity the considerations are restricted to the two-dimensional case although all statements can be transferred to the n-dimensional case without any problems.

In the following let $\mathbb{E}_1$ and $\mathbb{E}_2$ be two topological spaces with favorable properties (more precisely, let $\mathbb{E}_1$ and $\mathbb{E}_2$ be two locally compact Hausdorff second countable spaces), and let $\mathscr{F}_i$ and $\mathscr{K}_i$ denote the families of closed and compact subsets of $\mathbb{E}_i$, respectively, $i = 1,2$. Let $X_i : \Omega \to \mathscr{F}_i$, $i = 1,2$, be two random closed sets defined on the probability space $(\Omega, \Sigma, \mathrm{P})$. The joint distribution of the two random sets is a probability measure on the product space $\mathscr{F}_1 \times \mathscr{F}_2 = \{(F_1, F_2) : F_i \in \mathscr{F}_i\}$ equipped with the product-σ-algebra $\mathscr{B}(\mathscr{F}_1) \otimes \mathscr{B}(\mathscr{F}_2)$.

Motivated by the one-dimensional case one could define a set function on $\mathscr{K}_1 \times \mathscr{K}_2$ by

$$(K_1, K_2) \mapsto \mathrm{P}\big(X_1 \cap K_1 \neq \emptyset, X_2 \cap K_2 \neq \emptyset\big) = \mathrm{P}\big(X_1^-(K_1) \cap X_2^-(K_2)\big) = \mathrm{P}\left((X_1, X_2) \in \mathscr{F}_{K_1} \times \mathscr{F}_{K_2}\right)$$

i.e., pairs of compact sets (K_1, K_2) is assigned the probability of (X_1, X_2) belonging to the cylindrical event $\mathscr{F}_{K_1} \times \mathscr{F}_{K_2}$. In fact, this approach is not fruitful. It is more favorable to assign to pairs of compact sets the probability of the union of their upper inverses.

Proposition 1 ([14]). *Let $X_i : \Omega \to \mathscr{F}_i$, $i = 1,2$, be random closed sets on a probability space $(\Omega, \Sigma, \mathrm{P})$. Then*

$$\psi : \mathscr{K}_1 \times \mathscr{K}_2 \to [0,1], (K_1, K_2) \mapsto \mathrm{P}\big(X_1^-(K_1) \cup X_2^-(K_2)\big)$$

has the following properties:

(1) $\psi(\emptyset,\emptyset)=0$

(2) ψ *is jointly completely alternating, i.e., for all* $k\geq 2$, $1\leq j\leq k$, $K_i^j\in\mathscr{K}_i$, $i=1,2$ *it holds that*

$$\psi\left(\bigcap_{j=1}^{k}K_1^j,\bigcap_{j=1}^{k}K_2^j\right)\leq\sum_{\emptyset\neq J\subseteq\{1,\ldots,k\}}(-1)^{|J|+1}\,\psi\left(\bigcup_{j\in J}K_1^j,\bigcup_{j\in J}K_2^j\right).$$

(3) ψ *is jointly continuous from above, i.e., for all decreasing sequences* $\{K_i^k\}_{k=1}^{\infty}\subseteq\mathscr{K}_i$ *(i.e.,* $K_i^{k+1}\subseteq K_i^k$*),* $i=1,2$*, it holds that*

$$\lim_{k\to\infty}\psi(K_1^k,K_2^k)=\psi\left(\bigcap_{k=1}^{\infty}K_1^k,\bigcap_{k=1}^{\infty}K_2^k\right)$$

Note that for each component the set function ψ simultaneously satisfies the conditions of a capacity functional. Thus a set function satisfying the above three conditions shall be called a *joint (or multivariate) capacity functional*. Furthermore, note that

$$\psi(K_1,K_2)=\mathrm{P}\big((X_1,X_2)\in(\mathscr{F}_{K_1}\times\mathscr{F}_2)\cup(\mathscr{F}_1\times\mathscr{F}_{K_2})\big)$$

i.e., ψ assigns to the pair (K_1,K_2) the probability of the event

$$(\mathscr{F}_{K_1}\times\mathscr{F}_2)\cup(\mathscr{F}_1\times\mathscr{F}_{K_2})=\{(F_1,F_2):F_1\cap K_1\neq\emptyset\text{ or }F_2\cap K_2\neq\emptyset\}.$$

Similar to the one-dimensional case a joint probability distribution of two (or more generally finitely many) random sets is completely determined by a joint capacity functional.

Proposition 2 ([14]). *Let* $\psi:\mathscr{K}_1\times\mathscr{K}_2\to[0,1]$ *be a joint capacity functional (i.e., a set function satisfying Conditions (1) - (3) of Proposition 1). Then there exists a unique probability measure* $\Pi:\mathscr{B}(\mathscr{F}_1)\otimes\mathscr{B}(\mathscr{F}_2)\to[0,1]$ *such that for all* $(K_1,K_2)\in\mathscr{K}_1\times\mathscr{K}_2$ *it holds that*

$$\psi(K_1,K_2)=\Pi\Big(\big(\mathscr{F}_{K_1}\times\mathscr{F}_2\big)\cup\big(\mathscr{F}_1\times\mathscr{F}_{K_2}\big)\Big)$$

The dual set function of ψ is defined on $\mathscr{K}_1^{\mathsf{c}}\times\mathscr{K}_2^{\mathsf{c}}$ and is given by

$$\begin{aligned}\widetilde{\psi}(K_1^{\mathsf{c}},K_2^{\mathsf{c}})&=1-\psi(K_1,K_2)=1-\mathrm{P}\big(X_1^-(K_1)\cup X_2^-(K_2)\big)\\&=\mathrm{P}\big(X_1\cap K_1=\emptyset,X_2\cap K_2=\emptyset\big)=\mathrm{P}\big((X_1,X_2)\in\mathscr{F}^{K_1}\times\mathscr{F}^{K_2}\big)\end{aligned}$$

By duality, $\widetilde{\psi}$ has the following properties and can be called a *joint (or multivariate) containment functional*:

(1) $\widetilde{\psi}(\mathbb{E}_1,\mathbb{E}_2)=1$

(2) $\widetilde{\psi}$ is jointly completely monotone, i.e., for all $k\geq 2$, $1\leq j\leq k$, $L_i^j\in\mathscr{K}_i^{\mathsf{c}}$, $i=1,2$ it holds that

$$\widetilde{\psi}\left(\bigcup_{j=1}^{k} L_1^j, \bigcup_{j=1}^{k} L_2^j\right) \geq \sum_{\emptyset \neq J \subseteq \{1,\ldots,k\}} (-1)^{|J|+1} \widetilde{\psi}\left(\bigcap_{j\in J} L_1^j, \bigcap_{j\in J} L_2^j\right).$$

(3) $\widetilde{\psi}$ is jointly continuous from below, i.e., for all increasing sequences $\{L_i^k\}_{k=1}^{\infty} \subseteq \mathscr{K}_i^{\mathsf{c}}$ (i.e., $L_i^k \subseteq L_i^{k+1}$), $i = 1,2$, it holds that

$$\lim_{k\to\infty} \widetilde{\psi}(L_1^k, L_2^k) = \widetilde{\psi}\left(\bigcup_{k=1}^{\infty} L_1, \bigcup_{k=1}^{\infty} L_2\right)$$

By duality and Proposition 2, one can conclude that a probability measure Π on $\mathscr{B}(\mathscr{F}_1) \otimes \mathscr{B}(\mathscr{F}_2)$ is uniquely determined by a joint containment functional $\widetilde{\psi}$ via the relation

$$\begin{aligned}\widetilde{\psi}(K_1^{\mathsf{c}}, K_2^{\mathsf{c}}) = 1 - \psi(K_1, K_2) &= 1 - \Pi\Big(\big(\mathscr{F}_{K_1} \times \mathscr{F}_2\big) \cup \big(\mathscr{F}_1 \times \mathscr{F}_{K_2}\big)\Big) \\ &= \Pi\Big(\big(\mathscr{F}^{K_1} \times \mathscr{F}_2\big) \cap \big(\mathscr{F}_1 \times \mathscr{F}^{K_2}\big)\Big) = \Pi\big(\mathscr{F}^{K_1} \times \mathscr{F}^{K_2}\big)\end{aligned}$$

It should be pointed out that joint distributions of random (closed) sets, i.e., probability measures on $\mathscr{F}_1 \times \mathscr{F}_2$ can also be characterized by set functions defined on certain classes of subsets of $\mathbb{E}_1 \times \mathbb{E}_2$. The canonical but misleading way would be to consider capacity functionals on $\mathscr{K}(\mathbb{E}_1 \times \mathbb{E}_2)$, i.e., the family of compact subsets of the product space $\mathbb{E}_1 \times \mathbb{E}_2$. Application of the Choquet theorem would lead to a probability measure on the family of closed subsets of $\mathbb{E}_1 \times \mathbb{E}_2$, i.e., $\mathscr{F}(\mathbb{E}_1 \times \mathbb{E}_2) = \{F \subseteq \mathbb{E}_1 \times \mathbb{E}_2 \text{ closed}\}$. The latter is obviously not the same as $\mathscr{F}_1 \times \mathscr{F}_2 = \{(F_1, F_2) : F_i \in \mathscr{F}_i\}$. It has been shown in [14] that the family $\hat{\mathscr{K}}_{\cup}^2 = \{K_1 \times \mathbb{E}_2 \cup \mathbb{E}_1 \times K_2 : K_i \in \mathscr{K}_i\}$ is suitable.

Proposition 3 ([14]). *Let $\phi : \hat{\mathscr{K}}_{\cup}^2 \to [0,1]$ be a capacity functional, i.e., $\phi(\emptyset) = 0$, ϕ is completely alternating and continuous from above for sets from $\hat{\mathscr{K}}_{\cup}^2$. Then there exists a unique probability measure $\Pi : \mathscr{B}(\mathscr{F}_1) \otimes \mathscr{B}(\mathscr{F}_2) \to [0,1]$ such that*

$$\phi\Big(\big(K_1 \times \mathbb{E}_2\big) \cup \big(\mathbb{E}_1 \times K_2\big)\Big) = \Pi\Big(\big(\mathscr{F}_{K_1} \times \mathscr{F}_2\big) \cup \big(\mathscr{F}_1 \times \mathscr{F}_{K_2}\big)\Big)$$

for all $K_i \in \mathscr{K}_i$. If, in addition, it holds that

$$\sup\{\phi(K_1 \times \mathbb{E}_2) : K_1 \in \mathscr{K}_1\} = 1 \text{ and } \sup\{\phi(\mathbb{E}_1 \times K_2) : K_2 \in \mathscr{K}_2\} = 1$$

then $\Pi\big((\mathscr{F}_1 \setminus \{\emptyset\}) \times (\mathscr{F}_2 \setminus \{\emptyset\})\big) = 1$ and for all $L \in \hat{\mathscr{K}}_{\cup}^2$ it holds that

$$\phi(L) = \Pi\big(\{(F_1, F_2) \in \mathscr{F}_1 \times \mathscr{F}_2 : F_1 \times F_2 \cap L \neq \emptyset\}\big).$$

The additional condition implies that an empty set only appears with probability zero, and in this case the relation between ϕ and Π is very similar to that in the Choquet theorem. Of course, one can again give a dual formulation of the above proposition. To this end, consider the dual set function of ϕ which is defined on the complements of $\hat{\mathscr{K}}_{\cup}^2$. The latter are the cylindrical sets whose components are

complements of compact sets, i.e., $(\mathscr{K}^c)^2_\times = \{K_1^c \times K_2^c : K_i \in \mathscr{K}_i\}$. Thus, a probability measure Π on $\mathscr{F}_1 \times \mathscr{F}_2$ is uniquely determined by a containment functional ϕ defined on the cylindrical sets $(\mathscr{K}^c)^2_\times$ via the relation $\widetilde{\phi}(K_1^c \times K_2^c) = \Pi\left(\mathscr{F}^{K_1} \times \mathscr{F}^{K_2}\right)$.

In the case of finite spaces $\mathbb{E}_1$ and $\mathbb{E}_2$ one speaks of belief and plausibility functions, again. A *joint belief function* is a set function ϕ defined on $2^{\mathbb{E}_1} \times 2^{\mathbb{E}_2} = \{(A_1,A_2) : A_i \subseteq \mathbb{E}_i\}$ such that $\widetilde{\phi}(\emptyset,\emptyset) = 0$, $\widetilde{\phi}(\mathbb{E}_1,\mathbb{E}_2) = 1$ and which is jointly completely monotone. By duality, a *joint plausibility function* is a set function ϕ defined on $2^{\mathbb{E}_1} \times 2^{\mathbb{E}_2}$ such that $\widetilde{\phi}(\emptyset,\emptyset) = 0$, $\widetilde{\phi}(\mathbb{E}_1,\mathbb{E}_2) = 1$ and which is jointly completely alternating in each component. Proposition 2 implies that given a joint belief function it is possible to find a probability distribution on $2^{\mathbb{E}_1} \times 2^{\mathbb{E}_2}$. As in the one-dimensional case it is even possible to compute a joint density (basic probability assignment), i.e., a set function m on $2^{\mathbb{E}_1} \times 2^{\mathbb{E}_2}$ that satisfies

(1) $\mathrm{m}(\emptyset,\emptyset) = 0$
(2) $\sum_{A_1 \subseteq \mathbb{E}_1, A_2 \subseteq \mathbb{E}_2} \mathrm{m}(A_1,A_2) = 1.$

The following proposition shows how to obtain a joint belief function from m and how to compute a joint basic probability assignment from a joint belief function. Formula (5) can be seen as a two-dimensional Moebius inversion formula.

Proposition 4 ([14]). *Let* $\mathrm{m} : 2^{\mathbb{E}_1} \times 2^{\mathbb{E}_2} \to [0,1]$ *be a joint basic probability assignment. Then*

$$\widetilde{\psi} : 2^{\mathbb{E}_1} \times 2^{\mathbb{E}_2} \to [0,1], (A_1,A_2) \mapsto \sum_{B_1 \subseteq A_1, B_2 \subseteq A_2} \mathrm{m}(B_1,B_2) \tag{4}$$

is a joint belief function. On the other hand, if $\widetilde{\psi} : 2^{\mathbb{E}_1} \times 2^{\mathbb{E}_2} \to [0,1]$ *is a joint belief function then*

$$\mathrm{m}(A_1,A_2) = \sum_{B_1 \subseteq A_1, B_2 \subseteq A_2} (-1)^{|A_1 \setminus B_1| + |A_2 \setminus B_2|} \widetilde{\psi}(B_1,B_2) \tag{5}$$

is the unique joint basic probability assignment such that Equation (4) holds.

4 Random Sets and Copulas

This section is addressed to the question how copulas can be used to describe the relation between joint distributions of random sets and its marginal distributions. Again, the considerations are restricted to the case of two random sets.

Copulas are a useful tool for modeling dependence of random variables. A bivariate *copula* C is a function $C : [0,1]^2 \to [0,1]$ satisfying the following properties:

(1) $C(u_1,0) = 0$ and $C(0,u_2) = 0$ for all $u_1,u_2 \in [0,1]$
(2) $C(u_1,1) = u_1$ and $C(1,u_2) = u_2$ for all $u_1,u_2 \in [0,1]$
(3) $C(v_1,v_2) - C(v_1,u_2) - C(u_1,v_2) + C(u_1,u_2) \geq 0$ for all $u_i \leq v_i$, $i = 1,2$

A bivariate *subcopula* is a function whose domain is a subset D of $[0,1]^2$ and which satisfies the preceding three conditions on D.

The joint distribution of two random variables ξ_1, ξ_2 is fully determined by the joint distribution function ($x_1, x_2 \in \mathbb{R}$)

$$F(x_1,x_2) = \mathrm{P}(\xi_1 \le x_1, \xi_2 \le x_2) = \mathrm{P}\big((\xi_1,\xi_2) \in (-\infty,x_1] \times (-\infty,x_2]\big)$$

Hence, the rectangular events $(-\infty,x_1] \times (-\infty,x_2]$ completely determine the distribution of the random vector (ξ_1,ξ_2). The marginal distribution functions of ξ_1 and ξ_2 are given by

$$F_i(x_i) = \mathrm{P}\big(\xi_i \le x_i\big) = \mathrm{P}\big(\xi_i \in (-\infty,x_i]\big)$$

The well-known Sklar theorem [20] says that given a joint distribution function F there exists a copula C that links F to its marginals F_1, F_2 by

$$F(x_1,x_2) = C\big(F_1(x_1),F_2(x_2)\big)$$

The copula C is uniquely determined if F_1 and F_2 take all values between 0 and 1. On the other hand, given two marginal distribution functions F_1, F_2 and a copula C, the mapping $(x_1,x_2) \mapsto C\big(F_1(x_1),F_2(x_2)\big)$ yields a joint distribution function. Note that in the formulation of Sklar's theorem there is no reference to the underlying random variables, it only requires the knowledge of the distribution functions. For further information on copulas the reader is referred to the monograph [20].

The question arises whether copulas can be used to describe dependence of random sets in a similar manner. The role of distribution functions should be played by (joint) capacity or containment functionals since the latter uniquely determine the (joint) distribution of random sets (see Theorem 1 and Proposition 2). The desired result would thus be that a joint capacity or containment functional can be related to its marginals by a copula. The marginals of a joint capacity functional $\psi : \mathscr{K}_1 \times \mathscr{K}_2 \to [0,1]$ are given by

$$\varphi_1(K_1) = \psi(K_1,\emptyset), \quad \varphi_2(K_2) = \psi(\emptyset,K_2)$$

This can be seen by using Theorem 1 and Proposition 2

$$\psi(K_1,\emptyset) = \Pi\Big(\big(\mathscr{F}_{K_1} \times \mathscr{F}_2\big) \cup \big(\mathscr{F}_1 \times \mathscr{F}_\emptyset\big)\Big) = \Pi\big(\mathscr{F}_{K_1} \times \mathscr{F}_2\big) = \Pi_1\big(\mathscr{F}_{K_1}\big) = \varphi_1(K_1)$$

where Π_1 is the marginal distribution of Π with respect to the first component. By duality, the marginals of a joint containment functional $\widetilde{\psi} : \mathscr{K}_1^{\mathsf{c}} \times \mathscr{K}_2^{\mathsf{c}} \to [0,1]$ are given by

$$\widetilde{\varphi}_1(K_1^{\mathsf{c}}) = \widetilde{\psi}(K_1^{\mathsf{c}},\mathbb{E}_2), \quad \widetilde{\varphi}_2(K_2^{\mathsf{c}}) = \widetilde{\psi}(\mathbb{E}_1,K_2^{\mathsf{c}})$$

Scarsini [13] has presented a generalization of Sklar's theorem which applies to probability measures on very general spaces. More precisely, he considers Polish spaces (i.e., completely metrizable separable spaces) and shows that when choosing in each component an increasing family of subsets, i.e., a family of subsets which can be linearly ordered by set inclusion, there exists a unique subcopula that links

the joint distribution to the marginals when the latter are restricted to the chosen increasing subclasses. Since the families $\mathscr{F}_1$ and $\mathscr{F}_2$ of closed subsets of $\mathbb{E}_1$ and $\mathbb{E}_2$ are Polish spaces (when equipped with the so-called Fell topology - see [2] for details) one can make use of [13, Theorem 3.1] to obtain the following Proposition.

Proposition 5. *Let $\widetilde{\psi} : \mathscr{K}_1^{\mathsf{c}} \times \mathscr{K}_2^{\mathsf{c}} \to [0,1]$ be a joint containment functional and let $\widetilde{\varphi}_1$ and $\widetilde{\varphi}_2$ denote its marginal containment functionals. Furthermore, let $\mathscr{I}_1$ and $\mathscr{I}_2$ denote increasing families of subsets of $\mathscr{K}_1^{\mathsf{c}}$ and $\mathscr{K}_2^{\mathsf{c}}$, respectively. Then there exists a unique subcopula $C^{\mathscr{I}_1,\mathscr{I}_2}$ on $\widetilde{\varphi}_1(\mathscr{I}_1) \times \widetilde{\varphi}_2(\mathscr{I}_2)$ such that*

$$\widetilde{\psi}\big(K_1^{\mathsf{c}}, K_2^{\mathsf{c}}\big) = C^{\mathscr{I}_1,\mathscr{I}_2}\Big(\widetilde{\varphi}_1\big(K_1^{\mathsf{c}}\big), \widetilde{\varphi}_2\big(K_2^{\mathsf{c}}\big)\Big) \tag{6}$$

for each $K_1^{\mathsf{c}} \in \mathscr{I}_1$, $K_2^{\mathsf{c}} \in \mathscr{I}_2$.

Proof. By Theorem 1 and Proposition 2 and by duality $\widetilde{\psi}$, $\widetilde{\varphi}_1$ and $\widetilde{\varphi}_2$ uniquely determine probability measures Π, Π_1 and Π_2 on $\mathscr{F}_1 \times \mathscr{F}_2$, $\mathscr{F}_1$ and $\mathscr{F}_2$, respectively. The families $\mathscr{F}_1$ and $\mathscr{F}_2$ of closed subsets are Polish spaces when equipped with the so-called Fell topology [2], and thus $\mathscr{F}_1 \times \mathscr{F}_2$ is also Polish. Since $\mathscr{I}_i$ are increasing subclasses of $\mathscr{K}_i^{\mathsf{c}}$, $i = 1,2$, the families $\mathscr{A}_i = \{\mathscr{F}^{K_i} : K_i^{\mathsf{c}} \in \mathscr{I}_i\}$ are increasing subclasses of $\mathscr{F}_i$, $i = 1,2$. Thus, [13, Theorem 3.1] implies that there exists a unique subcopula C on

$$\Pi_1(\mathscr{A}_1) \times \Pi_2(\mathscr{A}_2) = \widetilde{\varphi}_1(\mathscr{I}_1) \times \widetilde{\varphi}_2(\mathscr{I}_2) = \big\{(\widetilde{\varphi}_1(K_1^{\mathsf{c}}), \widetilde{\varphi}_2(K_2^{\mathsf{c}})) : K_i \in \mathscr{I}_i\big\}$$

such that for all $K_i^{\mathsf{c}} \in \mathscr{I}_i$, $i = 1,2$ it holds that

$$\Pi(\mathscr{F}^{K_1} \times \mathscr{F}^{K_2}) = C\big(\Pi_1(\mathscr{F}^{K_1}), \Pi_2(\mathscr{F}^{K_2})\big)$$

But this implies Equation (6) since $\Pi(\mathscr{F}^{K_1} \times \mathscr{F}^{K_2}) = \widetilde{\psi}(K_1^{\mathsf{c}}, K_2^{\mathsf{c}})$ and $\Pi_i(\mathscr{F}^{K_i}) = \widetilde{\varphi}(K_i^{\mathsf{c}})$, $i = 1,2$, by Proposition 2 and Theorem 1.

The proposition suggests that in contrast to the classical case (of random variables), a single (sub-) copula is not enough to link a joint containment functional to its marginal containment functionals. In fact, it seems that a whole family of copulas is necessary to completely describe the relation. It is enough, though, to consider increasing families $\mathscr{I}_i$ in $\mathscr{K}_i^{\mathsf{c}}$ and the so induced increasing families $\{\mathscr{F}^{K_i} : K_i \in \mathscr{I}_i\}$ in $\mathscr{F}_i$. This is due to the fact that the family $\{\mathscr{F}^{K_i} : K_i \in \mathscr{K}_i\}$ is closed under finite intersections, i.e., if $K_i, L_i \in \mathscr{K}_i$ then $\mathscr{F}^{K_i} \cap \mathscr{F}^{L_i} \in \{\mathscr{F}^{K_i} : K_i \in \mathscr{K}_i\}$ since $\mathscr{F}^{K_i} \cap \mathscr{F}^{L_i} = \mathscr{F}^{K_i \cup L_i}$. Furthermore, $\{\mathscr{F}^{K_i} : K_i \in \mathscr{K}_i\}$ is a generator of the Effros-σ-algebra $\mathscr{B}(\mathscr{F}_i)$. Thus, the probability measures Π_i induced by the containment functionals $\widetilde{\varphi}_i$ are completely determined by their values on $\{\mathscr{F}^{K_i} : K_i \in \mathscr{K}_i\}$. In a similar manner, the joint distribution Π (induced by $\widetilde{\psi}$) is completely determined by its values on $\{\mathscr{F}^{K_1} \times \mathscr{F}^{K_2} : K_i \in \mathscr{K}_i, i = 1,2\}$. It remains as a topic for further research if all increasing subclasses of $\mathscr{K}_i^{\mathsf{c}}$ have to be considered to obtain the complete relation between a joint containment functional and its marginals or if it is

enough to restrict oneself to certain subfamilies of $\mathscr{K}_i^{\mathsf{c}}$. Furthermore, it has to be investigated if the (sub-) copulas induced by the increasing subclasses have to satisfy some kind of compatibility conditions.

It should be noted that there have already been attempts to use copulas for modeling the dependence of random sets. Alvarez [1] has presented an approach which can be applied to random sets of indexable type (see Example 2) whose underlying probability space is the interval $(0,1]$, and which directly deals with random sets instead of their capacity (or containment) functionals. The approach makes use of the fact that associated with a copula C one can define a measure μ_C on $(0,1]^2$ by

$$\mu_C\big((u_1,v_1]\times(u_2,v_2]\big) = C(v_1,v_2) - C(v_1,u_2) - C(u_1,v_2) + C(u_1,u_2)$$

where $u_i \le v_i$, $u_i, v_i \in [0,1]$, $i = 1,2$. The idea in [1] is to define a joint random (closed) set X from marginal random (closed) sets X_1 and X_2. The underlying probability space is $\Omega = (0,1]^2$ equipped with the measure μ_C and the focal sets of X are defined by $X(u_1,u_2) = X_1(u_1)\times X_2(u_2)$. One can define the following set function which is a joint containment functional by Proposition 1:

$$\begin{aligned}\widetilde{\psi}(K_1^{\mathsf{c}},K_2^{\mathsf{c}}) &= \mu_C\big(\{(u_1,u_2) : X(u_1,u_2)\subseteq K_1^{\mathsf{c}}\times K_2^{\mathsf{c}}\}\big)\\ &= \mu_C\big(\{(u_1,u_2) : X_1(u_1)\subseteq K_1^{\mathsf{c}}, X_2(u_2)\subseteq K_2^{\mathsf{c}}\}\big) \quad (7)\end{aligned}$$

Of course, this approach is also applicable for random sets in finite spaces $\mathbb{E}_1$, $\mathbb{E}_2$, if they are represented as random sets of indexable type as explained in Example 3. However, $\widetilde{\psi}$ defined in (7) depends on the order of the subdivisions of the intervals $(0,1]$ since μ_C not only depends on the lengths of the subintervals but also on their bounds. This has recently been noted by Nguyen [11] who has also shown that (in case of finite sets $\mathbb{E}_1$, $\mathbb{E}_2$) every joint belief function can be represented in the form (7) with X_1 and X_2 being random sets constructed from the marginal belief functions and their densities.

Note that Alvarez approach differs from the approach given in Proposition 5 not only by the fact that it directly uses random sets, but also by the fact that the copula is used on the basic probability spaces instead of the image spaces. When regarding random sets with basic probability space $(0,1]$ the question arises if there exists a (simple) relation between the copula used in (7) and the family of copulas from Proposition 5. The following example suggests that this is not the case.

Example 4. Let $\mathbb{E}_1 = \{x_1,x_2,x_3\}$, $\mathbb{E}_2 = \{y_1,y_2\}$, let X_1 be the random set from Example 1 and let Y_2 be the random set with focal sets $\{y_2\}$, $\{y_1,y_2\} = \mathbb{E}_2$, $\{y_1\}$ and probability weights q_1, q_2, q_3, respectively. Furthermore, let $\widetilde{\varphi}_1$, $\widetilde{\varphi}_2$ be the belief functions associated with X_1, X_2, let C be a copula and let $\widetilde{\psi}$ be the joint belief function defined by Equation (7). The following table lists all combinations of sets for which $\widetilde{\psi}$ yields non-trivial values.

A_1	A_2	$\widetilde{\varphi}(A_1)$	$\widetilde{\varphi}(A_2)$	$\widetilde{\psi}(A_1,A_2)$
$\{x_3\}$	$\{y_1\}$	p_2	q_3	$p_2 - C(p_1+p_2,q_1+q_2) + C(p_1,q_1+q_2)$
$\{x_3\}$	$\{y_2\}$	p_2	q_1	$C(p_1+p_2,q_1) - C(p_1,q_1)$
$\{x_1,x_2\}$	$\{y_1\}$	p_1	q_3	$p_1 - C(p_1,q_1+q_2)$
$\{x_1,x_2\}$	$\{y_2\}$	p_1	q_1	$C(p_1,q_1)$
$\{x_2,x_3\}$	$\{y_1\}$	p_2+p_3	q_3	$1-p_1-q_1-q_2+C(p_1,q_1+q_2)$
$\{x_2,x_3\}$	$\{y_2\}$	p_2+p_3	q_1	$q_1 - C(p_1,q_1)$
$\{x_1,x_3\}$	$\{y_1\}$	p_2	q_3	$p_2 - C(p_1+p_2,q_1+q_2) + C(p_1,q_1+q_2)$
$\{x_1,x_3\}$	$\{y_2\}$	p_2	q_1	$C(p_1+p_2,q_1) - C(p_1,q_1)$

Consider the increasing classes $\mathscr{I}_1 = \{\emptyset, \{x_3\}, \{x_2,x_3\}, \mathbb{E}_1\}$, $\mathscr{I}_2 = \{\emptyset, \{y_2\}, \mathbb{E}_2\}$. By Proposition 5 there exists a subcopula $C^{\mathscr{I}_1,\mathscr{I}_2}$ defined on $\{0, p_2, p_2+p_3, 1\} \times \{0, q_1, 1\}$ whose only non-trivial values are

$$C^{\mathscr{I}_1,\mathscr{I}_2}(p_2,q_1) = C(p_1+p_2,q_1) - C(p_1,q_1)$$
$$C^{\mathscr{I}_1,\mathscr{I}_2}(p_2+p_3,q_1) = q_1 - C(p_1,q_1)$$

These equations suggest that there does not exist a trivial relation between the copulas C and $C^{\mathscr{I}_1,\mathscr{I}_2}$.

5 Conclusion and Outlook

After reviewing the most important facts about random sets generalizations of the Choquet theorem and the Moebius inversion formula to the multidimensional case have been presented. It has been demonstrated that joint distributions of random sets can be characterized by multivariate analogues of containment or capacity functionals (Proposition 2) or by usual containment or capacity functionals whose domain is restricted to particular cylindrical subsets of the product space or their complements, respectively (Proposition 3).

Section 4 was devoted to the question how copulas can be used to describe dependence of random sets. Proposition 5 has been derived from a more general result from [13] and says that the relation between a joint containment functional and its marginals can be described by a family of copulas depending on increasing classes of subsets. In addition, an approach from [1] has been reviewed that uses copulas for modeling dependence of random sets of indexable type, and which seems to be fundamentally different from the approach of Proposition 5.

As a topic for further research the relation between the two approaches should be studied in more detail. Furthermore, it should be investigated if (and how) Equation (6) can be used to model dependence of random sets (or containment functionals). One open question is whether the copulas for each increasing class can be chosen arbitrarily or whether they have to satisfy some compatibility conditions. Moreover, it should be investigated if (7) can be used in general to describe the

dependence of random sets by a single copula. Recently, Nguyen [11] has given a positive answer for the case of finite sets.

References

1. Alvarez, D.A.: A Monte Carlo-based method for the estimation of lower and upper probabilities of events using infinite random sets of indexable type. Fuzzy Sets and Systems 160, 384–401 (2009)
2. Beer, G.: Topologies on Closed and Closed Convex Sets. Kluwer Academic Publishers, Dordrecht (1993)
3. Dempster, A.P.: Upper and lower probabilities induced by a multivalued mapping. Annals of Mathematical Statistics 38, 325–339 (1967)
4. Kruse, R., Meyer, K.D.: Statistics with vague data. D. Reidel Publishing Company, Dordrecht (1987)
5. Matheron, G.: Random Sets and Integral Geometry. Wiley (1975)
6. Miranda, E., Couso, I., Gil, P.: Approximations of upper and lower probabilities by measurable selections. Information Sciences 180, 1407–1417 (2010)
7. Molchanov, I.: Theory of random sets. Springer, London (2005)
8. Nelsen, R.B.: An Introduction to Copulas. Springer (2006)
9. Nguyen, H.T.: An Introduction to Random Sets. Chapman & Hall/CRC (2006)
10. Nguyen, H.T.: On random sets and belief functions. J. Math. Anal. Appl. 65, 531–542 (1978)
11. Nguyen, H.T.: Combining dependent evidence. Research Notes (2013)
12. Oberguggenberger, M., Fellin, W.: Reliability bounds through random sets: nonparametric methods and geotechnical applications. Computers and Structures 86, 1093–1101 (2008)
13. Scarsini, M.: Copulae of probability measures on product spaces. J. Mult. Anal. 31, 201–219 (1989)
14. Schmelzer, B.: Characterizing joint distributions of random sets by multivariate capacities. Journal of Approximate Reasoning 53, 1228–1247 (2012)
15. Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press, Princeton (1976)
16. Walley, P.: Statistical Reasoning with Imprecise Probabilities. Chapman & Hall, London (1991)

Vine Copulas As a Way to Describe and Analyze Multi-Variate Dependence in Econometrics: Computational Motivation and Comparison with Bayesian Networks and Fuzzy Approaches

Songsak Sriboonchitta, Jianxu Liu, Vladik Kreinovich, and Hung T. Nguyen

Abstract. In the last decade, vine copulas emerged as a new efficient techniques for describing and analyzing multi-variate dependence in econometrics; see, e.g., [1, 2, 3, 7, 9, 10, 11, 13, 14, 21]. Our experience has shown, however, that while these techniques have been successfully applied to many practical problems of econometrics, there is still a lot of confusion and misunderstanding related to vine copulas. In this paper, we provide a motivation for this new technique from the computational viewpoint. We show that other techniques used to described dependence – Bayesian networks and fuzzy techniques – can be viewed as a particular case of vine copulas.

1 Copulas – A Useful Tool in Econometrics: Motivations and Descriptions

Need for Studying Dependence in Econometrics

Many researchers have observed that economics is more complex than physics. In physics, many parameters, many phenomena are independent. As a result, we can

Songsak Sriboonchitta · Jianxu Liu
Department of Economics, Chiang Mai University, Chiang Mai, Thailand
e-mail: songsakecon@gmail.com, liujianxu1984@163.com

Vladik Kreinovich
Department of Computer Science, University of Texas at El Paso, 500 W. University, El Paso, TX 79968, USA
e-mail: vladik@utep.edu

Hung T. Nguyen
Department of Mathematical Sciences, New Mexico State University, Las Cruces, New Mexico 88003, USA, and Department of Economics, Chiang Mai University, Chiang Mai, Thailand
e-mail: hunguyen@nmsu.edu

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_11,

observe (and thoroughly study) simple systems which can be described by a small number of parameters. Based on these simple systems, we can separately determine the laws that describe mechanics, electrodynamics, thermodynamics, etc., and then combine these laws to describe more complex phenomena.

In contrast, in economics, most phenomena are interrelated. Thus, to numerically describe economic phenomena, we need to take into account several dependent parameters. So, in econometrics, studying dependence is of utmost importance.

Statistical Character of Economic Phenomena

An additional complexity of economics – as compared to physics – is that while most physical processes are deterministic, in economics, we can only make statistical predictions. If we repeatedly drop the same object from the Leaning Tower of Pisa (as Galileo did), we will largely observe the exact same behavior every time. In contrast, if several very similar restaurants open in the same area, some of them will survive and some will not, and it is practically impossible to predict which will survive – at best, we can predict the probability of survival. We can deterministically predict the future trajectory of a spaceship, but we can, at best, make statistical predictions about the future values of a stock index.

Conclusion: We Need to Study Dependence between Random Variables

Because of the statistical character of economic phenomena, each parameter describing the economics is a random variables. Thus, the need to study dependence means that we need to study dependence between random variables.

Simplest Case When Random Variables Are Independent: Reminder

In order to analyze how to describe dependence of random variables, let us recall how *in*dependent random variables can be described.

In general, a random variable X_i can be described by its cumulative distribution function $F_i(x_i) \stackrel{\text{def}}{=} \text{Prob}(X_i \le x_i)$. If two random variables X_1 and X_2 are independent, this means that their joint distribution function $F(x_1,x_2) \stackrel{\text{def}}{=} \text{Prob}(X_1 \le x_1 \,\&\, X_2 \le x_2)$ is equal to the product of the marginal distributions $F_1(x_1)$ and $F_2(x_2)$: $F(x_1,x_2) = F_1(x_1) \cdot F_2(x_2)$.

Towards Describing Dependence between Two Random Variables: The Notion of a Copula

In the independent case, general, the joint distribution function $F(x_1,x_2)$ of two random variables X_1 and X_2 is equal to the product $F_1(x_1) \cdot F_2(x_2)$ of the marginal distributions. In general, when the random variables X_1 and X_2 are dependent, the joint distribution function $F(x_1,x_2)$ is different from the product $F_1(x_1) \cdot F_2(x_2)$. It is reasonable to describe this general joint distribution in such a way that we will clearly see how different is the joint distribution from the independent case. In the independent case, $F(x,x_2)$ is the product of the marginal distributions $F_1(x_1)$ and

$F_2(x_2)$; to describe deviations from this product, it make sense to consider more general combination functions, i.e., to consider expressions of the type

$$F(x_1,x_2) = C(F_1(x_1),F_2(x_2)). \tag{1}$$

Such combination functions $C(a,b)$ are known as *copulas*; see, e.g., [19, 26] (see also [1, 2, 3, 7, 9, 10, 11, 13, 14, 21]).

The independence case corresponds to the product combination function $C(a,b) = a \cdot b$. The more the combination function $C(a,b)$ is different from the product, the more dependent are the random variables X_1 and X_2.

Probability Density Function in Terms of the Copula

The expression for the probability density function $f(x_1,x_2) = \dfrac{\partial^2 F(x_1,x_2)}{\partial x_1 \partial x_2}$ in terms of the copula can be obtained by differentiating the above formula with respect to x_1 and x_2. As a result, we get the expression

$$f(x_1,x_2) = c(F_1(x_1),F_2(x_2)) \cdot f_1(x_1) \cdot f_2(x_2), \tag{2}$$

where $c(a,b) \stackrel{\text{def}}{=} \dfrac{\partial^2 C(a,b)}{\partial a\, \partial b}$ and $f_i(x_i) \stackrel{\text{def}}{=} \dfrac{dF_i(x_i)}{dx_i}$ are probability densities of the marginal distributions.

Can Copulas Describe All Possible Dependencies?

The expression (1) is a natural generalization of the independence case. At first glance, it may sound that such expressions describe some special class of dependent variables. However, it can be shown that this expression is general enough to capture the general dependence between random variables. Namely, for continuous distributions, e.g., for distributions with well-defined probability density functions, once we know the joint distribution function $F(x_1,x_2)$ and marginal distributions $F_1(x_1)$ and $F_2(x_2)$, we can get the representation (1) if we take $C(a,b) = F(F_1^{-1}(a),F_2^{-1}(b))$, where $F_i^{-1}(a)$ denotes a function which is inverse to the function $F_i(x)$.

Computational Advantage of Copulas

In many applications of econometrics, it is important not only to have the right models for describing the corresponding phenomena, it is also extremely important to have efficient algorithms which use these models for predicting future values of the corresponding quantities. For example, if several agents have access to the models that can predict the increase in the price of a certain stock, but one of the agents has a faster algorithm for this prediction, then this agent can learn about this future increase before everyone else. This computational advantage will give this agent the opportunity to buy the about-to-increase stock for the current price, and thus, earn a profit when the price of this stock actually increases.

From this viewpoint, it should be noticed that a copula representation indeed speeds up computations. To explain this speed-up, let us start with the case of a single random variable. For a single variable X_i, we can use its observations $x_{i1},\dots,x_{iN}$ to estimate the corresponding probability distribution. For example, we can use a histogram distribution, i.e., approximate the probability by the corresponding frequency: $F_i(x_i) = \text{Prob}(X_i \le x_i) \approx \dfrac{1}{N} \cdot \#\{j : x_{ij} \le x_i\}$.

Comment

In practice, we rarely use the histogram distribution. Usually, we find a smooth distribution which is sufficient close to the histogram one (e.g., in the sense of the Kolmogorov-Smirnov criterion), so that this smooth distribution is statistically possible, and use the corresponding smooth distribution.

For two random variables X_1 and X_2, we can, in principle, also use the corresponding pairs of observations (x_{1j},x_{2j}), $1 \le j \le N$, and estimate the probability $F(x_1,x_2) = \text{Prob}(X_1 \le x_1 \,\&\, X_2 \le x_2)$ as the corresponding frequency $\dfrac{1}{N} \cdot \#\{j : x_{1j} \le x_1 \,\&\, x_{2j} \le x_2\}$. From the computational viewpoint, this would mean, however, that we need to process all N pairs (x_{1j},x_{2j}) (i.e., all $2N$ numbers x_{1j} and x_{2j}) to find each of the values $F(x_1,x_2)$. Usually, we have a large amount of economic data, so the need to process all the data all the time makes computations longer.

If instead of representing the unknown distribution by its joint distribution function $F(x_1,x_2)$, we use a copula representation, in which a distribution is represented by two marginals $F_1(x_1)$, $F_2(x_2)$, and a copula $C(a,b)$, then, to find each of the marginals $F_i(x_i)$, we only need to process N values x_{ij} $(j = 1,\dots,N)$ (and we only need to process all $2N$ real values to determine the copula $C(a,b)$). This decrease in the number of inputs speed up computations.

Case of Three of More Variables

As we have mentioned, to adequately describe economic phenomena, we need to use several random variables

$$X_1,\dots,X_n, \quad n \gg 2.$$

Each such random tuple can be described by its probability distribution

$$F(x_1,\dots,x_n) = \text{Prob}(X_1 \le x_1 \,\&\, \dots \,\&\, X_n \le x_n). \tag{3}$$

Similarly to the case of two variables, when all the random variables are independent, the joint distribution is equal to the product of all the marginal distributions:

$$F(x_1,\dots,x_n) = F_1(x_1) \cdot \ldots \cdot F_n(x_n).$$

Similarly to the two-variables case, the general distribution can be obtained by applying an appropriate combination function (copula) $C(a_1,\ldots,a_n)$ to the marginals:

$$F(x_1,\ldots,x_n) = C(F_1(x_1),\ldots,F_n(x_n)). \tag{4}$$

To prove that such a representation is possible for a given joint distribution $F(x_1,\ldots,x_n)$, we can take

$$C(a_1,\ldots,a_n) = F(F_1^{-1}(a_1),\ldots,F_n^{-1}(a_n)). \tag{5}$$

2 From General Copulas to Vine Copulas: Motivations and Descriptions

From the Computational Viewpoint, Additional Speed-Up Is Needed

Similarly to the two-variables case, the use of multi-dimensional copulas decreases the computation time. However, this decreased computation time still exponentially increases with the dimension n.

Indeed, a full knowledge about a function $f(x)$ of one variable defined on an interval $[0,1]$ would mean that we know infinitely many values of this function, corresponding to infinitely many real numbers $x \in [0,1]$. In practice, we can only store finitely many values. So, to describe a function in a computer, we select a small step h and only consider $\frac{1}{h}$ values

$$f(0), f(h), f(2h), \ldots, f(k\cdot h), \ldots, f(1), \;\; k = 1,2,\ldots,\frac{1}{h}. \tag{6}$$

Similarly, to describe a copula $C(a_1,\ldots,a_n)$, we need to store values

$$C(k_1\cdot h,\ldots,k_n\cdot h)$$

corresponding to all possible combinations of integers $k_1,\ldots,k_n$ corresponding to $k_i = 1,\ldots,\frac{1}{h}$. For each of n variables k_i, we have $\frac{1}{h}$ possible values. Thus, the total number of tuples $(k_1,\ldots,k_n)$ is equal to $\frac{1}{h^n}$.

Each of these values needs to be estimated and processed. Thus, the resulting computation time is proportional to $\frac{1}{h^n}$ and hence, exponentially grows with the number of variables n. For large n, this computation time becomes unrealistically large (see, e.g., [22]) – especially in view of the above-mentioned fact that in econometrics, we need computations to be as fast as possible. Thus, an additional speed-up is needed.

We already know that for two variables, a copula-based description – which only uses functions of two variables – is realistic and practically useful. From this viewpoint, it is desirable to only use functions of two variables in our description of

multi-variate distributions. Such a description is possible if we use *vine copulas*. Let us explain how the corresponding vine copula techniques naturally emerge from the analysis of our problem.

Main Idea: Using Conditional Probabilities

Our objective is to represent dependence. To arrive at the copula techniques, we started with the description of independence, and we used this description to come up with a general copula-based description of dependence. From the mathematical viewpoint, this copula-based description is sufficient to describe an arbitrary dependence. However, from the computational viewpoint, we need to go beyond the general copula-based formula. To move forward, let us go back to the independence case, and see if there are some other independence-related techniques that we can generalize to the general dependence case.

Our previous analysis was based on the fact that independence between random variables can be described in terms of the product of the corresponding probabilities: $F(x_1,x_2) = F_1(x_1) \cdot F_2(x_2)$. There is, however, an equivalent (and probably more intuitive) description of independence, a representation in term of conditional probabilities: $F_{1|2}(x_1\,|\,x_2) = F_1(x_1)$, where

$$F_{1|2}(x_1\,|\,x_2) \stackrel{\text{def}}{=} \text{Prob}(X_1 \le x_1\,|\,X_2 = x_2). \tag{7}$$

To relate this representation to the previous one, let us describe the conditional probability in terms of the copula. By definition of the conditional probability, we have

$$\begin{aligned} F_{1|2}(x_1\,|\,x_2) &= \text{Prob}(X_1 \le x_2\,|\,X_2 = x_2) = \\ \lim_{\varepsilon\to 0} \text{Prob}(X_1 \le x_2\,|\,x_2 - \varepsilon \le X_2 \le x_2 + \varepsilon) &= \\ \lim_{\varepsilon\to 0} \frac{\text{Prob}(X_1 \le x_2 \,\&\, x_2 - \varepsilon \le X_2 \le x_2 + \varepsilon)}{\text{Prob}(x_2 - \varepsilon \le X_2 \le x_2 + \varepsilon)}. \end{aligned} \tag{8}$$

The probability in the numerator $\mathscr{N}$ of the corresponding fraction can be described as

$$\begin{aligned} \mathscr{N} = \text{Prob}(X_1 \le x_1 \,\&\, X_2 \le x_2 + \varepsilon) - \text{Prob}(X_1 \le x_1 \,\&\, X_2 \le x_2 - \varepsilon) = \\ F(x_1, x_2 + \varepsilon) - F(x_1, x_2 - \varepsilon). \end{aligned} \tag{9}$$

In terms of the corresponding copula $C_{12}(a,b)$ and the marginals $F_1(x_1)$ and $F_2(x_2)$, we get

$$\mathscr{N} = C_{12}(F_1(x_1), F_2(x_2 + \varepsilon)) - C_{12}(F_1(x_1), F_2(x_2 - \varepsilon)). \tag{10}$$

Since ε is small, we get

$$\mathscr{N} \approx 2\varepsilon \cdot \frac{\partial C_{12}(F_1(x_1), F_2(x_2))}{\partial x_2} = 2\varepsilon \cdot C_{1|2}(F_1(x_1), F_2(x_2)) \cdot f_2(x_2), \tag{11}$$

where we denoted $C_{1|2}(a,b) \stackrel{\text{def}}{=} \dfrac{\partial C_{12}(a,b)}{\partial b}$, and $f_2(x_2) = \dfrac{dF_2(x_2)}{dx_2}$ is the probability density of the second marginal distribution.

Similarly, the denominator $\mathscr{D}$ has the form

$$\mathscr{D} = \text{Prob}(X_2 \le x_2 + \varepsilon) - \text{Prob}(X_2 \le x_2 - \varepsilon) = F_2(x_2 + \varepsilon) - F(x_2 - \varepsilon). \tag{12}$$

Since ε is small, we get

$$\mathscr{N} \approx 2\varepsilon \cdot f_2(x_2).$$

Thus, the ratio $F_{1|2}(x_1\,|\,x_2)$ is equal to:

$$F_{1|2}(x_1\,|\,x_2) = C_{1|2}(F_1(x_1), F_2(x_2)). \tag{13}$$

The corresponding conditional probability density $f_{1|2}(x_1\,|\,x_2)$ can be obtained by differentiating both sides of this equation with respect to x_1:

$$f_{1|2}(x_1\,|\,x_2) = c_{12}(F_1(x_1), F_2(x_2)) \cdot f_1(x_1), \tag{14}$$

where

$$c_{12}(a,b) = \frac{\partial C_{1|2}(a,b)}{\partial a} = \frac{\partial}{\partial a}\left(\frac{\partial C_{12}(a,b)}{\partial b}\right) = \frac{\partial^2 C_{12}(a,b)}{\partial a\,\partial b}.$$

There are two ways to use conditional probabilities to speed up our computations. Let us illustrate both of them on the example of trivariate distributions.

First Idea: D-vine Copulas

We know how to describe bivariate distributions in terms of copulas: namely, each pair of random variables X_1 and X_2 with a joint distribution $F(x_1,x_2)$ can be represented as $F(x_1,x_2) = C_{12}(F_1(x_1), F_2(x_2))$. We would like to use this idea to describe *three* random variables X_1, X_2, and X_3. A natural idea is to fix the value x_3, and to consider corresponding *conditional* distributions. For each x_3, we can have a similar representation of the corresponding conditional distribution

$$F_{12|3}(x_2,x_2\,|\,x_3) \stackrel{\text{def}}{=} \text{Prob}(X_1 \le x_1 \,\&\, X_2 \le x_2\,|\,X_2 = x_3) =$$

$$C_{12|3}(F_1(x_1\,|\,x_3), F_2(x_2\,|\,x_3), x_3). \tag{15}$$

In general, for different values x_3, we can have different copulas $C(a,b) = C_{12|3}(a,b,x_3)$. These copula describe the dependence between X_1 and X_2. In many practical situations, it makes sense to assume that the dependence between X_1 and X_2 does not depend on the value of X_3. In such situations, the copula $C_{12|3}(a,b)$ which describes this dependence does not depend on x_3: $C_{12|3}(a,b,x_3) = C_{12|3}(a,b)$. Then, the formula (14) takes the simplified form

$$F_{12|3}(x_1,x_2\,|\,x_3) = C_{12|3}(F_{1|3}(x_1\,|\,x_3), F_{2|3}(x_2\,|\,x_3)). \tag{16}$$

We already know how to describe conditional distributions $F_{1|3}(x_1\,|\,x_3)$ and $F_{2|3}(x_2\,|\,x_3)$ in terms of bivariate copulas and marginals: specifically, we can use the formula (13). Thus, we can describe the conditional probabilities $F_{12|3}(x_1,x_2\,|\,x_3)$ in terms of bivariate copulas and marginals.

Our goal is to compute the distribution function $F(x_1,x_2,x_3)$. To describe the corresponding probabilities $F(x_1,x_2,x_3)$ in terms of conditional probabilities $F_{12|3}(x_1,x_2\,|\,x_3)$, we can use the formula of total probability:

$$F(x_1,x_2,x_3)=\int_{-\infty}^{x_3} F_{12|3}(x_1,x_2\,|\,z)\cdot f_3(z)\,dz. \tag{17}$$

Combining formulas (13), (16), and (17), we get the following expression of the multivariuate distribution in terms of bivariate copulas and marginal distributions:

$$F(x_1,x_2,x_3)=\int_{-\infty}^{x_3} C_{12|3}(F_1(x_1\,|\,z),F_{2|3}(x_2\,|\,z))\,dz, \tag{18}$$

where

$$F_{1|3}(x_1\,|\,z)=C_{1|3}(F_1(x_1),F_3(z)),\ \ F_{2|3}(x_2\,|\,z)=C_{2|3}(F_2(x_2),F_3(z)), \tag{19}$$

$C_{1|3}(a,b)\stackrel{\text{def}}{=}\dfrac{\partial C_{13}(a,b)}{\partial b}$, and $C_{2|3}(a,b)\stackrel{\text{def}}{=}\dfrac{\partial C_{23}(a,b)}{\partial b}$. This description is a particular case of a *D-vine copula*.

Second Idea: C-vine Copulas

The idea behind C-vine copulas comes from considering not directly probabilities and conditional probabilities (as for D-vine copulas), but rather probability *densities* and conditional probability *densities*. A multivariate probability density can be described in terms of conditional probability densities, as

$$f(x_1,x_2,x_3)=f_{1|23}(x_1\,|\,x_2,x_3)\cdot f_{23}(x_2,x_3). \tag{20}$$

The probability density $f_{23}(x_2,x_3)$ can also be similarly represented as $f_{2|3}(x_2\,|\,x_3)\cdot f_3(x_3)$, so we conclude that

$$f(x_1,x_2,x_3)=f_{1|23}(x_1\,|\,x_2,x_3)\cdot f_{2|3}(x_2\,|\,x_3)\cdot f_3(x_3). \tag{21}$$

We know, from the formula (14), that

$$f_{2|3}(x_2\,|\,x_3)=c_{23}(F_2(x_2),F_3(x_3))\cdot f_2(x_2). \tag{22}$$

For dependence $f_{1|23}(x_1\,|\,x_2,x_3)$, we have a similar formula for each x_3:

$$f_{1|23}(x_1\,|\,x_2,x_3)=c_{12|3}(F_{1|3}(x_1\,|\,x_3)),F_{2|3}(x_2\,|\,x_3),x_3)\cdot f_{1|3}(x_1\,|\,x_3). \tag{23}$$

In general, the corresponding copula $c_{12|3}$ depends on x_3. However, in many practical situations, it makes sense to assume that this copula – describing the dependence – does not depend on x_3, i.e., that we have

$$f_{1|23}(x_1\,|\,x_2,x_3) = c_{12|3}(F_{1|3}(x_1\,|\,x_3)),F_{2|3}(x_2\,|\,x_3)) \cdot f_{1|3}(x_1\,|\,x_3). \tag{24}$$

We already know how to describe conditional distributions $F_{1|3}(x_1\,|\,x_3)$ and $F_{2|3}(x_2\,|\,x_3)$ and conditional probability density $f_{1|3}(x_1\,|\,x_3)$ in terms of bivariate copulas and marginals: specifically, we can use the formulas (13) and (14). Thus, we can describe the conditional probability density $f_{1|23}(x_1\,|\,x_2,x_3)$ in terms of bivariate copulas and marginals. By combining the formulas (21), (22), and (24), we get

$$f(x_1,x_2,x_3) = c_{12|3}(F_{1|3}(x_1\,|\,x_3)),F_{2|3}(x_2\,|\,x_3)) \cdot f_{1|3}(x_1\,|\,x_3) \cdot$$
$$c_{23}(F_2(x_2),F_3(x_3)) \cdot f_2(x_2) \cdot f_3(x_3), \tag{25}$$

where

$$F_{1|3}(x_1\,|\,x_3) = C_{1|3}(F_1(x_1),F_3(x_3));\;\; F_{2|3}(x_2\,|\,x_3) = C_{2|3}(F_2(x_2),F_3(x_3));$$
$$f_{1|3}(x_1\,|\,x_3) = c_{13}(F_1(x_1),F_3(x_3)) \cdot f_1(x_1). \tag{26}$$

This description is a particular case of a *C-vine copula*.

Comment

Similar expressions can be obtained for any number of variables. To get such an expression, we need to make some assumptions about copula independence. Depending on which assumptions we make, we get different expressions. For example, the above expression (25)–(26) corresponds to the case when we assume that the copula combining:

- the conditional dependence $F_{1|3}(x_1\,|\,x_3)$ of x_1 on x_3 and
- the conditional dependence $F_{2|3}(x_2\,|\,x_3)$ of x_2 on x_3

into a conditional joint dependence $F_{12|3}(x_1,x_2\,|\,x_3)$ of x_1 and x_2 on x_3 does not depend on x_3. Alternatively, we could assume that the copula combining:

- the conditional dependence $F_{2|1}(x_2\,|\,x_1)$ of x_2 on x_1 and
- conditional dependence $F_{3|1}(x_3\,|\,x_1)$ of x_3 on x_1

into a conditional joint dependence $F_{23|1}(x_2,x_3\,|\,x_1)$ of x_2 and x_3 on x_1 does not depend on x_1; this would lead to a different expression of the type (25)–(26).

How do we select a model? In some cases, from the econometric context, we know which dependencies are independent in each variables. In many practical situations, however, such an information is not available. In such situations, out of models corresponding to different dependencies, we need to select the model which is the best fit for the observations.

3 Comparing Vine Copulas with Other Techniques for Describing Multi-Variate Dependence

Vine Copulas vs. General Copulas

Vine copulas are a practically important class of copulas: they only use bivariate functions to describe a multi-variate dependence and are, thus, computationally easier (and more feasible) to implement.

It is important to remember, however, that vine copulas *do not* describe a general dependence. As we have mentioned earlier, vine copulas are based on certain *independence* assumptions: e.g., that the copula that transforms the conditional distributions $F_{1|3}(x_1\,|\,x_3)$ and $F_{2|3}(x_2\,|\,x_3)$ into a joint conditional distribution $F_{12|3}(x_1,x_2\,|\,x_3)$ does not depend on the value x_3.

It is worth mentioning that vine copulas' inability to represent a general function of three or more variables is not a drawback of any particular scheme, but rather a general property of smooth (differentiable) functions. Namely, as part of the work on D. Hilbert's 13th problem – one of the famous 23 problems presented in 1900 as a challenge to 20 century mathematics – a Russian mathematician A. G. Vitushkin proved that for any given integer N, it is not possible to represent (or even approximate) a general smooth function of three (or more) variables as a composition of functions of two or fewer variables; see, e.g., [5, 16, 27, 28, 29].

Vine Copulas vs. Bayesian Networks

Another approach actively used in applications to represent multivariate dependence is the approach of Bayesian networks, initiated by Judea Pearl; see, e.g., [18, 23, 24, 25]. Bayesian newtorks are based on the assumption that for some variables, the corresponding conditional distributions are independent. For example, for the case of three variables, a typical assumption is that the conditional distributions $F_{1|3}(x_1\,|\,x_3)$ and $F_{2|3}(x_2\,|\,x_3)$ are independent, i.e., that

$$F_{12|3}(x_1,x_2\,|\,x_3) = F_{1|3}(x_1\,|\,x_3)\cdot F_{2|3}(x_2\,|\,x_3). \tag{27}$$

One can easily see that the resulting formula is a particular case of the vine copula formula (16), corresponding to $C_{1|2}(a,b) = a\cdot b$. Thus, the Bayesian network approach can be viewed as a particular case of the general vine copula approach.

Vine Copulas vs. Fuzzy Techniques

Another practically successful approach for describing and analyzing multivariate dependence is an approach of fuzzy techniques; see, e.g., [12, 20, 30].

One of the main ideas behind fuzzy techniques is that

- while we can extract, from the experts, their degrees of confidence (= subjective probability) in different possible statements $S_1, S_2, \ldots, S_n$ about their domain of expertise,

- it is not realistically possible to extract, from the users, their degrees of confidence in different logical combinations of such statements, such as $S_i \& S_j$ or $S_i \& S_j \& S_k$ – since there are, in general, exponentially many (2^n) such combinations.

Since we cannot elicit all the values, we need to estimate the degree of confidence in a statement $S \& S'$ based on the known degrees of confidence $d(S)$ and $d(S')$ in component statements S and S'. The algorithm $f_\&(a,b)$ which transforms the known degrees $a = d(S)$ and $b = d(S')$ into an estimate $f_\&(d(S),d(S'))$ for the desired degree $d(S \& S')$ is known as an *"and"-operation* or a *t-norm*.

From the mathematical viewpoint, there are many possible t-norms. In practice, a t-norm is selected empirically, based on the cases when we do elicit the expert's degree of confidence $d(S \& S')$ in the composite statement $S \& S'$. Once these values are known, we select a function $f_\&(a,b)$ for which $f_\&(d(S),d(S')) \approx d(S \& S')$ for all such pairs of statements.

The resulting "and"-operation depends on the domain. Such an empirical determination was first implemented for the world's first practically successful expert system, a medical expert system MYCIN intended for diagnosing rare blood diseases; see, e.g., [6]. It is worth mentioning that the authors of the corresponding empirical study initially thought that the resulting "and"-operation is a general description of human reasoning. Alas, when they applied their idea to geophysics, it turned out that the medically best "and"-operation is not appropriate for geophysics at all. After the fact, it makes sense: e.g., in search for oil, it makes sense to start drilling a well once there is a reasonable expectation that this well will be productive – and it is OK that a large portion of these wells do not produce, as long as on average, we are successful. In contrast, in medicine, we do not want to perform a serious surgery on a patient unless we are absolutely sure about the diagnosis. In short, in medicine, experts use very conservative estimates, while in geophysics, they use more optimistic ones. As a result, different application domains use different "and"-operations – but the same "and"-operation is useful for all statements within a given application domain.

The main problem that we solve by using copulas can be described in similar terms. Namely, we have two statements $S =$ "$X_1 \le x_1$" and $S' =$ "$X_2 \le x_2$", whose probabilities are values of the marginal distributions $d(S) = F_1(x_1)$ and $d(S') = F_2(x_2)$. The logical combination $S \& S'$ is the statement

$$X_1 \le x_1 \,\&\, X_2 \le x_2$$

whose probability is equal to $F(x_1,x_2)$. Our objective is to transform the known degrees $d(S) = F_1(x_1)$ and $d(S') = F_2(x_2)$ into an estimate $f_\&(d(S),d(S')) = f_\&(F_1(x_1), F_2(x_2))$ for $F(x_1,x_2)$:

$$F(x_1,x_2) \approx f_\&(F_1(x_1),F_2(x_2)). \tag{28}$$

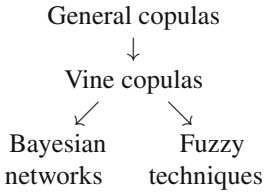
From this viewpoint, the copula is an "and"-operation.

The fuzzy approach can be viewed as a particular case of the vine copula approach, The main difference between fuzzy approach and the general vine copula approach is that:

- in the fuzzy case, the same "and"-operation is used to combine the probabilities corresponding to different variables, while
- in the general vine copula approach, we can use different copulas to combine the probabilities of different pairs of variables.

Summarizing Our Analysis

Vine copulas are a particular case of general copulas, and Bayesian network and fuzzy approaches can be viewed as particular cases of the vine copula approach:



Vine Copula Approach Combines Advantages of Bayesian and Fuzzy Approaches

Both Bayesian networks and fuzzy techniques have numerous successful applications. The very fact that both techniques have been successful means that for each of these techniques, there is an application areas where this particular technique works well. The fact that both techniques co-exist seems to indicate that for each of these techniques, there are application areas where the other technique works better.

In other words, each of these techniques has its own advantages and limitations. Numerous researchers have expressed the desire to come up with a new technique that would combine the advantages of both techniques – and have none of their limitations. From this viewpoint, the vine copula approach, an approach of which both Bayesian network and fuzzy techniques are particular cases, seems like the desired combination:

- in contrast to Bayesian techniques, vine copula can handle dependence between variables, not just independence;
- in contrast to fuzzy techniques, where the same "and"-operation (t-norm) is applied for combining all pieces of information, the vine copulas allow the use of different "and"-operations (copulas) to combine information about different variables.

4 How Vine Copulas Are Used in Econometrics

Main Challenge: Econometric Processes Are Dynamic

Vine copulas describe dependence between a few random *variables* $X_1, \ldots, X_n$. In econometrics, however, processes are highly dynamic, so what we have is random *processes* $X_1(t), \ldots, X_n(t)$, not random variables. How can we use vine copulas to describe the dependence between random processes?

Main Idea: Use Known Models to Describe the Dynamics of Each Variable

For each of the econometric dynamic variables $r_t \stackrel{\text{def}}{=} X_i(t)$, there are known ways to describe its dynamics. One of the most (and probably *the* most) adequate models for such a dynamics are described by an appropriate combination of the Auto-Regressive Moving-Average Model (ARMA) and the Glosten-Jagannathan-Runkle (GJR) form of a Generalized Auto-Regressive Conditional Heteroskedasticity (GARCH) model [4]; see, e.g., [8, 15]. The corresponding ARMA(p,q)-GJR(k,ℓ) model has the form

$$r_t = c + \sum_{i=1}^{p} \varphi_i \cdot r_{t-i} + \varepsilon_i \sum_{j=1}^{q} \psi_j \cdot \varepsilon_{t-j}, \tag{29}$$

$$\varepsilon_t = h_t \cdot \eta_t, \tag{30}$$

$$h_t^2 = \omega + \sum_{i=1}^{k} \alpha_i \cdot \varepsilon_{t-i}^2 + \sum_{i:\, \varepsilon_{t-i}<0} \gamma_i \cdot \varepsilon_{t-i}^2 + \sum_{j=1}^{\ell} \beta_j \cdot h_{t-j}^2, \tag{31}$$

where ε_t and h_t are auxiliary variables, c, φ_i, ψ_j, ω, α_i, and β_j are real-valued constants (which need to be determined based on the observations), and residuals η_t corresponding to different moments of time t are independent identically distributed random variables.

The distribution of the residuals is usually assumed to be distributed according to skewed student-t or skewed Generalized Error Distribution (GED). A skewed t-distribution means that we combine, with fixed weights, t-distributions $f_1(x)$ and $f_2(x)$ with different scalar parameters limited to, correspondingly, positive and negative values x_i: $f(x) = w_1 \cdot f_1(x)$ when $x \geq 0$ and $f(x) = w_2 \cdot f_2(x)$ when $x < 0$.

A GED distribution is a distribution with a probability density proportional to $\exp\left(-\dfrac{|x|^\nu}{\sigma^\nu}\right)$; it generalizes Gaussian distribution – which corresponds to $\nu = 2$. A skewed GED distribution is a combination of two GED distributions $f_1(x)$ and $f_2(x)$ corresponding to different values σ (but the same value ν): $f(x) = w_1 \cdot f_1(x)$ when $x \geq 0$ and $f(x) = w_2 \cdot f_2(x)$ when $x < 0$, where w_i are appropriate weights.

Resulting Solution: Copula Describes the Joint Distribution of Residuals

Copulas in general (and vine copulas in particular) are a good technique for describing the dependence between several random variables $X_1, \ldots, X_n$. In the dynamical case, instead of n variables $X_1, \ldots, X_n$, we have, in effect, a much larger number of dependent random variables $X_i(t)$ corresponding to different values i and different moments of time t. Not only are variables $X_i(t)$ and $X_j(t)$ corresponding to the same moment of time depending on each other, the values $X_i(t)$ and $X_i(t')$ corresponding to different moments of time also depend on each other – and thus, we also have dependence between $X_i(t)$ and $X_j(t')$.

We have already observed, in our motivation for the use of vine copulas, that the larger the number of dependent variables to consider, the more computationally complex the resulting problem, the more computation time it takes to process this data. We have econometric data corresponding to dozens of years, hundreds of months, thousands of days, so we have thousands of dependent quantities corresponding to different values of i and t. Thus, to be able to describe and process the dependence between different econometric quantities within a reasonable amount of computation time, we need to be able to reduce this dependence between thousands of variables to a dependence between a much smaller number of variables.

Good news is that such a reduction is possible: for such a reduction, we can use the above dynamical equations. Indeed:

- while the values $X_i(t)$ and $X_i(t')$ of the original quantity at different moments of time t and t' are, in general,
- the *residuals* η_t and $\eta_{t'}$ corresponding to different moments of time are *independent* (so all the dependence between $X_i(t)$ and $X_i(t')$ is described by the dynamical equations themselves).

Since residuals corresponding to different moments of time are independent of each other, it is sufficient to consider, for each moment of time t, the dependence between n residuals corresponding to this moment of time; see, e.g., [17]. Thus, for each t, we use a multi-variate copula to describe the dependence between the n residuals corresponding to the original n quantities $X_1, \ldots, X_n$.

Acknowledgments. The authors are greatly thankful to the Faculty of Economics of Chiang Mai University for the financial support.

This work was also supported in part by the National Science Foundation grants HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and DUE-0926721, by Grants 1 T36 GM078000-01 and 1R43TR000173-01 from the National Institutes of Health, and by a grant N62909-12-1-7039 from the Office of Naval Research.

References

1. Aas, K., Czado, C., Frigessi, A., Bakken, H.: Pair-copula constructions of multiple dependence. Insurance: Mathematics and Economics 44, 182–198 (2009)
2. Bedford, T., Cooke, R.M.: Monte Carlo simulation of vine dependent random variables for applications in uncertainty analysis. In: Proceedings of European Safety and Reliability Conference, ESREL 2001, Turin, Italy (2001)
3. Bedford, T., Cooke, R.M.: Vines-a new graphical model for dependent random variables. Annals of Statistics 30(4), 1031–1068 (2002)
4. Bollerslev, T.: Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307–327 (1986)
5. Browder, F.E. (ed.): Mathematical Developments Arising from Hilbert Problems. American Mathematical Society, Providence (1976)
6. Buchanan, B.G., Shortliffe, E.H.: Rule Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project. Addison-Wesley, Reading (1984)
7. Czado, C.: Pair-copula constructions of multivariate copulas. In: Jaworski, P. (ed.) Copula Theory and Its Applications. Lecture Notes in Statistics, vol. 198, pp. 93–109. Springer, Heidelberg (2010)
8. Glosten, L.R., Jagannathan, R., Runkle, D.E.: On the relation between the expected value and the volatility of the nominal excess return on stocks. The Journal of Finance 48(5), 1779–1801 (1993)
9. Joe, H., Hu, T.: Multivariate distributions from mixtures of max-infinitely divisible distributions. Journal of Multivariate Analysis 57(2), 240–265 (1996)
10. Joe, H.: Dependence comparisons of vine copulae with four or more variables. In: Kurowicka, D., Joe, H. (eds.) Dependence Modeling: Vine Copula Handbook. World Scientific, Singapore (2010)
11. Joe, H., Li, H., Nikoloulopoulos, A.K.: Tail dependence functions and vine copulas. Journal of Multivariate Analysis 101, 252–270 (2010)
12. Klir, G.J., Yuan, B.: Fuzzy Sets and Fuzzy Logic. Prentice Hall, Upper Saddle River (1995)
13. Kurowicka, D., Cooke, R.M.: Uncertainty Analysis with High Dimensional Dependence Modelling. Wiley, New York (2006)
14. Kurowicka, D., Joe, H. (eds.): Dependence Modeling: Vine Copula Handbook. World Scientific, Singapore (2010)
15. Ling, S.: Self-weighted and local quasi-maximum likelihood estimators for ARMA-GARCH/IGARCH models. Journal of Econometrics 140, 849–873 (2007)
16. Lorenz, G.G.: Approximation of Functions. American Mathematical Society, Providence (1966)
17. Manner, H., Reznikova, O.: A survey on time-varying copulas: Specification, simulations and application. Econometric Reviews 31(6), 654–687 (2012)
18. Neapolitan, R.E.: Learning Bayesian networks. Prentice Hall, Upper Saddle River (2004)
19. Nelsen, R.B.: An Introduction to Copulas. Springer, New York
20. Nguyen, H.T., Walker, E.A.: First Course In Fuzzy Logic. CRC Press, Boca Raton (2006)
21. Nikoloulopoulos, A.K.: Vine copulas with asymmetric tail dependence and applications to financial return data. Computational Statistics and Data Analysis 56, 3659–3673 (2012)
22. Papadimitriou, C.H.: Computational Complexity. Addison Wesley, San Diego (1994)
23. Pearl, J.: Probabilistic Reasoning in Intelligent Systems. Morgan Kaufmann, San Francisco (1988)

24. Pearl, J.: Causality: Models, Reasoning, and Inference. Cambridge University Press, Cambridge (2000)
25. Pourret, O., Naim, P., Marcot, B.: Bayesian Networks: A Practical Guide to Applications. Wiley, Chichester (2008)
26. Sklar, A.: Fonctions de répartition á *n* dimensions et leurs marges. Publ. Inst. Statist. Univ. Paris 8, 229–231 (1959)
27. Vitushkin, A.G.: On Hilbert's thirteenth problem. Soviet Math. Doklady (Dokl. Akad. Nauk SSSR) 96, 701–704 (1954)
28. Vitushkin, A.G.: Estimating Complexity of the Tabulation Problem. Fizmatgiz, Moscow (1959) (in Russian)
29. Vitushkin, A.G.: On Hilbert's thirteenth problem and related questions. Russian Mathematical Surveys 51(1), 11–25 (2004)
30. Zadeh, L.A.: Fuzzy sets. Information and Control 8, 338–353 (1965)

Part III
Applications

Extreme Value Copula Analysis of Dependences between Exchange Rates and Exports of Thailand

Chakorn Praprom and Songsak Sriboonchitta

Abstract. This study aims to investigate a correlation of the dependence structure between USD/THB exchange rate and exports of Thailand, using extreme value copula by combining the bivariate Generalized Pareto Distribution (GPD) extreme value theory and copula. Maximum likelihood method was adopted to fit a parameter estimation based on the GPD extreme value model, and a behavior of dependence was determined by the dependence function. The procedure is suggested for the measurement of the copula function to recover the joint tail distribution by comparing four extreme value copulas. The results of this analysis denote that the Tawn copula analysis is the most appropriate method to best fit extreme value copula because the AIC of this method is the lowest when compared with the other copulas. We applied Value at Risk (VaR) to calibrate the probability of the joint tail that may occur over the threshold. We found that Tawn copula stands the maximum risk of exceeding the threshold. This result could be beneficial for exporters and policy makers to predict the possibility of extreme economical fluctuation in the future.

1 Introduction

After it had to undergo the Asian financial Crisis (Tom Yam Kung Crisis) in 1997, Thailand changed its currency exchange system from the fixed exchange rate system to the floating exchange rate system since July 2, 1997. This change has immensely affected the economy of Thailand as well as of other countries in Asia, resulting in

Chakorn Praprom
Faculty of Economics, Chiang Mai University, Chiang Mai 50200 Thailand,
Department of Social sciences, Faculty of Humanities and Social Sciences,
Prince of Songkla University, Pattani Campus, Thailand
e-mail: chakornpraprom@gmail.com

Songsak Sriboonchitta
Faculty of Economics, Chiang Mai University, Chiang Mai 50200 Thailand
e-mail: songsak@econ.cmu.ac.th

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_12,

high volatility of the economy throughout the region, which we are able to observe in the dynamics of the various economic indicators such as exchange rate, interest rate, exports, imports, GDP, inflation, etc. Due to the floating exchange rate system, the fluctuating exchange rate directly affects the exports and imports. Theoretically, the exports expansion slows down or accelerates depending on whether the currency exchange rate is on an appreciation or a depreciation inclination, respectively. In addition, the fluctuation of the exchange rates has a chain reaction effect on Current Account as well as Balance of Payment, and will eventually affect the countrys GDP. Therefore, this study focused on the relationship between the currency exchange rate and the exports in order to see how it would be if the currency exchange rate were to change. The study laid emphasis on the data of changes in the exchange rates during the period from July 1997 to December 2012, and the data with extreme values or excess values above threshold that would affect the exports. The theory called Extreme Value Copula was applied for this economic research.

Extreme value theory is a branch of statistics dealing with the analysis of data with extreme values, whether high or low. The theory is widely used to analyze general data and excess data above threshold, as well as the behavior of a process at unusually large or small levels. Specifically, extreme value analysis usually requires an estimation of the probability of events that are more extreme than any that have been observed (Cole, [5]). Generally, extreme value theory is widely used in many fields of work, ranging from finance, insurance, hydrology, and environment.

In probability theory and statistics, a copula is a kind of distribution function used to describe the dependence between random variables. Copula is also widely accepted and has been used extensively to measure or evaluate the relationship of dependences between two or more variables. Copula connects marginals to obtain possible joint distributions. It is obvious that they provide the most general way to build multivariate statistical model marginals for the various applications of statistical sciences (Hung, [8]). The study of the relationship (correlation) between random variables using the copula method can provide more details than the traditional method (linear correlation) because the relationship between the two variables can be characterized as a kind of skewness and kurtosis, or as asymmetric and symmetric. From the information mentioned above, we used the extreme value copula to investigate the correlation of excess data above threshold in our research, while the variables used in this research were the USD/THB exchange rate and the exports of Thailand. The difference between extreme value copula and vine copulas is last one emphasize to study behavior of multivariate variable in general but extreme value copula explicit focus risk assessment unlike vine copulas. In this study, in addition to the extreme value theory, we used the theory of Value at Risk (VaR) to analyze the probability that two variable values were likely to exceed the threshold at a confidence level of 95%. We hope that after this study, we can use the trend of the changing exchange rates for adaptation for the use of exporters and policy makers.

The remainder of this paper is organized as follows: Section 2 presents literature review Section 3 presents the definition of bivariate extreme value and its Generalized Pareto Distribution (GPD) which is used to estimate models of distribution of maximum series, the concept of copulas, extreme value copulas, and joint tail

estimation, which have been employed in this study. Section 4 reveals the empirical findings. Lastly, section 5 presents the conclusion.

2 Review Literature

There are very few papers studying exports, imports, and economic growth by combining the extreme value theory and the copula method. Most of the studies applied either the extreme value or the copulas, separately, for investigation. Liu et al. [15] applied the copula-based GARCH model to analyze the volatility of the two dependences: tourist arrivals from China to Thailand and Singapore. They assumed that both the dependences were skewed-t distributions, and so, ARMA-GARCH was adopted to fit it. This paper shows 15 classes of copulas which are fit statics copula between two margins; the study used time-varying copulas to describe the dynamic Kendalls tau process. To ascertain whether the goodness of fit of static and dynamic copula are suitable or not, the AIC and BIC were used as the statistical criteria application in order to select the copula. The results of this analysis indicate that the AIC and BIC of the time-varying Gaussian copula had the highest explanatory power, more than the other copulas. Lu et al. [10] examined the extreme value copula analysis of the risk dependence for the exchange rate. This study observed the monthly maxima or minima (negative maxima) of daily USD/GBP, USD/EUR foreign return data, and applied the Block Maxima Model (BMM) Generalized Extreme Value Distribution (GEV) extreme value to analyze these excess dependences. In addition, copula was adopted to study the correlation between these exchange rates as well as to emphasize the joint tail dependence and the joint tail risk, based on the extreme value copulas. The results, as stated in this paper, show that three copula families, namely, Gumbel, Galambos, and Hsler-Reiss could be suitable for measuring the tail risk of their empirical market variables. Regarding the VaR estimation, there is a risks opportunity of about 0.004 of exceeding the threshold, in all of the copula families. Chuangchid et al. [4] utilized the extreme value copula to analyze palm oil prices and also used the GEV extreme value copula to study the dependence structure between the returns on palm oil future prices in three palm oil future markets, namely, Singapore Exchange Derivatives Trading Limited (SCX-DT), Dalian Commodity Exchange (DCE), and Malaysian future markets (KLSE). The Gumbel and Hsler-Reiss copulas were adopted to examine the extreme dependences. The results showed that the Gumbel and Hsler-Reiss copula parameters of KLSE and SGX-DT have dependence in the extreme, at 3.034 and 2.287, respectively. However, the returns on palm oil future prices between KLSE and DCE, and SGX-DT and DCE did not show any dependence. Velayoudoum et al. [18] applied the extreme value and Value at Risk (VaR) to explain the link between the oil price in the markets and the economic indicators. This paper emphasizes the need to measure risk at a given probability level, which is very important in risk management. In addition, extreme value was naturalized to compare with conventional models such as GARCH, Filtered Historical, and Historical Simulation. It was found that the extreme value theory and the Filtered Historical Simulation procedures symbolize a

major improvement over the non-parametric and parametric methods. Moreover, it holds true that the GARCH(1,1)-t model gives good results which are comparable to the combined procedures.

3 Data and Model Specification

3.1 Model

The data was restrictive since they were not daily data. However, the export of Thailand is secondary data which were collected monthly. Before 1997, the exchange rate of Thailand was fixed under fixed exchange rate system. For this reason, the data used in this study have been used since July 1997, after the Asian financial crisis. Due to such a restriction, we had only 186 observations; with such few data, it could not be divided into the Block Maxima (BMM) GEV extreme value. Therefore, the General Pareto Distribution (GPD) was selected to analyze the dependence structure. The GPD extreme value method is the best choice to study correlation between two margins. This method is more flexible and advantageous than the GEV (Block Maxima) method because the GPD has not demonstrated significantly whether the dependence is independent and identically distributed (i.i.d.) or not. This is unlike in the case of GEV that $M_n = (x_1, x_2, ..., x_n)$ is sample maxima which must be an i.i.d. random variable in $\mathbb{R}$. If $M_n = (x_1, x_2, ..., x_n)$ are not i.i.d., we cannot use this variable to estimate the dependence structure by using the GEV extreme value, according to Fisher and Tippett [28].

3.2 Bivariate Extreme Value and Generalized Pareto Distribution

In this paper, we employed the Extreme Value Theory (EVT) to define the relationship between the excess data of the USD/THB exchange rates and the exports of Thailand. The EVT is a concept of modeling and measuring extreme events which occur with a very small probability (see Brodin and Kluppelberg, [2]). There are two principal kinds of models for extreme values, the Generalized Extreme Value Distribution (GEV) and the Generalized Pareto Distribution (GPD). The GEV distribution, also known as the Block Maxima Model (BMM), was provided by Fisher and Tippett [28]. This method is the oldest method for analyzing extreme data which consist of the largest or the smallest values during a certain period. GPD was developed by Pickands [12], and it focuses on the behavior of large data exceeding the higher threshold in the sample. For the GPD method, there are two ways for determining excess data surpass threshold. Given u_x and u_y are the thresholds for each of the margins,X,Y are the distribution of the excess values of x and y,respectively. If $(X,Y) \sim F(x,y)$ then

$$P((X - u_x, Y - u_y) \leq (x,y) \mid (X,Y) \geq (u_x, u_y)) \tag{1}$$

$$P((X-u_x, Y-u_y) \leq (x,y) \mid (X,Y) \nleq (u_x,u_y)) \tag{2}$$

The asymptotic distributions are called Bivariate Generalized Pareto Distribution (BGPD). The (1) definition is called BGPD Type I and the (2) definition is called BGPD Type II. In this study, we employed the BGPD Type I for analyzing the joint exceedance to the bivariate extreme value of our data. In the univariate case, it is as follows:

$$P(X-u<x \mid X>u) \to 1-(1+\gamma\frac{x}{\sigma})^{-\frac{1}{\gamma}} \tag{3}$$

According to Cole [5], we can estimate the tail of X by

$$G(x) = 1-\eta_u\{1+\frac{\gamma(x-u)}{\sigma}\}^{-\frac{1}{\gamma}}, x>u \tag{4}$$

$$\eta_u = P(X>u)$$

For the bivariate distribution, an arbitrary joint distribution $F(X,Y)$ on the region of the form $x>u_x, y>u_y$, for large enough u_x and u_y has to be evaluated. Suppose $(x_1,y_1),...,(x_n,y_n)$ are independent realizations of a random variable (X,Y) with joint distribution function F. For suitable thresholds u_x and u_y, each of the marginal distributions of F has an approximation of the form (4)(Cole, [5]). By approximating the tail of $F(x,y)$ for $x>u_x, y>u_y$, we can estimate the tail of X for $x>u_x$ with $G(x:\eta_x,\sigma_x,\gamma_x)$ and the tail of Y for $y>u_y$ with $G(y:\eta_y,\sigma_y,\gamma_y)$. The result for the bivariate extreme value distribution suggests that

$$G(\tilde{x},\tilde{y}) = exp\Big(-(\frac{1}{x}+\frac{1}{y})\mathbb{A}(\frac{\tilde{x}}{\tilde{x}+\tilde{y}})\Big)$$

where $\tilde{x}$ and $\tilde{y}$ are distributions with Frechet margins. We use the approximates of the tails of X,Y and then transform them to unit Frechet.

$$\tilde{x} = -\Big(ln\{1-\eta_x[1+\frac{\gamma_x(x-u_x)}{\sigma_x}]^{-\frac{1}{\gamma_x}}\}\Big)^{-1}$$

$$\tilde{y} = -\Big(ln\{1-\eta_y[1+\frac{\gamma_y(y-u_y)}{\sigma_y}]^{-\frac{1}{\gamma_y}}\}\Big)^{-1}$$

If $x>u_x$ and $y>u_y$, then

$$F(x,y) \approx G(x,y) = e^{-V(\tilde{x},\tilde{y})}$$

where γ,σ are the shape and the scale parameter, respectively. V is the homogeneity property. The results of the research by Balkema and de Haan [1] and Pickands [12] state that the distribution of excesses may be estimated using the GPD method by selecting γ and σ and by setting a high threshold u. The GPD can be estimated using many different methods (Ramazan et al., [13]). There is evidence that the maximum likelihood normality conditions were met and that the maximum likelihood estimates were asymptotical, normal distributions (Hosking and Wallis, [7]). For our paper, we used the maximum likelihood estimation because this method

can approximate the standard error for the estimators of γ and σ which can be obtained using the maximum likelihood estimation (Ramazan et al, [13]). The bivariate peaks over the threshold models were fitted by maximizing the censored likelihood (Cole, [5]).

3.3 Extreme Value Copulas and Joint Tail Estimation

After our study about BGPD, we proceeded to learn about the concept of copulas and extreme value copulas, which indicates correlations between the margins which is given in this section. Furthermore, we pointed out the class of extreme value copula that can identify the excess dependence structure of the growth data.

3.4 Copula Function and Extreme Value Copulas

In this analysis, we applied copulas to calibrate the correlation between the USD/THB exchange rates and the exports of Thailand because copulas are flexible and provide more details as compared to other tools. Copulas are more suitable for use with non-linear marginal distributions as against the Pearson correlation which is appropriate for use with linear distributions. In addition, the approximation between the USD/THB exchange rates and the exports of Thailand was simultaneously estimated using the extreme value as well as the copulas. This method gives parameter values closer to facts, better than if we were to estimate the extreme value and the copula separately. The copula theory was first proposed by Sklar [25]. If x,y are real value random variables, then the commonly known marginal distribution functions are $F(\cdot)$ and $G(\cdot)$ of X and Y, respectively, which is not sufficient for our study. So, we take the joint distribution function $B(\cdot)$ of (x,y). If B is an arbitrary bivariate distribution function with marginal distribution functions F and G, then B is of the form

$$B(x,y) = C(F(x),G(y))$$

where C is a copula, $x,y \in \mathbb{R}$. Moreover, the copula is a parameter of the correlation between the two marginal distribution functions that forms the joint distribution. Therefore, the copula equation can be rewritten as

$$C(u,v) = B(F^{-1}(x),G^{-1}(y))$$

where F,G are the marginal distributions of x,y, respectively, and $F^{-1}(\cdot)$ is the quantile of the function $F(\cdot)$. The theory of BGPD extreme value can be reproduced in terms of extreme value copulas, which is a branch of the class of copulas. The bivariate extreme distribution B can be connected by the extreme value copula as

$$B(x,y) = C_0\Big(H_x(x : u_x,\gamma_x,\sigma_x),H_y(y : u_y,\gamma_y,\sigma_y)\Big)$$

where u,γ,σ are parameters of BGPD and H, the GPD margin. The unique copula shows that

$$C_0(u^t, v^t) = c_0^t(u,v), t > 0$$

Specifically, the bivariate copula is an extreme value copula if and only if it is converted into the form

$$C_0(u,v) = P(B_x(x) \le u, B_y(y) \le v) = exp\{ln(uv)A\Big(\frac{ln\, v}{ln(uv)}\Big)\}$$

where $A(t) = \int_0^1 max[(1-t)x, t(1-x)]dB(x)$.

The function $A(t)$ is known as the dependence function B on $[0,1]$. $A(t)$ must satisfy the following properties: $max(t, 1-t) \le A(t) \le 1$ for $0 \le t \le 1$. If $A(t) = 1$ then (x,y) are strongly dependent. Extreme value copulas allow the modeling of the dependence between the components of a random couple that represents two of the largest values observed over the same time period (Cebrian et al., [3]). In this paper, we chose four families of copula for practice, which are as follows:

Gumbel Copula

$$C(u,v) = exp\Big(-[(-lnu)^\theta + (-lnv)^\theta]^{\frac{1}{\theta}}\Big)$$

The dependence function is

$$A(t) = (t^\theta + (1-t)^\theta)^{\frac{1}{\theta}}$$

where θ is the Gumbel copula parameter, and $\theta \in [1, +\infty)$. Gumbel copula is the only copula that belongs to both the extreme value family and the Archimedean family. Complete dependence is obtained in the limit as θ approaches zero. Independence is obtained when $\theta = 1$ (Stephenson, [16]).

Galambos Copula

$$C(u,v) = uv\, exp[(-lnu)^{-\theta} + (-lnv)^{-\theta}]^{-\frac{1}{\theta}}$$

The dependence function is

$$A(t) = 1 - (t^{-\theta} + (1-t)^{-\theta})^{-\frac{1}{\theta}}$$

where θ is the Galambos copula parameter, $\theta \in [0, +\infty)$. If $\theta = 0$, these dependences will be independent; also, complete dependence is obtained as θ is led to infinity.

Hsler-Reiss Copula

$$C(u,v) = exp\{-\tilde{u}\Phi(\frac{1}{\theta} + \frac{1}{2} + \theta\, ln(\frac{\tilde{u}}{\tilde{v}})) - \tilde{v}\Phi(\frac{1}{\theta} + \frac{1}{2}\theta\, ln(\frac{\tilde{u}}{\tilde{v}}))\}$$

The dependence function is

$$A(t) = t\Phi(\theta^{-1} + \frac{1}{2}\theta\ ln(\frac{t}{1-t})) + (1-t)\Phi(\theta^{-1} - \frac{1}{2}\theta\ ln(\frac{t}{1-t}))$$

where $\theta \in [0, +\infty)$, $\tilde{u} = -ln\ u, \tilde{v} = -ln\ v$ and Φ is the standardized normal distribution.

Tawn Copula

$$C(u,v) = exp\{ln\ u^{1-\delta} + ln\ v^{1-\rho} - [(-\delta ln\ u)^{\theta} + (-\rho ln\ v)^{\theta}]^{\frac{1}{\theta}}\}$$

The dependence function is

$$A(t) = [\delta^{\theta}(1-t)^{\theta} + \rho^{\theta}t^{\theta}]^{\frac{1}{\theta}} + (\delta - \rho)t + 1 - \delta$$

where $\theta \in [1, +\infty)$, $\delta \in [0,1]$ and $\rho \in [0,1]$. This is an asymmetric extreme value copula that becomes exchangeable when $\delta = \rho$. The Gumbel copula corresponds to $\delta = \rho = 1$.

3.5 Joint Tail Estimation

In this part, we endeavor to assess risk in its various forms: financial risk, credit risk, etc., and how to measure, assess, and manage market risk. The Value at Risk (VaR) is a popular method to measure and evaluate a risk. From the equation

$$B(x,y) = C_0\Big(H_x(x : u_x, \gamma_x, \sigma_x), H_y(y : u_y, \gamma_y, \sigma_y)\Big)$$

we obtained the joint tail estimation of the BGPD, corresponding to relative Value at Risk under other confident levels p.

$$B_p = B(VaR_x(p), VaR_y(p)) = C_0\Big(H_x(VaR_x(p)), H_y(VaR_y(p))\Big) = C_0(p,p)$$

The above equation is useful for applying copula functions to determine the joint probability of the two returns that do not exceed specific VaR. It is assigned as a quantile of the distribution of return of the portfolio in question. Then, VaR can be computed as

$$VaR(p) = u + [\sigma\frac{(-ln\ p)^{-\gamma}}{\gamma - 1}]$$

where p is the probability value of VaR, and $0 < p < 1$, which means that higher values of VaR correspond to higher levels of risk. From the identity

$$p(X > x, Y > y) = 1 - H_x(x) - H_y(y) + B(x,y)$$

specified as a joint survival function, we can get the joint tail exceeding approximate for the two returns.

3.6 Data

The data on monthly USD/THB exchange rates and the exports of Thailand, beginning from July 1997 to December 2012 were used in this study, and the observations were 186. The data were obtained from the Bank of Thailand.

4 Empirical Results

To decrease the problem of data being non-stationary, each monthly data was converted into log-difference, and then the following function was obtained as

$$Y_t = 100 * (log(g_t) - log(g_{t-1}))$$

where g_t are USD/THB exchange rates or exports of Thailand at period t.

4.1 Parameter Estimation of Bivariate Generalized Pareto Distribution (BGPD) Model

In the GPD model, we lay emphasis on the statistical behavior of exceedance over the threshold. The first step is the determination of the threshold. In our paper, we define the threshold of the two dependences as 90%. Therefore, the threshold points of the USD/THB exchange rate returns and the export returns of Thailand are 0.1293 and 0.0235, respectively. Table 1 shows the estimator results of the Gumbel, Galambos, and Hsler-Reiss classes which have two parameters of either dependence, namely, σ and γ of the BGPD model, based on the maximum likelihood method, except the Tawn class which has four parameters, namely,σ, γ, t_1, t_2. In all of the classes, the element in the bracket corresponds to the standard deviation. The standard deviation estimates of the exchange rate returns were lower than the standard deviation estimates of the export returns of Thailand.

4.2 Results of Parameter Estimation of Copulas and Related Dependence Function

Table 2 shows the parameters of the various copulas (θ) including the AIC and $A(\frac{1}{2})$ which presented the information of tail dependence between margins. All four parameters of all the copulas can represent the dependence structure of the empirical exceedence over the threshold. To ascertain whether the goodness of fit of the copula is suitable or not, the Akaikes Information Criterion (AIC), which is the statistical criteria application for selecting the copula, is used. It has been found that the AIC of the Tawn copula shows the best explanatory ability compared to the other copulas because the AIC of the Tawn copula is 64.7684 which is the lowest value among all the copulas. Therefore, it was concluded that the two variables are correlated at 1.4552 by the Tawn copula. To explain the dependence function, as given in Table 2,

Table 1 Parameter Estimation Results Using Maximum Likelihood Method Based on BGPD Model

Variable	Kind of Copula	$\sigma_{USD/THB}$	$\gamma_{USB/THB}$	$\sigma_{Exports}$	$\gamma_{Exports}$	δ	ρ
Maxima of USD/THB exchange rates and exports of Thailand	Gumbel	0.0339	-0.1155	0.0155	0.8244		
		(0.0157)	(0.4165)	(0.0076)	(0.4839)		
	Galambos	0.0347	-0.1115	0.0153	0.7507		
		(0.0159)	(0.4238)	(0.0075)	(0.4546)		
	Hsler-Reiss	0.0349	-0.1066	0.0153	0.7339		
		(0.0160)	(0.4258)	(0.0074)	(0.4482)		
	Tawn	0.03354	-0.1660	0.0166	0.8200	0.2599	0.9997
		(0.0156)	(0.4043)	(0.0078)	(0.4819)		

Source: Computation.
Note: the element in the brackets correspond to the standard errors.

the estimated value of $A(\frac{1}{2})$ for the exchange rates and the exports of Gumbel and Galambos do not show any sharp contrast, whereas those of Tawn and Hsler-Reiss are rather skewed to the right. We can see that the parameters of the dependence function of all the types of copula are convex in function. All of them equal 0.9, which shows the asymptotic dependence of the data. Figure 1 presents the plot of estimation of the dependence function in each copula. If the upper $A(\frac{1}{2}) = 1$, they are both independent, and if the lower bound $A(\frac{1}{2}) = 0.5$, it indicates perfect dependence. This figure shows the sharp of the four dependence functions of all the four copulas; it shows that the Tawn copula is located in the bottom curve compared to the other copulas, and that its line is rather skewed to the right. In addition, $A(\frac{1}{2})$ of the Tawn copula is 0.9175, which is the lowest value among all the dependence functions. Thus, the dependence function value can categorically confirm that the Tawn copula is the appropriate-fitting copula.

4.3 Results of Joint Tail Estimation

VaR is the probability of joint tail risk that can occur over the threshold at different confidence levels. Table 3 shows the results of VaR under diverse copulas. For example, 0.0226 involved the Tawn copula representing the joint tail probability of the USD/THB exchange rates over the threshold 0.0792 (VaR estimate of the exchange rates at the 0.95 confidence level). Likewise, 0.0226 is the probability of the exports of Thailand exceeding over the threshold 0.2156 (VaR estimate of the exchange rates at the 0.95 confidence level). We found that the Tawn copula has the maximum risk of exceeding over the threshold because there is a probability of VaR of 0.0226. Besides, the VaR of Hsler-Reiss is 0.0107. This means that the probability of joint tail risk of Hsler-Reiss that can occur over the threshold of the USD/THB exchange rates and the exports of Thailand are 0.8947 and 0.1673, respectively. Gumbel and Galambos provide results similar to each others for VaR.

Table 2 Estimation of Copula Parameters and Dependence Function

Variable	Kind of Copula	θ	AIC	$A(\frac{1}{2})$
Maxima of USD/THB exchange rates and exports of Thailand	Gumbel	1.0950 (0.0709)	65.8518	0.9414
	Galambos	0.3319 (0.1144)	65.3474	0.9225
	Hsler-Reiss	0.6586 (0.1552)	65.2267	0.9337
	Tawn	1.4552 (0.0654)	64.7684	0.9175

Source: Computation.
Note: the element in the brackets correspond to the standard errors.

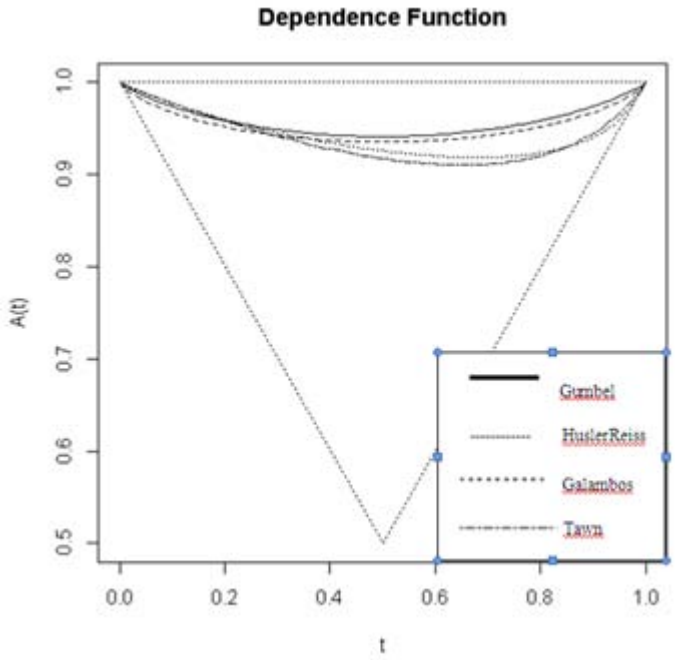


Fig. 1 A(t) estimation for the USD/THB exchange rates and exports of Thailand

Table 3 Joint Tail Exceeding Probability at 0.95 Confidence Level

VaR estimation of Maxima of USD/THB exchange rates and exports	0.95 Confidence level	Joint tail exceeding probability
Gumbel	(0.0859,0.2031)	0.0101
Hsler-Reiss	(0.8947,0.1673)	0.0107
Galambos	(0.0885,0.1733)	0.0104
Tawn	(0.0792,0.2156)	0.0226

Source: Computation.

5 Conclusion

This paper investigates the correlation between two dependence structures, the USD/THB exchange rates and the exports of Thailand. It is of interest that the extreme events in the future can be predicted by using the dependances over the threshold. We used a combination of the extreme value theory and the copulas to explain the relationship between the two dependence structures. This study confirms that the maximum likelihood method can be appropriately employed to estimate the GPD extreme value in each of the copula approaches. The main results confirm that the Tawn copula is the most suitable, best-fitting copula because we obtained the best parameter for this copula as against the other classes. In addition, we measured the dependence function in order to recover the tail dependence properties in comparison with all of the copula classes. The Tawn dependence function is in the bottom curve compared to the other copulas; this line is rather skewed to the right, and has the lowest value among all the dependence functions. Thus, the dependence function value can surely confirm that the Tawn copula is the appropriate-fitting copula. In addition, we applied Value at Risk (VaR) to calibrate the probability of the joint tail that may occur over the threshold. We found that the Tawn copula has the maximum risk of exceeding the threshold because there exists a probability of VaR equal to 0.0226. The results of this research are beneficial to policy makers and exporters for use in economic risk management. However, the clear relationship between the two variables would possibly be more noticeable if the data after 2000 were applied. Because Thailand had to request financial help from International Monetary Fund (IMF) in 1997 during the time of the economic crisis, the Thai policies were concise directed by the IMF which might not have been suitable for Thailand.

Acknowledgments. The authors wish to express their gratitude, particularly, to Prof. Nader Tajvidi, Mathematical Statistics Lund University, Sweden, and Miss Jaruchat Busaba, Faculty of Science, Mahasarakham University, Thailand, for their helpful suggestions and comments. We acknowledge the financial support received from the Prince of Songkla University Scholarship for Charkorn Praproms PhD study. We are grateful to Dr. Chanagun Chitmanat for reviewing the manuscript.

References

1. Balkema, A.A., de Haan, L.: Residual lifetime at great age. Annal of Probability 2, 792–804 (1974)
2. Brodin, E., Kluppelberg, C.: Extreme Value Theory in Finance. In: Encyclopedia of Quantitative Risk Analysis and Assessment (2008), http://dx.doi.org/10.1002/978047
3. Cebrian, A., Denuit, M., Philippe, L.: Analysis of bivariate tail dependence using extreme value copulas: An application to the SOA medical large claims database. Begian Actuarial Bulletin 3(1), 33–41 (2003)
4. Chuangchid, K., Wiboonpongse, A., Sriboonchitta, S., Chiaboonsri, C.: Application of Extreme value copulas to palm oil prices analysis. Business Management Dynamics 2(1), 25–31 (2012)
5. Cole, S.: An Introduction to Statistical Modeling of Extreme values. Springer Series in Statistics (2001)
6. Fisher, R.A., Tippett, L.H.C.: Limiting forms of the frequency distribution of largest or smallest member of a sample. Proceeding of the Cambridge Philosophical Society 24, 180–190 (1928)
7. Hosking, J.R.M., Wallis, J.R.: Parameter and quantile estimation for generalized Pareto distribution. Technometrics 29, 339–349 (1987)
8. Hung, T.N.: A Tutorial on copulas for correlation analysis in financial economics. Faculty of Economics, Chiang Mai University. Lecture note (2011)
9. Liu, J., Sriboonchitta, S.: Analysis of Volatility and Dependence between the Tourist Arrivals from China to Thailand and Singapore: A Copula-Based GARCH Approach. In: Huynh, V.N., Kreinovich, V., Sriboonchitta, S., Suriya, K. (eds.) Uncertainty Analysis in Econometrics with Applications. AISC, vol. 200, pp. 285–296. Springer, Heidelberg (2013)
10. Lu, J., Tian, W.J., Zhang, P.: The Extreme Value Copula Analysis of Risk Dependence for The Foreign Exchange Data. Wireless Communications, Networking and Mobile Computing, 1–6 (2008)
11. Nelsen, B.R.: An Introduction to copulas. Springer, New York (1999)
12. Pickands, J.: Statistical inference using extreme order statistics. Annals of Statistics 3, 119–131 (1975)
13. Ramazan, G., Faruk, S., Abduraahman, U.: EVIM: A Software Package for Extreme Value Analysis in MATLAB. Studies in Nonlinear Dynamics and Econometrics 5(3), 213–239 (2001)
14. Segers, J., Gudendoft, G.: Extreme-value Copulas (2009), http://arxiv.org/pdf/0911.1015.pdf
15. Sklar, A.: Fonctions de repartition a n dimentions et leurs marges. Publ. Inst. Statist. Univ. Paris 8, 229–231 (1959)
16. Stephenson, A.: Function for extreme value distributions. Packageevd. Version 2.2-4 (2011)
17. Thomas, M., Teiss, R.D.: Statistical Analysis of Extreme Values: with Applications to Insurance, Finance, Hydrology and Other Fields. Birkhauser (2007)
18. Velayoudoum, M., Bechir, R., Abdelwahed, T.: Extreme Value Theory and Value at Risk: Application to oil market. Energy Economics 31, 519–530 (2009)

Analysis of Volatility of and Dependence between Exchange Rate and Inflation Rate in Lao People's Democratic Republic Using Copula-Based GARCH Approach

Tongvang Xiongtoua and Songsak Sriboonchitta

Abstract. This paper aims to conduct a study of the volatility and dependence between the exchange rate and inflation rate in Laos. The results of the study show that the ARMA (1, 1) - GARCH (1, 1) models were appropriate for two random variables. The KS and Box-Ljung tests for skewed-t distribution and autocorrelation performed in the study found that the two margins were skewed-t distribution and had no autocorrelation. The modeling of the best-fit copula from the testing process found that the time-varying t copula was the best of all static copulas and time-varying copulas in terms of the AIC and the BIC, which means that it has the highest explanatory power of all the dependence structures. In addition, we can see that the indicator of the correlation (dependence parameter:) between the growth rates of the exchange rate and the inflation rate describes a high correlation in the long term, and also evinces that the dependence between the growth rates of the exchange rate and the inflation rate was positive, meaning that when the US Dollar appreciates, the inflation rate increases as well. Thus, this model as the time-varying t copula can help policy makers become more aware of what is likely to happen in the future.

1 Introduction

The inflation rate is a key problem for macroeconomic systems. A low rate of inflation is generally considered to be a good target because a high inflation rate often discourages investment and leads to lower long-term growth. As high and volatile inflation creates uncertainty and confusion about future prices and costs, investments tend to get reduced, leading to lower rates of growth of the economy, and this translates into frequent prices reductions, which means losses incur red. Also, high rates of inflation may call for frequent wage negotiations with trade unions,

Tongvang Xiongtoua · Songsak Sriboonchitta
Faculty of Economics, Chiang Mai University, Chiang Mai 50200 Thailand
e-mail: ttongvang@yahoo.com, songsakecon@gmail.com

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_13,

which constantly struggling to maintain the employees real wages; this can be costly for a manufacturing firm. High inflationary growth is often unsustainable. Reducing inflation often requires painful readjustments such as higher interest rates and deflationary fiscal policy, which again lead to lower growth rates of the economy. Therefore, countries going through high inflation could be susceptible to a period of recession in the near future. Exchange rate volatility can have an impact on the inflation rate. The effect of exchange rate fluctuations on domestic inflation has been an issue of concern, and has been discussed in the contemporaneous economics literature of Farah, Asma, and Khalid [14]; high exchange rate volatility contributes to higher exchange rate pass-through to inflation. For the Lao Peoples Democratic Republic (Lao PDR), during 1995 - 2012, there was high economic growth, of about 5.9 - 8.3%, but it was observed that Lao PDR had gone a high price phase during the period 1995 - 2005. The headline inflation stayed at more than 7% during this period, and the highest inflation rate in 1999 was 134%. During 2006 - 2012, the inflation rate decreased by about 0.03 - 7%, and the lowest rate was observed in 2009 as 0.03% (from the economic report of the Bank of Lao PDR). The exchange rate showed a tendency to increase during 1995 - 2002; the volatility in this period was so high that the average growth rate of the exchange rate dropped to 28.77% in, 62% in 1998, and 46.8% in 1999. It showed a decreasing trend during 2003 - 2012, with such low volatility in this period that the highest growth rate of the exchange rate was -5.69% in 2008 (from the annual economic re-port of Bank of Lao PDR). This study aims to investigate the volatility of and dependence between exchange rate and inflation rate in Lao Peoples Democratic Republic.

2 Literature Review

In some literature, while discussing the relationship between exchange rate volatility and inflation, exchange rate volatility was shown to have no connection to macroeconomic variables; an example for such is the literature of Flooda and Rose [21], who studied the fixing of the exchange rates, in their work. PARSLEY and WEI [19] studied the explanation that the border effects were due to the roles of exchange rate variability, shipping costs, and geography. ROGOFF [22] studied the perspectives on exchange rate volatility, as well. DUARTE and STOCKMAN [12] also studied the comments on exchange rate pass-through, exchange rate volatility, and exchange rate disconnect. However, the literature of the other authors who studied the connection between the exchange rate and macroeconomic variable volatilities shows that it was a closed correlation; this can be seen in the work of Luis Carranza [17], who studied the exchange rate and inflation dynamics in dollarized economies. Augustine and Srinivas [5] studied the variations in the exchange rates and inflation in 82 countries by conducting an empirical investigation. PARSLEY and WEI [11] also studied the border effects by analyzing the roles of exchange rate variability, shipping costs, and geography. Stephen Morris [27] studied the inflation dynamics and the parallel market for foreign exchange. Carlos, Jorge, and Scott [8] researched on how much the inflation targeters should care about the exchange rate. In this

study, we use the copula-based GARCH (generalized autoregressive conditional heteroskedasticity) approach to investigate the volatility and dependence between the exchange rates and the inflation rates. Of late, the copula-based GARCH model has been very popular in the financial field as it can be used to analyze the volatilities and dependence structure. Chih-Chiang Wu [9] studied the economic value of the co-movement between oil prices and exchange rates by using copula-based GARCH models, and Songsak Sriboonchitta [24] researched the modeling volatility and dependency of agricultural price and production indices of Thailand using a static versus time-varying copulas approach. Jianxu Liu and Songsak Sriboonchitta [16] also analyzed the volatility and dependence between tourist arrivals from China to Thailand and Singapore by using the copula-based GARCH approach.

3 Econometric Model

3.1 Models for Marginal Distribution

The growth rates of both the exchange rates and the inflation rates have the characteristics of heteroscedasticity, volatility, skewness, etc. Then, we use the ARMA-GARCH model in which standardized residuals satisfy the skewed-t distribution. Bollerslev (1986) proposed the GARCH (generalized autoregressive conditional heteroskedasticity) model that has replaced the ARCH model in application; the GARCH model has since been widely used in econometrics, economics, etc. In accordance with the findings of Songsak [24], the ARMA (p, q)-GARCH (k, l) model can be formulated as

$$r_t = c + \sum_{i=1}^{p} \phi_i r_{t-i} + \sum_{i=1}^{q} \psi_i \varepsilon_{t-i} + \varepsilon_t \tag{1}$$

$$\varepsilon_t = h_t \bullet \eta_t \tag{2}$$

$$h_t = \varpi + \sum_{i=1}^{k} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{l} \beta_i h_{t-i} \tag{3}$$

where $\sum_{i=1}^{p} \phi_i < 1, \omega > 0, \alpha_i \geq 0, \beta_i \geq 0$and $\sum_{i=1}^{k} \alpha_i + \sum_{i=1}^{l} \beta_i < 1$. The values of $\alpha_i$and β_i indicate the effect of short run shock and the persistence of volatility. When the values of $\alpha_i$are large, the short-term effects have greater influence. If the values of β_i are large, then the impact of unexpected shock on volatility is of longer duration.

3.2 Skewed Student-t Distribution

The skewed student-t distribution displays both flexible tails and possible skewness, each entirely controlled by a separate scalar parameter. The formula of skewed-t distribution, as taken from JianxuLiu [16], is shown as

$$P(x_i|v,\gamma) = \frac{2}{(\gamma+1/\gamma)}\left\{f_v(x_i/\gamma)I_{[0,\infty]}(x_i) + f_v(\gamma x_i)I_{[-\infty,0]}(x_i)\right\} \tag{4}$$

where $f_v(.)$ is the density function of the student-t distribution. The parameter v represents the degree of freedom, and γ is the skewness parameter that is defined from 0 to ∞ ; I denotes the indicator function.

3.3 Copula Functions

If there are uniform univariate marginal distribution functions in a multivariate distribution function, then the multivariate distribution function is a copula. A copula is a function that can link two or more marginal distributions together to form a joint distribution.

Ever since Sklar (1959), in the very beginning, described the original theorem of the relationship between a joint distribution and its marginal distributions, copula has been widely analyzed and applied in statistics.

Let F be an n-dimensional distribution function with marginal$F_1,...,F_n$. Then, there exists an n-copula C such that for all X inR_n

$$H(x_1,x_2,...,x_n) = C(F_1(x_1),F_2(x_2),...,F_n(x_n)) \tag{5}$$

If $F_1,...,F_n$are all continuous, then C is uniquely defined. The vice versa holds true for every copula C and for all types of distributions$F_1,...,F_n$. Sklar's theorem shows that the probability density function of any multivariate probability distribution can be represented by a marginal distribution and a dependence structure as follows:

$$\begin{aligned} f(x_1,...,x_n) &= \frac{\partial F(x_1,,...,x_n)}{\partial x_1,,...,\partial x_n} \\ &= \frac{\partial F(x_1,,...,x_n)}{\partial x_1,,...,\partial x_n} \times \prod \frac{\partial F(x_i)}{\partial (x_i)} \\ &= c(u_1,...,u_n) \times \prod f_i(x_i) \end{aligned} \tag{6}$$

This paper has two random variables, namely, the growth exchange rate, X_t, and the inflation rate, Y_t, with $H(x,y)$ as their joint probability distribution, which is a two-dimensional distribution function, as follows:

$$H(x,y) = C(F_x(X),F_y(Y)) \tag{7}$$

(*a*) Static Copulas

This study employed a variety of parametric copulas. Copulas include the Gaussian, t, Gumbel (rotated), Clayton (rotated), Frank, Joe (rotated), BB1 (rotated), BB6 (rotated), BB7 (rotated), and BB8 (rotated) copulas.

1. Gaussian copula: We follow the Patton [1] formula as follows:

$$C_{Ga}(u,v|\rho) = \int_{-\infty}^{\phi^{-1}(u)} \int_{-\infty}^{\phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{X^2 - 2\rho XY + Y^2}{2(1-\rho^2)}\right\} dXdY \tag{8}$$

where u and v are cumulative distribution functions or empirical cumulative functions of standardized residuals subjected to a uniform distribution between 0 and 1, the correlation coefficient ρ is Pearson's linear correlation.

2. t-copula: The formula for t-copula is as follows (from Songsak) [24]:

$$C_T(u,v|) = \int_{-\infty}^{T^{-1}(u)} dX \int_{-\infty}^{T^{-1}(v)} dY \frac{1}{2\pi\sqrt{1-\rho^2}} \left\{1 + \frac{X^2 - 2\rho XY + Y^2}{v(1-\rho^2)}\right\}^{\frac{-(v+2)}{2}} \tag{9}$$

where, $T_v(x) = \int_{-\infty}^{x} \frac{\Gamma((v+1)/2)}{\sqrt{\pi v}\Gamma(v/2)}(1+\frac{Z^2}{v})^{\frac{-(v+2)}{2}}$

$T$is the student-t distribution with degrees of freedom v and Pearson's correlation ρ, which is still linear. In comparison with the Gaussian copula, the biggest advantage of the t-copula is that it can capture tail dependence.

3. Archimedean copulas
   Archimedean copulas include the Clayton, Frank, and Gumbel copulas. The different copulas have different properties and applications, such as: The Clayton copula is an asymmetric Archimedean copula exhibiting greater dependence in the negative tail than in the positive. The Frank copula is a symmetric Archimedean copula, and the Gumbel copula is an asymmetric Archimedean copula exhibiting greater dependence in the positive tail than in the negative.
4. BBX copulas
   BBX copulas are two-parameter copulas in that BB6 and BB8 can capture the upper tail dependence, and BB1 and BB7 can reflect both the upper tail and the lower tail dependences.
5. Rotated copulas
   Many copulas cannot display negative tail dependences (e.g., the Gumbel, Clayton, Joe, and BBX copulas). Once the bivariate random variable has negative dependence, these copulas do not fit. However, these copulas may then be "rotated" 90°, 180°, and 270°, and applied again.

(*b*) Time-Varying Copulas

Time-varying copulas can be considered as the dynamic generalizations of a Pearson correlation or Kendall's tau. Patton [1] pointed out that it is still difficult to find causal variables to explain such dynamic characteristics, as did Songsak [24]. In practice, time-varying copulas are often assumed to follow the autoregressive moving average ARMA (p, q) process. The following are some formulas of time-varying copulas:

1. Time-varying Gaussian copula

$$\rho_t = \tilde{\Lambda}(\omega_N + \beta_N \rho_{t-1} + ... + \beta_{Np}\rho_{t-p} + \alpha_N \frac{1}{q}\sum_{j=1}^{q} \Phi^{-1}(u_{t-j})\Phi^{-1}(v_{t-j})) \quad (10)$$

where $\tilde{\Lambda}$ is a logistic transformation, which is defined as$\tilde{\Lambda} = (1-e^{-x})(1+e^{-x})^{-1}$.

2. Time-varying t-copula

$$\rho_t = \tilde{\Lambda}(\omega_T + \beta_{T1}\rho_{t-1} + ... + \beta_{Tp}\rho_{t-p} + \alpha_T \frac{1}{q}\sum_{j=1}^{q} T^{-1}(u_{t-j};DoF) \bullet T^{-1}(v_{t-j};DoF)) \quad (11)$$

where $T^{-1}$is the inverse function of the student t-distribution with the given degrees of freedom (DoF).

3. Time-varying (rotate) Gumbel copula

$$\tau_t = \Lambda(\omega_G + \beta_{G1}\tau_{t-1} + ... + \beta_{Gp}\tau_{t-p} + \alpha_G \frac{1}{q}\sum_{j=1}^{q} \left|u_{t-j} - v_{t-j}\right|) \quad (12)$$

where$\Lambda = (1+e^{-x})^{-1}$. This guarantees that Kendall's tau will be between -1 and 1.

4. Time-varying (rotate) Clayton copula

$$\tau_t = \Lambda(\omega_C + \beta_{C1}\tau_{t-1} + ... + \beta_{Cp}\tau_{t-p} + \alpha_{C1}\left|u_{t-1} - v_{t-1}\right| + ... + \alpha_{Cq}\left|u_{t-q} - v_{t-q}\right|. \quad (13)$$

3.4 Goodness-of-Fit Tests

We can consider the best-fit copula from two steps, as follows:

1. Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), and the formula as taken from Brechmann [13]. We have

$$AIC := -2\sum_{t=1}^{T} \ln[c(\hat{u}_t, \hat{v}_t; \theta)] + 2k \quad (14)$$

$$BIC := -2\sum_{t=1}^{T} \ln[c(\hat{u}_t, \hat{v}_t; \theta)] + \ln(T)k \quad (15)$$

where k=1 for one-parameter copulas, and k=2 for two-parameter copulas.

2. Two tests based on Kendall's transform.
Using Cramer-von Mises and Kolmogorov-Smirnov statistics, we estimate the p-values by using bootstrapping. In accordance with the research findings of Songsak [24], Genest and Rivest [7], and Genest [10], the formula can be written as follows:

$$S_n = \int_0^1 |\kappa_n(t)|^2 \, dK_{\theta n}(t) \tag{16}$$

$$T_n = \sup_{0 \le t \le 1} |\kappa_n(t)| \tag{17}$$

4 Descriptive Data and Empirical Results

The estimations in this part include the data description and statistics, the estimation of the marginal distribution of two random variables, the KS and Box-Ljung tests, static copulas, goodness-of-fit test for static copulas, and time-varying copulas.

4.1 Data Descriptions and Statistics

The statistics, as given in Table 1, show that the data for the growth rates of the exchange rate and the inflation rate are not normally distributed; the Jarque-Bera test is rejected at the 0.01 significance level. The distribution of the growth rate data is positively skewed. Therefore, we guess that these two marginal distributions were skewed-t distributions.

Table 1 Data Description and Statistics

Statistics	Growth Rate of Exchange Rate	Growth Rate of Inflation Rate
Mean	0.009072	0.012984
Median	0.00045	-0.0001
Maximum	0.2325	0.6223
Minimum	-0.1296	-0.4564
Std. Dev.	0.049376	0.164491
Skewness	1.416947	0.530071
Kurtosis	7.456132	4.016922
Jarque-Bera	320.7128	24.81737
Probability	0	0.000004
Observation	276	276

4.2 Estimates of Marginal Distribution of Growth Rates of Exchange Rate and Inflation Rate

For both the marginal distributions of the growth rates of the exchange rate and the inflation rate, it was found that the ARMA (1, 1)-GARCH (1, 1) skewed-t distributions, which are corresponding residuals, satisfy the i.i.d(0,1). The omega (ω) and alpha (α) are not significant, as shown in Table 2, which means that the constant as well as the residuals had no impact on the exchange rate. For both the distributions, λ is statistically significantly different from 1, implying that the skewed-t

distribution is necessary for these data sets; also, the $\beta = 0.81$ in Table 2 and the $\alpha + \beta = 0.94$ in Table 3, respectively, illustrate the growth rate of the exchange rate and the inflation rate, and also that there is a long-run persistence of volatility.

Table 2 Results of Growth Rate of Exchange Rate in ARMA (1,1)-GARCH (1,1) Skewed-t

	Estimate	Std. Error	t-value	Pr(>|t|)
ar1	0.9740	0.0142	68.4000	<2e-16 ***
ma1	-0.9620	0.0192	-50.2000	<2e-16 ***
ω	0.0000	0.0001	0.2360	0.814
α	1.0000	1.2200	0.8210	0.412
β	0.8130	0.0706	11.5000	<2e-16 ***
λ	1.0700	0.0460	23.3000	<2e-16 ***
Dof	2.1300	0.1560	13.7000	<2e-16 ***

Source: Computation.
Note: Signif. codes: 0.01 '***', 0.05 '**', 0.1 '*'.

Table 3 Results of Growth Rate of Inflation Rate in ARMA (1, 1)-GARCH (1, 1) Skewed-t

	Estimate	Std. Error	t-value	Pr(>|t|)
ar1	0.9850	0.0186	53.1000	<2e-16 ***
ma1	-0.9740	0.0235	-41.4000	<2e-16 ***
ω	0.0015	0.0010	1.5000	0.1347
α	0.1240	0.0522	2.3800	0.0172 **
β	0.8200	0.0717	11.4000	<2e-16 ***
λ	1.2500	0.1170	10.7000	<2e-16 ***
Dof	10.0000	4.1800	2.3900	0.0167 **

Source: Computation.
Note: Signif.Codes: 0.01 '***', 0.05 '**', 0.1 '*'.

4.3 KS and Box-Ljung Tests

Table 4 shows the margin values $\hat{u}_i$ and $\hat{v}_i$ of the growth rates of the exchange rate and the inflation rate. The results as obtained by the KS test are very clear: each series accepts the null hypothesis, which means that both the margins have uniform distribution. The Box-Ljung test for autocorrelation found the serial independence of the first four moments, and all of them accept the null hypothesis at the 0.10 level, implying that there is no autocorrelation from the first to the fourth moments.

Table 4 KS Test for Uniform Distribution and Box-Ljung Test for Autocorrelation

KS Test of Both Margins for Uniform Distribution			
	Statistic	P-value	Hypothesis
$\hat{u}_i$	0.0018	1	0 (acceptance)
$\hat{v}_i$	0.0018	1	0 (acceptance)
Box-Ljung Test of Both Margins for Autocorrelation			
		X-squared	P-value
$\hat{u}_i$	First moment	4.1075	0.534
	Second moment	1.7182	0.8866
	Third moment	6.6791	0.2456
	Fourth moment	3.0073	0.6989
$\hat{v}_i$	First moment	6.3313	0.7867
	Second moment	9.697	0.4675
	Third moment	13.326	0.206
	Fourth moment	9.9248	0.4471

Source: Computation.

4.4 Static Copulas

The results from the estimation of the static copulas with one parameter, as shown in Table 5, demonstrate that for each family of copulas, the best-fit copula fits the data with $P < 0.05$. From the AIC and BIC perspective, the Gumbel copula was the best among the one-parameter static copulas. For two-parameter static copulas, as illustrated in Table 6, the results showed that the only Rotated BB6 copula (180°) was not significant. The best copula model for the two-parameter static copula was the t-copula. However, we take into consideration the Gumbel copula and the t-copula by choosing from the AIC and the BIC. The t-copula was the best of all static copulas as it could capture tail dependence as well, as it has upper tail and low tail correlation. The value of the lower tail dependence is the same as the upper tail dependence, which is 0.48.

4.5 Goodness-of-Fit Test

Table 7 illustrates the results of the goodness-of-fit tests. In this study, we used the two tests based on the Kendall's transform for investigation, which are the Cramer-von Mises (CvM) test and the Kolmogorov Smirnov (KS) test. The results of the goodness-of-fit test, by providing the probability values for CvM and KS, show that half of the copulas did not reject the null hypothesis, that is, copulas such as the Gaussian copula, T copula, Joe copula, BB1 copula, BB7 copula, Rotated Joe

Table 5 Results for Copula Models of One Parameter

Copula	Parameter	Kendall's tau	AIC	BIC
Gaussian	0.6922***	0.4867	-177.185	-173.565
	0.0258			
Clayton	1.3462***	0.4023	-136.995	-133.374
	0.4109			
Gumbel	1.9833***	0.4957	-191.348	-187.728
	0.0991			
Frank	5.9672***	0.5124	-173.704	-170.083
	0.4789			
Joe	2.3091***	0.4166	-160.283	-156.663
	0.1419			
Rotated Clayton (180°)	1.4673***	0.4231	-162.022	-158.401
	0.1441			
Rotated Gumbel (180°)	1.9521***	0.4877	-180.306	-176.686
	0.0975			
Rotated Joe (180°)	2.191***	0.3946	-136.674	-133.054
	0.1371			

Note: Signif. codes: 0.01 '***', 0.05 '**', 0.1 '*'.

Table 6 Results for Copula Models of Two Parameters

Copula	δ	θ	Kendall's tau	AIC	BIC
t copula	0.71176***	2.6914***	0.5041	-205.98	-198.74
	0.0345	0.6671			
BB1	0.2952**	1.7686***	0.5073	-195.32	-188.07
	0.1305	0.1251			
BB6	1.001***	1.9819***	0.4957	-189.32	-182.08
	0.1251	0.4011			
BB7	2.0044***	0.8489***	0.484	-191.98	-184.74
	0.1532	0.1669			
BB8	4.0109***	0.8513***	0.4978	-178.96	-171.72
	1.0024	0.0962			
Rotated BB1 (180°)	0.5722***	0.5722***	0.5058	-197.75	-190.51
	0.1496	0.1167			
Rotated BB6 (180°)	1.001	1.9508**	0.4876	-178.27	-171.03
	0.7316	0.9481			
Rotated BB7 (180°)	1.7509***	1.1491***	0.8488	-195.45	-188.21
	0.1471	0.1656			
Rotated BB8 (180°)	6***	0.6568***	0.4917	-165.19	-157.95
	2.1038	0.1467			

Note: Signif. codes: 0.01 '***', 0.05 '**', 0.1 '*'.

copula (180°), Rotated BB1 (180°), and Rotated BB7 (180°), which means that they can appropriately model the dependency structure. Other copula candidates have rejected the null hypothesis, which were no fit for these data sets. As for which copula is the best-fit copula, we still choose the t-copula.

Table 7 Results of Goodness-of-Fit Test of Copula Models

Copula	CvM	KS		CvM	KS
Gaussian	0.29	0.39	Gumbel	0.05	0.1
T copula	0.6	0.39	Clyton	0	0
Joe	1	0.92	Frank	0	0.02
BB1	0.43	0.36	Rotated Clayton (180°)	0.03	0.02
BB7	0.62	0.71	Rotated Gumbel (180°)	0	0.01
Rotated Joe (180°)	1	1	BB6	0.09	0.12
Rotated BB1 (180°)	0.38	0.34	BB8	0	0
Rotated BB7 (180°)	0.43	0.34	Rotated BB6 (180°)	0	0
			Rotated BB8 (180°)	0	0

Source: Computation.

4.6 Time-Varying Copulas

As for the time-varying copulas, we still selected the AIC and the BIC as the criteria for choosing the best time-varying copulas. Table 8 shows the results of the time-varying copula analysis, and it displays the parameter values of the time-varying copulas, the standard error, the AIC, and the BIC. This study found that it is only the time-varying t-copula whose parameters are significant at the 0.01 level; also, the AIC and BIC values are smaller than those for the family of static copulas, and the autoregressive parameter β in the time-varying t copula was 1.42, suggesting that there was a high degree of persistence pertaining to the dependence structure between the growth rates of the exchange rate and the inflation rate. Therefore, the time-varying t copula is the best-fit copula for the dependence structure; also, we can see from Figure 1 that the Kendall's tau value of the time-varying t copula has a nonlinear, positive correlation, meaning that it can capture upper tail dependence. The smallest Kendall's tau was about 0.28, and the highest was about 0.88. In this study, we chose the time-varying t copula for policy implication.

Table 8 Results of Time-varying Copula Analysis

	Ω	β	α	AIC	BIC
Time-varying t copula	1.3934*** (0.2852)	1.4211*** (0.2281)	-2,4861*** (0.2272)	- 226.4208	- 237.2820

Source: Computation.
Note: Signif. codes: 0.01 '***', 0.05 '**', 0.1 '*'.

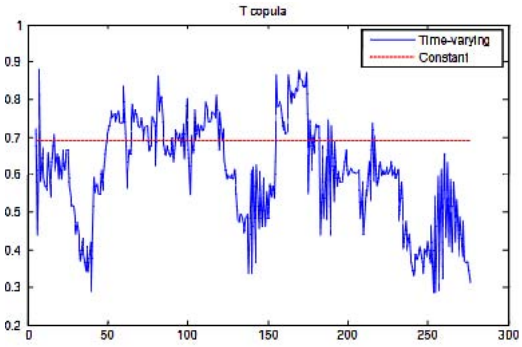


Fig. 1 Dependence parameter (ρ_t) from the time-varying t copula

5 Policy Implication

The testing process of the best-fit copula shows that the exchange rate and the inflation rate have strong nonlinear correlation such as that in the dynamic Kendalls tau, which explains the dependence between the growth rates of the exchange rate and the inflation rate; thus, obviously, it can be used to predict the next periods dependence. For the positive dependence structure, what this implies is that when the US Dollar is on an appreciation swing, it causes the inflation rate to increase. In addition, the time-varying t copula can help policy makers become aware of what is likely to happen in the future. Then the policy makers concerned with exchange rate policy should issue stronger policies for managing inflation.

6 Conclusion

This study modeled volatility and dependence. The marginal density shows that the ARMA (1, 1)-GARCH (1, 1) model was appropriate for the analysis of the growth rates of the exchange rate and the inflation rate. In addition, the Kolmogorov-Smirnov and Box-Ljung tests found that the two margins were skewed-t distributions, and that there was no autocorrelation for these data sets. The family of static copulas was used to analyze the dependence between the exchange rate and the inflation rate. Another point is that we applied time-varying copulas that explained the dynamic Kendall's tau. The empirical results show that the time-varying t copula was the best among the several copula candidates in terms of the AIC and BIC values, and that it has the highest explanatory power of all the dependence structures between the exchange rate and the inflation rate.

References

1. Patton, A.J.: Modeling asymmetric exchange rate dependence. International Economics Review 47, 527–556 (2006)
2. Patton, A.J.: Estimation of multivariate models for time series of possibly different lengths. Journal of Applied Econometrics 21, 147–173 (2006b)
3. Tsui, A.K., Yu, Q.: Constant conditional correlation in a bivariate GARCH model, evidence from the stock markets of China. Mathematics and Computers in Simulation 48, 503–509 (1999)
4. Charles, A.: Are Unit Root Tests Useful in the Debate over the (Non) Stationarity of Hours Worked? Research, University of Nantes (2011)
5. Arize, A.C., Malindretos, J., Nippani, S.: Variations in exchange rates and inflation in 82 countries: an empirical investigation. North American Journal of Economics and Finance 15, 227–247 (2004)
6. Arnold, B.C., Groeneveld, R.A.: Measuring skewness with respect to the mode. The American Statistician 49, 34–38 (1995)
7. Genest, C., Rivest, L.-P.: Statistical inference procedures for bivariate Archimedean copulas. Journal of the American Statistical Association 88(423), 1034–1043 (1993)
8. Garcia, C.J., Restrepo, J.E., Roger, S.: How much should inflation targeters care about the exchange rate. Journal of International Money and Finance 30, 1590–1617 (2011)
9. Wu, C.-C., Chung, H., et al.: The economic value of co-movement betweenoil price and exchange rate using copula-based GARCH models. Energy Economics 34(1), 270–282 (2012)
10. Genest, C., Quessy, J.-F., Rmillard, B.: Goodness-of-fit procedures for copula models based on the probability integral transformation. Scand. J. Statist. 33(2), 337–366 (2006)
11. Parsleya, D.C., Wei, S.-J.: Explaining the border effect: the role of exchange rate variability, shipping costs, and geography. Journal of International Economics 55, 87–105 (2001)
12. Duarte, M., Stockman, A.C.: Comment on: Exchange rate pass-through, exchange rate volatility, and exchange rate disconnect. Journal of Monetary Economics 49(5), 941–946 (2002)
13. Brechmann, E.C.: Truncated and simplified regular vines and their applications. Diploma thesis, Technische Universitaet Muenchen (2010)
14. Naz, F., Mohsin, A., Zaman, K.: Exchange rate pass-through in to inflation: New insights in to the cointegration relationship from Pakistan. Journal of Economic Modelling 29, 2205–2221 (2012)
15. Manner, H., Reznikova, O.: A survey on time-varying copulas: Specification, simulations and application. Econometric Reviews 31(6), 654–687 (2012)
16. Liu, J., Sriboonchitta, S.: Analysis of Volatility and Dependence between the Tourist Arrivals from China to Thailand and Singapore: A Copula-based GARCH Approach (2012)
17. Carranza, L., Galdon-Sanchez, J.E., Gomez-Biscarri, J.: Exchange rate and inflation dynamics in dollarized economies. Journal of Development Economics 89, 98–108 (2009)
18. Vogiatzoglou, M.: Dynamic copula toolbox (2010), `http://www.downloadplex.Com/Scripts/Matlab/Development-Tools/dynamic-copula-toolboxscripts_388595.html`
19. Parsley, D., Wei, S.: Explaining the border effect: the role of exchange rate variability, shipping costs and geography. NBER Working Paper 7836 (2000)
20. Engle, R.F.: Dynamic conditional correlation: a simple class of multivariate GARCH models. Journal of Business and Economic Statistics 20(3), 339–350 (2002)

21. Flooda, R.P., Rose, A.K.: Fixing exchange rates: A Virtual Quest for Fundamentals. Research Department, International Monetary Fund, Washington DC 20431 (1997)
22. Rogoff, K.: Perspectives on exchange rate volatility. In: Feldstein, M. (ed.) International Capital Flows, pp. 441–453. University of Chicago Press, Chicago (2001)
23. Chung, S.-K.: Bivariate mixed normal GARCH models and out-of-sample hedge performances. Finance Research Letters 6, 130–137 (2009)
24. Sriboonchitta, S.: Modeling Volatility and Dependency of Agricultural Price and Production Indices of Thailand: Static versus Time-Varying Copulas. Research of Faculty of Economics, Chiang Mai University (2013)
25. Verheggen, S.G.J.: Modeling stock return dependence between main American financial institutions during the financial crisis using a copula approach. Master thesis (2009)
26. Ling, S.: Self-weighted and local quasi-maximum likelihood estimators for ARMA-GARCH/IGARCH models. Journal of Econometrics 140, 849–873 (2007)
27. Morris, S.: Inflation dynamics and the parallel market for foreign exchange. Journal of Development Economics 46(1995), 295–316 (1990)
28. Fountas, S., Karanasos, M., Kim, J.: Inflation and output growth uncertainty and their relationship with inflation and output growth. Economics Letters 75, 293–301 (2002)
29. Chuang, W.-I., Liu, H.-H., Susmel, R.: The bivariate GARCH approach to investigating the relation between stock returns, trading volume, and return volatility. Global Finance Journal 23, 1–15 (2012)
30. Fang, W., Lai, Y., Miller, S.M.: Does exchange rate risk affect exports asymmetrically? Asian evidence. Journal of International Money and Finance 28, 215–239 (2009)
31. Wang, W., Wells, M.T.: Model selection and semiparametric inference for bivariate failure-time data. Journal of the American Statistical Association 95(449), 62–72 (2000)

Modeling Dependence of Accident-Related Outcomes Using Pair Copula Constructions for Discrete Data

Jirakom Sirisrisakulchai and Songsak Sriboonchitta

Abstract. This paper investigates the relationship between accident-related outcomes and per capita income, and explores the interdependency between them by using vine pair copula constructions. Equations for number of accidents, number of fatalities, and number of people injured are estimated using a provincial level data of Thailand in 2011. We discovered that there exists an inverted U-shaped relationship between accident-related outcomes and per capita income. Moreover, it was found that the accident-injury pair had stronger concordance and tail dependence, whereas the accident-fatality and fatality-injury pairs had weaker concordance and tail dependence. Our findings provide useful insight and information to policymakers who can then use the same to select appropriate road safety measures.

1 Introduction

The World Health Organization (WHO) reports that over 1.2 million people die and about 20 to 50 million people sustain non-fatal injuries from road accidents each year. WHO also estimates that road traffic injuries were the ninth leading cause of global mortality and burden of disease in 2004. Projections to 2030 show that road traffic injuries will move up in rank to become the fifth leading cause of global mortality, resulting in an estimated 2.4 million people per year. This problem is more severe in developing countries than in developed countries. Over 90 percent of the worlds traffic fatalities occur in developing countries which have only 48 percent of the worlds vehicles (WHO, 2009).

In Thailand, road accidents have killed approximately 130,000 people and injured nearly 500,000 people in the past decade (Source: Royal Thai Police). This has made road accidents the second leading cause of death in the country (Bundhamcharoen et al., 2011). These losses of human lives and ability to work have caused

Jirakom Sirisrisakulchai · Songsak Sriboonchitta
Faculty of Economics, Chiang Mai University, Chiang Mai 50200 Thailand
e-mail: sirisrisakulchai@hotmail.com, songsakecon@gmail.com

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_14,

substantial damage to the economy. It was estimated that the economic loss due to road accidents was about 232 billion Baht (approximately, 8 billion US dollars), or about 2.8 percent of the countrys GDP (Taneerananon, 2008). When compared with the recent estimates of national economic loss due to road traffic injuries across the country, which shows the range as from 1 to 2 percent (Jacob et al., 2000), Thailand has suffered more than the average. This may be attributed directly to rapid economic development, higher motorization levels, and increased complexity in the traffic patterns.

Tanaboriboon et al. (2005) roughly explained the fluctuation in the accident trends using the economic business cycle of Thailand. During the economic recession period (from 1984 to 1986), the fatality rate was rather stable where the average number of accident cases was around 19,000 per year, with about 5 fatalities and 17 injuries for every 100,000 of the population. In the economic recovery period (from 1987 to 1992), it was not just that the number of accidents rose dramatically about three times, but the severity of these accidents was also very distressing as the fatality and injury rates leaped to about 15 persons and 41 persons per 100,000 of the population, respectively. In the bubble economy period (from 1993 to 1996), road accidents shot up to about two to three folds of the previous period. The fatality rate also increased to about two times that of the previous period, at approximately 25 deaths per 100,000 of the population. In the economic crisis period (from 1997 to 2000), there was a positive trend in the road accident situation in Thailand. Road accidents were reduced to about 70,000 cases per year, with a similar downward trend in the fatality rate, but the injury rate showed a reverse trend with a high rate of 86 persons for every 100,000 of the population, in 1998. As expected, when the economy began the re-recovery era in the period from 2001 to 2003, the number of road accidents started to rise again, as did the fatality rate. However, the explanations for the same were based on intuition and did not use a statistical model.

Empirical literature on the relationship between traffic fatalities and economic development establishes the biphasic relationship, with the fatalities rising as a country develops and falling once the income levels exceed a certain threshold echelon (Law et al., 2009 and 2011; Bishai et al., 2006; Garg and Hyder, 2006; Koptis and Cropper, 2005; Beeck et al., 2000). Even though these studies used different data and statistical methods, they concluded that there is an inverted U-shaped relationship between traffic fatalities and per capita income. This result is similar to the curve Kuznets found as existing between income inequality and economic growth, which is well-known to environmental economists as the environmental Kuznets curve (Kuznets, 1955). Economists (Bishai et al., 2006) explain this relationship as follows: At low levels of income, an increase in income levels leads to higher motorization levels, which increases the risk to road users and causes traffic fatalities to rise. When income and economic development reach a certain level, general concern about road safety issues is more, which makes the government and the road users agree to and comply with the new institutions and regulations in order to enhance road safety. The more advanced stages of economic development are a prerequisite for the new institutions and regulations to successfully deal with road safety problems. Moreover, at low levels of income, the government is rational enough to under invest in

road safety and take the higher risk transport alternatives in order to promote higher income levels, which can be used to address the other public health issues. In an early epidemiological transition, curbing infections and nutritional health risks is of more concern as regards public health, and the government usually gives the first priority to these issues. In advanced stages of development, it becomes rational for the government to invest more resources in road safety and advanced medical aid and technology for road trauma victims. Due to the above-mentioned reasons, traffic fatalities gradually decline with higher levels of income.

The analysis of factors affecting accident-related outcomes is very often performed by studying only a single outcome. However, in practice there are several accident-related outcomes such as the number of accidents, the number of fatalities, the number of injuries, and so on. Moreover, road safety measures usually do not have the same effect on each outcome. A lack of knowledge of these interdependencies between the outcomes may be one reason for selecting inappropriate policies for reducing those accident-related outcomes. Since these outcomes show some interdependencies, their multivariate analysis is required to take into account the entirety of these interdependencies.

Cameron and Trivedi (1998, p. 252) state that applications of multivariate count models are relatively uncommon. Practical experience has been restricted to some special computationally tractable cases. However, a multivariate model for multiple count variables such as accident-related outcomes, as discussed in our paper, is one example of practical application.

Most of the multivariate count models start from the multivariate Poisson model (see Johnson et al., 1997). The multivariate Poisson distributions only allow for positive correlation. Chib and Winkelmann (2001) and Karlis and Xekalaki (2005) extend the multivariate Poisson model based on mixtures to allow for more flexible correlation structure and overdispersion on marginal distribution. The limitation of this approach is that the possible choices of mixing distribution are limited and sometimes they lead to the very specific marginal distribution. On the other side, Winkelmann (2000) constructed the bivariate negative binomial model and the model based on conditional distributions can be also constructed for bivariate count model (Berkhout and Plug, 2004). However, these models suffer from the difficulty to generalize to other families of marginal distributions.

The more innovative and flexible models are based on copula functions. Nikoloulopoulos and Karlis (2010) present the copula-based model for bivariate negative binomial regression. So et al., (2011) proposed an alternative bivariate zero-inflated negative binomial model based on a copula that allows for heterogeneous dispersion, negative correlations, and a more general zero-inflation structure. The advantages of the copula-based bivariate count models are that they allow for both flexible dependence structure and marginal distributions. Nikoloulopoulos and Karlis (2009) proposed the multivariate count models by using the multivariate parametric family of copulas. In contrast, we use a pair-copula construction for discrete margins, as proposed by Panagiotelis (2012). Panagiotelis et al. (2012) recently proposed a copula modeling framework for multivariate discrete data that is flexible, easy to estimate, and applicable in high dimensions. This framework fits very well

with our empirical work and makes it tractable to perform the multivariate count model.

This paper has two objectives. The first objective is to identify the relationship between accident-related outcomes and economic development as measured by per capita gross provincial product (GPP) in Thailand. The second objective is to perform the multivariate count models for accident-related outcomes in order to get a better understanding of the nature of interdependency between those outcomes.

The rest of this paper is organized as follows. In Section 2, we describe the data used. In Section 3, we provide a brief discussion of negative binomial regression and vine pair copula constructions (PCC). In Section 4, we discuss our empirical results, followed by discussion and conclusion in Section 5.

2 Statistical Models

2.1 Negative Binomial Regression

We proxy the factors affecting accident-related outcomes with the economic development of each province, as measured by the per capita gross provincial product (GPP). Negative binomial distribution is used to model the relationship between accident-related outcomes and gross provincial products (GPP). We use a quadratic specification to test for the inverted U-shaped relationship. In our analysis, the risk of a casualty accident varies across provinces depending on the level of exposure, such as the population numbers. For example, a province with a higher population should have more crashes, given that all other characteristics are held constant. Thus, we use the population number to normalize the effect of risk exposure on the accident-related outcomes. This can be done by using population as an offset variable in our specification. The model specification is

$$\log(Y_i) = \alpha_i + \log(population_i) + \beta_1 \log(GPP_i) + \beta_2 [\log(GPP_i)]^2 + \varepsilon_i$$

where Y stands for crash counts, that is, traffic fatalities, injuries, and numbers of traffic accidents, GPP_i is per capita income (measured at current market prices), $\alpha_i$is an intercept, and $\varepsilon_i$is an error term.

2.2 Copula

We perform the multivariate analysis of the accident-related outcome by using the copula model. In recent years, copula modeling has been extensively applied in many fields of application. Introduction and standard reference on copula theory can be found in Joe (1996) and Nelsen (2006). The foundation of copula is based on the theorem of Sklar (1959), which states that there is a copula function C such that

$$F(y_1, y_2, ..., y_m) = C(F_1(y_1), F_2(y_2), ..., F_m(y_m)) \quad (1)$$

where $y = (y_1, y_2, ..., y_m)$ is the realization of an m-dimensional random vector $Y = (Y_1, Y_2, ..., Y_m)$. $F_j(y_j)$ is the marginal distribution function of the j^{th} margin for $j = 1, 2, ..., m$ and F is a joint distribution function. Sklar's theorem establishes the link between the multivariate distribution function (copula) and their univariate margins.

In the continuous case, we can obtain the joint density $f(y_1, y_2, ..., y_m)$ by taking the derivative of both the sides of equation (1), which is

$$f(y_1, y_2, ..., y_m) = c(F_1(y_1), F_2(y_2), ..., F_m(y_m)) f_1(y_1) f_2(y_2) ... f_m(y_m),$$

where $f_j(y_j)$ is the marginal density function of the j^{th} margin, f is a joint density function, and c is a copula density function. The copula function is unique for the continuous random vector Y. For a discrete random vector, the copula function is unique only over the Cartesian product of the ranges of the marginal distribution function (Genest and Neslehova, 2007; Panagiotelis et al., 2012). However, Genest and Neslehova (2007) demonstrated evidence to show that parametric modeling of discrete variables by copula acquires dependence properties in a way that is similar to the continuous case.

In general, there are two approaches to compute the probability mass function for discrete variables. Both are evaluated by taking the difference of the copula function. Consider the case in which the discrete variables are non-negative integers; the joint probability mass function (pmf) of Y would be

$$\Pr(Y = y) = \sum_{i_1=0,1} \cdots \sum_{i_m=0,1} (-1)^{i_1+...+i_m} C(F_1(y_1 - i_1), ..., F_m(y_m - i_m)) \quad (2)$$

To compute this pmf, we have to evaluate 2^m times of the copula functions. The second approach is based on vine pair copula constructions (PCC), which we will briefly discuss, both in the continuous and in the discrete cases, in the next section.

2.3 Vine Pair Copula Constructions

Vine pair copula constructions (PCC) were initially proposed by Joe (1996) and developed in more detail in the works of Bedford and Cook (2001, 2002), and Kurowicka and Cooke (2006). Aas et al. (2009) provided a principle for constructing multivariate copula from the product of bivariate pair copula and described the statistical inference techniques for the specific vines which were called canonical (C-) vines and drawable (D-) vines.

Vine PCCs are tree-like constructions with bivariate copula as the building blocks. A vine is characterized by $m - 1$ trees. Each tree is made up of nodes, and edges joining these nodes. The most popular and specific structures of the vine are the C-vines and D-vines. C-vine trees have a star structure and D-vine trees have a path structure. As for three-dimensional cases, there are six ways of permuting the three variables but only three give different decompositions. Moreover, these decompositions are both C-vine and D-vine.

For continuous$Y$, a PCC is derived, starting with the factorization of the joint density function into the conditional density function and the marginal density function, as follows:

$$f(y_1, y_2, ..., y_m) = f_{1|2,...,m}(y_1 | y_2, ..., y_m)\, f_{2|3,...,m}(y_2 | y_3, ..., y_m) \ldots f_m(y_m) \tag{3}$$

By Sklar's theorem, it can be shown that the conditional density function on the right hand side of equation (3) can be decomposed into the product of a bivariate copula density and a univariate conditional density. This can be done recursively to each of the terms on the right hand side of equation (3) until $f(y_1, y_2, ..., y_m)$is decomposed into the product of m(m-1)/2 bivariate copulas (Panagiotelis, 2012). The decomposition in this manner can be done in several ways. For a summary of the different ways to decompose the joint density function, the readers are requested to refer to Bedford and Cooke (2001, 2002).

For discrete margins, we can decompose a pmf, as follows:

$$\begin{aligned} \Pr(Y_1 = y_1, ..., Y_m = y_m) = \Pr(Y_1 = y_1 | Y_2 = y_2, ..., Y_m = y_m) \times \\ \Pr(Y_2 = y_2 | Y_3 = y_3, ..., Y_m = y_m) \times ... \times \Pr(Y_m = y_m) \end{aligned} \tag{4}$$

And then, in a similar manner as the continuous case, each term on the right hand side of equation (4) can be decomposed into the product of a bivariate copula. However, in contrast with the continuous case, for each m(m-1)/2 bivariate copula function in discrete pmf, we have to evaluate four different values. Therefore, there is a total of 2m(m-1) evaluation values at a single point.

We now derive a three-dimensional discrete margin PCC in our crash counts model. For $m = 3$,

$$\begin{aligned} &\Pr(Y_1 = y_1, Y_2 = y_2, Y_3 = y_3) \\ &\quad = \Pr(Y_1 = y_1 | Y_2 = y_2, Y_3 = y_3) \times \Pr(Y_2 = y_2 | Y_3 = y_3) \times \Pr(Y_3 = y_3), \end{aligned} \tag{5}$$

where

$$\begin{aligned} &\Pr(Y_1 = y_1 | Y_2 = y_2, Y_3 = y_3) \\ &\quad = \frac{\left\{\Sigma_{i_1=0,1} \Sigma_{i_2=0,1} (-1)^{i_1+i_2} C_{12|3}\left(F_{1|3}(y_1 - i_1 | y_3), F_{2|3}(y_2 - i_2 | y_3)\right)\right\}}{\Pr(Y_2 = y_2 | Y_3 = y_3)} \end{aligned} \tag{6}$$

And the arguments in the copula function are

$$F_{1|3}(y_1 - i_1 | y_3) = \frac{C_{13}(F_1(y_1 - i_1), F_3(y_3)) - C_{13}(F_1(y_1 - i_1), F_3(y_3 - 1))}{\Pr(Y_3 = y_3)},$$

and

$$F_{2|3}(y_2 - i_2 | y_3) = \frac{C_{23}(F_2(y_2 - i_2), F_3(y_3)) - C_{23}(F_2(y_2 - i_2), F_3(y_3 - 1))}{\Pr(Y_3 = y_3)}$$

Since the dominator of equation (2) cancels with the second term on the right hand side of equation (1), the full expression for the pmf of the three-dimensional discrete margin PCC is

$$\begin{aligned}&\Pr(Y_1 = y_1, Y_2 = y_2, Y_3 = y_3)\\ &\quad = \left\{\sum_{i_1=0,1}\sum_{i_2=0,1}(-1)^{i_1+i_2}C_{12|3}\left(\frac{C_{13}(F_1(y_1-i_1),F_3(y_3))-C_{13}(F_1(y_1-i_1),F_3(y_3-1))}{F_3(y_3)-F_3(y_3-1)},\right.\right.\\ &\qquad \left.\left.\frac{C_{23}(F_2(y_2-i_2),F_3(y_3))-C_{23}(F_2(y_2-i_2),F_3(y_3-1))}{F_3(y_3)-F_3(y_3-1)}\right)\right\}[F_3(y_3)-F_3(y_3-1)]\end{aligned}$$

A major advantage of PCC, when compared with the first approach, is the greater flexibility to model a large range of dependence structures, which greatly reduces the computational cost of evaluating the higher dimension of pmf (Panagiotelis, 2012).

In this paper, we construct the marginal models by negative binomial distribution with the covariate that proxy economic development by GPP, as discussed in section 2.1. Then we can estimate the parameters of marginal models and dependence parameters by using maximum likelihood estimation.

3 Data

The accident-related outcomes considered in this study consist of the number of accidents, the number of fatalities, and the number of injuries. The data used in this analysis is a cross-sectional data for 77 provinces of Thailand in 2011. Data on traffic fatalities, injuries, and accidents were obtained from the Royal Thai Police, Ministry of Public Health (MPH), and Department of Highway (DOH). For the data on population and gross provincial product, we approached the National Economic and Social Development Office.

Ponboon and Tanaboriboon (2005) were skeptical of the accuracy of the traffic fatalities data in Thailand. They showed the difference between the numbers of traffic fatalities as reported by the police and as per hospital records, and the difficulty in predicting the trend because of the under-reported data. Since there is no distinct definition of fatality in Thailand, the fatality reports by the police only include death at the scene. The number of accidents and people injured as reported by the police is the number of cases to be in lawsuit. As far as MPH data is concerned, we retrieved the data directly from the trauma accident database system. MPH receives the report on the number of people injured from hospitals. These numbers include minor injuries as well as serious injuries. Therefore, the number of people injured as reported by MPH is ten times higher than the number reported by the police. As far as police data is concerned, they tend to report only the serious injury cases. The police may never hear of a crash in a rural area involving a single vehicle and with slightly injured people. MPH data also reports the fatality cases in which the victims were admitted to hospitals and died there. As for DOH, they report only those accident-related outcomes that take place on the national highways. About 14 percent of the total fatalities reported by the police occurred on the national highways.

4 Empirical Results

4.1 Negative Binomial Regressions

The descriptive statistics of accident-related outcomes for three different data sets are reported in Table 1. The province Amnat Charoen, the poorest, had a per capita income of about 30,231 Baht, and Rayaong, the richest, had a per capita income of about 1,235,694 Baht in 2011. That is, the per capita income of the richest province is about 40 times that of the poorest province. This disparity should explain some of the variations in traffic fatalities between the provinces.

Table 2 presents the results of estimation of the negative binomial (NB) regression for the number of accidents, traffic fatalities, and injuries. The likelihood ratio (LR) test is used to test for over-dispersion in the data (Cameron and Trivedi, 1998). The LR test results show the existence of over-dispersion. This makes the negative binomial model a better choice than the Poisson model[1]. The structure of Table 2 consists of three sets of three columns each. The first one shows the results from using the police data set, the second, those from using the MPH, and the last one, those from using the DOH. The three columns of each set consist of negative binomial regression for the number of accidents, traffic fatalities, and injuries, respectively. The standard errors presented in Table 2 are robust standard errors.

The results from the NB model show the presence of the inverted U-shaped pattern of traffic fatalities, and so, it cannot be statistically rejected for the MPH and DOH data sets. As for the police data, it also has an inverted U-shaped pattern, but the coefficient estimates are not statistically significant. In contrast to the finding of Bishai et al. (2006) for international data, we found the presence of the inverted U-shaped pattern for traffic accidents and injuries in the MPH and DOH data sets. These results suggest that the drop in the traffic fatality rate for the richer provinces may be due to fewer injuries and crashes. The interdependencies between these outcomes shall be discussed in the next section.

Since the MPH data covers a broad range of accident-related outcomes, a discussion on only the coefficient interpretation of the MPH data will be presented here. The income level at which the traffic fatalities first decline is 301,608 Baht. This was the approximate income level attained by provinces such as Pathumthani, Chachoengsao, and Phuket in 2011. There are 10 provinces[2] which have income levels higher than this. Figure 1 plots the predicted number of traffic fatalities as a function of per capita income to give a complete picture of the model results.

Figure 1 also gives the details of the predicted number of accidents and injuries. The pattern and shape of the traffic accidents and injuries are almost the same. The income level at which the number of accidents and injuries first declines is about 160,000 Baht. This is the approximate income level attained by provinces such as

[1] Note that our data sets do not have a problem of excessive zero observations, so we are not taking into consideration a zero-inflated negative binomial regression.

[2] These provinces comprise Pathumthani, Chachoengsao, Phuket, Ayudthaya, Prachin Buri, Samut Prakan, Bangkok, Chonburi, Samut Sakorn, and Rayong.

Trat, Phang Nga, and Nonthaburi. As can be seen from Figure 1, the number of accidents and injuries first starts to decline when the income reaches a number around 160,000 Baht and then, later, at an income level of about 300,000 Baht. One possible reason for this may be that economic development empowers more road users to switch their mode of transport to the safer cars instead of the more dangerous modes of transport such as motorcycles.

The regression models in Table 3 introduce some control variables, namely, vehicles per capita, number of drunk driving cases, and number of speed limit exceeding cases. However, the introduction of these variables does not change the inverted U-shaped pattern in the MPH and DOH data sets. Law et al. (2011) found the indication that political and institutional development, and medical care and technology improvement are the main sources of the Kuznets relationships. Upon controlling these variables, the inverted U-shaped pattern vanished for the international data sets. However, the same variables could not be found for the provincial data in Thailand.

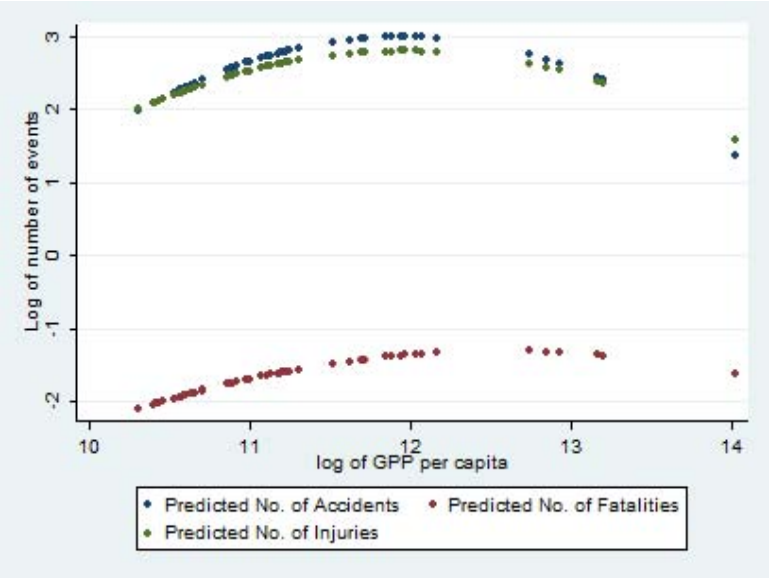


Fig. 1 The predicted number of accident-related outcomes and GPP per capita for the MPH data

4.2 Discrete Vine PCC Results

Before estimating the vine PCC model, we test for marginal model misspecification using the Kolmogorov-Smirnov (KS) test. If the marginal distributions are correctly specified, then the probability transformations should be independent and identically distributed uniformly (0,1). Table 4 shows the result of the KS-test. We cannot reject

Table 1 Descriptive Statistics of Accident-related Outcomes

Data set	Variable	Obs.	Mean	Std. Dev.	Min	Max
Police	accident	77	890.69	4088.24	54	35947
	death	77	119.55	75.17	16	360
	injury	77	284.64	894.38	16	7923
MPH	accident	52	10800.90	9666.69	105	51867
	death	52	168.19	139.22	20	792
	injury	52	9959.17	8067.68	44	35916
DOH	accident	77	137.75	187.24	10	1455
	death	77	16.77	14.59	0	83
	injury	77	116.49	99.12	1	437
	GPP (Baht)	77	140765.9	175818.1	30231	1235695
	pop (1000 person)	77	877.9	846.9	192	6859
	No. of drunk cases	70	815.4286	1586.023	4	8875
	No. of speeding cases	65	2842.262	5509.052	1	24153
	No. of vehicles	77	393791.4	778788.9	14616	6885080

Table 2 Parameter Estimates for Negative Binomial Regression

	police			mph			doh		
	accident	death	injury	accident	death	injury	accident	death	injury
	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.
lGPP	1.345	2.367	2.549	**9.084***	**3.886***	**6.901***	**6.175***	**5.863***	**5.861***
lGPPsq	-0.036	-0.095	-0.094	**-0.380***	**-0.154***	**-0.288***	**-0.237***	**-0.236***	**-0.232***
_cons	-11.239	**-16.540***	-18.203	**-51.299***	**-25.751***	**-38.524***	**-41.356***	**-39.807***	**-38.424***
log(pop)	1	1	1	1	1	1	1	1	1
alpha	**0.391***	**0.127***	**0.258***	**0.599***	**0.126***	**0.659***	**0.356***	**0.560***	**0.481***
LL	-517.13	-387.92	-446.38	-524.92	-275.58	-520.6	-418.02	-284.63	-426.45
turning point	1.06E+08	257273	773368	155225	301608	159676	454715	248106	306038
No of province		77			52			77	

*indicate statistical significance at the 5 percent level

the null hypothesis that the probability transformations of the margins are a uniform distribution.

For the three-dimensional model of accident-related outcomes, we estimate all three combinations of the ordering variables. We consider using the Gaussian, Student-t, Clayton, Gumbel, Frank, Joe, BB1, BB6, BB7, and BB8 copulas as bivariate copula building blocks[3]. These copulas were chosen because they give different shapes and they exhibit different tail dependence, namely, no tail dependence, lower tail dependence, upper tail dependence, symmetric tail dependence, and asymmetric tail dependence.

[3] For details of the copula functional form, see Joe (1997).

Table 3 Parameter Estimates for Negative Binomial Regression (with some control variables)

	Police			MPH			DOH		
	accident	death	injury	accident	death	injury	accident	death	injury
	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.
lgppcap	2.397	1.691	**2.926***	**7.280***	**3.018***	5.312	**5.268***	**4.913***	**5.769***
lgppsq	-0.101	-0.066	-0.117	**-0.311***	**-0.120***	-0.225	**-0.201***	**-0.199***	**-0.226***
lcar	**0.598***	**-0.191***	**0.279***	0.591	0.253	-0.338	-0.219	**-0.648***	-0.181
ldrink	**0.136***	0.002	-0.003	-0.055	-0.035	-0.038	-	-	-
lspeed	-	-	-	-	-	-	-0.003	-0.014	0.019
_cons	**-18.318***	**-13.603***	**-21.177***	**-43.291***	**-21.756***	-30.519	**-37.102***	**-37.668***	**-39.232***
lpop	1	1	1	1	1	1	1	1	1
alpha	0.209	0.074	0.174	0.462	0.089	0.522	0.341	0.420	0.436
LL	-440.36	-335.62	-389.32	-502.79	-259.04	-499.59	-343.87	-232.26	-353.53
No of province		70			50			65	

*indicate statistical significance at the 5 percent level.

Since all the models in our analysis have the same number of parameters, we select the best-fit model by the highest value of log-likelihood. This strategy corresponds to selecting the model by popular information criteria such as the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). We also investigate further on the improvement of the goodness of fit of our three combinations of the ordering variables by performing Vuong (1989) and Clarke (2007) tests. Both Vuong and Clarke tests analyze the null hypothesis that both the models can explain the data equally well as against the argument that the model with the higher log-likelihood is to be favored.

As for the police data, the Vuong test cannot reject the null hypothesis of no difference between those three models. However, the Clarke test selects the model in which the second tree is conditional on the number of fatalities. As far as the MPH data is concerned, neither the Vuong test nor the Clarke test can reject the null hypothesis of no difference between those three models. Therefore, we present only model 1, which has the same structure as the model fitted on the police data, for comparison. As for the DOH data, the Clarke test selects the model in which the second tree is conditional on the number of accidents.

There are two popular approaches for the estimation of unknown parameters in discrete vine PCC, namely, the inference function for margins (IFM) and full maximum likelihood (FML). The results show a slight improvement of the FML over the IFM estimations in terms of log-likelihood. Table 4 shows the parameter estimates and the corresponding Kendall's tau measures for three-dimensional vine PCC models by the FML method.

The dependence structure between traffic accident, fatality, and injury has the same direction for all the three different data sets. The accident–injury pair has stronger concordance measure (Kendall's tau) when compared with the accident–fatality pair and the fatality–injury pair. However, for the best-fit models, there exists a different pattern of tail dependence (dependence in extreme values) for the three

different data sets. For example, the MPH data exhibits weak lower tail dependence in both the accident–fatality pair and the fatality–injury pair, but strong upper tail dependence in the accident–injury pair.

Table 4 Goodness of Fit Test for Margins by Kolmogorov-Smirnov Test

	Police	MPH	DOH
	p-value		
Accident	0.07	0.08	0.79
Fatality	0.15	0.97	0.65
Injury	0.86	0.06	0.99

Table 5 Parameter Estimates for Discrete Margins PCC of Accident-related Outcomes

Police			MPH			DOH		
Copula	Parameter	Value	Copula	Parameter	Value	Copula	Parameter	Value
t	par1	0.172	Clayton	par1	0.567	BB7	par1	1.262
	par2	2.446		par2	0.000		par2	1.125
c12	tau	0.110	c12	tau	0.221	c12	tau	0.408
	upper tail	0.205		upper tail	0		upper tail	0.268
	lower tail	0.205		lower tail	0.295		lower tail	0.540
t	par1	0.434	Clayton	par1	0.465	t	par1	0.826
	par2	4.237		par2	0.000		par2	3.701
c23	tau	0.286	c23	tau	0.189	c13	tau	0.619
	upper tail	0.207		upper tail	0		upper tail	0.535
	lower tail	0.207		lower tail	0.225		lower tail	0.535
clayton	par1	1.493	Gumbel	par1	3.030	BB7	par1	1.639
	par2	0.000		par2	0.000		par2	0.319
c13c2	tau	0.427	c13c2	tau	0.670	c23c1	tau	0.338
	upper tail	0.000		upper tail	0.743		upper tail	0.474
	lower tail	0.629		lower tail	0		lower tail	0.114

Note: 1 = number of accidents, 2 = number of fatalities, 3 = number of injuries.

5 Discussion and Conclusion

In the current paper, we demonstrate the application of a vine PCC to model the dependency between the various accident-related outcomes according to the provincial data in Thailand. The marginal models are negative binomial regressions. The results show a pattern of inverted U shape between the accident-related outcomes and the per capita income for the MPH and DOH data sets. The improvements in medical care and technology are hypothesized to impact the accident-related outcomes. In future research, the inclusion of these proxy variables to the model could give a better understanding of the factor underlying the Kuznets relationship. The

vine PCCs are performed to find out the dependence model between these accident-related outcomes.

Based on the strong concordance measure of the accident–injury pair, the policymaker should consider the effect of road safety measures on both the number of accidents and the number of people injured, taken together. Road safety measures aimed at reducing the number of accidents might have the same effect on reducing the number of people injured, but may not necessarily have the same effect on reducing traffic fatalities. The vine PCC frameworks used here can be employed to model the decomposition effect of road safety measures on accident-related outcomes simultaneously. This model should be able to give a better understanding of the decomposition effect of road safety measures than single outcome models.

References

1. World Health Organization. Global status report on road safety: time for action, Geneva (2009), http://whqlibdoc.who.int/publications/2009/9789241563840_eng.pdf
2. Bundhamcharoen, K., Odton, P., Phulkerd, S., Tangcharoensathien, V.: Burden of disease in Thailand: changes in health gap between 1999 and 2004. BMC Public Health 11, 53 (2011)
3. Taneerananon, P.: The Cost of Road Accidents in Thailand, Technology and Innovation for Sustainable Development Conference (TISD). Faculty of Engineering, KhonKaen University, Thailand (January 28-29, 2008)
4. Tanaboriboon, Y., Satiennam, T.: Road Accidents in Thailand. Journal of the International Association of Traffic and Safety Sciences, IATSS Research 29(1), 88–100 (2005)
5. Jacobs, G., Aeron-Thomas, A., Astrop, A.: Estimating global road fatalities: Crowthorne: Transport Research Laboratory (2000)
6. Law, T.H., Noland, R.B., Evans, A.W.: Factors associated with the relationship between motorcycle deaths and economic growth. Accident Analysis and Prevention 41, 234–240 (2009)
7. Law, T.H., Noland, R.B., Evans, A.W.: The sources of the Kuznets relationship between road fatalities and economic growth. Journal of Transport Geography 19, 355–365 (2011)
8. Koptis, E., Cropper, M.: Traffic fatalities and economic growth. Accident Analysis and Prevention 37(1), 169–178 (2005)
9. Garg, N., Hyder, A.A.: Exploring the relationship between development and road traffic injuries: a case study from India. Eur. J. Pub. Health 16(5), 487–491 (2006)
10. Beeck, E.F.V., Borsboom, G.J.J.M., Mackenbach, J.P.: Economic development and traffic accident mortality in the industrialized world, 1962–1990. International Journal of Epidemiology 29(3), 503–509 (2000)
11. Bishai, D., Quresh, A., James, P., Ghaffar, A.: National road casualties and economic development. Health Economics 15(1), 65–81 (2006)
12. Kuznets, S.: Economic growth and incomes inequality. The American Economic Review 45(1), 1–28 (1955)
13. Cameron, A.C., Trivedi, P.K.: Regression Analysis of Count Data. Econometric Society Monograph, No. 30. Cambridge University Press (1998)
14. Ponboon, S., Tanaboriboon, Y.: Development of Road Accident Reporting Computerized System in Thailand. Journal of the Eastern Asia Society for Transportation Studies 6, 3453–3466 (2005)

15. Panagiotelis, A., Czado, C., Joe, H.: Pair Copula Constructions for Multivariate Discrete Data. Journal of the American Statistical Association 107(499), 1063–1072 (2012)
16. Joe, H.: Multivariate Models and Dependence Concepts. Chapman & Hall, London (1997)
17. Joe, H.: Families of m-variate distributions with given margins and m(m - 1)/2 bivariate dependence parameters. In: Ruschendorf, L., Schweizer, B., Taylor, M.D. (eds.) Distributions with Fixed Marginals and Related Topics, pp. 120–141 (1996)
18. Bedford, T., Cooke, R.M.: Probability density decomposition for conditionally dependent random variables modeled by vines. Annals of Mathematics and Artificial Intelligence 32, 245–268 (2001)
19. Bedford, T., Cooke, R.M.: Vines – a new graphical model for dependent random variables. Annals of Statistics 30(4), 1031–1068 (2002)
20. Nelsen, R.B.: An Introduction to Copulas, 2nd edn. Springer, New York (2006)
21. Sklar, A.: Fonctions de repartition a n dimensions et leurs marges. Publications de l'Institut de Statistique de l'Universite de Paris 8, 229–231 (1959)
22. Aas, K., Czado, C., Frigessi, A., Bakken, H.: Pair-Copula Constructions of Multiple Dependence. Insurance, Mathematics and Economics 44, 182–198 (2009)
23. Genest, C., Neslehova, J.: A Primer on Copulas for Count Data. The Astin Bulletin 37, 475–515 (2007)
24. Kurowicka, D., Cooke, R.: Uncertainty Analysis With High Dimensional Dependence Modelling. Wiley Series in Probability and Statistics. Wiley, Chichester (2006)
25. Vuong, Q.: Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses. Econometrica 57, 307–333 (1989)
26. Clarke, K.: A Simple Distribution Free Test for Non-Nested Model Selection. Political Analysis 13, 347–363 (2007)
27. Berkhout, P., Plug, E.: A bivariate Poisson count data model using conditional probabilities. Statistica Neerlandica 58, 349–364 (2004)
28. Chib, S., Winkelmann, R.: Markov chain Monte Carlo analysis of correlated count data. Journal of Business & Economic Statistics 19(4), 428–435 (2001)
29. Johnson, N., Kotz, S., Balakrishnan, N.: Discrete Multivariate Distributions. Wiley, New York (1997)
30. Karlis, D., Xekalaki, E.: Mixed Poisson distributions. International Statistical Review 73, 35–58 (2005)
31. Winkelmann, R.: Seemingly unrelated negative binomial regression. Oxford Bulletin of Economics and Statistics 62(4), 553–560 (2000)
32. Nikoloulopoulos, A.K., Karlis, D.: Modeling Multivariate Count Data Using Copulas. Communications in Statistics - Simulation and Computation 39(1), 172–187 (2009)
33. Nikoloulopoulos, A.K., Karlis, D.: Regression in a copula model for bivariate count data. Journal of Applied Statistics 37(9) (2010)
34. So, S., Lee, D.-H., Jung, B.C.: An alternative bivariate zero-inflated negative binomial regression model using a copula. Economics Letters 113(2), 183–185 (2011)

Dependence Analysis of Exchange Rate and International Trade of Thailand: Application of Vine Copulas

Chakorn Praprom and Songsak Sriboonchitta

Abstract. This paper aims to investigate the correlation of multivariate dependences between the international trade of Thailand and the USD/THB exchange rate using vine copulas, including canonical (C-vine) and drawable (D-vine) vine copulas which are very flexible dependency structures. Another advantage is that these methods overcome limitations and complex dependency models. Before we built the pair-copula constructions of the vine models, ARMA(1,1)-GARCH(1,1) was adopted to remove time dependence in each of the marginal time series. Furthermore, we got the various standardized residuals to transform into appropriate uniform margins $[0,1]$. The results can be seen for C-vine case, Gaussian, Rotated Joe, and BB1 which are suitable bivariate copula families for each pair-copula construction. On the other hand, D-vine case, Gaussian, and Rotated Joe are appropriate copula families for the pair-copula construction. In addition, the sequential log-likelihood is quite close to the one obtained by joint maximization; it means that both the vine models are appropriate-fit models. In order to confirm that it is not possible to distinguish between the two models, we employed the Vuong and Clarke tests to verify the suitability of the non-nested model. These tests confirm that the C-vine and D-vine copulas are not distinguishable. It can be concluded that our pair constructions of the time-varying Gaussian copula could be appropriate fits, better than those of the static copula. This study will help policy makers take action to combat the exchange rate volatility.

Chakorn Praprom
Faculty of Economics, Chiang Mai University, Chiang Mai 50200 Thailand,
Department of Social Sciences, Faculty of Humanities and Social Sciences,
Prince of Songkla University, Pattani Campus, Thailand
e-mail: chakornpraprom@gmail.com

Songsak Sriboonchitta
Faculty of Economics, Chiang Mai University,
Chiang Mai 50200 Thailand
e-mail: songsak@econ.cmu.ac.th

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_15,

1 Introduction

Over the past two decades, many countries have been turning their attention to the issue of economic stability more and more because of the occurrence of several economic crises repeatedly, such as the Asian financial crisis in 1970, the subprime mortgage crisis in the United States in 2007, the European sovereign-debt crisis in 2009, and, lastly, the Cyprus crisis. These crises introduced adverse economic impacts that spread widely from country to country and, subsequently, reached global levels. All nations around the world have suffered from this economic impact in one way or the other, in terms of international trade, foreign investment, international financial market, foreign capital movement, stock index, and foreign exchange rate market. These impacts affected each and every countrys economic goals, and, ultimately, it is bound to influence the Gross Domestic Production (GDP) growth.

In our paper, the study will focus on the relationship between the volatility of the foreign exchange rate, especially the USD/THB exchange rate, and the imports/exports of Thailand to examine how the fluctuating exchange rate would have a direct effect on the changes in the imports and exports of Thailand. Theoretically, if the exchange rate appreciates, the exports of Thailand will decrease because of the higher prices. Foreign purchasers of the countrys products will turn to alternative producers. On the other hand, the imports of Thailand will increase in volume because foreign goods are cheaper, and this will result in a trade deficit. However, a depreciation in the currency exchange rate will provide a good chance for Thailand exporters to export more products. Meanwhile, it will adversely affect the imports because imported goods become much more expensive and, eventually, this will lead to a trade surplus. The traditional volatility of the currency exchange rate also affects the ability to compete with the National Competitive Advantage. In this study, we will pay more attention to the analysis of the relationship between the USD/THB exchange rate and the imports/exports scenario of Thailand, using a model called vine copula, consisting of two types of vine copulas: the canonical vine (C-vine) copula and the drawable vine (D-vine) copula.

Vine copula is a branch of copula which is a probabilistic construction of high-dimensional or multivariate distributions based on bivariate copulas that are building-block so-called pair-copulas. The dependency structure is assigned by the bivariate copulas and the nest set of tree (see Kramer [13]). Moreover, this one is more flexible as we can choose bivariate copula from various families. Presently, vine copula was widely applied to study a pair-copula construction in economics and financial modeling, namely, application of exchange rate or stock index from different countries. Hence, the primary purpose of our study was to investigate the relationship between the USD/THB exchange rate and international trade; if these trivariate relations are quite well, it would be very useful for policy makers. In case of exchange rate volatility, the policy makers would need to take some action to prevent these problems.

The remainder of this article is organized as follows: Section 2 presents literature review. Section 3 presents the definition of C-vine and D-vine copulas and reviews the concept of estimating both the vine copulas and the evaluation of the vine copula

models which were employed in this study. Section 4 demonstrates the empirical findings. Lastly, section 5 presents our conclusion.

2 Review Literature

However, in recent years, many researchers have popularly studied the relationship financial modeling by using vine copulas; however, this method was adopted by very few for studying international trade or economic growth. Most researchers applied other methods, for example, VAR-copula, CGE, Panel method, PML-IV method, etc. Most studies related to economic growth have also confirmed that international trade is crucial and necessary for the growth and development of all countries.

There are very few papers available to investigate on the subject of the relationship of exchange rate with international trade by using vine copula because most papers applied vine copula to study financial modeling. The first paper, Czado et al. [10], clearly explains a pair-copula construction for modeling exchange rate dependence. They employed two possible ways, including bivariate t-copula building blocks of PCC between a pair of exchange rates and the directed acyclic graph (DAG) embedded special PCC to estimate the various exchange rates. Exchange rates were observed for many currencies, including the British pound, US dollar, Malaysia ringgit, Swiss franc, Japanese yen, Danish crona, and Swedish krona. The results of this paper indicate that the US dollar becomes the first node of C-vine and regular vine, while the R-vine specification prefers the DAG specification. Moreover, this indicates that C-vine and R-vine are not distinguishable at $\alpha = 0.5$ and that the C-vine specification is the best fit. Aas et al. [23] clearly proposed a pair-copula construction of the multiple dependences. Four stock indices, including the Norwegian stock index (TOTX), the MSXI world stock index, the Norwegian bond index (BRIX), and the SSBWG hedged bond index were subjects of study in this research. D-vine was applied to the building structure, and then, a students t-copula, which was the best-fitting data set, was used to estimate a pair-copula. Also, Czado et al. [9] had successfully denoted a maximum likelihood estimation of mixed C-vine with application to exchange rate. They have been employing the mixed C-vines model with the maximum likelihood estimation to approximately eight principal time series of US currency exchange rates with different countries such as the European countries, the United Kingdom, Canada, Australia, Brazil, Japan, Switzerland, and India. They found that Gaussian t-copula, Clayton, Gumbel, and Frank were required for the data sets. Furthermore, mixed C-vine (sequential selection without independence test) and Gaussian copula (same as mixed C-vine, but all pair-copulas are Gaussian copulas) are the best chosen models when it comes to the selection. The Voung and Clarke test with Schwarz correction was used to compare the two models.

3 Data and Model Specification

Our paper endeavors to study the relationship between three variables, namely, the USD/THB exchange rate, the imports of Thailand and the exports of Thailand, and whether each pair has a relationship or not, via the vine copula method. While studying the relationship between these multivariate variables, Czado [7] states, the researchers may face many problems, such as different multivariate data structure, non-symmetric dependencies between some pairs of variables, heavy tail dependencies between some pairs of variables, etc. Normally, multivariate variables cannot be modeled with standard parametric distributions such as the Gaussian or multivariate t distribution. On the other hand, the copula approach allows the modeling of dependencies and marginal distributions separately. However, standard multivariate copula models such as the elliptical and Archimedean copulas do not allow for different dependency models between pairs of variables. In order to solve these problems, we have been adopting vine copula which is a part of copula for our study because vine copula can overcome all of these troubles.

3.1 Model

(*a*) Vine Copula

In this section, we present vine copulas that were first introduced by Joe [29] and developed in more detail by Bedford and Cooke([2], [3]) who explained the details about the regular vine (R-vine). This one is a type of vine copulas which can assess any structure to the needs of those who want to study. Kurowicka and Cooke [9] denoted that vines are flexible graphical models for explaining the multivariate copula make-up using a cascade of bivariate copulas, the so-called pair-copulas. In addition, Aas, Czado, Frigessi, and Bakken [23] bring the regular vines (R-vine) to continuous development and proposed the statistical inference techniques for the two classes of regular vines, that is, the canonical vine (C-vine) and the drawable vine (D-vine).

Vines are graphical representations of the so-called pair-copula constructions (PCCs). If we assume that these are three dimensional, we can obviously show an illustration of the PCCs, as shown in Figure 1. Let $X = (X_1, ..., X_3) \sim G$ be with the marginal distribution function G_1, G_2, G_3 and the corresponding densities; we can write

$$g(x_1, x_2, x_3) = g(x_1)g(x_2|x_1)g(x_3|x_1, x_2)$$

By Sklar's Theorem [25], we denote that

$$g(x_2|x_1) = \frac{g(x_1, x_2)}{g_1(x_1)} = \frac{c_{1,2}\Big(G_1(x_1)G_2(x_2)\Big)g_1(x_1)g_2(x_2)}{g_1(x_1)}$$

$$= c_{1,2}\Big(G_1(x_1), G_2(x_2)\Big)g_2(x_2)$$

$$g(x_3|x_1,x_2) = \frac{g(x_2,x_3|x_1)}{g(x_2|x_1)} = \frac{c_{2,3|1}\Big(G(x_2|x_1),G(x_3|x_1)\Big)g(x_2|x_1)g(x_3|x_1)}{g(x_2|x_1)}$$
$$= C_{2,3}\Big(G(x_2|x_1),G(x_3|x_1)\Big)g(x_3|x_1)$$
$$= c_{2,3|1}\Big(G(x_2|x_1),G(x_3|x_1)\Big)c_{1,3}\Big(G_1(x_1),G_3(x_3)\Big)g_3(x_3)$$

where $c_{1,2}, c_{1,3}$, and $c_{2,3|1}$ are the pair-copulas, and $C_{2,3|1}$ is independent of the conditional variable X_1 in facilitating the inference (see Aas et al., [23] and Hobaek et al., [11]). As for d-dimensional, there are $\frac{d(d-1)}{2}$ vine arranges for the pair-copulas and $d-1$ vine arranges for the trees. If these are considered as three dimensional, then, we have 3 pair-copulas and 2 trees. In the case of the C-vines, the first root node is designed by applying a bivariate copula for each pair. Conditioned on this variable, pair-wise dependencies with respect to the second variable are modeled, which is the second root node. Moreover, the trees of the C-vines also have a patterned star structure (see Figure 1, left panel). A root node is chosen in each tree, and all the pair-wise dependencies with respect to this node are modeled conditioned on all the previous root nodes (see Brechmann et.al, [5]). Then the general expression of the C-vines can be given as follows:

$$g(x) = \prod_{k=1}^{d} g_x(x_k) \times \prod_{i=1}^{d-1}\prod_{j=1}^{d-i} c_{i,i+j|1:(i-1)}\Big(G(x_i|x_1,...,x_{i-1}),G(x_{i+j}|x_1,...,x_{i-1})|\theta_{i,i+J|1:(i-1)}\Big)$$

where $g_k, k = 1,,d$ show the marginal densities and $\theta_{i,i+j|1:(i-1)}$ is the parameter(s) of the bivariate copula densities $(c_{i,i+j|1:(i-1)})$ (see Aas et al., [23] and Hobaek et al., [11]).

Identical to this, the D-vines are of regular vine distributions, but there exists no node in any trees which is connected to more than two edges. It is similarly constructed by selecting a specific order of the variable. We assume these to be three dimensional, in the first tree of this one (see Figure 1, right panel). The dependence of the first and the second variables, of the second and the third, and so on, is modelled using pair-copulas in the second tree as well, with conditional dependence of the first and the third, given the second variable (the pair $13|2$). Similarly, the general expression of the D-vines is as follows:

$$g(x) = \prod_{k=1}^{d} g_x(x_k) \times \prod_{i=1}^{d-1}\prod_{j=1}^{d-i} c_{j,j+i|(j+1):(j+i-1)}\Big(G(x_j|x_{j+1},...,x_{j+i-1}),$$
$$G(x_{j+i}|x_{j+1},...,x_{j+i-1})|\theta_{j,j+1|(j=1):(j+i-1)}\Big)$$

(*b*) Estimation of C-vine or D-vine Copulas and Evaluation of Vine Copula Models

From our analysis, we learned somewhat about C-vine and D-vine that each tree node includes bivariate copula pairs. In this part, we progressed toward building a

structure and selecting tools for the bivariate exploratory analysis. Upon filtering copula, it can be seen that there are 31 families, including the Archimedean and elliptical copulas. Each family has different properties; for example, student-t copula is needed for the degree of freedom or some families have lower and upper tail coefficients.

To analyze the bivariate copula selection, first we have to provide the structure of the data. There are various alternatives from which structure can be selected. Many papers applied matrix of the empirical Kendalls tau to select a variable which will become the first node. However, other alternatives can also be used, such as manual selection, using expert knowledge, and observing that which is implied by the structure of the data, for deciding the structure of the C-vine and the D-vine. After we successfully selected some variables to become the first node, pair-copula families were selected for each pair of the variables. There are two methods to analyze the bivariate copula: graphical and analytical tools. In our paper, we used analytical tools to investigate the bivariate copula. Each copula selection was conducted tree by tree, from the first tree, the second tree, and so on, which depended on the specification of the previous tree according to the h-function, as follows (see Brechmann et al. [5]):

$$h(x,v,\theta) = G(x,v) = \frac{\partial C_{x.v}(x,v,\theta)}{\partial v}$$

For BB1, the h-function is

$$h = \left(1+((x^{-\theta}-1)^{\delta}+(v^{-\theta}-1)^{\delta})^{\frac{1}{\delta}}\right)^{-\frac{1}{\theta}-1} \times \left((x^{-\theta}-1)^{\delta}+(v^{-\theta}-1)^{\delta}\right)^{\frac{1}{\delta}-1}(v^{-\theta}-1)^{\delta-1}v^{-\theta-1}$$

where θ is a parameter of the bivariate copula of the joint distribution function of x and v. v always corresponds to the conditioning variable and δ is another parameter of the BB1 copula family. $h^{-}(u,v,\theta)$ is the inverse of the h-function. Thereafter, the preliminary C-vine and D-vine copula models are fitted by proceeding repeatedly tree by tree. In addition, we applied the maximum likelihood estimation (MLE) to estimate the parameter of each pair-copula. Subsequently, as these approximates provide good fits, it is then naturally interesting to maximize the log-likelihood of the vine copula specification for the observation $u = (u_{k,j}), k = 1,,N, j = 1,,d$.

The canonical vine (C-vine) copula log-likelihood with parameter set θ_{CV} is given by

$$l_{CV}(\theta_{CV}|u) = \sum_{k=1}^{N}\sum_{i=1}^{d-1}\sum_{j=1}^{d-j} log[c_{i,i+j|1:(i-1)}(G_{i|1:(i-1)}, G_{i+j|1:(i-1)}|\theta_{i,i+j|1:(i-1)})]$$

where $G_{j|i_1:i_m} := G(u_{k,j}|u_{k,i1},...,u_{k,im})$ and the marginal distribution are uniform, at $g_k(u_k) = 1_{[0,1]}(u_k)$. Note that $G_{j|i_1:i_m}$ depends on the parameters of the pair-copula terms in tree 1 up to tree i_m.

The drawable vine (D-vine) copula log-likelihood with parameter set θ_{DV} is given by

$$l_{DV}(\theta_{DV}|u) = \sum_{k=1}^{N}\sum_{i=1}^{d-1}\sum_{j=1}^{d-j} log[c_{j,j+i|(j+1):(j+i-1)}(G_{j|(j+1):(j+i-1)}, G_{j+i|(j+1):(j+i-1)}|\theta_{j,j+1|(j+1):(j+i-1)})]$$

(*c*) Evaluation of Vine Copula Models

First, we have to choose an appropriate bivariate copula family for the given observations by using the analytical tools; we have to then obtain copula families as well as the parameters of each pair of the C-vine and the D-vine. Next, we have to calculate the log-likelihood of these models for comparison with the log-likelihood of the joint distribution parameter which is obtained from the previous step. The parameter of log-likelihood has to be quite close to the parameter of log-likelihood of the joint distribution (see Brechmann and Schepsmeier, [4]).

Let $u = (u_1^{'}, ..., u_N^{'})$ be the d-dimensional observation with $u_i = (u_{i,1}, ..., u_{i,d})^{'} \in [0,1]^d, i = 1, ..., N$.

The log-likelihood of C-vine copula is given by

$$loglik = l_{CVine}(\theta|u) = \sum_{i=1}^{N}\sum_{j=1}^{d-1}\sum_{k=1}^{d-j} ln[c_{j,j+k|1,...,j-1}]$$

where $c_{j,j+k|1,...,j-1} := c_{j,j+k|1:(j-1)}(G(u_{i,j}|u_{i,1}, ..., u_{i,j-1}),$
$G(u_{i,j+k}, ..., u_{i,j-1})|\theta_{j,j+k|1,...,j-1})$ shows pair-copulas with the parameter $\theta_{j,j+k|1,...,j-1}$.

Identically, the log-likelihood of the d-dimensional D-vine copula is

$$loglik = l_{DVine}(\theta|u) = \sum_{i=1}^{N}\sum_{j=1}^{d-1}\sum_{k=1}^{d-j} ln[c_{k,k+j|k+1,...,k+j-1}]$$

again with pair-copula densities shown by

$c_{k,k+j|k+1,...,k+j-1} := c_{k,k+j|k+1,...,k+j+1}(G(u_{i,k}|u_{i,k+1}, ..., u_{i,k+j-1}),$
$G(u_{i,k+j}|u_{i,k+1}, ..., u_{i,k+j-1})|\theta_{k,k+j|k+1,...,k+j-1})$.

Normally, Akaikes information criterion (AIC) and Bayesian information criterion (BIC) are used to decide whether these models are appropriate or not. For the C-vine and D-vine, there is another function that is used to compare these models to identify the better-fitting vine copula model for the data set. We perform a Vuong test and a Clarke test to compare the two models.

(*d*) Vuong Test

This test was introduced by Vuong [31] and it can be used for comparing non-nested models. Let d_1 and d_2 be the two compared vines which are compared in terms of their densities and with the estimated parameters set as $\hat{\theta}_1$ and $\hat{\theta}_2$. The log difference of their point-wise likelihood $r_i := log[\frac{c_1(u_i|\hat{\theta}_1)}{c_2(u_i|\hat{\theta}_2)}]$ for observations $u_i = \in [0,1], i = 1, ..., N$.

$$statistic := v = \frac{\frac{i}{n}\sum_{i=1}^{N} r_i}{\sqrt{\sum_{i=1}^{N}(r_i - \bar{r})^2}}$$

where v is asymptotically standard normal and the null-hypothesis

$$H_0 : E[r_i] = 0 \forall_i = 1, ..., N$$

Then, we would prefer vine model 1 to vine model 2 at level α if

$$v > \Phi^{-1}(1 - \frac{\alpha}{2}),$$

where Φ^{-1} shows the inverse of the standard normal distribution function. If $v < -\Phi^{-1}(1 - \frac{\alpha}{2})$, then we would select model 2. If $|v| \leq -\Phi^{-1}(1 - \frac{\alpha}{2})$, no decision on which model to choose is possible.

(*e*) Clarke Test

Another option is Clarke test which was proposed by Clarke [32]. The null hypothesis of the statistical indistinguishability of the two models is

$$H_0 = P(r_i > 0) = 0.5 \forall_i = 1, ..., N$$

Then, under statistical equivalence of the two models, the log-likelihood ratios of the single observations are uniformly distributed around 0, and if it is expected that 50% of the log-likelihood ratios are greater than 0, then the statistic is

$$statistic := C = \sum_{i=1}^{N} 1_{(0,\infty)}(r_i),$$

If A is not significantly different from the expected value $N_p = \frac{N}{2}$, then it is an indication that model 1 can be comprehended as statistically equivalent to model 2.

(*f*) Time Varying in the Conditional Copula

In the previous part, we discussed static copula in a clear manner. In this part, we will deliberate on time-varying copula that Patton [11] proposed in his study, parameterizing time variation in the conditional copula. He stated that the function form of the copula remains fixed over the sample whereas the parameters vary according to some evolutional equation. Moreover, it is very difficult to know what factors might influence them to change. Patton [11] applied the Gaussian and SJC copulas to estimate and investigate the upper and lower tail dependence parameters following something akin to a restricted $ARMA(p,q)$ process. According to Pattons paper, it was found that it was an autoregressive term and a forcing variable. In our paper, we introduced specifically the time-varying Gaussian copula, the evolution equation, as

$$\rho_t = \tilde{A}\Big(\omega_p + \beta_{N1} \cdot \rho_{t-1} + \ldots + \beta_{Np} \cdot \rho_{t-p} + \alpha_N \cdot \frac{1}{q} \sum_{j=1}^{q} \Phi^{-1}(u_{t-j}) \cdot \Phi^{-1}(v_{t-j})\Big)$$

where $\tilde{A}(x)$ is the logistic transformation which is determined as note: $\tilde{A}(x) \equiv (1 - e^{-x})(1 + e^{-x})^{-1}$, ρ_t is the correlation coefficient and $\rho_t \in (-1, 1)$ at all times.

3.2 Data

We applied three time series to estimate the vine copulas; these were the monthly USD/THB exchange rates, the import figures, as well as the export data of Thailand, starting August 1997 to December 2012. We made a total of 185 observations. The total data were taken from the Bank of Thailand. To reduce the difficulty of data being non-stationary, each monthly data was converted into log-difference, and then calculated as

$$Y_t = 100 * (log(g_t) - log(g_{t-1}))$$

where g_t stands for the USD/THB exchange rates, the import figures, or the export figures of Thailand at the period of t, respectively.

4 Empirical Results

Before the log-returns were used, they were filtered by the ARMA(1,1)-GARCH(1,1) model via skew-t distributions with the maximum likelihood estimation (MLE). Afterward, we had to test the goodness of fit of each univariate distribution using the Kolmogorov-Smirnov (K-S) test to ascertain whether it was appropriate or not. Moreover, we used Ljung-Box tests to indicate the independence of the standardized residuals. When we tested all process already, standard residuals of each univariate distribution are transformed to approximately i.i.d. uniform data on $[0, 1]$. Czado [10] state, across margins these unit interval variables are dependent and we model their dependence structure using the pair-copula constructions base vine copulas that copula-GARCH have to qualify namely, standard residuals are i.i.d. with $E(\eta_t) = 0$ and $Var(\eta_t) = 1$

4.1 Specifications of C-vine and D-vine Copula Models

Next, we had to build pair-constructions that are marginally uniform. In the last section, we learned that there are many ways to build pair-constructions, for instance, manually, by using expert knowledge, as implied by structure, etc. Normally, it can be observed that the empirical Kendalls tau matrix could be applied for structure building. But in our analysis, we opted to build pair-constructions by choosing the expert knowledge method based on the economics theory because our data was not of the same type that is, some of our data were exchange rates, some were economic growth variables, and so on and so forth. According to the economics theory,

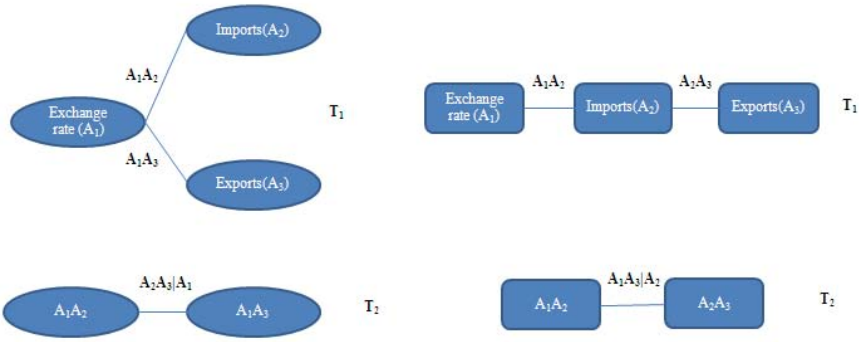


Fig. 1 Pair-constructions of the C-vine (left panel) and the D-vine (right panel) with edge indices

imports and exports depend on the variations in the exchange rate. For the C-vine and the D-vine, the building structure could be drawn as follows:

From Figure 1, in the C-vine model, the USD/THB exchange rate was selected to be the first node of the first tree. This exchange rate determined the exports and the imports in the second root node and the third root node, respectively. Similarly, in the D-vine model, the order of the variable of the first tree had to be chosen after taking the first node of the C-vine as the USD/THB exchange rate, the second and third variables being imports and exports, respectively. Moreover, in the second tree of the D-vine, there is conditional dependence of the USD/THB exchange rate and exports given the imports.

4.2 Estimation of C-vine and D-vine Copula Models

After we built the structure successfully, the next step in the process was the estimation of the C-vine and the D-vine copula models which are ordinarily fitted sequentially by proceeding iteratively tree by tree. All the parameters of the pair-copula were estimated by the maximum likelihood estimation (MLE) method or by inversion of Kendall's τ. Selection results for the C-vine and the D-vine are summarized in Table 1 and Table 2, respectively. The corresponding C-vine and D-vine tree representations are given in Figure 2. Subsequently, we employed the following abbreviations: A_1 for USD/THB exchange rate, A_2 for the import figures, and A_3 for the export figures of Thailand.

Table 1 shows the results of the pair-constructions of the canonical vine, or the C-vine. We found that Gaussian, Rotated Joe, and BB1 are appropriate bivariate copula families from a set of possible copula families for A_1A_2, A_1A_3, and $A_2A_3|A_1$, respectively. In addition, each pair-copula construction proposed appropriate copula families, and both the parameters estimated by the MLE method concluded that θ is the parameter of the bivariate copula and the standard error; also, the Akaike's

Table 1 Results of Pair-constructions of Canonical Vine, or C-vine

Variable	Copula Family	θ	δ	Pair AIC	Pair BIC	Kendall's τ
A_1A_2	Gaussian	0.1267 (0.0722)	0.0000 (0.0000)	-0.9522	2.2682	0.0809
A_1A_3	Rotated Joe (90)	-0.9045 (0.0716)	0.0000 (0.0000)	-1.0737	2.1466	-0.0578
$A_2A_3\|A_1$	BB1	0.2585 (0.1331)	1.1577 (0.0817)	-25.1640	-18.7233	0.2350
Log-likelihood	17.5955					
CDVineLogLik	17.5950					

Source: Computation.
Note: the element in the brackets correspond to the standard errors.

information criterion (AIC) and the Bayesian information criterion (BIC) statistics of the pair-constructions and Kendall's τ statistic for this model are given in this table. For example, the third row indicates that BB1 is the appropriate-fitting copula model for $A_2A_3|A_1$. There exists two parameters of this, pair-wise, which are 0.2585 and 1.1577, and the AIC and BIC statistics of this one are 25.1640 and -18.7233, respectively. The Kendall's τ is equal to 0.2350. Moreover, an extremely important result is that the sequential log-likelihood and the log-likelihood of joint maximization are 17.5950 and 17.5955, respectively, which means that the C-vine was appropriate fitting because the sequential log-likelihood was quite close to the one obtained by joint maximization.

Identically, from the results of the drawable vine, or the D-vine, in Table 2, each pair-copula construction also indicated suitable copula families, the parameters of the pair-wise, the AIC and BIC statistics, and Kendall's statistic. This shows that Gaussian is an appropriate choice for A_1A_2 and A_2A_3 while Rotate Joe is the appropriate copula family for $A_1A_3|A_2$. Like in the C-vine, we found that the sequential log-likelihood was quite close to the one obtained by joint maximization. Thus, both the log-likelihoods establish that the D-vine model is an appropriate fit, just like the C-vine model.

Normally, two models are compared to find out which model would fit. We usually employ the Akaike's information criterion (AIC) and the Bayesian information criterion (BIC) to justify that one of the two models is better. If the AIC and BIC statistics of any model is the least, it means that that model which has the least AIC and BIC would be the suitable fit. Table 3 gives the statistics of the AIC and the BIC of the C-vine and the D-vine. From this table, it can be seen that the AIC and BIC statistics of the C-vine are -27.1900 and -14.3085, respectively. The AIC and BIC statistics of the D-vine are -27.3849 and -17.7239, respectively. It was found that the statistic values of the D-vine are less than those of the C-vine; this

Table 2 Results of Pair-constructions of Drawable vine, or D-vine

Variable	Copula Family	θ	δ	Pair AIC	Pair BIC	Kendall's τ
A_1A_2	Gaussian	0.1350 (0.0722)	0.0000 (0.0000)	-0.9522	2.2682	0.0809
A_2A_3	Gaussian	0.3524 (0.0603)	0.0000 (0.0000)	-23.7302	-20.5098	0.2360
$A_1A_3\|A_2$	Rotated Joe (90)	-0.9389 (0.0761)	0.00000 (0.0000)	-2.7025	0.5178	-0.0730
Log-likelihood	16.7109					
CDVineLogLik	16.6925					

Source: Computation.
Note: the element in the brackets correspond to the standard errors.

Table 3 Akaikes Information Criterion (AIC) and Bayesian Information Criterion (BIC) Statistics of C-vine and D-vine

Statistic	C-vine	D-vine
Akaike's Information Criterion (AIC) statistic	-27.1900	-27.3849
Bayesian Information Criterion (BIC) statistic	-14.3085	-17.7239

Source: Computation.

Table 4 Results of Comparison of Two Non-nested Parametric Models Using Vuong and Clarke Tests

Test	tatistic	statistic Akaike	statistic Schwarz	p-value	p-value Akaike	p-value Schwarz
Vuong Test	0.5281	-0.0688	-1.0300	0.5975	0.9451	0.3030
Clarke Test	106	96	80	0.0564	0.6592	0.0774

Source: Computation.

means that the D-vine is the preferred model, and so, the D-vine copula is the best-fitting model. In addition to applying the AIC and BIC statistics to compare the two models for multivariate uniform, Vuong and Clarke tests can also be used to compare the two models.

Table 4 shows the results of the Vuong and Clarke tests. According to this table, the two vines specification cannot be distinguished statistically at $\alpha = 0.05$. We used both the tests to investigate whether the two vines are distinguishable or not. If the p-values of the Vuong and Clarke tests are greater than 0.05, it means that

we cannot distinguish between these two vines statistically. From the above table, it is clear that the p-values of the Vuong and Clarke tests are 0.5975 and 0.0564, respectively, which are larger than 0.05, thus confirming that these vines are not distinguishable.

4.3 Estimation Using Time-Varying Gaussian Copula of All Pair Constructions

In this estimation, we focused on time-varying copula, following the findings of Patton [11]. He presented time-varying copula which contained time-varying Gaussian copula, time-varying rotated Gumbel copula, and time-varying SJC copula. In our study, we used time-varying Gaussian copula to estimate all the pair constructions of both the vines because the various static pair constructions used by us were efficient fitted by the Gaussian copula. From the estimated results given in Table 5, we can safely assume that our main pair constructions of the time-varying Gaussian copula are appropriate and fit better than the static copula because the AIC and BIC statistics of most pairs in the time-varying Gaussian copula are smaller than those in the static copula, except $A_1A_3|A_2$. However, it needs to be mentioned that this study still provided conflicting results.

Table 5 Estimated Results of Time-varying Gaussian Copula

Pairs	parameter	ω	α	β	AIC	BIC
A_1A_2	ρ	0.4893	-0.7022	-1.2682	-3.4557	-3.4383
	Std error	0.0145	0.0826	0.0607		
A_1A_3	ρ	-0.1617	-1.4383	-2.0713	-3.6946	-3.6772
	Std error	0.0068	0.0529	0.0048		
$A_2A_3\|A_1$	ρ	1.5810	-0.0183	0.0377	-58.9220	-58.9202
	Std error	0.0588	0.0150	0.0596		
A_2A_3	ρ	0.5554	-0.4590	1.1014	-30.5995	-30.5821
	Std error	0.0262	0.0225	0.0509		
$A_1A_3\|A_2$	ρ	0.1837	-1.4275	-1.5259	-1.5350	-1.5176
	Std error	0.0182	0.0792	0.0294		

Source: Computation.

5 Conclusion

In this paper, we discussed an application in multivariate copulas, including the use of the C-vine and D-vine copulas based on pair-copula constructions. Before the application of each pair-copula construction for the multivariate copula, we applied $ARMA(1,1)$-$GARCH(1,1)$ in order to eliminate the time dependence in each of the margins of this time series. In addition, the Ljung-Box test was adopted to

verify whether all marginal distributions are independent and identically distributed (i.i.d) or not. In addition to this, we effected appropriate uniform margins in the standardized residual transforming $[0, 1]$. After that, we built the C-vine and D-vine structures. For the C-vine, there exists Gaussian, Rotated Joe, and BB1 as suitable bivariate copula families forA_1A_2, A_1A_3, and $A_2A_3|A_1$, respectively. As for the D-vine, Gaussian is an appropriate choice for A_1A_2 and A_2A_3, while Rotated Joe is an appropriate copula family for $A_1A_3|A_2$. After both the vine copulas were estimated, many parameters were successfully obtained. We needed to approximate the AIC and BIC statistics in order to compare between the C-vine and D-vine models. Moreover, it was found that the sequential log-likelihood of both the models is quite close to the one obtained through joint maximization; this means that both the vine models are appropriate fit models.

However, to confirm that the two models are not distinguishable, we applied the Vuong and Clarke test to verify their suitability for a non-nested model. From the results, it was evident that the C-vine and the D-vine were not corroboratively distinguishable. In addition, it was found that it could be safely assumed that our pair constructions of the time-varying Gaussian copula were more appropriate and better fitting than those of the static copula. This is because the values of the AIC and BIC statistics of the time-varying Gaussian copula were less than those of the static copula. In future studies, we plan to apply a regular (R-vine) for better interpretation.

Acknowledgments. We acknowledge the financial support from the Prince of Songkla University Scholarship for Charkorn Praprom's PhD study. We are grateful to Dr. Chanagun Chitmanat for reviewing the manuscript.

References

1. Aas, K., Czado, C., Frigessi, A., Bakken, H.: Pair-copula constructions of multiple dependence. Insurance: Mathematics and Economics 44(2), 182–198 (2009)
2. Bedford, T., Cooke, R.: Probability density decomposition for conditionally dependent random variable modeled by vine. Annals of Mathematics and Artificial Intelligence 32, 245–268 (2001)
3. Bedford, T., Cooke, R.: Vine-a new graphical model for dependent random variables. Annual of Statistics 10, 1031–1068 (2002)
4. Brechmann, E.C., Schepsmeier, U.: Package CDVine (2012), `http://cran.r-project.org/web/packages/CDVine/index.html`
5. Brechmann, E.C., Schepsmeier, U.: Modeling Dependence with C- and D-Vine Copulas: The R Package CDVine. Journal of Statistical Software 52(3), 1–27 (2013)
6. Clarke, K.A.: A simple Distribution-Free test for Nonnested Model Selection. Political Analysis 15, 347–363 (2007)
7. Czado, C.: The world of vines (2011), `http://www-m4.ma.tum.de/fileadmin/w00bdb/www/veranstaltungen/vine_world.pdf`
8. Czado, C., Schepsmeier, U., Min, A.: Maximum likelihood estimation of mixed C-vines with application to exchange rates. Statistical Modelling 12(3), 229–255 (2012)
9. Czado, C., Schepsmeier, U., Min, A.: A Mixed Copula Model for Insurance Claims and Claim Sizes. Scandinavian Actuarial Journal 4, 278–305 (2012)

10. Czado, C., Min, A., Baumann, T., Dakovic, R.: Pair-copula constructions for modeling exchange rate dependence (2008), http://www.m4.ma.tum.de/Papers/index.html
11. Hobaek, H.I., Aas, K., Frigessi, A.: On the simplified pair-copula constructions. Bernoulli 19, 462–491 (2010)
12. Joe, H.: Families of m-variate distributions with given margin and m(m-1)/2 bivariate dependence parameters. In: Ruschendorf, L., Schweizer, B., Taylor, M.D. (eds.) Distributions with Fixed Marginals and Related Topics, pp. 120–141 (1996)
13. Kramer, N., Schepsmeier, U.: Introduction to vine copulas (2011), http://www-m4.ma.tum.de/fileadmin/w00bdb/www/veranstaltungen/Vines.pdf
14. Kurowicka, D., Cooke, R.M.: Uncertainty Analysis with High Dimentional Dependence Modelling. John Wiley & Sons, Chichester (2006)
15. Nikoloulopoulos, A.K., Joe, H., Li, H.: Vine copulas with asymmetric tail dependence and applications to financial return data. Computational Statistics and Data Analysis 56(11), 3659–3673 (2012)
16. Patton, A.J.: Modelling Asymmetric Exchange Rate Dependence. International Economic Review 47, 527–556 (2006)
17. Schepsmeier, U., Stoeber, J., Brechmann, E.C.: Package VinceCopula (2013), http://cran.r-project.org/web/packages/VineCopula/index.html
18. Sklar, A.: Fonctions de repartition a n dimentions et leurs marges. Publ. Inst. Statist. Univ. Paris 8, 229–231 (1959)
19. Vuong, Q.H.: Ratio tests for model selection and non-nested hypothesis. Econometrica 57(2), 307–333 (1989)

A Vine Copula Approach for Analyzing Financial Risk and Co-movement of the Indonesian, Philippine and Thailand Stock Markets

Songsak Sriboonchitta, Jianxu Liu, Vladik Kreinovich, and Hung T. Nguyen

Abstract. This paper aims at analyzing the financial risk and co-movement of stock markets in three countries: Indonesia, Philippine and Thailand. It consists of analyzing the conditional volatility and test the leverage effect in the stock markets of the three countries. To capture the pairwise and conditional dependence between the variables, we use the method of vine copulas. In addition, we illustrate the computations of the value at risk and the expected shortfall using Monte Carlo simulation with copula based GJR-GARCH model. The empirical evidence shows that all the leverage effects add much to the capacity for explanation of the three stock returns, and that the D-vine structure is more appropriate than the C-vine one for describing the dependence of the three stock markets. In addition, the value at risk and ES provide the evidence to confirm that the portfolio may avoid risk in significant measure.

1 Introduction

Southeast Asia has emerged as the new Asian tiger at a time when China's economic growth is on the wane. Even if the global economy takes a downturn, as before, the IMF has constantly forecast that the economic growth will be about 6.1% in 2013 for Indonesia, Malaysia, the Philippines, Thailand, and Vietnam.

Songsak Sriboonchitta · Jianxu Liu
Faculty of Economics, Chiang Mai University, Chiang Mai 50200 Thailand
e-mail: songsakecon@gmail.com, liujianxu1984@163.com

Vladik Kreinovich
Computer Science Department, University of Texas at El Paso, Texas, USA
e-mail: vladik@utep.edu

Hung T. Nguyen
Department of Mathematical Sciences,
New Mexico State University, New Mexico, USA
e-mail: hunguyen@nmsu.edu

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_16,

Regardless of economic downturn or economic prosperity, the Southeast Asian countries maintain consistency for instance, the GDPs of Thailand, the Philippines, and Indonesia decreased by 40.0%, 83.4%, and 37.3%, respectively, during the Southeast Asian financial crisis. Even though in recent years, the growth in the Southeast Asian countries has been impressive for example, the GDPs of Thailand, the Philippines, and Indonesia was on a year-on-year increase of 5.9%, 6.6%, and 6.1%, respectively, in 2012. Southeast Asias booming economy has also led to the prosperity of the stock market. In 2012, the Philippine benchmark stock index rose 29.8%, Indonesias stock market rose 12.6%, and Thailands stock market was up 30%. In addition, the Thailand SET Index earnings per share forecast growth of up to 24%, and return on equity of up to 19.2%, higher than the 16.9% of India and 16.8% of China. Thus, the Southeast Asian countries have been growing according to, or above, expectations; in particular, Thailand, Indonesia, and the Philippines have been very strong over the past year, and they displayed a wave of strong co-movement and interdependence. Thus, it is evident that the study of the Southeast Asian stock market is of practical significance for investors, businesses, and governments.

In addition, a detailed survey of the ASEAN stock market is relevant because of the increased economic cooperation in accordance with the ASEAN agreement, the successful financial reforms, the current booming economy, and the distinguished structure of the emerging stock markets. Moreover, there is a dearth of research material and literature that focus on their dependence structure. A noteworthy exception to this is the study done by Sharma and Wongbangpo [1] who analyze the degree of the long-term and short-term co-movements in the stock markets of the five ASEAN countries, Indonesia, Malaysia, Singapore, Thailand, and the Philippines. Their results revealed that there exists a long-run relationship among the stock markets of Indonesia, Malaysia, Singapore, and Thailand, but the Philippine market does not share this relationship. Of course, in recent years, there has emerged some literature that focuses on the dependence patterns of the Asian stock market, as well. For example, Ning and Wirjanto [2] used the copula approach to examine the extreme returnvolume relationship in six countries, Taiwan, Singapore, Malaysia, Thailand, Indonesia, and Korea. The study applied Clayton, survival Clayton, Frank and Gumbel copulas to fit asymmetric returnvolume dependence at extremes for these markets. Lim et al. [3] applied a battery of nonlinearity tests to re-examine the weak-form efficiency of 10 emerging Asian stock markets that include China, India, Indonesia, South Korea, Malaysia, Pakistan, the Philippines, Sri Lanka, Taiwan, and Thailand. Sharma [4] studied the correlation between emerging Asian markets and the United States. The study found that the linear positive correlation between Malaysia and the Philippines reaches up to 0.976. Although there are few researchers who studied the co-movement or correlations between ASEAN countries, they focus on pair dependences (see Sharma [4], Ning and Wirjanto [2]) and the degree of the long-term and short-term co-movement (see Sharma and Wongbangpo [1]). Or more accurately, there are not studies of multivariate dependence structure and tail dependence in ASEAN stock market so far to date.

Since Bedford and Cooke [6] [7] introduced pair-copula construction (PCC) of mutivariate distribution, vine copulas have been widely developed and used in econometrics and finance. Especially, Aas et al. [12] developed standard maximum likelihood (ML) estimation for Canonical vine (C-vine) and Drawable vine (D-vine) copulas, where the challenge was to provide a good starting point for the required high dimensional optimization. Compared vine copulas with standard multivariate copulas, standard multivariate copulas, such as multivariate normal and multivariate-t copulas, become inflexible in high dimensions because of never allowing for different dependency structures between pairs of variables. On the contrast, vine approach is more flexible, as we can select bivariate copulas from a wide range of (parametric) families. Additionally, copula approach may capture the upper and lower tail dependence, which is more precise to calculate value at risk (VaR) and expected shortfall (ES).

This paper applies the vine copula approach to study the stock return co-movement and tail dependence, especially to shed new light on the dependence between three countries: Indonesia, Philippine and Thailand. Moreover, on the basis of this approach, we investigate the value at risk (VaR) and the expected shortfalls (ES). The main contributions of the paper are as follows: (1) This paper describes the conditional volatility and the leverage effect in Indonesia, the Philippines, and Thailand; (2) The study makes use of vine copulas to analyze the co-movement and conditional dependences, and tail dependences; (3) The paper combines vine copula with the Monte Carlo simulation method, thus enabling the estimation of value at risk and expected shortfall.

The paper is organized as follows: Section 2 describes the methodology used in the investigation. Section 3 discusses the empirical results. Section 4 provides the results of economic application for risk measure. Lastly, Section 5 offers conclusions.

2 Methodology

Copulas are functions that join multivariate marginal distribution functions to form joint distribution functions. If $X = (X_1, X_2, \ldots, X_n)$ is a random vector with joint distribution function H and marginal distributions $F_1, F_2, \ldots, F_n$, then there exists a copula C, such that

$$H(x_1, x_2, \ldots, x_n) = C(F(x_1), F(x_2), \ldots, F(x_n)) \tag{1}$$

In the light of formula (1), the copula function can be expressed as:

$$C(u_1, u_2, \ldots, u_n) = H(F^{-1}(u_1), F^{-1}(u_2), \ldots, F^{-1}(u_n)) \tag{2}$$

So, we need to find the appropriate marginal distributions for the copula model. Taking into consideration the characteristics of stock returns, which are generally non-normal, volatility clustering, and asymmetric, we employ the

Glosten-Jagannathan-Runkle (GJR) model with the skewed student-t and skewed generalized error distribution (SGED) to capture the time-varying volatility and leverage effect, and to fit the marginal distributions for the copula model.

2.1 A GJR Model for Marginal Distributions

Glosten, Jagannathan, and Runkle [14] extended the Generalized Auto-Regressive Conditional Heteroskedasticity (GARCH) model. Subsequently, it was named GJR-GARCH model; it includes leverage terms for modeling asymmetric volatility clustering. The form of the ARMA (P, Q)-GJR (K, L) model can be expressed as

$$r_t = c + \sum_{i=1}^{p} \phi_i r_{t-i} + \sum_{i=1}^{q} \psi_i \varepsilon_{t-i} + \varepsilon_i \tag{3}$$

$$\varepsilon_t = h_t \eta_t \tag{4}$$

$$h_t^2 = \omega + \sum_{i=1}^{k} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{k} \gamma_i I[\varepsilon_{t-i} < 0] \varepsilon_{t-i}^2 + \sum_{i=1}^{l} \beta_i h_{t-i}^2 \tag{5}$$

where $\sum_{i=1}^{p} \phi_i < 1, \omega > 0, \alpha_i >= 0, \beta_i >= 0$, $\alpha_i + \gamma_i >= 0$, and $\sum_{i=1}^{k} \alpha_i + \sum_{i=1}^{l} \beta_i + \frac{1}{2} \sum_{i=1}^{k} \gamma_i < 1$. The formulas (3) and (5) are called mean equation and variance equation, respectively; the formula (4) describes the residuals ε_t is consist of standard variance h_t and standardized residuals η_t; the leverage coefficient γ_j is applied to negative standardized residuals, giving negative changes additional weight. In addition, the standardized residuals are assumed to be the skewed student-t or skewed generalized error distribution in this study, and the cumulative distributions of standardized residuals are formed to plug into copula model.

2.2 Vine Copulas

Regarding vine copulas, it is worth taking a moment to understand the development process. Joe and Hu [5] gave the first pair-copula construction (PCC) of a multivariate copula, the construction of which is dependent on distribution functions. Bedford and Cooke [6] [7]expressed these constructions in terms of densities, and organized these constructions in a graphical way involving a sequence of nested trees, which are called regular vines. They also proposed two subclasses of the PCC: we call them C-vine and D-vine copulas. Furthermore, C-vine and D-vine copulas have been made use of in analyzing the conditional dependence for finance asset return, as they are more flexible than some multivariate copulas. For example, multivariate normal copula does not have tail dependence; multivariate t-copula has only a single degree of freedom parameter and symmetric tail dependence, while the nested Archimedian copulas and Hierarchical Archimedian copulas require additional parameter restrictions and thus result in reduced flexibility for modeling

dependence structures (see Joe [15]; Savu and Trede [16]; Czado [17]). Various studies demonstrate the properties, classifications, structures, and merits of vine copulas (Nikoloulopoulos et al. [8]; Kurowicka and Cooke [9]; Joe et al. [10]; Joe [11]; Aas et al. [12]).

Compared to the above-mentioned multivariate copulas, the vine copulas are more flexible in high dimensions because vine copulas allow for different dependency structures between the pairs of variables. C-vine and D-vine copulas are subclasses of the vine copula. They possess all the characteristics of the vine copula, and find applications far and wide. Let us consider the three-dimensional structures of the C-vine and D-vine copulas, the trivariate distribution, and the density function, which can be expressed as

$$F_{123}(x_1,x_2,x_3) = \int_{-\infty}^{x_1} C_{23|1}(F_{2|1}(x_{2|z}),F_{3|1}(x_{3|1}))dF_1(z) \tag{6}$$

$$f_{123}(x_1,x_2,x_3) = c_{12}(F_1,F_2) \times c_{13}(F_1,F_3) \times c_{23|1}(F_{2|1},F_{3|1})) \times \prod_{i=1}^{3} f_i(x_i) \tag{7}$$

and

$$F_{123}(x_1,x_2,x_3) = \int_{-\infty}^{x_2} C_{13|2}(F_{1|2}(x_{1|z}),F_{3|2}(x_{3|z}))dF_2(z) \tag{8}$$

$$f_{123}(x_1,x_2,x_3) = c_{12}(F_1,F_2)) \times c_{23}(F_2,F_3) \times c_{13|2}(F_{1|2},F_{3|2})) \times \prod_{i=1}^{3} f_i(x_i) \tag{9}$$

respectively. The formulas (6) and (7) reflect the structure of the three-dimensional C-vine copula, and the formulas (8) and (9) reflect that of the D-vine copula. In formulas (6) and (7), $C_{23|1}(\cdot,\cdot)$ and $C_{13|2}(\cdot,\cdot)$ are the dependency structure of the bivariate conditional distribution, while $c_{ij}(\cdot,\cdot)$ is a bivariate copula density in formulas (7) and (9). The marginal conditional distribution in the C-vine and D-vine is in the form $F(r_t \mid \upsilon)$, which can be written as

$$F(r_t \mid \upsilon) = \frac{\partial C_{r,\upsilon_j|\upsilon_{-j}}(F(r \mid \upsilon_{-j}),F(\upsilon_j \mid \upsilon_{-j}))}{\partial F(\upsilon_j \mid \upsilon_{-j})} \tag{10}$$

where $C_{r,\upsilon_j|\upsilon_{-j}}$ is the dependency structure of the bivariate conditional distribution of r and υ_j conditioned on υ_{-j}, and the vector υ_{-j} is the vector υ excluding the component υ_j (see Aas et al. [12]). For a univariate υ, we use the function $h(r,\upsilon;\theta)$ to represent the conditional distribution function when r and υ are uniform, i.e. $f(r) = f(\upsilon) = 1$, $F(r) = r$ and $F(\upsilon) = \upsilon$. This special marginal conditional distribution is given by

$$h(r,\upsilon;\theta) = F(r|\upsilon) = \frac{\partial C_{r,\upsilon}(r,\upsilon)}{\partial \upsilon} \tag{11}$$

where θ is the parameter set of $C_{r,v}$. We employ different methods to order the sequences of variables in the C-vine and D-vine models. For C-vine, we calculate the sum of empirical Kendall's tau $S_\tau^i = \sum_{j=1,i\neq j}^{n} \tau_{i,j}$ for each variable i, and select

the maximum one as the first variable. After that, we record the remainder of the variables and repeat the process of calculating the sum of Kendall's tau, thus finding out the second and third variables. For example, there are three variables in our study. So,$S_\tau^1 = \tau_{12} + \tau_{13}$, $S_\tau^2 = \tau_{21} + \tau_{23}$ and $S_\tau^3 = \tau_{31} + \tau_{32}$, if S_τ^2 is the biggest value, then the order should be 2, 1, 3 or 2, 3, 1. For D-vine, we determine the order that satisfies the maximization of the sum of empirical Kendall's tau $S_\tau = \sum_{i=1}^{n-1} \tau_{i,i+1}$, e.g., the S_τ of the order 2, 1, 3 is the biggest, then the preferable order should be 2, 1, 3 or 3, 1, 2.

2.3 Parameter Estimation Method

Generally, we use the two-stage estimation method that is called inference function margins (IFM) to estimate our model. This point means that we first estimate GJR-GARCH model thereby getting the marginal distributions, and then plug the marginal distributions into copula model for estimated parameters of vine copulas. Joe [15] [18] showed that this estimator is close to and asymptotically efficient to the maximum likelihood estimator under some regularity conditions. Hence, the two-stage estimation method can efficiently compute the estimator without losing any real information. In the process of parameter estimation of vine copulas, we turn to sequential maximum likelihood estimation method for obtaining initial values of vine copulas, and then use maximum likelihood estimation to estimate the parameters of C- and D-vine copulas. Aas et al. [12], Czado et al. [13] introduced detailed calculate process. A brief process of sequential maximum likelihood estimation can be described as follows.

First, using maximum likelihood estimation to estimate parameters of each non-conditional copula; second, computing observations by using conditional distribution function (formula (11)) and known non-conditional copulas in the first step; third, we estimate the parameters of the copulas conditional on one variable; fourth, computing observations for copulas given two variables by using formula (10); at last, we estimate copulas given two variables using observations from the fourth step. Through these five steps we can get initial values of 4 dimensional vine copulas. If there are more 4 dimensional variables, observations may be gotten by using formula (10) again. We only use the first three steps for getting starting values of each copula in our study.

In this paper, we use Gaussian copula, T copula, Clayton copula, Frank copula, Gumbel copula, Joe copula, BB1 copula, BB6 copula, BB7 copula, BB8 copula, and the rotate copulas to analyze the co-movement. Further details regarding this, which include their properties and characteristics, are discussed in Liu and Sriboonchitta [19], Sriboonchitta et al. [20] and Brechmann and Schepsmeier [21]. We should note is that this study applies Akaike information criterion (AIC) and Bayesian information criterion (BIC) to select a fitting pair-copula family, where both information criteria correspond to the results of sequential maximum likelihood estimation.

3 Empirical Results

We investigate, in this, study, the interactions between three major stock market indices, namely, the Philippine SE (Composite Index in the Philippines), Jakarta SE (Composite Index in Indonesia), and SET (SET Index in Thailand). Our sample covers the period from January 2, 2008, to April 30, 2013. The index returns are calculated by using the differences between the logarithms of the close prices of each index.

The data description and statistics for three index returns are detailed in Table 1. Obviously, the three series are very similar. They all have heavy tails, are skewed to the left, especially the Philippines, and have kurtosis greater than three. In addition, they do not follow normal distribution. So we assume that the margins are skewed student-t and skewed GED, which are appropriate.

Table 1 Data Description and Statistics on Daily Returns

	Indonesia	Philippines	Thailand
Mean	0.0005	0.0006	0.0005
Median	0.0013	0.0008	0.0013
Maximum	0.1032	0.0706	0.0861
Minimum	-0.1095	-0.1309	-0.1109
Std. Dev.	0.0169	0.0147	0.0153
Skewness	-0.5488	-0.9897	-0.4079
Kurtosis	11.0660	11.6427	9.2490
Jarque-Bera	3249	3855	1947
Probability	0.0000	0.0000	0.0000

Table 2 shows the results of the marginal assumption of the skewed student-t distribution performed with the GJR-GARCH model for the three stock returns. First, it can be concluded that all the leverage effects add much to the capacity for explanation of the three stock returns, since each leverage effect parameter γ is significant. Second, this paper calculates the AIC and BIC when the margin is the skewed GED, and the AIC and BIC are -5.9702 and -5.9357, -5.8278 and -5.8020, -5.8839 and -5.8537, respectively, for the Philippines, Indonesia, and Thailand. When compared with the skewed student-t distributions assumption, the AIC and BIC are smaller, as shown in Table 2. Therefore, the GJR-GARCH model with the skewed student-t marginal distribution is the better performing in terms of AIC and BIC.

There exists a precondition for using any copula, which is that the marginal distribution must be uniform (0, 1); if it is not, the wrongly specified model for the marginal distribution may cause incorrect fit copulas. We use Box-Ljung and Kolmogorov-Smirnov (KS) tests to test the validity of the models, and the test results obtained are given in Table 3. None of the KS tests rejects the null hypothesis, and at 5% level, none of the Box-Ljung tests rejects the null hypothesis. Therefore, it can be clearly seen that all the series satisfy the condition of iid uniformity (0, 1).

Table 2 Results of ARMA-GARCH Model

	Indonesia		Philippines		Thailand
C	0.0007*	—	—	—	—
	(0.0004)				
Ar1	0.1122***	—	—	Ar1	0.0348
	(0.0311)				(0.0287)
ω	0.831e-05***	ω	0.7e-05**	ω	0.7e-05**
	(0.3e-06)		(0.2e-05)		(0.2e-05)
α	0.0601*	α	0.0435*	α	0.0475**
	(0.0254)		(0.0193)		(0.0182)
β	0.8185***	β	0.8496***	β	0.8454***
	(0.0341)		(0.0283)		(0.0304)
γ	0.1675***	γ	0.1847***	γ	0.1645***
	(0.0472)		(0.0529)		(0.0482)
Skew	0.9433***	Skew	0.8509***	Skew	0.9018***
	(0.0412)		(0.0320)		(0.0370)
ν	7.1698***	ν	5.0109***	ν	7.2489***
	(1.3808)		(0.7243)		(1.4606)
LM-test	0.3958	LM-test	0.8162	LM-test	0.6271
LogL	3527.4720	LogL	3445.0980	LogL	3473.1870
AIC	-5.9804	AIC	-5.8438	AIC	-5.8899
BIC	-5.9459	BIC	-5.8180	BIC	-5.8597

Note: Signif. codes are as follows: 0 *** 0.001 ** 0.01 * 0.05 0.1. The numbers in the parentheses are the standard deviations.

In the light of the maximum value of the empirical Kendall's tau, the sequence for the C-vine copula is Indonesia, the Philippines, and Thailand, and the sequence for the D-vine copula is Thailand, Indonesia, and Philippines. Thus, we see that C-vine and D-vine have the same structure, both of which calculate the dependence between the Philippines and Thailand, conditional to Indonesia. Since there are only three variables, it is easy to implement, and comprehensive analysis is possible to study the dependences conditional to each variable. Therefore, we use C-vine to estimate the dependence conditional to Indonesia under maximum empirical Kendall's tau, and for others, we make use of D-vine. Table 4 and Table 5 present the estimated parameters of the C-vine and D-vine copulas, respectively. According to the minimum AIC and BIC principle, the optimal choices of the C-vine copula are BB1, Survival BB1, and BB7 copula, in that order, while the same for the D-vine copula are Survival BB1, BB1, and T copula when the selected in the order Indonesia, Thailand, and the Philippines; the other best choices of the D-vine copula are survival BB1, BB1, and T copula. First and foremost, it is evident that the D-vine structure for Thailand is more appropriate than the C-vine one because the sum values of the AIC and BIC are the smallest for D-vine. Second, all the market pairs have significant co-movement and tail dependence especially so for the Indonesian and Thailand markets which possess the greatest dependence, which includes

Table 3 KS Test for Uniform and Box-Ljung Test for Autocorrelation

KS Test			
	Statistic	P value	Hypothesis
$u_{1,t}$	0.0167	0.8969	0 (acceptance)
$u_{2,t}$	0.0239	0.5099	0 (acceptance)
$u_{3,t}$	0.0330	0.1538	0 (acceptance)
Box-Ljung Test			
	Moments	X-squared	P-value
$u_{1,t}$	First moment	5.5303	0.3546
	Second moment	7.2354	0.2037
	Third moment	5.8187	0.3243
	Fourth moment	5.8543	0.3207
$u_{2,t}$	First moment	10.5864	0.0602
	Second moment	2.5125	0.7746
	Third moment	8.7818	0.1181
	Fourth moment	1.0282	0.9603
$u_{3,t}$	First moment	2.4138	0.7894
	Second moment	9.4736	0.0916
	Third moment	10.5063	0.0621
	Fourth moment	9.1190	0.1044

Note: $u_{1,t} = F_{skt}(x_{phi,t})$, $u_{2,t} = F_{skt}(x_{indo,t})$, and $u_{3,t} = F_{skt}(x_{thai,t})$

their upper tail (0.6013) and lower tail (0.3369), among these three country markets. Third, the Kendall's tau of $C_{PT|I}$ and $C_{T,P}$ are 0.1147 and 0.2709, and their upper tail and lower tail dependence are 0.1234 and 0.0080, and 0.2035 and 0.1591, respectively. So, if the Indonesian market is given as the condition, the Kendall's tau falls by 57.66%; the lower tail dependence almost becomes independent, while the upper tail dependence decreases 39.36%. In addition, if we compare $C_{I,T}$ with $C_{IT|P}$, the dependence structure can be observed to undergo a change, when the Philippine market is given as the condition. Moreover, the Philippine market has been seen to have a more profound effect on the tail dependence of Indonesia and Thailand. Last, when the Philippine market is given as the condition, the lower and upper tail dependences between the Thailand and Indonesian markets are seen to become symmetric and tiny. From the above-mentioned results, we can conclude that the information of Indonesia stock market has the effective influence to the lower dependence between Philippine and Thailand, which means the information make investors reduce the probability of high loss simultaneously. On the contrast, the information of Philippine stock market contributes to reduce the possibilities of high loss and profitability at the same time. The information of Thailand plays the same role as Philippines.

Table 4 Results of C-vine Copulas and Kendall's tau

Copulas	parameters	standard error	Lower and upper tail dependence	Kendall'tau	AIC	BIC
BB1 ($C_{I,P}$)	0.3164***	0.0563	0.1574	0.2712	-245.1503	-235.0088
	1.1847***	0.0337	0.2049			
Survival BB1($C_{I,T}$)	0.2573***	0.0555	0.3369	0.3498	-389.3588	-379.2173
	1.3627***	0.0448	0.6013			
BB7($C_{T,P\|I}$)	1.1011***	0.0304	0.0080	0.1147	-51.9699	-41.8285
	0.1436***	0.0389	0.1234			
sum					-686.479	-656.0546

Table 5 Results of D-vine Copulas Conditional to Thailand and the Philippines

Copulas	parameters	standard error	Lower and upper tail dependence	Kendall'tau	AIC	BIC
Survival BB1 ($C_{I,T}$)	0.2651***	0.0559	0.3314	0.3478	-389.3458	-379.2044
	1.3538***	0.0441	0.5993			
BB1($C_{T,P}$)	0.3187***	0.0565	0.1591	0.2709	-245.1539	-235.0124
	1.1831***	0.0331	0.2035			
T($C_{I,P\|T}$)	0.1958***	0.0292	0.0010	0.1255	-47.3063	-37.1648
	20.4644	14.1012	0.0010			
sum					-681.806	-651.3816
BB1($C_{T,P}$)	0.2644***	0.0539	0.1023	0.2319	-181.6724	-171.531
	1.1499***	0.0309	0.1728			
BB1 ($C_{P,I}$)	0.3695***	0.0625	0.2330	0.3446	-384.3979	-374.2564
	1.2878***	0.0390	0.2870			
T($C_{I,T\|P}$)	0.2854***	0.0290	0.0686	0.1843	-116.2577	-106.1162
	6.9649***	1.5714	0.0686			
sum					-682.328	-651.9036

4 Economic Application of Risk Measures

Copulas have attracted much attention in the computation of value at risk, expected shortfall for risk measure, as pointed out by Kole et al. [22], Junker and May [23], Ouyang et al. [24], etc. In order to strengthen the practical applicability of the empirical results, we make use of the Monte Carlo simulation and the estimation results of the vine copula to calculate the VaR and ES of equally weighted portfolio.

The detailed procedures that we propose to evaluate the risk consist of four steps: first, we generate 1117 random numbers of $C_{I,P}$ (BB1) and $C_{I,T}$ (Survival BB1); second, the standardized residual can be got from the inverse function of the skewed student-t distribution which is an assumption of the marginal distribution in the GJR-GARCH model; third, the next period stock returns can be forecasted through the mean equations of the GJR-GARCH models; fourth, we distribute equal weights to each stock return, and then we get the returns after the weighting; finally, the VaR and ES can be calculated at the 5%, 2%, and 1% levels. The four processes can be repeated 1000, 2000, and 5000 times to get the convergence values.

Table 6 presents the results of the VaR and ES of equally weighted portfolio. As can be seen in Table 6, the VaR converges to -1%, -1.35%, and -1.61% at the 5%, 2%, and 1% levels, respectively, and -1.41%, -1.78%, and -2.08% for the ES. Table 7 provides the VaR and ES of each stock market and the average value at the 5%, 2%, and 1% levels. First, there is no doubt that portfolio may successfully avoid risk, as can be seen by comparing the results as given in Table 6 with those in Table 7. The VaR and ES of Thailand are the least, which means that the Thailand stock market is at more risk. At the same time, this illustrates that Indonesia is at less risk, and that the Philippines is at medium risk.

Table 6 VaR and ES of Equally Weighted Portfolio

VaR	5%	2%	1%
1000 times	-0.01002	-0.01353	-0.01607
2000 times	-0.01004	-0.01349	-0.01608
5000 times	-0.01003	-0.01351	-0.01608
ES			
1000 times	-0.01412	-0.01777	-0.02081
2000 times	-0.01408	-0.01778	-0.02082
5000 times	-0.01408	-0.01777	-0.02080

Table 7 VaR and ES for Each Stock Market

VaR (5000 times)	Indonesia	Philippines	Thailand	Average
5%	-0.0138	-0.0159	-0.0167	-0.0155
2%	-0.0194	-0.0217	-0.0225	-0.0212
1%	-0.0238	-0.0260	-0.0270	-0.0256
ES (5000 times)				
5%	-0.0205	-0.0225	-0.0234	-0.0221
2%	-0.0269	-0.0287	-0.0295	-0.0284
1%	-0.0324	-0.0336	-0.0344	-0.0335

5 Conclusions

This paper depicts a model for estimating conditional volatility, dependency, VaR, and ES through a vine copula based GJR-GARCH model, in which the empirical evidence shows that there do exist leverage effects in these three country stock markets, and that all appropriate margins are skewed student-t distributions; given these, the optimal choices of the C-vine copula are BB1, Survival BB1, and BB7 copula, in that order, while the same for the D-vine copula are Survival BB1, BB1, and T copula. Another significant observation is that the D-vine structure is more appropriate than the C-vine one, as a whole. In addition, the Indonesian and Thailand markets show the greatest dependence, which includes their upper tail (0.6013) and lower tail (0.3369) in these three country markets. Also, the Philippine market has a significant effect on the tail dependence between Indonesia and Thailand. As a final note, it needs to be emphasized that the vine copula based GJR-GARCH model captures the VaR and ES successfully.

References

1. Sharma, S.C., Wongbangpo, P.: Long-term trends and cycles in ASEAN stock markets. Review of Financial Economics 11, 299–315 (2002)
2. Ning, C., Wirjanto, T.S.: Extreme return-volume dependence in East-Asian stock markets: A copula approach. Finance Research Letters 6, 202–209 (2009)
3. Lim, K.P., Brooks, R.D., Hinich, M.J.: Nonlinear serial dependence and the weak-form efficiency of Asian emerging stock markets. Int. Fin. Markets, Inst. and Money 18, 527–544 (2008)
4. Sharma, P.: Asian Emerging Economics and United States of America: Do they offer a diversification benefit. Australian Journal of Business and Management Research 1(4), 85–92 (2011)
5. Joe, H., Hu, T.: Multivariate distributions from mixtures of max-infinitely divisible distributions. Journal of Multivariate Analysis 57(2), 240–265 (1996)
6. Bedford, T., Cooke, R.M.: Monte Carlo simulation of vine dependent random variables for applications in uncertainty analysis. In: Proceedings of ESREL 2001, Turin, Italy (2001)
7. Bedford, T., Cooke, R.M.: Vines-a new graphical model for dependent random variables. Annals of Statistics 30(4), 1031–1068 (2002)
8. Nikoloulopoulos, A.K.: Vine copulas with asymmetric tail dependence and applications to financial return data. Computational Statistics and Data Analysis 56, 3659–3673 (2012)
9. Kurowicka, Cooke, R.M.: Uncertainty Analysis with High Dimensional Dependence Modelling. Wiley, New York (2006)
10. Joe, H.: Dependence comparisons of vine copulae with four or more variables. In: Kurowicka, D., Joe, H. (eds.) Dependence Modeling: Vine Copula Handbook. World Scientific, Singapore (2010)
11. Joe, H., Li, H., Nikoloulopoulos, A.K.: Tail dependence functions and vine copulas. Journal of Multivariate Analysis 101, 252–270 (2010)
12. Aas, K., Czado, C., Frigessi, A., Bakken, H.: Pair-copula constructions of multiple dependence. Insurance: Mathematics and Economics 44, 182–198 (2009)

13. Czado, C., Schepsmeier, U., Min, A.: Maximum likelihood estimation of mixed C-vines with application to exchange rates. Statistical Modelling 12, 229–255 (2012)
14. Glosten, L.R., Jagannathan, R., Runkle, D.E.: On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. The Journal of Finance 48(5), 1779–1801 (1993)
15. Joe, H.: Multivariate Models and Dependence Concepts. Chapman and Hall, London (1997)
16. Savu, C., Trede, M.: Hierarchical Archimedean copulas. In: International Conference on High Frequency Finance, Konstanz, Germany (2006)
17. Czado, C.: Pair-copula constructions of multivariate copulas. Copula Theory and Its Applications, 93-109 (2010)
18. Joe, H.: Asymptotic efficiency of the two-stage estimation method for copulabased models. Journal of Multivariate Analysis 94, 401–419 (2005)
19. Liu, J., Sriboonchitta, S.: Analysis of Volatility and Dependence between the Tourist Arrivals from China to Thailand and Singapore: A Copula-Based GARCH Approach. In: Huynh, V.N., Kreinovich, V., Sriboonchitta, S., Suriya, K. (eds.) Uncertainty Analysis in Econometrics with Applications. AISC, vol. 200, pp. 285–296. Springer, Heidelberg (2013)
20. Sriboonchitta, S., Nguyen, H.T., Wiboonpongse, A., Liu, J.: Modeling volatility and dependency of agricultural price and production indices of Thailand: Static versus timevarying copulas. International Journal of Approximate Reasoning 54, 793–808 (2013)
21. Brechmann, E.C., Schepsmeier, U.: Modeling Dependence with C- and D-Vine Copulas: The R Package CDVine. Journal of Statistical Software 52(3), 1–27 (2013)
22. Kole, E., Koedijk, K., Verbeek, M.: Selecting copulas for risk management. Journal of Banking and Finance 31, 2405–2423 (2007)
23. Junker, M., May, A.: Measurement of aggregate risk with copulas. Econometrics Journal 8, 428–454 (2005)
24. Ouyang, Z., Liao, H., Yang, X.: Modeling dependence based on mixture copulas and its application in risk management. Appl. Math. J. Chinese Univ. 24(4), 393–401 (2009)

Studying Volatility and Dependency of Chinese Outbound Tourism Demand in Singapore, Malaysia, and Thailand: A Vine Copula Approach

Jianxu Liu, Songsak Sriboonchitta, Hung T. Nguyen, and Vladik Kreinovich

Abstract. This paper investigates the volatility and dependence of Chinese tourism demand for Singapore, Malaysia, and Thailand (SMT) destinations, using the vine copula based auto regression moving average-generalized autoregressive conditional heteroskedasticity (ARMA-GARCH) model. It is found that a jolt to the tourist flow can have long-standing ramifications for the SMT countries. The estimation of the vine copulas among SMT show that the Survival Gumbel, Frank, and Gaussian copulas are the best copulas for Canonical vine (C-vine) or Drawable vine (D-vine) among the possible pair-copulas. In addition, this paper illustrates the making of time-varying Frank copulas for vine copulas. Finally, there is a discussion on tourism policy planning for better managing the tourism demand for the SMT countries. We suggest tour operators and national tourism promotion authorities of SMT collaborate closely in the marketing and promotion of joint tourism products.

1 Introduction

Outbound tourism in China is growing rapidly, and has become a significant contributor to international tourism. In 2011, the total volume of the outbound tourists made more than 70 million trips, with an increase by 22.4% year by year. By 2015,

Jianxu Liu · Songsak Sriboonchitta
Faculty of Economics, Chiang Mai University, Chiang Mai 50200 Thailand
e-mail: liujianxu1984@163.com, songsakecon@gmail.com

Hung T. Nguyen
Department of Mathematical Sciences,
New Mexico State University, New Mexico, USA
e-mail: hunguyen@nmsu.edu

Vladik Kreinovich
Computer Science Department, University of Texas at El Paso, Texas, USA
e-mail: vladik@utep.edu

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_17,

the China National Tourism Administration (CNTA) forecasts, the Chinese international passengers will increase to 100 million, a quantity that will put China in the numero uno position in the international tourism source market.

With the immense increase in Chinese outbound tourism, the number of Chinese visitors and their expenditure has become the most important and the most potential passengers' market. As one of their classic, routine travel destinations, SMT (Singapore, Malaysia, and Thailand), which is a very popular tourist destination for the Chinese tourists, is witnessing more changes shining through. For example, China has become the second largest passenger source country for Singapore, only behind Indonesia, in 2011; China is Malaysia's third largest source of tourists, following Indonesia and Singapore, and in the first half of 2012, 871,959 Chinese tourists visited Malaysia, up 53.4% year by year; for Thailand, The Tourism Authority of Thailand (TAT) forecasts that China will become their largest passenger source market in 2014. In addition, it is easy to see, as shown in Figure 1, that the outbound tourism to Singapore, Malaysia, and Thailand maintains the overall upward trend and reaches peaks and troughs simultaneously. In view of this scenario, we infer that the volatilities of the outbound tourism to Singapore, Malaysia, and Thailand may be similar to each other, as well, and that their dependence should be also quite high.

The tourism industries in Singapore, Malaysia, and Thailand possess their own prominent places. First, the total contribution of Travel and Tourism to the GDP of these countries (SMT) accounts for 10.6%, 15.8%, and 16.3%, respectively, of their total GDP, in 2011. Second, the total contribution of Travel and Tourism to

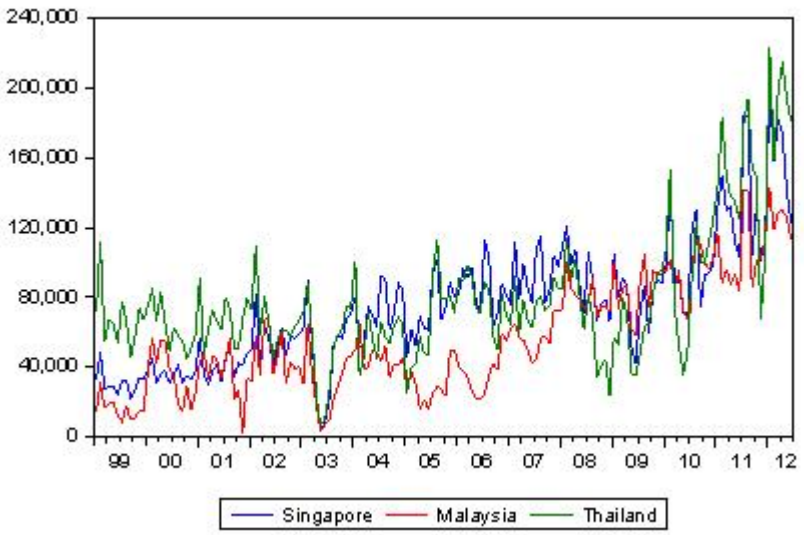


Fig. 1 Tourist flows to SMT from China

the employment sector was 266,500 jobs (8.6% of total employment) for Singapore, 1,587,000 jobs (13.8% of total employment) for Malaysia, and 4,468,500 jobs for Thailand (11.4% of total employment), in 2011. Third, in 2011, Singapore, Malaysia, and Thailand generated SGD 22.5 bn (3.3% of total exports), MYR 62.4 bn (8.4% of total exports) and THB 950.4 bn (11.4% of total exports) from visitor exports. It is thus clear that it is the tourism industry of Singapore, Malaysia, and Thailand that drives the development of the whole economy and the society of these countries.

It is worth mentioning, besides, that the volume of China outbound tourism is enormous; the great spending power ranked No. 2 (USD 59.52 bn) in the world in 2011. SMT may benefit from analyzing the international tourism demands from China. Hence, the analysis of volatility and dependence of tourism demand is essential for investigating the effects of shocks and co-movements in the SMT tourism demand from China. Furthermore, for tourism managers and travel corporations, it is important to evaluate the dependence structure of tourism demand and to discern attractive opportunities. They also need to figure out the implied threat caused by volatilities in tourism demand. Mastering the behaviors of volatility and dependence structure can help governments and tourism corporations adjust strategies for improving the profitability of the tourism industry and reducing adverse impacts such as political unrest, natural disaster, etc.

This study is organized as follows. Section 2 reviews the tourism research, the copula based GARCH model, and the vine copulas. The copula based ARMA-GARCH model is discussed in section 3 which includes ARMA-GARCH models for margins, copulas, and vines. Data description and empirical findings are presented in section 4. Policy planning is discussed in section 5, while some concluding remarks are given in section 6.

2 Literature Review

Recognizing the relevance and significance of tourism industry to SMT economy, a number of studies have been undertaken on various aspects of outbound tourism from China to SMT: for example, Li et al. [1] examined the Chinese tourists' expectations of outbound travel products, Lee [2] studied the dynamic interactions between hotel room rates and international inbound tourists in Singapore, and Chang et al. [3] forecasted tourism demand from East Asia to Thailand. However, none of the literature relates to the dependencies and volatility of tourism demand from China to SMT except just one paper which is by Liu and Sriboonchitta [4] who studied the volatility and dependence between tourist arrivals from China to Thailand and those to Singapore. It was found that the Gaussian copula fitted very well, and that the Kendall's tau was 0.5737. It is thus clear that the dependence of Chinese outbound tourism demand between the destinations of Singapore and Thailand is very high.

Many scholars studied tourism demand by applying econometric and statistical tools to analyze the volatilities and relationships of inbound or outbound tourism demand. Kim and Wong [5], and Song et al. [6] used the univariate

autoregressive conditional heteroskedasticity (GARCH) model to analyze the volatility of tourism demand. Chan et al. [7] used the symmetric constant conditional correlation-multivariate generalized autoregressive conditional heteroskedasticity (CCC-MGARCH) model and the symmetric vector ARMA-GARCH to model the multivariate international tourism demand and volatility among the four tourism source countries to Australia. Hoti et al. [8] made use of the VARMA-GARCH model to investigate international tourism and country risk spillovers for Cyprus and Malta. Seo et al. [9] analyzed the relationships of the Korean outbound tourism demand by using the MGARCH and Vector Error Correction (VEC) models. Lee [2] investigated the short-run and the long-run dynamic interactions using the cointegration and Granger causality test.

However, the above-mentioned papers always assumed the conditional correlation to be the linear Pearson's correlation and constant over time, which is a strong and strict assumption. Of late, the copula-GARCH model has been very popular in the financial field, as it can be used to analyze the volatilities and dependence structure. Patton [10] used this model to analyze the dynamic dependence between the exchange rates of YenUSD and DMUSD. Jondeau and Rockinger [11] assumed the marginals of the copula-GARCH model to be a skewed student-t distribution in order to capture heavy tail information regarding the international stock market. Lee and Long [12] proposed copula based multivariate GARCH model with uncorrelated dependent errors, which are generated through a linear combination of dependent random variables. Wu [13] also researched the economic value of comovement between oil prices and exchange rates using copula-based GARCH models. Wang et al. [14] studied the dynamic dependence between the Chinese market and other international stock markets using the time-varying copula approach. But the above-mentioned studies all used the bivariate copula-GARCH model to study the dependence structure.

To study the multivariate dependence structure, Joe [15] gave the first pair-copula construction (PCC) of a multivariate copula, the construction of which is dependent on distribution functions. Bedford and Cooke [17] [18] expressed these constructions in terms of densities, and organized these constructions in a graphical way involving a sequence of nested trees, which are called regular vines. They also proposed two subclasses of PCC, which we call the C-vine and D-vine copulas. Note that the C-vine and D-vine copulas have been widely used in finance asset returns and other data by many researchers, such as Aas et al. [20], Min and Czado [19], and Czado [21]. For studying the dependence of outbound tourism demand in the three countries (SMT), we make use of the C-vine and D-vine copulas instead of the bivariate copula in the copula-GARCH model. In other words, we use the ARMA-GARCH model to fit the marginals, and then transform the standardized residuals into specified distributions; finally, we make use of the C-vine and D-vine copulas to capture the dependence structure.

To sum up, the main contributions of this study are as follows: (1) we introduce the C-vine and D-vine copula based ARMA-GARCH model into tourism demand research; (2) we propose time-varying Frank copula to capture the dynamic Kendall's tau for vine copulas; (3) we investigate the impact of the short-run and

long-run Chinese outbound tourism demand for SMT; (4) we compare the results of the bivariate copulas with those of the vine copulas, thus finding out the dominant country among SMT; and (5) finally, through this study, we provide inferences that are applicable for competitive destination strategies and policy development.

3 Copula Based ARMA-GARCH Model

This paper utilizes copula based ARMA-GARCH model to analyze the volatility and dependence of Chinese outbound tourism demand to SMT destinations. We filter growth rate data using the ARMA-GARCH model with appropriate distributions, for example, the skewed student-t, skewed GED, and skewed normal, and transform standardized residuals to copula data ($u_1 = F_1(x_1)$, $u_2 = F_2(x_2)$, and $u_3 = F_3(x_3)$) by using appropriate distribution functions. After that, we estimate the bivariate copula and the C-vine and D-vine copulas using the maximum likelihood estimation method.

3.1 ARMA-GARCH Model for Margins

Bollerslev [23] proposed the GARCH (generalized autoregressive conditional heteroskedasticity) model, which has replaced the ARCH model in application and has since been widely used in econometrics, economics, etc. In accordance with the findings of Ling [24], the ARMA (p, q)-GARCH (k, l) model can be formed as

$$r_t = c + \sum_{i=1}^{p} \phi_i r_{t-i} + \sum_{i=1}^{q} \psi_i \varepsilon_{t-i} + \varepsilon_i \tag{1}$$

$$\varepsilon_t = h_t \eta_t \tag{2}$$

$$h_t^2 = \omega + \sum_{i=1}^{k} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{l} \beta_i h_{t-i}^2 \tag{3}$$

where $\sum_{i=1}^{p} \phi_i < 1, \omega > 0, \alpha_i >= 0, \beta_i >= 0, \sum_{i=1}^{k} \alpha_i + \sum_{i=1}^{l} \beta_i < 1$, and the time series r_t are return data; the formulas (1) and (3) are known as the conditional mean equation and conditional variance equation, respectively; the formula (2) shows that these return residuals are split into a stochastic piece η_t and a time dependent standard deviation h_t.The values of α_i and β_i indicate the presence or absence of short-run shock and persistence of volatility, respectively. If the value of α_i is larger, then the short-term unexpected factors affecting the volatility have greater influence. If the value of $\sum_{i=1}^{k} \alpha_i + \sum_{i=1}^{l} \beta_i$ is larger, then the impact of unexpected shock to volatility has the longer duration. η_t is the standardized residual, which can be assumed for any distribution. In this study, we assume the distribution of the standardized residuals to be the skewed student-t distribution or skewed-generalized error distribution (GED), both of which can capture the characteristics of heavy tail

and asymmetry, anyway. The standardized skewed-t and skewed-GED distributions can be expressed as

$$f_{skt}(x_i|v,\gamma) = \frac{2}{(\gamma+\gamma^{-1})}\{f_v(x_i/\gamma)I_{[0,\infty]}(x_i) + f_v(\gamma\, x_i)I_{[\infty,0]}(x_i)\} \tag{4}$$

$$f_{sged}(x_i|v,\gamma) = v(2\theta\Gamma(1/v))^{-1} \times \exp(-\frac{|x_i-\delta|^v}{(1-sign(x_i-\delta)\gamma)^v\theta^v}) \tag{5}$$

where

$$A = \Gamma(2/v)\Gamma(1/v)^{-0.5}\Gamma(3/v)^{-0.5} \tag{6}$$

$$S(\gamma) = \sqrt{1+3\gamma^2-4A^2\gamma^2} \tag{7}$$

$$\delta = 2\gamma A \times S(\gamma)^{-1} \tag{8}$$

$$\theta = \Gamma(3/v)^{-0.5}\sqrt{\Gamma(1/v)}S(\gamma)^{-1} \tag{9}$$

where $f_v(.)$ is the density of the student t-distribution, the parameter v represents the number of degrees of freedom, γ is the skewness parameter ranging from 0 to ∞, I denotes the indicator function, and "sign" is the sign function.

3.2 Copulas

Copulas [25] have long been recognized and developed in various fields like econometrics, economics, financials, etc. If $X = (X_1, X_2, \ldots, X_n)$ is a random vector with joint distribution function H and marginal distributions $F_1, F_2, \ldots, F_n$, then there exists a function C called copula, such that

$$H(x_1, x_2, \ldots, x_n) = C(F_1(x_1), F_2(x_2), \ldots, F_n(x_n)) \tag{10}$$

The copula C is extracted from the joint H and marginals $F_1, F_2, \ldots, F_n$ as

$$C(u_1, u_2, \ldots, u_n) = H(F_1^{-1}(u_1), F_2^{-1}(u_2), \ldots, F_n^{-1}(u_n)) \tag{11}$$

where $F_i^{-1}(u_i) = inf\{x \in \Re : F_i(x) \geq u_i\}$. If F_i is absolutely continuous and strictly increasing, then

$$\begin{aligned} f(x_1, x_2, \ldots, x_n) &= \frac{\partial H(x_1, x_2, \ldots, x_n)}{\partial x_1 \cdots \partial x_n} \\ &= \frac{\partial C(u_1, u_2, \ldots, u_n)}{\partial u_1 \cdots \partial u_n} \times \prod \frac{F(x_i)}{\partial x_i} \\ &= c(u_1, u_2, \ldots, u_n) \times \prod f_i(x_i) \end{aligned} \tag{12}$$

The joint distribution H contains all the statistical information about $X = (X_1, X_2, \ldots, X_n)$. In particular, the marginal distributions of the components are derived as

$$F_i(x_i) = H(\infty, \infty, \cdots, x_i, \infty, \infty) \tag{13}$$

In this study, the Gaussian copula, T copula, Clayton copula, Frank copula, Gumbel copula, Joe copula, BB1 copula, BB6 copula, BB7 copula, BB8 copula, and rotated copulas have been utilized to analyze the dependence structure (see Brechmann and Schepsmeier [16]).

3.3 Vines

A bivariate copula vine specification is called a pair-copula construction or a vine copula. Compared to some multivariate copulas, vine copulas are more flexible in the high dimensions. For example, multivariate normal copula does not have tail dependence; multivariate t-copula has only a single degree of freedom parameter and symmetric tail dependence. For three variables, we can assume that 12 is the first pair, then the second pair is either 13 or 23, and the third pair should be 23|1 or 13|2. The structure 12, 13, 23|1 is the standard form of Canonical vine copula (C-vines), and the other 12, 23,13|2 is called Drawable vine copula (D-vines).

To use the C-vine and D-vine constructions to represent dependency structure through copulas, we assume that there are three univariate marginals that are uniform in [0, 1]. Note that these univariate marginals correspond to cumulative distribution functions of the standardized residuals by generating from ARMA-GARCH model. Here, we concentrate on the C-vine and D-vine representations with three variables. The densities of the C-vine and D-vine copulas can be expressed as

$$c(u_1, u_2, u_3) = c_{12}(u_1, u_2) \cdot c_{13}(u_1, u_3) \cdot c_{23|1}(F(u_2|u_1), F(u_3|u_1)) \text{for C-vine copula} \tag{14}$$

and

$$c(u_1, u_2, u_3) = c_{12}(u_1, u_2) \cdot c_{23}(u_2, u_3) \cdot c_{13|2}(F(u_1|u_2), F(u_3|u_2)) \text{for D-vine copula} \tag{15}$$

where

$$F(u_2|u_1) = \frac{\partial C_{12}(u_1, u_2)}{\partial u_1} \tag{16}$$

$$F(u_3|u_1) = \frac{\partial C_{13}(u_1, u_3)}{\partial u_1} \tag{17}$$

$$F(u_1|u_2) = \frac{\partial C_{12}(u_1, u_2)}{\partial u_2} \tag{18}$$

$$F(u_3|u_2) = \frac{\partial C_{23}(u_2, u_3)}{\partial u_2} \tag{19}$$

There are two things need to be finished before we estimate vine copulas model. One is the selection of a specific ordering; the other is the choice of pair-copula families. In this study, we employ different methods to select the orderings of the variables in the C-vine and D-vine models. For the C-vine model, we calculate the sum of the empirical Kendall's $S_\tau^i = \sum_{j=1, i\neq j}^{n} \tau_{i,j}$ (see Czado et al. [22]) for each variable i, and select the maximum one as the first variable. After that, we reorder the remaining variables and repeat the process of calculating the sum of Kendall's tau, thus finding out the second and third variables. For the D-vine model, we just determine the order that satisfies the maximization of the sum of the empirical Kendall's tau $S_\tau = \sum_{i=1}^{n-1} \tau_{i,i+1}$.To choose the appropriate pair-copula families, we firstly estimate all possible copula families for C_{12}, C_{13} (C-vine), and for C_{12}, C_{23} (D-vine) by using maximum likelihood method. Then, we determine the required observations for $C_{23|1}$ and $C_{13|2}$ through the formulas (16)-(19). Thus, both C-vine and D-vine copulas can be estimated, and the Akaike information criterion (AIC) and Bayesian information criterion (BIC) can be calculated as well. Thereby, the copula families corresponding to the minimum values of AIC and BIC are selected among many copula families.

To improve efficiency of estimation of vine copula models, we need consider two-step maximum likelihood estimation method (see Aas et al. [20], Czado et al. [22]). The purpose of the first step is to obtain the starting values of the appropriate copula families, thus this step has been implemented in calculating AIC and BIC process. In the last step, all parameters of C-vine and D-vine copulas are estimated by the full maximum likelihood. The corresponding log-likelihood can be constructed by using the formulas (10) and (11). The log-likelihood functions of C-vine and D-vine copulas can be written as

$$\begin{aligned} L_C(u_1, u_2, u_3; \theta) = \sum_{i=1}^{n} [&\log c_{12}(u_{1,i}, u_{2,i}; \theta_1) + \log c_{13}(u_{1,i}, u_{3,i}; \theta_2) \\ &+ \log c_{23|1}(F(u_{2,i}|u_{1,i}), F(u_{3,i}|u_{1,i}; \theta_3))] \end{aligned} \tag{20}$$

and

$$\begin{aligned} L_D(u_1, u_2, u_3; \theta) = \sum_{i=1}^{n} [&\log c_{12}(u_{1,i}, u_{2,i}; \theta_1) + \log c_{23}(u_{2,i}, u_{3,i}; \theta_2) \\ &+ \log c_{13|2}(F(u_{1,i}|u_{2,i}), F(u_{3,i}|u_{2,i}; \theta_3))] \end{aligned} \tag{21}$$

where θ is the parameter vector that need to be estimated; θ_1, θ_2 and θ_3 represent the parameters corresponding to the appropriate copula families.

4 Empirical Results

4.1 Data

This paper models the time series of the difference between the logarithms of the monthly international arrivals (from January 1999 to June 2012) from China to Singapore, Malaysia, and Thailand. The data description and statistics are shown in Table 1: all the mean values are positive, the skewness values of both Singapore and Thailand are negative, and the kurtosis values are greater than 3. It is thus clear that the data show non-normality and that the data of Singapore and Thailand are skewed to the left. The results of the Jarque-Bera test reject the null hypothesis that the data are from a normal distribution, which are more convictive explanations for non-normality distributions, thereby implying that the skewed distribution is the more appropriate one for our study.

Table 1 Data Description and Statistics

	Singapore	Malaysia	Thailand
Mean	0.007982	0.012729	0.005707
Median	0.040844	0.007487	0.008301
Maximum	0.923611	2.848938	0.897066
Minimum	-1.750788	-2.645770	-1.281747
Std. Dev.	0.325849	0.466296	0.343930
Skewness	-0.888538	-0.020366	-0.652907
Kurtosis	7.744626	16.98058	4.556646
Jarque-Bera	172.1994	1311.199	27.69398
Probability	0.000000	0.000000	0.000001

4.2 Estimation Results of ARMA-GARCH Model

Table 2 presents the results of the ARMA-GARCH model with different assumptions of marginal distribution. To analyze the volatility of the China outbound tourist demand to STM, we employ ARMA (12, 4)-GARCH (1, 1) with skewed student-t distribution for Singapore, ARMA (6, 6)-GARCH (1, 1) with skewed student-t distribution for Malaysia, and ARMA (12, 4)-GARCH (1, 1) with skewed-GED distribution for Thailand. The values of the GARCH coefficient, or β, equal 0.5482, 0.8467, and 0.6993, and they are significant as well. These results indicate that a shock to the tourist arrival series has long-run persistence in all cases, and that Chinese tourists outbound to Malaysia have stronger long-run persistence. The estimated ARCH effect, or α, is significant only to the tourist arrival series from China to Thailand, and so tourist arrivals from China to both Singapore and Malaysia do not have short-run persistence. The values of the parameter γ equal 0.8862, 0.9083, and 0.6626 in each model of SMT, respectively, implying that the Chinese outbound

tourism demand in SMT are skewed to the left, and the series of the destination Thailand is more skewed to the left.

Since the parameters γ and v are significant, we transform the standardized residuals into standard skewed student-t distribution and skewed-GED distribution as margins. However, the margins must satisfy the condition of uniform distribution from 0 to 1. If it cannot satisfy this condition, then the misspecified model for the marginal distribution may cause incorrect-fit copulas. Thus, testing for marginal distribution model misspecification is a critical step in constructing multivariate distribution models using copulas. Therefore, we present the Box-Ljung test for evaluating the serial independence of the marginals, $F_{skt}(x_{sing,t})$, $F_{skt}(x_{malay,t})$, and $F_{sged}(x_{thai,t})$, and the Kolmogorov-Smirnov (K-S) test for the distribution specification. The results of the KS test and the Box-Ljung test are given in Table 3. It is very clear that each of the series accepts the null hypothesis, which means that all the three marginals are of uniform distribution. The second part of Table 3 shows the results of the Box-Ljung test, which evaluates the serial independence of the first four moments, and it can be observed that all of them accept the null hypothesis at the 0.10 level. Therefore, the marginals that we assumed satisfy the two preconditions: uniformity and serial independence.

4.3 Estimation Results of Vine Copulas

For the C-vine copula, the order is Thailand, Malaysia, and Singapore, so we need to estimate $C(F_{sged}(x_{thai,t}), F_{skt}(x_{malay,t}))$, $C(F_{sged}(x_{thai,t}), F_{skt}(x_{sing,t}))$, and $C(F_{skt}(x_{malay,t}), F_{skt}(x_{sing,t})|F_{sged}(x_{thai,t}))$. For the D-vine copula, the order is Singapore, Thailand, and Malaysia. As far as the structures of the C-vine and D-vine are concerned, the D-vine has the same pair-copulas as the C-vine. Therefore, in this case, we only need to do the calculation either for the C-vine or for the D-vine. Nevertheless, all the possible vine structures are calculated for a comprehensive analysis of the SMT inbound tourism from China.

The possible pair-copula families were the Gaussian copula, T copula, (Survival) Clayton copula, Frank copula, (Survival) Gumbel copula, (Survival) Joe copula, (Survival) BB1 copula, (Survival) BB6 copula, (Survival) BB7 copula, (Survival) BB8 copula, and rotated copulas. We make use of the AIC and BIC to choose the best copula for each pair. In Table 4, we present the results of the vine copulas and Kendall's tau for each best copula. The table shows that the Survival Gumbel, Frank, and Gaussian copulas are the best copulas for the C-vine or D-vine among the possible pair-copula families. First, we can find that China outbound tourism demand between Thailand and Malaysia has lower tail dependence equaling 0.2643, implying that negative influences may have simultaneous impact on these two countries' tourism industries. Second, China outbound tourism demand between Thailand and Singapore shows stronger dependence equaling 0.5285, but there does not exist tail dependence, which illustrates the fact that any positive and negative shocks cannot have an effect on these two countries' tourism industries at the same time. Third, the dependency parameter of the Gaussian copula between Singapore and Malaysia

Table 2 Results of ARMA-GARCH Model

	Singapore		Malaysia		Thailand
AR1	0.6281***	Constant	0.0198***	AR1	0.4041***
	(6.771e-05)		(2.857e-05)		(0.0024)
AR2	0.4762***	AR1	0.0495***	AR2	0.6583***
	(6.802e-05)		(5.772e-05)		(0.0023)
AR3	0.4621***	AR2	0.6758***	AR3	0.5453***
	(7.236e-05)		(5.678e-05)		(0.0039)
AR4	1.2550***	AR3	0.2839***	AR4	0.1460***
	(6.808e-05)		(5.774e-05)		(0.0059)
AR5	1.8250***	AR4	0.4870***	AR5	0.1723***
	(7.913e-05)		(6.088e-05)		(0.0025)
AR6	0.0595***	AR5	0.8201***	AR6	0.0143***
	(6.816e-05)		(6.284e-05)		(0.0020)
AR7	0.0231***	AR6	0.3124***	AR7	0.1865***
	(6.807e-05)		(5.553e-05)		(0.0018)
AR8	0.0088***	MA1	0.5572***	AR8	0.0586***
	(1.230e-03)		(9.567e-05)		(0.0022)
AR9	0.1827***	MA2	0.9847***	AR9	0.0047
	(1.261e-03)		(9.937e-05)		(0.0026)
Ar10	0.1512***	MA3	0.2598***	AR10	0.1907***
	(7.569e-05)		(1.016e-04)		(0.0028)
AR11	0.1025***	MA4	0.9287***	AR11	0.0952***
	(8.032e-05)		(1.038e-04)		(0.0024)
AR12	0.3090***	MA5	0.6446***	AR12	0.0329***
	(5.710e-05)		(1.062e-04)		(0.0027)
MA1	0.3453***	MA6	0.9127***	MA1	0.0952***
	(1.480e-04)		(1.052e-04)		(0.0035)
MA2	0.0516***	ω	0.0113	MA2	0.3289***
	(1.790e-04)		(0.0241)		(0.0030)
MA3	0.3352***	α	1.0000	MA3	0.2688***
	(1.866e-04)		(1.9370)		(0.0054)
MA4	0.7582***	β	0.8467***	MA4	0.7845***
	(1.758e-04)		(0.0614)		(0.0039)
ω	6.161e-04	γ	0.9083***	ω	0.0079***
	(9.575e-04)		(0.0597)		(0.0018)
α	1.0000	γ	2.1060***	α	0.2970***
	(0.4920)		(0.2460)		(0.0259)
β	0.5482*	LM-test	1.0000	β	0.6993***
	(0.1344)				(0.0253)
γ	0.8862***	LogL	15.2619	γ	0.6626***
	(0.0065)				(0.0138)
v	2.7080***	AIC	0.0340	v	1.0000***
	(0.3821)				(0.0428)
LM-test	0.9841	BIC	0.3785	LM-test	0.4531
LogL	65.4508			LogL	0.0170
AIC	0.5522			AIC	0.2611
BIC	0.1503			BIC	0.6630

Note: Signif. codes are as follows: 0 *** 0.001 ** 0.01 * 0.05 0.1. The numbers in the parentheses are the standard deviations.

Table 3 KS Test for Uniform and Box-Ljung Test for Autocorrelation

KS Test			
	Statistic	P value	Hypothesis
$u_{1,t}$	0.0976	0.0933	0 (acceptance)
$u_{2,t}$	0.0670	0.4650	0 (acceptance)
$u_{3,t}$	0.0652	0.5013	0 (acceptance)
Box-Ljung Test			
	Moments	X-squared	P-value
$u_{1,t}$	First moment	1.7570	0.8816
	Second moment	4.1513	0.5278
	Third moment	0.9194	0.9688
	Fourth moment	2.5339	0.7714
$u_{2,t}$	First moment	2.5915	0.7626
	Second moment	2.6553	0.7529
	Third moment	4.1255	0.5315
	Fourth moment	0.8972	0.9704
$u_{3,t}$	First moment	1.9974	0.8495
	Second moment	5.1570	0.3970
	Third moment	2.5577	0.7678
	Fourth moment	2.0434	0.8431

Note: $u_{1,t} = F_{skt}(x_{sing,t})$, $u_{2,t} = F_{skt}(x_{malay,t})$, and $u_{3,t} = F_{sged}(x_{thai,t})$.

conditional on Thailand is not significant, which means that the inbound tourism demand of Singapore and Malaysia from China is independent, given the Thailand inbound tourism from China as condition, while the inbound tourism demand of Singapore and Malaysia from China is not independent, and that the Survival Gumbel copula fits them very well. Thus, the Thailand inbound tourism from China may exert an influence on the inbound tourism demand of Singapore and Malaysia from China. In addition, both the best $C_{TM|S}$ and $C_{TS|M}$ are Frank copulas. The Kendall's tau of $C_{TM|S}$ and $C_{TS|M}$ are 0.1494 and 0.5062, respectively. If we compare $C_{S,T}$ with $C_{ST|M}$, it can be clearly observed that Malaysia inbound tourism from China makes little difference to the inbound tourism from China to Singapore and Thailand, whereas Thailand and Malaysia inbound tourism from China are partly affected by Singapore inbound tourism from China, corresponding to $C_{T,M}$ and $C_{MT|S}$.

4.4 Application of Dynamic Dependence Structure

Patton [10] proposed the time-varying copulas that include the Gaussian and symmetric Joe copulas. Manner [26], Wu [13], and Ng et al. [27] further researched the time-varying copulas. In our case, we follow our predecessors achievements to invent the time-varying Frank copula. The formula can be expressed as

$$\theta_t = \Lambda(\omega + \alpha\theta_{t-1} + \beta(u_{i,t-1} - 0.5)(u_{j,t-1} - 0.5)) \tag{22}$$

Table 4 Results of Vine Copulas and Kendall's tau

Copulas		parameters	tail dependence	Kendall'tau	AIC	BIC
C-vine or D-vine						
Survival ($C_{T,M}$)	Gumbel	1.2570***	0.2643	0.2045	-14.5042	-11.4228
		(0.0769)				
Frank($C_{T,S}$)		6.2800***	0	0.5285	-111.5832	-108.5018
		(0.6354)				
Gaussian($C_{SM\|T}$)		0.0454	0	0.0289	1.6484	4.7298
		(0.0769)				
Other pair copulas						
Survival ($C_{S,M}$)	Gumbel	1.1854***	0.2054	0.1564	-8.5131	-5.4316
		(0.0650)				
Frank($C_{MT\|S}$)		1.3697***	0	0.1494	-5.7991	-2.7177
		(0.4934)				
Frank($C_{ST\|M}$)		5.8502***	0	0.5062	-100.4416	-98.3602
		(0.6311)				
Dynamic copulas		ω	α	β	AIC	BIC
Frank($C_{MT\|S}$)		16.3448***	0.1852***	95.1764***	-1.4235	1.6578
		(0.4409)	(0.0032)	(2.2623)		
Frank($C_{ST\|M}$)		101.060***	0.4608***	1828.80***	-101.9672	-98.8858
		(7.6314)	(0.0039)	(11.4990)		

Note: Signif. codes are as follows: 0 *** 0.001 ** 0.01 * 0.05 0.1. The numbers in the parentheses are the standard deviations.

where $\Lambda(x) = ln(x)$ is the logistic transformation, $0 =< \alpha =< 1$. The formula for Kendall's tau derived for the Frank copula is

$$\tau = 1 - \frac{4}{\theta} + 4\frac{D_1(\theta)}{\theta} where D_1(\theta) = \frac{1}{\theta}\int_0^\theta \frac{x}{\exp(x)-1}dx \tag{23}$$

The second part of Table 4 reports the results of the time-varying copulas. In terms of AIC and BIC, the time-varying CST—M exhibits better explanatory ability than the static Frank copula, while it is the other way round for $C_{MT|S}$. We can see that the autoregressive parameter α in the time-varying Frank copula $C_{ST|M}$ equals to 0.4608, implying a low degree of persistence pertaining to the dependence structure between Thailand and Singapore inbound tourism demand from China, given Malaysia inbound tourism demand from China. The dynamic Kendall's tau from the Frank copula $C_{ST|M}$ is illustrated in Figure 2, and we can see that the smallest value of Kendall's tau is approximately 0.36, while the greatest is about 0.54, indicating that the nonlinear correlations are always shifting with time, and have seasonal volatility.

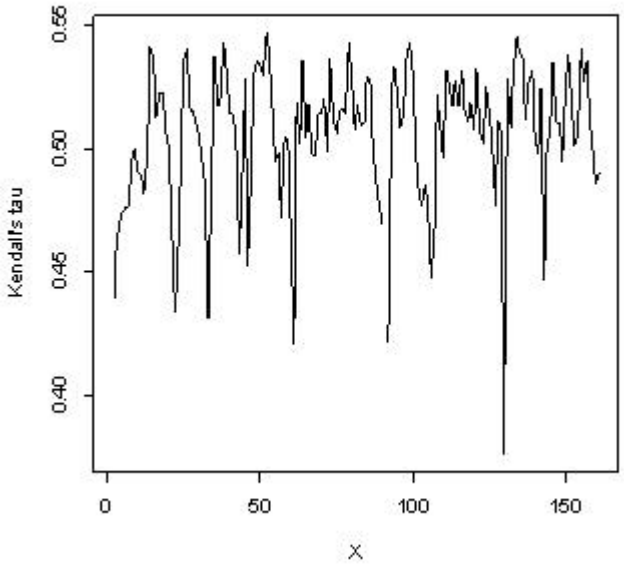


Fig. 2 Kendall's tau from the time-varying Frank copula $C_{ST|M}$

5 Policy Planning

The empirical findings of this study reveal that there exists obvious volatility and interdependency in the Chinese tourist flow to the countries of SMT. Hence, tour operators and national tourism promotion authorities of SMT should collaborate closely in marketing and promoting joint tourism ventures and products.

As for Thailand and Singapore, the Chinese tourist flow to both these countries is highly correlated, especially in December, January, and February every year. In this case, the travel agents and airlines in Thailand and Singapore should come to a strong mutual understanding, cooperate to form a powerful alliance, and launch tourism packages through different routes. Given that the correlation gets reduced in May and June every year, it becomes more important that a series of high quality and low cost travel programs are launched for attracting tourists; as for Thailand and Malaysia, there exists a lower tail correlation in the Chinese tourist flow to these countries, which explains why a negative impact will shock their inbound tourism demand. Therefore, the tourism authorities of Thailand and Malaysia should enhance awareness of prevention, and jointly deploy some tourism program for stimulating the development of the tourism market; as for Singapore and Malaysia, the relevance of Malaysia and Singapore is similar to that of Malaysia and Thailand, and is also lower tail related, so the two countries should also implement similar measures toward meeting unexpected needs.

On the whole, among the three countries, Thailand plays a crucial role. The tourist population of China traveling to Thailand directly impacts the dependency of Singapore and Malaysia, since $C_{SM|T}$ indicates that Malaysia and Singapore are independent in the case of a known tourist flow of Thailand.

6 Conclusions

This paper examined the vine copula-ARMA-GARCH model based on past tourist arrivals from China which is a major tourist source market for SMT. This paper applied separately the logarithm differences of the monthly tourist arrivals to SMT from China. The empirical findings of this study indicate that ARMA-GARCH with assumed skewed student-t distribution for standardized residuals is the best-fitting model to explain the volatility of the tourist flow to Singapore and Malaysia from China, while ARMA-GARCH with assumed skewed-GED distribution for standardized residual is the appropriate model for analyzing the tourist flow to Thailand from China. In addition, various diagnostic cheeks were also used. We discuss how traditional tests for marginal distribution, using the Kolmogorov-Smirnov and Box-Ljung tests, can be implemented to see if the underlying assumptions are satisfied. In addition, fifteen kinds of static copulas were used to analyze the dependence between the tourist flows to the SMT from China. Another point is that we applied the time-varying vine copulas that described the dynamic Kendall's tau. Finally, in the light of the empirical findings, we propose some constructive ideas and policy planning for the attention of the tourism authorities and travel agents.

References

1. Li, X., Lai, C., Harrill, R., Kline, S., Wang, L.: When east meets west: An exploratory study on Chinese outbound tourists travel expectations. Tourism Management 32, 741–749 (2011)
2. Lee, C.G.: The dynamic interactions between hotel room rates and international inbound tourists: Evidence from Singapore. International Journal of Hospitality Management 29, 758–760 (2010)
3. Chang, C., Sriboonchitta, S., Wiboonpongse, A.: Modelling and forecasting tourism from East Asia to Thailand under temporal and spatial aggregation. Mathematics and Computers in Simulation 79, 1730–1744 (2009)
4. Liu, J., Sriboonchitta, S.: Analysis of Volatility and Dependence between the Tourist Arrivals from China to Thailand and Singapore: A Copula-based GARCH Approach. In: Huynh, V.N., Kreinovich, V., Sriboonchitta, S., Suriya, K. (eds.) Uncertainty Analysis in Econometrics with Applications. AISC, vol. 200, pp. 285–296. Springer, Heidelberg (2013)
5. Kim, S.S., Wong, K.F.: Effect of news shock on inbound tourist demand volatility in Korea. Journal of Travel Research 44(4), 457–466 (2006)
6. Song, H., Romilly, P., Liu, X.: An empirical study of outbound tourism demand in the U.K. Applied Economics 32(5), 611–624 (2000)
7. Chan, F., Lim, C., McAleer, M.: Modelling multivariate international tourism demand and volatility. Tourism Management 26, 459–471 (2005)

8. Hoti, S., McAleer, M., Shareef, R.: Modeling international tourism and country risk spillovers for Cyprus and Malta. Tourism Management 28(6), 1472–1484 (2007)
9. Seo, J.H., Park, S.Y., Yu, L.: The analysis of the relationships of Korean outbound tourism demand: Jeju Island and three international destinations. Tourism Management 30, 530–543 (2009)
10. Patton, A.J.: Modelling asymmetric exchange rate dependence. International Economics Review 47(2), 527–556 (2006)
11. Jondeau, E., Rockinger, M.: The copula-GARCH model of conditional dependencies: an international stock market application. Journal of International Money and Finance 25, 827–853 (2006)
12. Lee, T., Long, X.: Copula-based multivariate GARCH model with uncorrelated dependent errors. Journal of Econometrics 150(2), 207–218 (2009)
13. Wu, C.C., Chung, H., Chang, Y.H.: The economic value of co-movement between oil price and exchange rate using copula-based GARCH models. Energy Economics 34(1), 270–282 (2012)
14. Wang, K., Chen, Y.H., Huang, S.W.: The dynamic dependence between the Chinese market and other international stock markets: a time-varying copula approach. International Review of Economics and Finance 20(4), 654–664 (2011)
15. Joe, H., Hu, T.: Multivariate distributions from mixtures of maxinfinitely divisible distributions. Journal of Multivariate Analysis 57(2), 240–265 (1996)
16. Brechmann, E.C., Schepsmeier, U.: Modeling Dependence with C- and D-Vine Copulas: The R Package CDVine. Journal of Statistical Software 52(3), 1–27 (2013)
17. Bedford, T., Cooke, R.M.: Monte Carlo simulation of vine dependent random variables for applications in uncertainty analysis. In: Proceedings of ESREL 2001, Turin, Italy (2001)
18. Bedford, T., Cooke, R.M.: Vines-a new graphical model for dependent random variables. Annals of Statistics 30(4), 1031–1068 (2002)
19. Min, A., Czado, C.: Bayesian inference for multivariate copulas using pair-copula constructions. Journal of Financial Econometrics 8(4), 511–546 (2010)
20. Aas, K., Czado, C., Frigessi, A., Bakken, H.: Pair-copula construction of multiple dependence. Insurance: Mathematics and Economics 44, 182–198 (2009)
21. Czado, C.: Pair-Copula Constructions of Multivariate Copulas. Copula Theory and Its Applications. Lecture Notes in Statistics, vol. 198, pp. 93–109. Springer, Heidelberg (2010)
22. Czado, C., Schepsmeier, U., Min, A.: Maximum likelihood estimation of mixed C-vines with application to exchange rates. Statistical Modelling 12, 229–255 (2012)
23. Bollerslev, T.: Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307–327 (1986)
24. Ling, S.: Self-weighted and local quasi-maximum likelihood estimators for ARMA-GARCH/IGARCH models. Journal of Econometrics 140, 849–873 (2007)
25. Sklar, M.: Fonctions de repartition an dimensions et leurs marges. Publ. Inst. Statist. Univ. Paris 8, 229–231 (1959)
26. Manner, H., Reznikova, O.: A survey on time-varying copulas: Specification, simulations and application. Econometric Reviews 31(6), 654–687 (2012)
27. Ng, W.L.: Modeling duration clusters with dynamic copulas. Finance Research Letters 5, 96–103 (2008)

Vine Copula-Cross Entropy Evaluation of Dependence Structure and Financial Risk in Agricultural Commodity Index Returns

Songsak Sriboonchitta, Jianxu Liu, and Aree Wiboonpongse

Abstract. Many studies used the empirical Kendall's tau to select a preferable ordering of vine copulas or to fix such a sequence. In this study, for high dimension vine copulas, we propose the vine copula based cross entropy method to figure out a more appropriate ordering of the vine copula. The goal of this study is to estimate the non-conditional, conditional, and tail dependences for agricultural price index returns by using the C-vine and D-vine copula based cross entropy model. In addition, we show that a framework uses the Monte Carlo simulation and the results of vine copula to estimate the expected shortfall (ES) of an equally weighted portfolio. The optimal portfolio allocations can also be estimated using global optimization with the differential evolution algorithm.

1 Introduction

Copulas have become an essential tool for measuring dependence structure in finance, economics, etc. Their use has also been extended to include tasks like forecasting dependence, evaluating risks, managing portfolios, and formulating sensible policies. For instance, Sriboonchitta et al. [1] used the time-varying copula based generalized autoregressive conditional heteroskedasticity (GARCH) model to forecast the agriculture price and policy implications, and Huang et al. [2] estimated the value at risk (VaR) of the portfolio by using the conditional copula-GARCH model. Moreover, copulas can be used in high dimension, as propounded by Liebscher [3], and Charpentier and Segers [4]-especially in the case of the pair-copula construction (PCC), also called vine copulas, which uses only bivariate copulas to

Songsak Sriboonchitta · Jianxu Liu
Faculty of Economics, Chiang Mai University, Chiang Mai 50200 Thailand
e-mail: songsakecon@gmail.com, liujianxu1984@163.com

Aree Wiboonpongse
Department of Agricultural Economics and Agricultural Extension, Faculty of Agriculture, Chiang Mai University, Chiang Mai 50200, Thailand

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_18,

construct multivariate copulas. PCC includes two special structures: C-vine and D-vine copulas. Vine copulas have more flexibility than the known classes of multivariate copulas like Gaussian and T. Vine copulas have found extensive application in finance asset returns and other data over the last five years or so, as pointed out by Nikoloulopoulos et al. [5], Aas et al. [6], Gugan and Maugis [7], etc.

There are many different orderings possible of the variables in the C-vine and D-vine models. So, it is crucial that we assign some principles to order the sequence of the variables. Research so far, however, has only been on two main methods which we can follow. One method consists of depending on the research that takes into consideration evidence for deciding the ordering of the variables. The other is of choosing the models with the high dependence in the bivariate condition distribution, as preferred by Aas et al. [6] and Czado et al. [8]. Also, it is possible for us to calculate all the structures of the C-vine or D-vine if there are three or four variables. So, we select the ordering of the vine copula structure only in the descending order of dependence because a high dependence between the variables may have a great impact on the other variables. Let the variable that possesses high correlation be taken as the condition variable, which means we can test the dependence under more effective information as known condition.

Aas et al. [6] proposed that we determine the optimal ordering of the vine copula by using the empirical Kendall's tau method; Aas and Berg [9] chose the most appropriate ordering according to the degree of freedom of the student-t copula, because a low number of degree of freedom indicates strong dependence. Regardless of whether it was the empirical Kendall's tau or the degree of freedom, both of them chose the optimal structure of the vine copula in accordance with the size of dependence. But the size of dependence only reflects the degree of interdependence in terms of rank. There is more information corresponding to the correlation between the two variables, such as lower and upper tail correlations, linear correlations, etc. Therefore, relying solely on the size of dependency while choosing the optimal structure of the vine copula cannot be entirely justified. However, the cross entropy method may measure the information theoretical distance between the two probability distributions of the variables, and the information theoretical distance completely ignores the oneness of the selection criterion. Moreover, the asymmetric property of information distance may enable us to understand which among the two variables plays a more important function.

Previous research, such as those conducted by Engle and Sheppard [10], Rombouts and Verbeek [11], and Chang et al. [12], commonly used the GARCH model to analyze the dependence structures and portfolio management with linear correlation, but it was premised on strict restriction to ensure a well-defined covariance matrix. Moreover, we usually assume that the financial returns follow a multivariate Gaussian or student-t distribution in the multivariate GARCH model, while most of the asset returns possess the characteristics of being skewed, and having high kurtosis and fat tails. Although multivariate student-t distributions can capture high kurtosis and fat tails, it is symmetric, and specifies the same degree of freedom for both or more financial returns. However, vine copulas may effectively measure rank correlation, and lower and upper tail dependence, as well as allow the assumption

of different marginal distributions for the asset returns. Thus, vine copulas may be preferable for measuring value at risk, expected shortfall, optimal portfolio weights, etc. Some relationships that exist between agricultural commodity prices are interesting. For instance, knowing which commodities are positively or negatively correlated with a given commodity is very important for gaining an understanding of the future directional movement of the commodity we propose to trade. Thus, we use a data set with 6 agricultural price indices to measure the dependence, expected shortfall, and portfolio weights.

The main contributions of the paper are three, and are as follows: (1) we provide a comprehensive solution to the quandary of the selection of vine copula orderings by drawing cross entropy into vine copula models; (2) we show how this framework can be used to estimate expected shortfalls and optimal portfolio weights using the results of the copulas and the Monte Carlo simulation method; (3) we construct the optimal portfolio weights of the selected assets under the minimum expected shortfall framework, allowing for global optimization via a Differential Evolution algorithm. The paper is organized as follows. Section 2 provides a brief review on C-vine and D-vine copulas, and introduces the applications of cross entropy, expected shortfall, and optimal portfolios in vine copulas. Section 3 conducts empirical analysis for agricultural commodities corresponding to vine copula based cross entropy model. Finally, section 4 offers conclusions.

2 Methodology

This study uses the auto regression moving average-generalized autoregressive conditional heteroskedasticity (ARMA-GARCH) model to estimate the marginals, and this leads to the formation of standardized residuals. It is worth mentioning that we assumed the marginals to follow the skewed generalized error distribution (SGED). In addition, the Gaussian copula, T copula, Clayton copula, Frank copula, Gumbel copula, Joe copula, BB1 copula, BB6 copula, BB7 copula, BB8 copula, and rotated copulas were candidates in the selection of the best one by using the model comparison criteria, Akaike information criteria (AIC) and Bayesian information criteria (BIC).

2.1 C-vine and D-vine Copulas

The whole idea of vine copulas is that we reduce generic functions of several variables to only functions of two variables, and vine copulas specify the dependence and conditional dependence of selected pairs of random variables and all marginal distribution functions. A few d-dimensional vine copulas are decomposed into d(d-1)/2 pair-copulas, and the densities of the vine copulas are factorized in terms of pair-copulas and marginals.

For the C-vine and D-vine copulas, the densities are, respectively (Aas et al. [6]),

$$f(x_1,x_2,\ldots,x_d)=\prod_{k=1}^{d}f(x_k)\times\prod_{j=1}^{d-1}\prod_{i=1}^{d-j}c_{j,i+j|1,\cdots,j-1}(F(x_j|x_\Phi),F(x_{i+j}|x_\Phi)) \quad (1)$$

and

$$f(x_1,x_2,\ldots,x_d)=\prod_{k=1}^{d}f(x_k)\times\prod_{j=1}^{d-1}\prod_{i=1}^{d-j}c_{i,i+j|i+1,\cdots,i+j-1}(F(x_i|x_\Psi),F(x_{i+j}|x_\Psi)) \quad (2)$$

where $\Phi=1,\ldots,j-1$ and $\Psi=i+1:i+j-1$.For d = 6, for example, the C-vine density (1) can be written as

$$\begin{aligned}f(x_1,\cdots,x_6)=&\prod_{i=1}^{6}f(x_i)\cdot c_{12}(F(x_1),F(x_2))\cdot c_{13}(F(x_1),F(x_3))\\&\cdot c_{14}(F(x_1),F(x_4))\cdot c_{15}(F(x_1),F(x_5))\cdot c_{16}(F(x_1),F(x_6))\\&\cdot c_{23|1}(F(x_{2|1}),F(x_{3|1}))\cdot c_{24|1}(F(x_{2|1}),F(x_{4|1}))\\&\cdot c_{25|1}(F(x_{2|1}),F(x_{5|1}))\cdot c_{26|1}(F(x_{2|1}),F(x_{6|1}))\\&\cdot c_{34|12}(F(x_{3|12}),F(x_{4|12}))\cdot c_{35|12}(F(x_{3|12}),F(x_{5|12}))\\&\cdot c_{36|12}(F(x_{3|12}),F(x_{6|12}))\cdot c_{45|123}(F(x_{4|123}),F(x_{5|123}))\\&\cdot c_{46|123}(F(x_{4|123}),F(x_{6|123}))\cdot c_{56|1234}(F(x_{5|1234}),F(x_{6|1234}))\end{aligned} \quad (3)$$

For d = 6, for example, the D-vine density (2) can be written as

$$\begin{aligned}f(x_1,\cdots,x_6)=&\prod_{i=1}^{6}f(x_i)\cdot c_{12}(F(x_1),F(x_2))\cdot c_{23}(F(x_2),F(x_3))\\&\cdot c_{34}(F(x_3),F(x_4))\cdot c_{45}(F(x_4),F(x_5))\cdot c_{56}(F(x_5),F(x_6))\\&\cdot c_{13|2}(F(x_{1|2}),F(x_{3|2}))\cdot c_{24|3}(F(x_{2|3}),F(x_{4|3}))\\&\cdot c_{35|4}(F(x_{3|4}),F(x_{5|4}))\cdot c_{46|5}(F(x_{4|5}),F(x_{6|5}))\\&\cdot c_{14|23}(F(x_{1|23}),F(x_{4|23}))\cdot c_{25|34}(F(x_{2|34}),F(x_{5|34}))\\&\cdot c_{36|45}(F(x_{3|45}),F(x_{6|45}))\cdot c_{15|234}(F(x_{1|234}),F(x_{5|234}))\\&\cdot c_{26|345}(F(x_{2|345}),F(x_{6|345}))\cdot c_{16|2345}(F(x_{1|2345}),F(x_{6|2345}))\end{aligned} \quad (4)$$

The vine copulas involve marginal conditional distributions that can be expressed by h function. Assume that u_1 and u_2 are uniform, i.e. $f(u_1)=f(u_2)=1, F(u_1)=u_1$ and $F(u_2)=u_2$. Then the univariate conditional distribution is

$$h(u_1|u_2;\theta):=F(u_1|u_2;\theta)=\frac{\partial C_{u_1,u_2}(u_1,u_2;\theta)}{\partial u_2} \quad (5)$$

where θ is the parameter vector for C_{u_1,u_2}. If there are two conditional variables, the h function can be written as

$$h(u_1|u_2,u_3;\theta) = F(u_1|u_2,u_3;\theta) = \frac{\partial C_{u_1,u_3|u_2}(F(u_1|u_2),F(u_3|u_2);\theta)}{\partial F(u_3)} \tag{6}$$

In general, we use the form $h(u|\upsilon;\theta)$ to represent the marginal conditional distributions. Joe [13] showed that

$$h(u|\upsilon;\theta) := F(u \mid \upsilon) = \frac{\partial C_{u,\upsilon_j|\upsilon_{-j}}(F(u \mid \upsilon_{-j}),F(\upsilon_j \mid \upsilon_{-j}))}{\partial F(\upsilon_j \mid \upsilon_{-j})} \tag{7}$$

where $C_{u,\upsilon_j|\upsilon_{-j}}$ is the dependency structure of the bivariate conditional distribution of u and υ_j conditioned on υ_{-j}, and the vector υ_{-j} is the vector υ excluding the component υ_j (see Aas et al. [6]).

In this study, we follow the estimation method of Aas et al. [6]-who used sequential estimates as starting values and then estimate the vine copulas through the maximum likelihood estimation method, again.

2.2 Minimum Cross Entropy

Entropy has been extended for application to the fields of econometrics, extreme value, etc., as done by Golan [14], Zellner and Tobias [15], Pandey [16], and others. Kullback [17] proposed an information-theoretic distance D between the two probability distributions. This information theoretic distance is known as directed divergence, or cross entropy.

Assume that the uniforms of a six-dimensional vine copula are U, V, W, X, Y, and Z that can be formed from the ARMA-GARCH model. Let $U_i' = U_i / \sum_{i=1}^{T} U_i$ be the probability distribution. In the same way, if we transform V, W, X, Y, and Z into V_i', W_i', X_i', Y_i' and Z_i', then the distance between U_i' and V_i' can be expressed as

$$D(U',V') = \sum_{i=1}^{T} U_i' \log(U_i'/V_i') \tag{8}$$

where $D(U',V') \neq D(V',U')$, so the cross entropy is asymmetric. Lind [18] proved that a consequence of minimum cross entropy is that minimum information is inferred from the prior distribution. Thus, we solve the minimum information theoretic distance D between one variable and the other variables, implying that the two variables are the closest. It is worth mentioning that the closest distance does not always mean the maximum empirical Kendall's tau, linear correlation, or tail dependence. In addition, the property of asymmetry explains that two markets or variables are never in the same status, and that the smaller has more impact on the other market or variable.

The process of selecting the ordering of the C-vine copula is similar to the maximum empirical Kendall's tau method. We calculate the sum of the cross entropies between one variable and the remaining variables, and select the minimum as the first variable. Similarly, we reorder the remaining variables and repeat the process of calculating the sum of the cross entropies, thus finding the second variable, third

variable, and so on. For the ordering of the D-vine copula, let the minimum pair be the first two variables; then we calculate the cross entropy between the first two variables and the other variables, using the minimum pair as the second and third variables. The rest of the procedure can be done in the same manner.

2.3 ES and Optimal Portfolio

To extend the economically useful application of the vine copula model, we use the Monte Carlo simulation for the vine copula model to calculate the expected shortfall of an equally weighted portfolio. After that, the optimal portfolio weights of the selected assets are constructed under a minimum expected shortfall framework, using global optimization with the differential evolution algorithm.

The method for calculating the expected shortfall consists of, to summarize, four steps. First, we use the best vine copula to generate the random number whose length is sample size N. Second, we plug the random number into inverse functions of the probability distributions of the random variables, such as the skewed generalized error distribution in this study, and employ the mean and variance equations of the ARMA-GARCH model to get the N values of each variable at period $t+1$. Third, we distribute equal weights to each variable, and get the totaling variable by adding them up. Last, the expected shortfall is calculated at 100%, 10%, 5%, 2%, and 1% levels, and then we repeat the first three steps 1000, 2000 and 5000 times for getting the convergence values.

Now consider an investor who wants to minimize the ES at 100%, 10%, 5%, 2%, and 1%, subject to achieving a particular expected return. Let w_i be the weight vector of the portfolio weights of the risky assets. The investor solves the following optimization problem:

$$\begin{aligned}
&\text{Min } ES = E[r|r \le r_\alpha] \\
&\text{subject to} \\
&r = w_1 \times r_{1,t+1} + w_2 \times r_{2,t+1} + \cdots + w_d \times r_{d,t+1} \\
&w_1 + w_2 + \cdots + w_d = 1 \\
&0 \le w_i \le 1,\ i = 1,2,\ldots,d
\end{aligned}$$

where r_α is the lower α-quantile, and $r_{i,t+1}$ represents the asset return of the variable i at period t+1. Global optimization can then solve this problem with maximum iterations to be 30 and with only 10 repetitions. Even at this small simulation scale, the estimated weights still converge.

3 Data and Empirical Results

3.1 Data, KS, and LM Tests

This study uses the Dow Jones-UBS subindices for the prominent agricultural commodities, including coffee, corn, cotton, soybean, sugar, and wheat. Our sample covers the period from January 1, 2008, to January 14, 2013, and, to eliminate the

spurious correlation arising from holidays, we drop those observations for any holidays associated with the least one index. The return series are 100 times the log-difference of the commodity indices. Table 1 contains descriptive statistics; these statistics show that the data for each variable are not a normal distribution, since the Jarque-Bera test rejects the null hypothesis for each index return. All skewness values are less than 0, and the values of kurtosis are greater than 3, thereby implying that all the variables are skewed to the left, high kurtosis, and fat-tailed.

For each data set, we use the ARMA-GARCH process to estimate the marginals, and assume that all the marginals are skewed generalized error distributions. The soybean, wheat, and sugar data sets follow the ARMA (0, 0)-GARCH (1, 1) process; the corn data set follows the ARMA (1, 1)-GARCH (1, 1) process; the coffee data set follows the ARMA (1, 0)-GARCH (1, 1) process; and the cotton data set follows the ARMA (0, 1)-GARCH (1, 1) process. Next, we have to ensure that all the marginals are uniform distributions and that the i.i.d lies between 0 and 1. Two kinds of tests, the Kolmogorov-Smirnov test (KS-test) and the Lagrange multiplier test (LM test), are performed. Table 2 shows that none of the marginals rejects the null hypotheses of the KS and LM tests at the 5% level. Thus, we can safely assume that the marginals are skewed generalized error distributions, which is feasible and appropriate.

Table 1 Data Description and Statistics

	Coffee	Corn	Cottton	Soybean	Sugar	Wheat
Mean	−0.0220	0.0104	0.0010	0.0239	0.0245	−0.0720
Median	0.0000	0.0000	0.0427	0.0181	0.0002	−0.0481
Maximum	7.5106	8.6624	6.9333	6.4351	8.1859	8.7966
Minimum	−11.2407	−8.1229	−7.1241	−7.3371	−12.3654	−9.9713
Std. Dev.	1.9145	2.1785	2.0603	1.7850	2.4379	2.4578
Skewness	−0.2111	−0.0113	−0.1550	−0.2829	−0.4248	−0.0559
Kurtosis	4.5772	4.2833	3.4610	4.6729	4.6770	4.2077
Jarque-Bera	143	88	16	168	190	79
Probability	0.0000	0.0000	0.0002	0.0000	0.0000	0.0000

3.2 The Ordering of Vine Copulas Based Cross Entropy

Table 3 and Table 4 show all the empirical Kendall's tau for each variable and the sums of each variable with the other variables; Table 5 and Table 6 are for the results of the cross entropy method. Apparently, the ordering of the C-vine copula in the methods of minimum cross entropy and maximum empirical Kendall's tau are V, U, W, X, Y, Z and V, U, X, W, Y, Z, respectively. For the D-vine copula, both the methods have the same ordering, that is, W, V, U, Y, X, and Z. We calculate the sum of the AIC and BIC for each ordering. The AIC and BIC are −2498.842 and −2364.519 as per the maximum empirical Kendall's tau method, while they

Table 2 KS Test and ARCH LM Test

	Coffee	Corn	Cottton	Soybean	Sugar	Wheat
KS test statistics	0.0221	0.0089	0.0180	0.0331	0.0235	0.0242
P value	0.5544	1	0.7937	0.1167	0.4714	0.4329
P value of LM test						
First moment	0.6422	0.7578	0.8158	0.6001	0.1715	0.1622
Second moment	0.8583	0.0948	0.3927	0.4839	0.3623	0.9671
Third moment	0.8655	0.0659	0.6133	0.4046	0.3221	0.8527
Fourth moment	0.8145	0.0738	0.7436	0.3263	0.2969	0.6921

Table 3 Empirical Kendall's tau for Ordering Sequences in Vine Copulas (1)

	U	V	W	X	Y	Z	SUM	SUMV
U	1	0.4388	0.3747	0.2093	0.2152	0.2039	2.4419	2.0031
V	0.4388	1	0.5141	0.2029	0.2112	0.2247	2.5917	—
W	0.3747	0.5141	1	0.1977	0.2063	0.1955	2.4882	1.9741
X	0.2093	0.2029	0.1977	1	0.1994	0.2339	2.0432	1.8403
Y	0.2152	0.2112	0.2063	0.1994	1	0.1903	2.0224	1.8112
Z	0.2039	0.2247	0.1955	0.2339	0.1903	1	2.0482	1.8235

Note:U, V, W, X, Y, and Z represent soybean, corn, wheat, coffee, cotton, and sugar, respectively. SUMV represents the sum cross entropy value that excludes the variable V.

Table 4 Empirical Kendall's tau for Ordering Sequences in Vine Copulas (2)

	W	X	Y	Z	SUMVU	SUMVUX	SUMVUXW
W	1	0.1977	0.2063	0.1955	1.5995	1.3978	—
X	0.1977	1	0.1994	0.2339	1.6310	—	—
Y	0.2063	0.1994	1	0.1903	1.5960	1.3966	1.1903
Z	0.1955	0.2339	0.1903	1	1.6196	1.3857	1.1903

Note: SUMVU,SUMVUX and SUMVUXW have the same meaning with SUMV.

equal to -2499.947 and -2365.651 as per the ordering of minimum cross entropy method. Therefore, it is explicitly evident that the minimum cross entropy method is more appropriate than and preferable to the maximum empirical Kendall's tau in selecting the sequence of the many different orderings of high dimension vine copulas. Moreover, the asymmetric differences in the first row are all positive, implying that soybean clearly has a more important status in comparison with the other agricultural commodities.

Table 5 Minimum Cross Entropy Values for Ordering Sequences in Vine Copulas (1)

	U	V	W	X	Y	Z	Sum	SumV
U	0	0.1798 (0.0027)	0.2229 (0.0046)	0.3308 (0.0010)	0.3290 (0.0110)	0.3416 (0.0188)	1.4042	1.2244
V	0.1771	0	0.1423 (0.0028)	0.3319 (−0.0101)	0.3326 (0.0091)	0.3180 (−0.0086)	1.3019	—
W	0.2183	0.1395	0	0.3334 (0.0266)	0.34090 (0.0045)	0.3443 (−0.0055)	1.3765	1.2370
X	0.3298	0.3420	0.3600	0	0.3465 (0.0105)	0.3124 (−0.0024)	1.6907	1.3487
Y	0.3180	0.3235	0.3364	0.3360	0	0.3595 (0.0014)	1.6734	1.3499
Z	0.3228	0.3266	0.3498	0.3148	0.3609	0	1.6749	1.3483

Note: SumV represents the sum cross entropy value that excludes the variable V. The numbers in the parentheses are the asymmetric differences.

Table 6 Minimum Cross Entropy Values for Ordering Sequences in Vine Copulas (2)

	W	X	Y	Z	SumVU	SumVUW	SumVUWX
W	0	0.3334	0.3409	0.3443	1.0187	—	—
X	0.3600	0	0.3465	0.3124	1.0189	0.6589	—
Y	0.3364	0.3360	0	0.3595	1.0319	0.6955	0.3595
Z	0.3498	0.3148	0.3609	0	1.0255	0.6757	0.3609

Note: SumVU,SumVUW and SumMVUWX have the same meaning with SumV.

3.3 Estimation Results

Table 7 and Table 8 report the parameter estimates for the C-vine and D-vine copulas, respectively. 1, 2, 3, 4, 5, and 6, in proper turn, represent the variables of V, U, W, X, Y, and Z. First, it has to be noted that the fitting pair-copula families of the C-vine copula structure are T, T, Gaussian, Survival BB1, T, BB7, T, T, T, Frank, Survival Gumbel, Frank, Survival BB8, Survival BB1, and BB8 in the many different copula families, and T, T, Survival BB7, T, BB1, BB7, Frank, T, Frank, Survival Gumbel, Gaussian, Survival BB8, Frank, T, and Survival Clayton for the D-vine copula structure. Second, when we compare the sum values of the AIC and BIC of the C-vine copula with those of the D-vine copula, we find that the C-vine copula structure offers a better performance. Third, corn and wheat have the maximum Kendall's tau, and corn and cotton have the greatest upper tail dependence; also, the lower and upper tail dependences between corn and soybean are symmetric, of which one has the greatest lower tail dependence among all the pair-copulas. There is more exact information in Table 7 and Table 8. Specifically, when we compare

Table 7 Results of C-vine Copulas and Kendall's tau

Copulas	parameters	Lower and upper tails	Kendall'tau	AIC	BIC
T (C_{12})	0.6424***	0.2671	0.4442	−736.1186	−725.7860
	(0.0165)				
	5.7249***	0.2671			
	(1.105)				
T (C_{13})	0.7196***	0.2496	0.5114	−955.5787	−945.2462
	(0.0125)				
	8.2569***	0.2496			
	(1.9489)				
Gaussian ($C_{1,4}$)	0.3292***	0	0.2136	−138.5584	−133.3922
	(0.0229)				
Survival BB1 (C_{15})	0.1324***	0.1997	0.2087	−152.8361	−142.5035
	(0.0460)				
	1.1789***	0.5554			
	(0.0304)				
T (C_{16})	0.3356***	0.0194	0.2179	−156.8095	−146.4769
	(0.0247)				
	13.0129***	0.0194			
	(4.3773)				
BB7 ($C_{23\|1}$)	1.1298***	0.0061	0.1244	−56.0286	−45.6961
	(0.0342)				
	0.1361***	0.1531			
	(0.03949)				
T ($C_{24\|1}$)	0.1782***	0.0022	0.1141	−39.6613	−29.3288
	(0.0277)				
	17.1412*	0.0022			
	(10.5062)				
T ($C_{25\|1}$)	0.1847***	0.0009	0.1183	−40.4455	−30.1229
	(0.0275)				
	20.4298	0.0009			
	(12.1983)				
T ($C_{26\|1}$)	0.1286***	0.0053	0.0821	−23.7662	−13.4377
	(0.0287)				
	13.0391***	0.0053			
	(5.1834)				
Frank($C_{34\|12}$)	0.4856***	0	0.0538	−6.7844	−1.6181
	(0.1641)				
Survival Gumbel($C_{35\|12}$)	1.0540***	0.0698	0.0512	−16.2294	−11.0632
	(0.0173)				
Frank($C_{36\|12}$)	0.3691***	0	0.0409	−3.3367	1.8295
	(0.1662)				
Survival BB8 ($C_{45\|123}$)	1.9190***	0	0.1309	−50.1434	−39.8109
	(0.5764)				
	0.6749***	0			
	(0.1980)				
Survival BB1 ($C_{46\|123}$)	0.0961***	0.1625	0.1625	−96.6443	−86.3118
	(0.0453)				
	1.1393***	0.5442			
	(0.0297)				
BB8 ($C_{56\|1234}$)	1.807***	0	0.0969	−27.0332	−16.7006
	(0.2839)				
	0.7841***	0			
	(0.1590)				
Sum				−2499.974	−2365.651

Note: Signif. codes are as follows: 0 *** 0.001 ** 0.01 * 0.05. The numbers in the parentheses are the standard deviations.

Table 8 Results of D-vine Copulas and Kendall's tau

Copulas	parameters	Lower and upper tails	Kendall'tau	AIC	BIC
T (C_{31})	0.7080*** (0.0131)	0.2389	0.5008	−954.7202	−944.3876
	8.2663*** (1.8606)	0.0.2389			
T (C_{12})	0.5660*** (0.0184)	0.1405	0.3830	−532.4336	−522.1010
	8.3539*** (2.0874)	0.1405			
Survival BB7 (C_{25})	1.1834*** (0.0349)	0.2036	0.1867	−137.8274	−127.4949
	0.2526*** (0.0432)	0.0643			
T (C_{54})	0.3054*** (0.0247)	0.0025	0.1976	−127.0945	−116.7620
	20.9189 (18.2367)	0.0025			
BB1 (C_{46})	0.2007*** (0.0450)	0.0441	0.1786	−114.9425	−104.6100
	1.1064*** (0.0277)	0.1289			
BB7 ($C_{23\|1}$)	1.3049*** (0.0453)	0.1385	0.2562	−241.3642	−231.0316
	0.3507*** (0.0530)	0.2991			
Frank($C_{15\|2}$)	0.8702*** (0.1725)	0	0.0960	−23.8892	−18.7229
T ($C_{24\|5}$)	0.1430*** (0.0282)	0.0003	0.0913	−22.4205	−12.0880
	22.3778*** (7.8860)	0.0003			
Frank($C_{56\|4}$)	1.1137*** (0.1728)	0	0.1222	−44.8028	−39.63653
Survival Gumbel($C_{35\|12}$)	1.1007*** (0.0207)	0.1228	0.0543	−34.5325	−29.3663
Gaussian ($C_{14\|25}$)	0.1458*** (0.0268)	0	0.0932	−29.9344	−24.7681
Survival BB8 ($C_{26\|45}$)	1.3410*** (0.3912)	0	0.0690	−13.1723	−2.8398
	0.7803*** (0.3201)	0			
Frank($C_{34\|125}$)	0.1964*** (0.1699)	0	0.1312	−48.4655	−43.2992
T ($C_{16\|245}$)	0.1832*** (0.0274)	0.0017	0.1173	−35.6728	−25.3403
	18.2518** (7.9812)	0.0017			
S-Clayton ($C_{36\|1245}$)	0.2482*** (0.0396)	0.0612	0.1104	−49.1264	−43.9602
Sum				−2410.399	−2286.409

Note: Signif. codes are as follows: 0 *** 0.001 ** 0.01 * 0.05. The numbers in the parentheses are the standard deviations.

Table 9 Expected Shortfall of Equally Weighted Portfolios

ES	100%	10%	5%	2%	1%
1,000 times	0.0025	−1.310	−1.572	−1.897	−2.129
2,000 times	0.0033	−1.309	−1.572	−1.897	−2.127
5,000 times	0.0032	−1.309	−1.572	−1.897	−2.127

Table 10 Optimal Portfolio Weights Based on Minimum ES with MC Simulation Given Copulas

	w_1	w_2	w_3	w_4	w_5	w_6	O.P.R
100%	0.1377	0.1716	0.2059	0.1260	0.2063	0.1525	0.0098
10%	0.1433	0.0833	0.1564	0.3153	0.1075	0.1942	−1.226
5%	0.1083	0.1299	0.1592	0.3334	0.1345	0.1347	−1.465
2%	0.1245	0.1430	0.1524	0.2908	0.1098	0.1795	−1.764
1%	0.1509	0.1372	0.1138	0.3106	0.1229	0.1646	−1.981

some dependences with the conditional dependence, the structures and families may be seen to have changed, such as C_{25} and $C_{25|1}$, C_{15} and $C_{15|2}$, etc.

We use the first tree of the C-vine copula to calculate the ES and optimal portfolio weights since the C-vine copula is the best-performing model in terms of both the information criteria. Table 9 presents the ES at levels of 100%, 10%, 5%, 2%, and 1%. We can see that the estimated ES converges to 0.003, −1.31, −1.57, −1.90, and −2.13 at period t+1 at 100%, 10%, 5%, 2%, and 1% levels, respectively. Table 10 is a report of the optimal portfolio weighting estimates at period t+1. As long as we invest in strict accordance with the optimal portfolio weights, the ES will mitigate risk by 6.41%, 6.69%, 7.16%, and 7.00% at 10%, 5%, 2%, and 1%, respectively. The ES at the 100% level represents the mean. Once we begin investing the optimal weights, the interest is bound to increase dramatically 2.27 times. This is clear evidence of the strategys hedging potential.

4 Conclusions

First and foremost, this paper proposes the vine copula based cross entropy model as more suitable and preferable to select the ordering and estimate dependence, conditional dependence, and tail dependence. In addition, we extend the application of vine copula to estimate the ES and optimal portfolios by using the Monte Carlo simulation and global optimization, which provides new and interesting risk management strategies for managers and investors working with high dimensional portfolios. Moreover, the empirical analysis of the agricultural commodities shows

that soybean evidently has a more important status in comparison with the other agricultural commodities. Another of the significant observations is that the C-vine copula structure performs better than the D-vine copula.

References

1. Sriboonchitta, S., Nguyen, H.T., Wiboonpongse, A., Liu, J.: Modeling volatility and dependency of agricultural price and production indices of Thailand: Static versus time-varying copulas. International Journal of Approximate Reasoning 54, 793–808 (2013)
2. Huang, J.J., Lee, K.J., Liang, H., Lin, W.F.: Estimating value at risk of portfolio by conditional copula-GARCH method. Insurance: Mathematics and Economics 45, 315–324 (2009)
3. Liebscher, E.: Construction of asymmetric multivariate copulas. Journal of Multivariate Analysis 99, 2234–2250 (2008)
4. Charpentier, Segers, J.: Tails of multivariate Archimedean copulas. Journal of Multivariate Analysis 100, 1521–1537 (2009)
5. Nikoloulopoulos, A.K., Joe, H., Li, H.: Vine copulas with asymmetric tail dependence and applications to financial return data. Computational Statistics and Data Analysis 56, 3659–3673 (2012)
6. Aas, K., Czado, C., Frigessi, A., Bakken, H.: Pair-copula constructions of multiple dependence. Insurance: Mathematics and Economics 44, 182–198 (2009)
7. Gugan, D., Maugis, P.A.: An Econometric Study of Vine Copulas. International Journal of Economics and Finance 2(5), 2–14 (2011)
8. Czado, C., Schepsmeier, U., Min, A.: Maximum likelihood estimation of mixed C-vines with application to exchange rates. Statistical Modelling 12, 229–255 (2012)
9. Kurowicka, D., Joe, H.: Dependence modeling, pp. 305–328. World Scientific Publishing, Printed in Singapore (2011)
10. Engle, R.F., Sheppard, K.: Theoretical and Empirical properties of Dynamic Conditional Correlation Multivariate GARCH. National Bureau of Economic Research (2001)
11. Jeroen, V.K., Rombouts, Verbeek, M.: Evaluating Portfolio Value-at-Risk using Semi-Parametric GARCH Models. Quantitative Finance, Taylor and Francis Journals 9(6), 737–745 (2009)
12. Chang, C.L., McAleer, M., Tansuchat, R.: Crude oil hedging strategies using dynamic multivariate GARCH. Energy Economics 33(5), 912–923 (2011)
13. Joe, H., Hu, T.: Multivariate distributions from mixtures of max-infinitely divisible distributions. Journal of Multivariate Analysis 57(2), 240–265 (1996)
14. Golan, A.: Information and EntropyEconometrics-Editors View. Journal of Econometrics 107, 1–15 (2002)
15. Zellner, A., Tobias, J.: Further Results on Bayesian Method of Moments Analysis of the Multiple Regression Model. International Economic Review 42(1), 121–140 (2001)
16. Pandey, M.D.: Minimum cross-entropy method for extreme value estimation using peaks-over-threshold data. Structural Safety 23, 345–363 (2001)
17. Kullback, S.: Information theory and statistics. Wiley, New York (1959)
18. Lind, N.C.: The information-theoretical methods to estimate a random variable. J. Environmental Management 49, 43–51 (1997)

A Study on Whether Economic Development and Urbanization of Areas Are Associated with Prevalence of Obesity in Chinese Adults: Findings from 2009 China Health and Nutrition Surveys

Jing Dai, Songsak Sriboonchitta, Cheng Zi*, and Yunjuan Yang

Abstract. China's economy has experienced rapid development in the past 20 years. In 2010, China's GDP was valued at $5.87 trillion, surpassing Japan's $5.47 trillion, and the nation became the world's second largest economy after the USA. People's incomes are also rapidly rising in all parts of the country. However, along with the prosperity seems to have come a malady that is the modern world's woe: obesity. In China, the prevalence of obesity has increased dramatically. Obesity and its related diseases lay a heavy burden on medical expenditure and constrain economic development. Therefore, it is urgent and imperative to identify those influencing factors related to obesity, and take some measures and make corresponding, appropriate policies to control its prevalence. The objective of this study is to identify the impact factors of obesity from different levels, and to evaluate whether the relationship between urbanization and obesity can be explained by individual socio-demographic, socioeconomic factors and lifestyle habits. Three-level logistic models are used in this paper to evaluate the relationship between each indicator and obesity.

Jing Dai
Kunming University of Science and Technology,
Yunnan, China, and Faculty of Economics,
Chiang Mai University, Chiang Mai, Thailand
e-mail: jdai3126@gmail.com

Cheng Zi
Kunming University of Science and Technology, Yunnan, China
e-mail: 15418080@qq.com

Songsak Sriboonchitta
Faculty of Economics, Chiang Mai University, Chiang Mai, Thailand
e-mail: songsak@econ.cmu.ac.th

Yunjuan Yang
Yunnan Center of Disease Control and Prevention, Yunnan, China
e-mail: yncdcyyj@126.com

* Corresponding author.

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_19,

Our findings show that there is a strong relationship between obesity and some of the individual socioeconomic factors and living habits, and that this relationship is mediated by the characteristics of households and areas, as well. The empirical findings of this study can be used to develop more effective ways of intervention and strategies for obesity prevention in the specific context of each of the different regions.

1 Introduction

China is currently undergoing rapid economic development, demographic transformation and large-scale urbanization. Since the reform and the opening up in 1978, the average living standards have experienced a sustained and rapid growth. The gross domestic product per capita has risen by about 220% with an increase in annual growth of around 7%. The rapid increase in productivity has resulted in higher incomes and an ample food supply. However, at the same time, some worrying trends have been detected, such as the fast pace of urbanization, dietary transition with an inclination toward Western-style, fewer fitness activities, serious food safety problems, and environmental pollution. All of these lead to many health problems, and more and more people suffer from overweight and obesity. Obesity and overweight have been recognized as major worldwide public health concerns, jeopardizing the health of adults as well as children [31]. Excess body fat is associated with increased morbidity, disability, and premature mortality from cardiovascular diseases, diabetes, cancers, and musculoskeletal disorders.

Nowadays, the obesity epidemic is a public health challenge not only for developed countries but also for developing countries, such as China, India, Mexico, and Thailand. In developing countries, although the obesity rate is lower compared to the same in the U.S and the Western countries, the increase in the absolute volume of obese population is striking. In China, the number of overweight and obese adults and children has continued to grow dramatically in the recent years. In 2010, more than 340 million adults were diagnosed as overweight, of whom at least 30 million were certified clinically obese. Among the Chinese adults aged from 18 to 75 years, the prevalence of overweight has surged from 14.6% in 1992 to 45.38% in 2011. In the same population, obesity has nearly tripled from approximately 5.2% in 1992 to 15.06% in 2011 [29]. It is now estimated that a fifth of the overweight and obese individuals in the world are located in China. The sudden onslaught of obesity has led to an increase in the prevalence of chronic diseases. According to some studies, in China, the direct and indirect economic loss of CNDs related to overweight and obesity is up to 23.5 billion Yuan in 2010, and these losses are continually on the increase [9]. Thus, the medical cost of overweight and obesity in China is enormous. If no actions are taken on obesity prevention, and the opportunities for the prevention of chronic disease are missed, even worse increases in medical costs can be expected in the near future [30].

In recent years, the problem of obesity has gradually drawn much more attention from the whole society. Identifying the risk factors at the individual level as well

as the regional level is becoming important, since with more information on it, we can guide people better to change their unhealthy lifestyle and diet habits to avoid becoming overweight. Meanwhile, at the community level, related policies can be chalked out and measures taken to change the related social environment in order to reduce the prevalence of obesity.

The objective of this paper is to identify the risky factors at the individual level, household level and community level, and evaluate whether the relationship between area-based socio-economic environment and obesity can be explained by individual socio-demographic, and socio-economic factors and living habits.

The remainder of the paper is organized as follows. Section 2 describes the data source we used. Section 3 discusses the multilevel analysis model. The estimated models and the empirical results are discussed in section 4. Finally, some concluding remarks are given in the last section.

2 Literature Review

The prevalence of obesity has increased dramatically worldwide, and its dangers and health hazards have been the focus of attention of many scholars. Manson suggested that obesity should be paid much more attention since there is a strong relationship between overweight and the development of chronic illnesses such as diabetes, cardiovascular disease, osteoarthritis, and some cancers [11]. Sturm found that the rising obesity rates may be a greater threat to public health than even the smoking or drinking related problems [23].

Upon reviewing relevant literatures, I found many of the previous studies on obesity come mainly from the fields of medicine, biology and public health. These studies reported various factors that were associated with obesity risk from the perspectives of medicine, biology, genetic, and epidemiology [2] [12] [16] [24] . However, the finer details and causes of mechanism of obesity and hypertension still remains unresolved.

With the participation of social scientists and economists, researches dedicated to explicating the multifaceted relationship between the socio-economic determinants of obesity prevalence have been on the risen. Philipson and Posner pointed out that an unhealthy lifestyle had negative influences on the Body Mass Index (BMI), for example, increased reliance on technology in people's daily lives has promoted sedentary lifestyles which, in turn, lead to weight gain [19]. Drewnowski and Specter found the highest obesity rates among the poorest and the least educated members of the industrialized societies [6]. Monteiro demonstrated that obesity is increasing the fastest in the sub-populations with low socio-economic status [18]. McLaren pointed out consistent evidence of a negative relationship between obesity and socio-economic status of women [15]. Yoon found that there was a strong positive relationship between high income and obesity in males, but not in females [27]. Ross pointed out that living in disadvantaged regions may affect BMI and obesity [21]. Lovasi et al also found that differences in area facilities, such as the availability and price of healthy food and the absence of parks and sports and recreational

facilities, may give rise to area differences in dietary intake and physical inactivity [10]. In Cranes research, social contagion models suggest that people's behaviour is influenced by the norms or values of those around them [5]. Matheson examined the impact of neighbourhood material deprivation on gender differences in BMI for urban Canadians, and found that living in neighbourhoods with higher material deprivation was associated with higher BMI [13]. Corsi used a multilevel perspective to investigate the importance of local geographical context in shaping the BMI in low and middle income countries, and the results showed that in countries with greater neighbourhood variation it is possible that the BMI is being influenced by local conditions more than in countries with lesser neighbourhood variation [4].

However, most of the studies have been conducted on the developed countries, such as thee US, Canada, Britain, etc., and very limited research has been done on the developing countries, whose economic and social environments are very different from the developed countries.In addition, we found that in China, most studies about obesity mainly focus on individual influencing factors and neglect the impacts from the particular region. This will lead to deficiencies in policy formulation in that the policies will only consider individual factors and overlook the diversifications at the regional level. Therefore, more studies which consider both micro-level and macro-level factors should be conducted to make the socio-economic determinants of obesity in China become clear. This will provide the policy makers with better and comprehensive information to formulate both macro and micro policies targeting the prevention of obesity, and the related chronic non-communicable diseases (CNDs).

To summarize, there are three main reasons for conducting this research. First, the quick prevalence of obesity is becoming an urgent problem challenging public health in China, and they place a heavy burden on the national medical expenditures and constrain the nation's socio-economic development. Second, there are limited researches about the socio-economic determinants of obesity in China. So the impacts of the socio-economic factors of obesity are not clear. Finally, most existing studies in China mainly focus on the determinants at an individual level and overlook the influence from the higher levels, such as the household level and the community level. This leads to inefficiency and inadequacy of policies that target the prevention of obesity.

3 Data Sets Introduction

The data we used are from the China Health and Nutrition Survey (CHNS) conducted in 2009. CHNS was conducted by an international team of researchers whose backgrounds include nutrition, public health, economics, sociology, Chinese studies, and demography. This survey was designed to examine the effects of the health, and nutrition, programs implemented by the national and local governments and to see how the social and economic transformation of the Chinese society is affecting the health and nutritional status of its population.

A total of nine provinces were covered in this survey. The combined population of these provinces accounts for approximately 56% of the total population [20]. All the provinces vary substantially in geography, economic development, public resources, and health indicators. A multi-stage, random cluster process was used to draw the samples surveyed in each of the provinces. The counties in these nine provinces were stratified on the basis of income (low, middle, and high), and a weighted sampling scheme was used to randomly select four counties in each province. In 2009, there were 218 primary sampling units consisting of 36 urban neighbourhoods, 37 suburban neighbourhoods, 37 towns, and 108 villages. In all, almost 16,000 individuals participated.

In the CHNS, all the questions related to individual activities, lifestyle, health status, marriage and birth history, body shape, mass media exposure, etc. were categorized into two sets of individual questionnaires: for adults aged 18 and older and for children and adolescents under age 18. The adults were made to undergo detailed physical examinations that included the measuring of weight, height, arm and head circumference, mid-arm skin-fold measurements, and blood pressure.

In this study, from the data set of 2009, we only used the data set of adults, and dropped the data of the participants who were younger than 18 years of age. Pregnant women were excluded, leaving a final sample of 10,931 respondents from 9 provinces; of these, 839 were from the Liaoning province, 821 were from the Heilongjiang province, 937 were from the Jiangsu province, 818 were from the Shandong province, 762 were from the Henan province, 792 were from the Hubei province, 699 were from theHunan province, 956 were from the Guangxi province, and 757 were from the Guizhou province. All 10931 respondents were nested in 4,423 households, within 218 communities.

Figure 1 shows the proportion of overweight and obesity in the nine provinces; we can see that the Shandong and the Henan provinces had higher proportions of

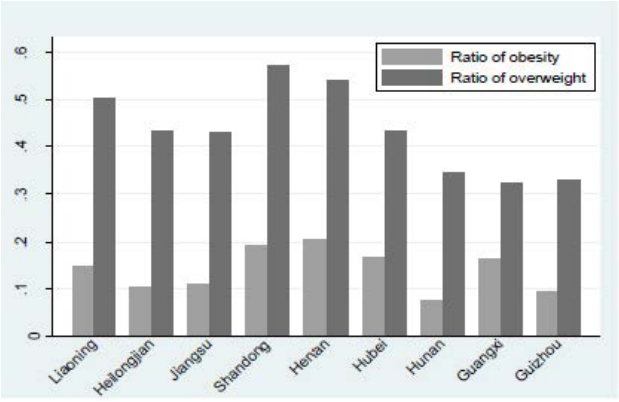


Fig. 1 Proportion of Overweight and Obesity in Nine Provinces

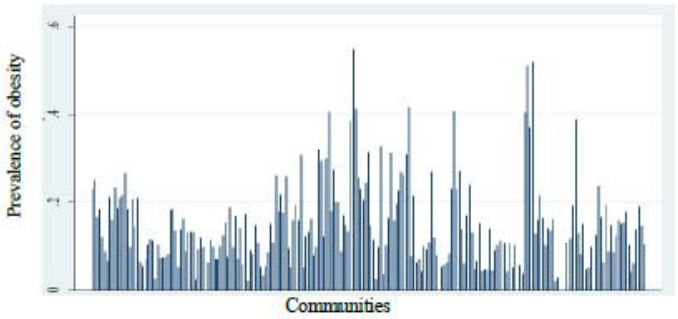


Fig. 2 Proportion of Obesity in 218 Communities

obesity and overweight than the others, and that the lowest ratio of obesity was in the Hunan province. Figure 2 shows the uneven distribution of the obesity prevalence in the 218 communities.

4 Methods

Multilevel statistical models allow for the estimation of contextual effects of higher level factors by accounting for the spatial clustering of individuals within a region. The Stata 12.0 software package was used to estimate multilevel logistics models.

In this research, since the dependent variable obesity is a binary variable, we begin with a brief review of logit regression model, and then proceed to formulate the multilevel logit regression model. In logistic regression for binary outcome measures in non-hierarchically structured data, the probability of "event" is usually converted to odds: $p/(1-p)$, the logarithm of the odds called the logit, is then treated as a linear function of a set of explanatory variables, resulting in the following logit model:

$$logit\,(p) = \log\left(\frac{p}{1-p}\right) = \beta_0 + \sum_{k=1}^{k} \beta_k x_k, \tag{1}$$

where β_k is the regression slope coefficient of the explanatory variable x_k. With this logit transformation of the outcome, the nonlinear relationship between the outcome and covariates is converted to a linear relationship. A regression coefficient can be interpreted as the amount of change in the log odds per unit change in the corresponding explanatory variable, controlling for other covariates.

Since the multi-stage random cluster process was used in CHNS to collect data, the dependence among the observations often comes from several levels of the hierarchy. In this case, the use of single-level statistical models is no longer valid and reasonable. This is because in traditional logistic regression, the assumptions of it require: (1) independence of the observations conditional on the explanatory variables and; (2) uncorrelated residual errors. In the nested dataset, these assumptions cannot always be satisfied [8]. Hence, in order to draw appropriate inferences and

conclusions from the multi-stage stratified clustered data we may require tricky and complicated modelling techniques like multilevel modelling. It allows the simultaneous examination of the effects of group level (cluster and division) and individual level variables on individual level outcomes while accounting for the non-independence of the observations within the groups.

Multilevel analyses allow for the estimation of variability within and between the groups by separating the variation at the individual level from that at the household level and the community level. According to Wang [8], the three-level logit model for obesity prevalence is explained as follows:

The individual-level (level-1) model can be written as:

$$\text{logit}\left\{Pr\left(y_{ijk}=1|\eta_{1jk},x_{\text{ni}jk}\right)\right\}=\eta_{1jk}+\beta_1 x_{1ijk}+\cdots+\beta_5 x_{5ijk} \tag{2}$$

where y is the binary outcome variable, if people are suffering from obesity, it is coded as 1, and x_{1ijk} to x_{5ijk} are the covariates at the individual level,and the intercept η_{1jk} varies between the families j and the communities k. Denoting the two covariates at the family level as ω_{1jk} to ω_{2jk}, the household-level (level-2) model for the intercept becomes:

$$\eta_{1jk}=\pi_{11k}+\pi_{12}\omega_{1jk}+\pi_{13}\omega_{2jk}+\varsigma_{jk}^{(2)}, \tag{3}$$

here only the intercept π_{11k} has a k subscript and therefore requires a community-level model:

$$\pi_{11k}=\gamma_{111}+\gamma_{112}v_{2k}+\gamma_{113}v_{3k}+\varsigma_k^{(3)}, \tag{4}$$

where v_{2k} is the covariate at the community level (level-3). Substituting the model for π_{11k} in the level-2 model and subsequently for η_{1jk} into the level-1 model, we obtain:

$$\begin{aligned}&Logit\left\{\Pr\left(y_{ijk}=1|x_{nijk},\varsigma_{jk}^{(2)},\varsigma_k^{(3)}\right)\right\}=\gamma_{111}+\beta_1 x_{1ijk}+\cdots+\beta_5 x_{5ijk}\\&+\pi_{12}\omega_{1jk}+\pi_{13}\omega_{2jk}+\gamma_{112}v_{2k}+\varsigma_{jk}^{(2)}+\varsigma_k^{(3)}\end{aligned} \tag{5}$$

The intra class coefficient (ICC) is defined as the amount of variation in the responses explained by the clustering variable. It informs us on the proportion of total variance in the outcome that is attributable to the area level. Based on above information, ICC can be estimated from the empty model. For the same community k but different households j and j' we obtain

$$\text{ICC}\left(\text{communities}\right)=\text{Cor}\left(y_{ijk}^{*},y_{i'j'k}^{*}|x_{ijk}x_{i'j'k}\right)=\frac{\psi^{(3)}}{\psi^{(2)}+\psi^{(3)}+\pi^2/3} \tag{6}$$

where y^* is the binary outcome variable, if people are suffering from obesity, it is coded as 1, x is the covariate at the individual level, and ψ is the variance at the different levels. Whereas for the same household j and the same community k, we get:

$$\text{ICC}\,(\text{household},\text{communities}) = \text{Cor}\left(y^*_{ijk}, y^*_{i'j'k} | x_{ijk}x_{i'j'k}\right) = \frac{\psi^{(3)} + \psi^{(2)}}{\psi^{(2)} + \psi^{(3)} + \pi^2/3} \tag{7}$$

In a three-level model, $\psi^{(2)} > 0$ and $\psi^{(3)} > 0$, and it follows that:

$$\rho\,(household, communities) > \rho\,(communities)\,, \tag{8}$$

this makes sense since the individuals of a given family are more similar than the individuals that belong to a given community but who are from different families. We can also quantify the unobserved heterogeneity by considering the median odds ratio (MOR). The MOR is defined as the median value of the odds ratio between the area at highest risk and the area at lowest risk when randomly picking out two areas. The MOR estimates the individual risk of obesity in median that can be attributed to the different level. In the three-level model for the CHNS data, comparing the individuals of the different families in the same community gives the median odds ratio as:

$$\text{MOR}(\text{community})_{median} = \exp\left\{\sqrt{2\psi^{(2)}}\Phi^{-1}(3/4)\right\}, \tag{9}$$

where Φ is the cumulative distribution function for a normal distribution. And comparing the individuals of the different families from the different communities gives:

$$\text{MOR}_{median} = \exp\left\{\sqrt{2(\psi^{(2)} + \psi^{(3)})}\Phi^{-1}(3/4)\right\} \tag{10}$$

The modeling strategy consists of sequential model estimation. Model 1 includes only a constant term that will allow the calculation of the ICC and the MOR. This intercept-only model predicts the probability of obesity. Then the individual independent variables, including the socio-demographic and socioeconomic factors and the lifestyle factors, are added to the following models. Finally, the family characteristics factors at level-2, and the urbanization index at level-3 are included in the model sequentially.

4.1 Dependent Variables

Body Mass Index (BMI) is the preferred standard for estimating the dependent variable obesity, since BMI is a population-based easure which has been found in clinical settings to be a good approximation for the assessment of total body fat for a majority of patients. The BMI is calculated from the respondents' weight (in kilograms) divided by their height in square meters.

$$\text{BMI} = \frac{weight(kg)}{height^2(m^2)} \tag{11}$$

Usually, according to the standards stipulated by the WHO, the BMI is classified into four categories: a $BMI < 20kg/m^2$ is called underweight, a BMI that is between 20 and 25 is defined as normal weight, a BMI of $25kg/m^2$ is termed as overweight,

and a BMI of $30kg/m^2$ is defined as obese. However, for the Asian population, there is much debate about the most appropriate BMI cut-off points to distinguish between normal weight, overweight and obesity. Many researchers tend to reduce the BMI cutoff points for the Asian population [17]. Quite a few researches revealed that those Asians with a lower BMI had the same major metabolic morbidities as Americans with a higher BMI [9]. On the basis of the previous studies and according to the standard proposed by the Ministry of Health of the People's Republic of China, this paper used a BMI standard value of $28kg/m^2$ for obesity. Obesity is classified as a binary variable, if the BMI value is greater than 28 then the obesity is coded as 1, otherwise the obesity is taken as 0.

4.2 Independent Variables

The individual independent variables cover many aspects. First, the aspects of socio-demography, age and gender are included. Age is a continuous variable, and it is centralized to make it meaningful in a multilevel model. Gender is a binary variable. Second, three measures of socio-economic positions were used: educational level, employment status and occupation type. Education was categorized into four levels: primary school or below (reference group), 6-9 years of education (middle school), 10-12 years of education (high school and technical school), and over 12 years of education (college, university and above). The employment status was a binary variable, with unemployment as the reference group. The respondents were of, mainly, four types of occupation: farmers, professionals, administrators, skilled workers, and service workers. Third, some lifestyle habits, such as smoking, drinking alcohol and participating in activities were included. All of these factors are binary variables, and the reference individuals are taken to be those who practice no smoking, no drinking and no activities

At the second level, the total net household income, which is adjusted by CPI, is included. The household income is classified into five categories. In addition, the cooking style in the households is also considered; this variable is classified into two groups, and the reference group is those who tend to process food by steaming and boiling, and another group is those prefer frying food. In the third level, the community level, urbanization index which was developed by Popkin is adapted [8].

5 Results

Table 1 shows the detailed descriptive statistics of the individual level socio-demographic characteristics, socioeconomic status factors, household level variables, and community level urbanization index of the sample. The total sample includes 10,931 adults. It can be seen that the average BMI is 23.29, 42.91% adults have been classified as overweight, and 14.67% of all the adults have been categorized as obese. Approximately 52% of the sample is female. At the time of the survey, the average age was 47.85 years old. In the aspect of education, about 40.66%

received only up to 6 years of education, 34% received 6 to 9 years of education, 18.99% received 9 to 12 years of education, and only 6.1% adults obtained more than 12 years of education. About 61% of the respondents belonged to the workforce. The occupation status of the sample is also diverse: 36.61% of the respondents were farmers, 14.72% were professionals, which included senior and junior professionals, 15.52% were administrators, officers, and office staff, 21.37% were skilled workers, and 12.05% were service workers. As far as their lifestyle is concerned, we can see that 27.85% of the respondents admitted to smoking at the time, about one third admitted to drinking alcohol, and about 60% of the people said they take part in activities, such as running, swimming, gymnasium, etc., at least twice a week. At the household level, we see that the average household income adjusted by CPI is about 38,476.23 RMB. In addition, we were also interested in the cooking method used in the households; we can see that nearly 58.82% families had the tendency to process food by steaming and boiling, and the others preferred to fry food. At the community level, the average urbanization index is 67.42, and the range of index is from 30 to 106.

Table 2 and Table 3 display the odds ratio (OR) and the 95% confidence intervals (CI) from the multilevel logistic regression models (the three-level model). From Model 1, which is called the empty model, we get the ICC for the household level and the community level as 27.17% and 11.26%, respectively. The MOR for the household level and the community level are 2.801 and 1.962, respectively, which indicates substantial cluster heterogeneity at the contextual level. Model 2 includes two individual demographic factors; however, these two factors, age and gender, are insignificant. After adding the socioeconomic status factors in Model 3, we note that people with 912 years of education are about 18.3% less likely to be obese compared to those with lower education. Respondents who are working currently have lower probability (65.2%) of becoming obese. We also see an interesting point which is that administrators and office staffs have the highest probability of suffering from obesity compared to the other work types. From Model 4, it can be seen that those who admitted to smoking at the time have lower odds (OR=71%) in comparison with those who did not smoke. Other lifestyle habits, such as alcohol consumption and activity participation, are insignificant. Model 5 considers the characteristics of households, and we see that the odds of obesity are getting higher with the increase of household income; people from families whose income is located in the 4th quintile have 28% higher possibility of becoming obese. However, people from the richest of families have only 16% higher odds when compared with the reference group, and this odds ratio is not significant. In Model 6, all the variables at the three levels are included. We see that a high urbanization index is associated with increased odds of obesity in adults. Meanwhile, after being adjusted using this index, the odds for respondents with more than 12 years of education (OR=0.757) appear significant at 10% level. In addition, people from families which tend to process food by the frying method have nearly 11% higher probability to suffer from obesity than those who prefer to cook by steaming and boiling.

Table 1 Descriptive Statistics of Variables (N=10931)

Variable	Definition	Mean	Std. Dev.
Dependent Variables			
BMI	BMI derived from weight(kg) by height (m)squared	23.292	3.499
Obesity	1 if BMI is equal or larger than 28 kg/m2, 0 if otherwise	14.67%	0.354
Overweight	1 if BMI is equal or larger than 24 kg/m2, 0 if otherwise	42.91%	0.495
Independent Variables			
Individual level			
Demographic factors			
AGE	The samples age is restricted to 18 years and older	47.845	15.845
GENDER			
Female	1 if gender is female	51.96%	0.499
Male	0 if gender is male	48.04%	0.427
Socioeconomic factors			
Education Level			
Low	1 if one has 0-6 years of education, 0 otherwise	40.66%	0.419
Medium	1 if one has 6-9 years of education, 0 otherwise	34.25%	0.475
Medium-to-high	1 if one has 9-12 years of education, 0 otherwise	18.99%	0.392
High	1 if one has more than 12 years of education, 0 otherwise	6.10%	0.239
Work status			
No jobs	0 if currently no working	38.85%	0.321
Have a job	1 if currently working	61.15%	0.493
Occupation types			
Farmers	1 if ones occupation belongs to farmers, 0 otherwise	36.61%	0.371
Professionals	1 if ones occupation belongs to professionals, 0 otherwise	14.72%	0.214
Administrator	1 if ones occupation belongs to officers/administrators, 0 otherwise	15.25%	0.224
Skilled worker	1 if ones occupation belongs to skilled workers, 0 otherwise	21.37%	0.291
Service worker	1 if ones occupation belongs to service workers, 0 otherwise	12.05%	0.326
Lifestyle			
Smoking			
No smoking	0 if one does not smoke currently	72.15%	0.418
Smoking	1 if one does smoke currently	27.85%	0.448
Alcohol			
No Alcohol	0 if one does not drink alcohol	66.87%	0.435
Alcohol	1 if one does drink alcohol	33.13%	0.471
Activity participation			
No activity	0 if one does not participate in any activities	42.68%	0.481
Activity	1 if one participates in any kinds of activities (such as running, etc) more than twice a week and over 40 min each time	58.32%	0.499
Household level variables			
hhinc_cpi	Household total net income adjusted by CPI	38476.2	46269.77
Cooking method			
Steaming	0 if family tends to process food by steaming and boiling	53.82%	0.325
Frying	1 if family tends to process food by frying and baking	46.18%	0.386
Community level variables			
Urbanization index	An index made from 12 dimensions to reflect the community urbanization level	67.42	19.46

Table 2 Odds Ratio Estimated from Random Intercept 3-level Multilevel Logistic Models of Obesity Prevalence (N=10931) (Models 1-3)

	Model 1		Model 2		Model2	
	OR	CI	OR	CI	OR	CI
Constant	0.146***	[0.131,0.163]	0.183***	[0.111,0.301]	0.197***	[0.120,0.326]
AGE			0.995	[0.976,1.014]	1.003	[0.983,1.023]
GENDER						
Male			1	[1.000,1.000]	1	[1.000,1.000]
Female			0.962	[0.858,1.078]	0.929	[0.824,1.047]
Education Level						
<6 years					1	[1.000,1.000]
6 – 9 years					0.896	[0.769,1.044]
9 – 12 years					0.817**	[0.675,0.989]
>12 years					0.825	[0.604,1.126]
Work status						
No jobs					1	[1.000,1.000]
Have a job					0.652***	[0.543,0.782]
Primary Occupation						
Farmers					1	[1.000,1.000]
Professionals					0.961	[0.656,1.408]
Administrator					1.876***	[1.389,2.533]
Skilled worker					1.373**	[1.076,1.752]
Service worker					1.625***	[1.303,2.025]
Lifestyle factors						
No smoking						
Smoking						
No Alcohol						
Alcohol						
No activity						
Activity						
Random Effect						
Level-2 Variance	0.662***		0.661***		0.640***	
Level-3 Variance	0.454***		0.459***		0.424***	
Model fit statistics						
AIC	8099.526		8087.79		7143.67	
BIC	8113.955		8111.842		7201.404	

Note:Exponentiated coefficients; 95% confidence intervals in brackets, * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table 3 Odds Ratio Estimated from Random Intercept 3-level Multilevel Logistic Models of Obesity Prevalence (N=10931) (Models 3-6)

	Model 4		Model 5		Model6	
	OR	CI	OR	CI	OR	CI
Constant	0.210***	[0.127,0.348]	0.294***	[0.167,0.517]	0.225***	[0.131,0.385]
AGE	1.006	[0.986,1.026]	1.007	[0.987,1.027]	1.006	[0.986,1.026]
GENDER						
Male	1	[1.000,1.000]	1	[1.000,1.000]	1	[1.000,1.000]
Female	0.829**	[0.712,0.967]	0.823**	[0.706,0.959]	0.822**	[0.705,0.958]
Education Level						
<6 years	1	[1.000,1.000]	1	[1.000,1.000]	1	[1.000,1.000]
$6-9$ years	0.887	[0.761,1.034]	0.895	[0.768,1.043]	0.885	[0.758,1.032]
$9-12$ years	0.807**	[0.667,0.977]	0.812**	[0.670,0.985]	0.792**	[0.651,0.963]
>12 years	0.785	[0.574,1.073]	0.786	[0.574,1.078]	0.757*	[0.550,1.042]
Work status						
No jobs	1	[1.000,1.000]	1	[1.000,1.000]	1	[1.000,1.000]
Have a job	0.683***	[0.536,0.871]	0.683***	[0.536,0.871]	0.709***	[0.553,0.908]
Primary Occupation						
Farmers	1	[1.000,1.000]	1	[1.000,1.000]	1	[1.000,1.000]
Professionals	0.938	[0.640,1.376]	0.948	[0.645,1.392]	0.908	[0.616,1.341]
Administrator	1.891***	[1.399,2.556]	1.872***	[1.382,2.536]	1.794***	[1.317,2.443]
Skilled worker	1.360**	[1.065,1.737]	1.380**	[1.079,1.764]	1.335**	[1.040,1.715]
Service worker	1.617***	[1.296,2.016]	1.618***	[1.296,2.020]	1.555***	[1.238,1.954]
Lifestyle factors						
No smoking	1	[1.000,1.000]	1	[1.000,1.000]	1	[1.000,1.000]
Smoking	0.710***	[0.604,0.836]	0.706***	[0.600,0.830]	0.707***	[0.601,0.832]
No Alcohol	1	[1.000,1.000]	1	[1.000,1.000]	1	[1.000,1.000]
Alcohol	1.113	[0.956,1.295]	1.109	[0.953,1.290]	1.111	[0.955,1.292]
No activity	1	[1.000,1.000]	1	[1.000,1.000]	1	[1.000,1.000]
Activity	0.962	[0.783,1.180]	0.961	[0.783,1.180]	0.963	[0.785,1.183]
Household level variables						
1st Quintile			1	[1.000,1.000]	1	[1.000,1.000]
2nd Quintile			1.059***	[0.901,1.288]	1.062***	[0.902,1.284]
3rd Quintile			1.180**	[0.913,1.316]	1.196**	[0.920,1.362]
4th Quintile			1.288**	[0.918,1.439]	1.285**	[0.946,1.405]
5th Quintile			1.161	[0.791,1.429]	1.165	[0.818,1.435]
Cooking method						
Steaming			1	[1.000,1.000]	1	[1.000,1.000]
Frying			1.102	[0.986,1.026]	1.109*	[0.978,1.220]
Urbanization index						
1st Quantile					1	[1.000,1.000]
2nd Quantile					1.069*	[0.890,1.267]
3rd Quantile					1.192**	[0.869,1.356]
4th Quantile					1.156**	[0.910,1.308]
Random Effect						
Level-2 Variance	0.621***		0.607***		0.579***	
Level-3 Variance	0.453***		0.421***		0.406***	
Model fit statistics						
AIC	7099.526		7067.79		7043.67	
BIC	7150.146		7139.842		7101.404	

Note:Exponentiated coefficients; 95% confidence intervals in brackets, * $p<0.1$, ** $p<0.05$, *** $p<0.01$

6 Discussion

This study has shown that regional and socioeconomic disparities exist in adult obesity in China. People living in high SES and more urbanized areas have higher BMI and higher odds of obesity than those living in lower SES and rural areas. From this study, we understand that people with high education have low odds of obesity, and that the possible reason for this phenomenon is that educated people have more knowledge about the harmfulness of obesity and tend to control their weight so that it stays in the normal range. Upon comparing all types of occupation, we find that the highest ratio of obesity exists in the group of administrators and office staffs. The possible reason for this is that most of these people spend long hours sitting in the office and, so, have less time to do exercises; in addition, more chances of social eating and social drinking also increase the risk of obesity among them. Interestingly, we found that the respondents who admitted to smoking had lower odds of obesity, and this finding is consistent with that of many other research studies [9], as some components of cigarette, such as nicotine, can effect weight loss. However, after the smokers quit smoking, the weight may respond and increase quickly. People from families with lowmedium to mediumhigh incomes tend to have more risk of being obese, and this result is consistent with many of the studies conducted in the developing countries[9], this finding is opposite to the results from the developed countries which find that obesity is more prevalent in families of low income [25]. Finally, this study finds that the prevalence of obesity is also linked to the industrialization and urbanization of China. The urbanization has changed peoples lives in many ways. The fast food outlets have grown dramatically in many areas during the last decade [14]. It has been reported that more than 18% of the people in big cities consume fast food regularly and frequently, and that about 60% of them are not aware of the fact that it is energy-dense food. In addition, people living in more urbanized areas own more televisions, video disc players, and computers, and so they spend more time watching them or playing them, and cut down on the time for doing exercises [22]. It also seems most probable that the increased use of automobiles, instead of bicycling or walking, in urban areas has contributed to the epidemic of obesity [20]. Therefore, all of the afore-mentioned influences coupled with urbanization have resulted in an increased overweight and obesity prevalence in these areas.

7 Concluding Remarks

Adult obesity is associated with both immediate and long-term health problems and psychosocial problems. It burdens the health care system, strains economic resources, and has far reaching social consequences. Much of the rise in healthcare costs today can be attributed to the increase in obesity-related diseases such as diabetes, hypertension, pulmonary conditions, and chronic back pain. Many of today‘s most commonly prescribed drugs are for obesity-related conditions, and as such, obese individuals spend two to four times more on prescription medications than

adults who are non-obese. Therefore, it is crucial for the government to identify those risk factors related to obesity at individual levels, household levels and community levels, and then make corresponding, appropriate policies and take effective measures to reduce the prevalence of obesity.

From this study, we notice an uneven distribution of obesity prevalence, the prevalence varied from 3.67% to 46.18% in the different areas. This study estimated the magnitude of these differences, and, to the best of my knowledge, it is the first study in China to determine the influence of the indicators at individual, household and community levels simultaneously. Information derived from this study can be used to develop more effective ways of intervention and strategies for obesity prevention in the specific context of the various regions.

In conclusion, China is a nation undergoing rapid economic development. Obesity, as an important health problem which accompanied economic development, industrialization and urbanization, should be paid enough attention. Intervention and strategy development for obesity prevention should be based on this specific context, targeting the high SES families in the more urbanized areas.

Acknowledgements. This research uses data from the China Health and Nutrition Survey (CHNS). We thank the National Institute of Nutrition and Food Safety, China Center for Disease Control and Prevention, Carolina Population Center, the University of North Carolina at Chapel Hill, the NIH (R01-HD30880, DK056350, and R01-HD38700), and the Fogarty International Center, NIH, for the financial support for the CHNS data collection and for the analysis files from 1989 to 2009. We also extend our gratitude to both the parties plus the ChinaJapan Friendship Hospital, Ministry of Health, for support for the CHNS 2009 survey and for future surveys.

References

1. Ahn, S., Zhao, H., Smith, M.L., Ory, M.G., Phillips, C.D.: BMI and lifestyle changes as correlates to changes in self-reported diagnosis of hypertension among older Chinese adults. J. Am. Soc. Hypertens. 5(1), 21–30 (2011)
2. Arita, Y., Kihara, S., Ouchi, N., Takahashi, M., Maeda, K., Miyagawa, J., Hotta, K., Shimomura, I., Nakamura, T., Miyaoka, K.: Paradoxical decrease of an adipose-specific protein, adiponectin, in obesity. Biochemical and Biophysical Research Communications 257(1), 79–83 (1999)
3. Chen, J.L., Weiss, S., Heyman, M.B., Lustig, R.: Risk factors for obesity and high blood pressure in Chinese American children: maternal acculturation and children's food choices. J. Immigr. Minor Health 13(2), 268–275 (2011)
4. Corsi, D.J., Finlay, J.E., Subramanian, S.V.: Weight of communities: a multilevel analysis of body mass index in 32,814 neighborhoods in 57 low- to middle-income countries (LMICs). Soc. Sci. Med. 75(2), 311–322 (2012)
5. Crane, J.: The epidemic theory of ghettos and neighbourhood effects on dropping out and teenage childbearing. American Journal of Sociology, 1226–1259 (1991)
6. Drewnowski, A., Specter, S.: Poverty and obesity: the role of energy density and energy costs. The American Journal of Clinical Nutrition 79(1), 6–16 (2004)
7. Griffiths, S.M.: Leading a healthy lifestyle: the challenges for China. Asia Pac. J. Public Health 22(suppl. 3), 110S–116S (2010)

8. Jones-Smith, J.C., Popkin, B.M.: Understanding community context and adult health changes in China: development of an urbanicity scale. Soc. Sci. Med. 71(8), 1436–1446 (2010)
9. Goyal, R.K., Shah, V.N., Saboo, B.D., Phatak, S.R., Shah, N.N., Gohel, M.C., Raval, P.B., Patel, S.S.: Prevalence of overweight and obesity in Indian adolescent school going children: its relationship with socioeconomic status and associated lifestyle factors. J. Assoc. Physician India 58, 151–158 (2010)
10. Lovasi, G.S., Neckerman, K.M., Quinn, J.W., Weiss, C.C., Rundle, A.: Effect of individual or neighborhood disadvantage on the association between neighborhood walkability and body mass index. Journal Information 99(2) (2009)
11. Manson, J.E., Skerrett, P.J., Willett, W.C.: Epidemiology of health risks associated with obesity. In: Eating Disorders and Obesity: A Comprehensive Handbook, pp. 422–432 (2002)
12. Mark, A.L., Correia, M., Morgan, D.A., Shaffer, R.A., Haynes, W.G.: Obesity-induced hypertension new concepts from the emerging biology of obesity. Hypertension 33(1), 537–541 (1999)
13. Matheson, F.I., Moineddin, R., Glazier, R.H.: The weight of place: a multilevel analysis of gender, neighborhood material deprivation, and body mass index among Canadian adults. Soc. Sci. Med. 66(3), 675–690 (2008)
14. Maruapula, S.D., Jackson, J.C., Holsten, J., Shaibu, S., Malete, L., Wrotniak, B., Ratcliffe, S.J., Mokone, G.G., Stettler, N., Compher, C.: Socio-economic status and urbanization are linked to snacks and obesity in adolescents in Botswana. Public Health Nutrition 1(1), 1–8 (2011)
15. McLaren, L.: Socioeconomic status and obesity. Epidemiologic Reviews 29(1), 29–48 (2007)
16. Miyawaki, K., Yamada, Y., Ban, N., Ihara, Y., Tsukiyama, K., Zhou, H., Fujimoto, S., Oku, A., Tsuda, K., Toyokuni, S.: Inhibition of gastric inhibitory polypeptide signaling prevents obesity. Nature Medicine 8(7), 738–742 (2002)
17. Misra, A., Khurana, L.: Obesity-related non-communicable diseases: South Asians vs White Caucasians. Int. J. Obes. (Lond.) 35(2), 167–187 (2011)
18. Monteiro, C.A., Moura, E.C., Conde, W.L., Popkin, B.M.: Socio-economic status and obesity in adult populations of developing countries: a review. Bulletin of the World Health Organization 82(12), 940–946 (2004)
19. Philipson, T.J., Posner, R.A.: The long-run growth in obesity as a function of technological change: National Bureau of Economic Research (1999)
20. Popkin, B.M., Adair, L.S., Ng, S.W.: Global nutrition transition and the pandemic of obesity in developing countries. Nutrition Reviews 70(1), 3–21 (2012)
21. Ross, C.E.: Walking, exercising, and smoking: does neighborhood matter? Social Science & Medicine 51(2), 265–274 (2000)
22. Siervo, M., Grey, P., Nyan, O., Prentice, A.: Urbanization and obesity in The Gambia: a country in the early stages of the demographic transition. European Journal of Clinical Nutrition 60(4), 455–463 (2005)
23. Sturm, R.: The effects of obesity, smoking, and drinking on medical problems and costs. Health Affairs 21(2), 245–253 (2002)
24. Trayhurn, P., Wood, I.: Signalling role of adipose tissue: adipokines and inflammation in obesity. Biochemical Society Transactions 33, 1078–1081 (2005)
25. Vernay, M., Malon, A., Oleko, A., Salanave, B., Roudier, C., Szego, E., Deschamps, V., Hercberg, S., Castetbon, K.: Association of socioeconomic status with overall overweight and central obesity in men and women: the French Nutrition and Health Survey 2006. BMC Public Health 9(1), 215–218 (2009)

26. Wang, J., Xie, H., Fisher, J.H.: Multilevel Models: Applications using SAS: De Gruyter (2011)
27. Yoon, Y.S., Oh, S.W., Park, H.S.: Socioeconomic status in relation to obesity and abdominal obesity in Korean adults: a focus on sex differences. Obesity 14(5), 909–919 (2012)
28. Wang, J., Xie, H., Fisher, J.H.: Multilevel Models: Applications using SAS: De Gruyter (2011)
29. World Health Report. Reducing risks, promoting healthy life. World Health Organization, Geneva (2011), http://www.who.int/whr/2011/
30. Zhao, W., Zhai, Y., Hu, J., Wang, J., Yang, Z., Kong, L., Chen, C.: Economic burden of obesityrelated chronic diseases in Mainland China. Obesity Reviews 9(s1), 62–67 (2008)
31. Paraponaris, A., Saliba, B., Ventelou, B.: Obesity, weight status and employability: empirical evidence from a French national survey. Economics and Human Biology 3(2), 241–258 (2005)

Statistical Analysis of Political Cycles in Australian Stock Market Returns

S.T. Boris Choy and Celestine M. Bond

Abstract. Political cycles in the Australian stock market from January 1901 to July 2011 are analysed through econometric volatility models. The stochastic volatility model with a skew t distribution for return and a Student-t distribution for volatility is proposed for analysis, estimated via Bayesian techniques. Evidence from the full period shows higher return under non-Labor governments while there is little evidence of election or length-of-term effects on market return. If we split the data before and after World War II, political cycles are non-existent. There is however clear evidence of positive skewness of returns before the war compared to negative skewness otherwise.

1 Introduction

Of interest to the broad Australian community of late has been the effect of particular political parties in power and their policies on financial, economic, industrial and consumer issues. Of importance to economists and the financial media is the effect of political outcomes and events on market behaviour and returns. For example, in the lead up to elections, politicians promise a combination of progress on consumer and industrial issues, tax benefits, subsidies and economic stimulation, all of which may result in varying effects in the stock market. Other effects may be due to key political strategies between governments such as stance on interest rates, inflationary measures and unemployment targets. It is the difference between political parties, their actions and consequent effects of policy decisions over time and political events which define political cycles.

There has been substantial research into this area in overseas stock markets, which indicates some key relationships between political variables and market returns. Herbst and Slinkman (1984) found evidence of political cycles in the US

S.T. Boris Choy · Celestine M. Bond
Discipline of Business Analytics, The University of Sydney, Australia
e-mail: boris.choy@sydney.edu.au

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_20,

market where the timing and magnitude of the cycles follow elections and other political events. In particular, political cycles in market returns peaked in the month and year where there was a presidential election. Evidence of this phenomenon in the US dates back to Niederhoffer *et al.* (1970) and on an international scale by Bialkowski *et al.* (2008) who found evidence for significantly higher stock market volatility due to election surprises in 27 Organisation for Economic Cooperation and Development (OECD) countries. More recently election effects on market returns have been described by the stage of ministerial term. This stream of research, investigated by several authors such as Allvine and O'Neill (1980), Huang (1985), Hensel and Ziemba (1995) and Booth and Booth (2003) find that US market returns are higher in the second half of a political term than in the first. The basis behind this is that political parties employ deflationary monetary and fiscal policies before and during elections, boosting the economy in an attempt to win or regain power for another term.

Other than evidence of election and political tenure effects in the market, a difference in policies between political parties may also have an effect on stock market returns. This is explained in a US context by Hibbs (1977) through "partisan theory" in which the Democratic Party has a more expansionist view on macroeconomic policies such as inflation and unemployment, leading to higher stock returns than under Republican rule. This is supported by Hensel and Ziemba (1995), Booth and Booth (2003), Santa-Clara and Valkanov (2003) and Wisniewski (2009). Santa-Clara and Valkanov (2003) examined the 'presidential puzzle' in which while there is an apparent preference of the market for right-of-centre parties (that is, the Republicans in the US), average excess market returns are higher for Democratic parties than under Republicans. However a critique by Powell *et al.* (2007) suggests that relationships found by Santa-Clara and Valkanov (2003) are spurious and hence insignificant.

Although there is evidence of higher stock market returns under left-leaning governments in the US, there is also significant evidence of a preference of the market for right-of-centre governments predominantly in international markets. Evidence against the 'presidential puzzle' in the US include work by Riley and Luksetich (1980) and Snowberg *et al.* (2007). A study by Bohl and Gottschalk (2006) looked at 15 countries and found that evidence for higher returns under left-leaning governments only existed in Denmark, Germany and the US, whereas there was very little evidence of a difference between right- and left-centred governments and their effects on Australian and New Zealand market returns. In comparison, Cahan *et al.* (2005) found that New Zealand stock market returns were lower under the left-leaning Labor government than National governments, indicating a preference for right-centred parties. Anderson *et al.* (2008) also found that stock markets performed better in Australia and New Zealand under right-leaning governments when inflation is lower.

The bulk of Australian work on political cycles has been performed by Worthington (2009) who examined the effect of political cycles in the Australian stock market from January 1901. His analysis examined the difference between market returns in non-Labor and Labor governments, whether returns vary during a party's time in office and whether an election results in an observable difference in returns compared

to the rest of the ministerial term. His research applied a generalized autoregressive conditional heteroskedasticity (GARCH) model, which captures the changing variance of stock returns in the Australian Stock Exchange (ASX). For the full period of data (1901 to 2005), there was weak evidence that non-Labor governments have a higher market return than Labor governments. However, from 1950 onwards, this relationship seems to have disappeared, which Worthington (2009) attributes to a reduction in bias towards a certain political party by businesses and investors.

Despite this thorough analysis of political cycles, Worthington's model is not without limitations. The purpose of this article is to extend on Worthington's work by using another type of time-varying variance model, the stochastic volatility (SV) model, to assess the relationship between the Australian federal political cycle and the Australian stock market return. In particular, a Bayesian analysis is emphasised, with model estimation via Markov chain Monte Carlo (MCMC) methods to obtain posterior inference. For an introduction to MCMC methods, see Smith and Roberts (1993), Gilks *et al.* (1996) and Andrieu *et al.* (2004). By using a Bayesian approach a variety of flexible error distributions, both symmetric and asymmetric for the return and volatility are able to be explored.

2 Methodology

2.1 Data and Variable Specification

The data and political variable choices are heavily influenced by Worthington (2009), since our analysis is mainly an exercise in improving on the original GARCH specification and to investigate whether these improved models lead to different conclusions about the effect of Australian political cycles in the stock market.

Table 1 summarises the governing terms of 32 Prime Ministers of Australia, the start and end date of their period in office and the average market return in the Australian All Ordinaries (AORD) Index. Note that the ministries are divided into two groups of parties: Labor and non-Labor. Labor refers to the Australian Labor Party while non-Labor refers to all other ministries such as the Liberal and National parties, as well as the Protectionist, Free Trade, Tariff Reform, Nationalist Labor, Nationalist, United Australia, Country, and Country Liberal parties which were especially active in the early 20th century.

The dependent variable used in this analysis is market returns based on the AORD Index retrieved from Wren Investment Advisers (http://www.wrenresearch.com.au/downloads). Market returns, r_t are defined as monthly returns $r_t = 100\ln(P_t/P_{t-1})$ where P_t is the closing value of the index at the end of month t. The time period analysed is from January 1901 to July 2011 using 1327 observations for r_t. Although Worthington (2009) used two additional dependent variables, that is, market returns in excess of inflation and interest rate, he found no significant evidence for a difference in excess returns between Labor and non-Labor governments. As such, an analysis of market return is deemed sufficient for our study.

Table 1 Australian Prime Minister Terms and Monthly Market Returns

No.	*Prime Minister*	*Party*	*Start date*	*End date*	*Term in office (months)*	*% Return*
1	Barton	N-L	Jan 1901	Sep 1903	32	0.1522
2	Deakin	N-L	Sep 1903	Apr 1904	7	1.2177
3	Watson	Labor	Apr 1904	Aug 1904	4	1.6396
4	Reid	N-L	Aug 1904	July 1905	11	0.4924
5	Deakin	N-L	July 1905	Nov 1908	40	0.6269
6	Fisher	Labor	Nov 1908	June 1909	7	0.4418
7	Deakin	N-L	June 1909	Apr 1910	10	0.8343
8	Fisher	Labor	Apr 1910	June 1913	38	0.1929
9	Cook	N-L	June 1913	Sep 1914	15	0.4717
10	Fisher	Labor	Sep 1914	Oct 1915	13	-0.2123
11	Hughes	Labor	Oct 1915	Feb 1923	88	0.4804
12	Bruce-Page	N-L	Feb 1923	Oct 1929	80	0.6647
13	Scullin	N-L	Oct 1929	Jan 1932	27	-1.6480
14	Lyons	N-L	Jan 1932	Apr 1939	87	0.7589
15	Page	N-L	Apr 1939	Apr 1939	1	-2.398
16	Menzies	N-L	Apr 1939	Aug 1941	27	0.0000
17	Fadden	N-L	Aug 1941	Oct 1941	2	2.2156
18	Curtin	Labor	Oct 1941	July 1945	45	0.2099
19	Forde	Labor	July 1945	July 1945	1	-0.2821
20	Chifley	Labor	July 1945	Dec 1949	52	0.5492
21	Menzies	N-L	Dec 1949	Jan 1966	193	0.4234
22	Holt	N-L	Jan 1966	Dec 1967	23	1.6199
23	McEwen	N-L	Dec 1967	Jan 1968	1	-3.8188
24	Gorton	N-L	Jan 1968	Mar 1971	38	0.3158
25	McMahon	N-L	Mar 1971	Dec 1972	21	0.8894
26	Whitlam	Labor	Dec 1972	Nov 1975	35	-1.1102
27	Fraser	N-L	Nov 1975	Mar 1983	88	0.7075
28	Hawke	Labor	Mar 1983	Dec 1991	105	1.1189
29	Keating	Labor	Dec 1991	Mar 1996	51	0.6198
30	Howard	N-L	Mar 1996	Dec 2007	141	0.7526
31	Rudd	Labor	Dec 2007	June 2010	30	-1.2462
32	Gillard	Labor	June 2010	July 2011	14	0.1414
All	-	-	Jan 1901	July 2011	1327	0.4667

Notes: N-L refers to non-Labor ministries. Information retrieved from the Australian Electoral Commission, `http://www.aec.gov.au/Elections/Australian_Electoral_History`, in August 2011.

The political cycle variables used in this analysis completely describe the dependent variable of market return. Following methods outlined in similar literature and relationships identified in the previous section, the political variables include dummy variable L_t, indicating whether Labor was in power in month t (equals zero

when non-Labor is in power), continuous variable T_t measuring the time in office in months (reset at the next prime minister's term in office) and dummy variable E_t indicating whether a federal election occurred during month t. The motivation for including the continuous variable T_t follows the idea that political influence on market return may depend on how long that political party has been in office. See, for example, Huang (1985) and Booth and Booth (2003). Including an election dummy variable E_t allows us to measure the effect of elections in the stock markets, including the effect of political promises in the lead up to elections or unexpected election results such as in Bialkowski *et al.* (2008).

Worthington (2009) used an additional dummy political variable NL_t, to indicate whether a non-Labor party is in power in month t and hence had no intercept term in the observation equation. In fact, the variables NL_t and L_t are perfectly multicollinear thus one of them should be removed from the model. We measure the effect of the political power only as changes in the coefficient of L_t in excess of non-Labor governments. For example, a positive coefficient on L_t would suggest that higher market returns are expected under Labor governance than non-Labor governance. We also consider an additional independent dummy variable W_t which is equal to zero when month t is before January 1946 and one on and after January 1946. The cut off point for this variable was chosen to be around the end of World War II (WWII), and allows us to measure the effects of political cycles pre- and post-war without having to perform three separate analysis (full period, pre- and post-war) on the dependent variable. Moreover, Labor and the Coalition have been the main political candidates since WWII.

2.2 Descriptive Analysis

Figure 1 plots the monthly market return of the AORD Index from January 1901 to July 2011. Time series of returns should typically be stationary, with constant mean and constant fluctuations (variance) around it. The series we analyse does not have constant variance, especially evident from the period before and after the 1960's. This motivates our use of volatility models which are able to account for fluctuating, time-varying volatility, such as GARCH and SV models in Section 3. There are also a few peculiar outliers corresponding to the market response to global events such as the Great Depression in the early 1930's, the global oil crisis in 1973, the stock market crash in October of 1987 and the Global Financial Crisis of 2008.

Table 2 provides descriptive statistics of monthly returns on the AORD Index for all ministries and by Labor and non-Labor groups. The returns are also divided into the full period before and after WWII. For the full period, return under non-Labor ministries is higher than for Labor ministries (0.61 compared to 0.17), as well as the period before WWII (0.60 for non-Labor and -0.16 for Labor). After WWII, Labor ministries have higher return (0.34 compared to 0.15). However, only in the period before WWII is this difference significant using a standard t-test for equality of means. It should also be noted that Labor ministries take the minimum return values for all three periods. The returns under Labor governments are also more

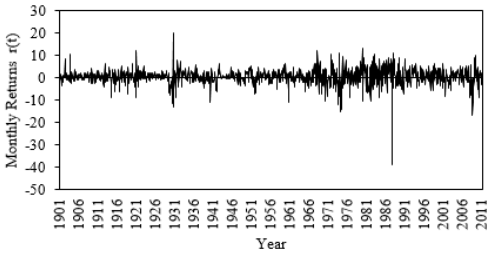


Fig. 1 Market Returns of the AORD Index from 1901 - 2011

variable than under non-Labor governments using the standard F-test for equality of variances rejecting equality at the 1% level of significance.

An important observation which may affect validity of estimates is the fact that market returns appear non-normally distributed for all time periods considered and under both ministries. In particular, there is slight negative skewness for the whole period, slightly positive skewness for the period before WWII and negative skewness in the period after WWII, especially under non-Labor ministries. In addition, the kurtosis under all ministries for all time periods exhibit tails which are heavier than the normal distribution. The non-normality of returns is supported by the Jarque-Bera test for normality, which is rejected even at the 1% level of significance for all time periods under either ministry.

3 Models

3.1 Generalised Autoregressive Heteroskedastic (GARCH) Models

Market returns commonly exhibit persistent high and low volatilities which can be captured by a volatility model. Proposed by Bollerslev (1986), the GARCH model and its derivatives have been widely used for modelling time-varying volatility. In GARCH models, the conditional variance of the error term is dependent upon past error terms and past variances. In addition, the volatility component can also be included in the mean component of the return equation which is known as the GARCH-in-mean (GARCH-M) model.

Consider the GARCH(1,1)-M model:

Return equation:

$$r_t = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \gamma_0 \sigma_t^2 + \varepsilon_t \tag{1}$$

$$\varepsilon_t | \Omega_{t-1} \sim N(0, \sigma_t^2) \tag{2}$$

Volatility equation:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \sigma_{t-1}^2 \tag{3}$$

where r_t is the market return at time t, x_i is a set of political variables which influence returns, σ_t^2 is the volatility of the return at time t and ε_t is the error term of the model which is distributed as $N(0, \sigma_t^2)$. The coefficients on the political factors β_i measure the effect of those variables on the stock market return. γ_0 measures the effect of the conditional variance on returns hence the 'in-mean' specification. The conditional variance σ_t^2 follows a GARCH specification and is dependent upon a time invariant coefficient α_0, past squared error term ε_{t-1}^2 and past volatility σ_{t-1}^2 whose effects are measured by α_1 and α_2 respectively. Note that by the nature of the conditional variance, $\alpha_1 \geq 0$, and $\alpha_2 \geq 0$ and for stationarity $\alpha_0 > 0$, and $\alpha_1 + \alpha_2 < 1$.

3.2 Stochastic Volatility (SV) Models

SV models are another commonly used volatility modelling which allow the conditional variance to follow a stochastic process. In particular, they have two noise processes making it more flexible than GARCH-type models. For a comprehensive overview of SV models, see Ghysels *et al.* (1996). SV models have also proven preferable to GARCH models for returns in an Australasian market context by Yu (2002).

Following the initial 'in-mean' specification used by Worthington (2009), we consider the SV-M model, which has been compared to the GARCH-M model in international markets by Koopman and Hol Uspensky (2002). The model is specified by:

Return equation:

$$r_t = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \gamma_0 \sigma_t^2 + \varepsilon_t, \quad \varepsilon_t | h_t \sim N(0, \sigma_t^2) \tag{4}$$

Volatility equation:

$$h_t = \ln \sigma_t^2 = \begin{cases} \mu + \phi(h_{t-1} - \mu) + \eta_t & t > 1 \\ \mu + \left(1 - \phi^2\right)^{-1/2} \eta_t & t = 1 \end{cases}, \quad \eta_t \sim N(0, \tau^2) \tag{5}$$

where h_t is the log-volatility, $\phi \in (-1, 1)$ is the persistence of the volatility, τ^2 is the variance of the log-volatility, η_t is the normally distributed error term of the volatility equation with $E(\eta_t) = 0$ and $V(\eta_t) = \tau^2$, and the variables and parameters in the return equation are interpreted as in the GARCH model.

3.3 Error Distributions

In our basic SV model, we assume that both the error terms in the return and volatility equations are normally distributed. Due to the flexibility of Bayesian MCMC methods, we are able to specify more robust error distributions for both the return and volatility equation.

There is substantial evidence supporting non-normality of stock returns, not only from our brief analysis in Section 2.2. In an Australian context some of the evidence which spans over the time period in which we are interested in includes work by Praet and Wilson (1978), Beedles (1986) and Gray and Kalotay (1998). By using more robust error distributions we are able to protect parameter estimates against data which exhibit outlying behaviour. Normal and Student-t error distributions for the volatility equation are used in our analysis. In addition to these symmetric distributions, a skew t distribution is utilised for the return equation.

Dermata and McNeil (2005) introduced a skew-t distribution (referred to ST_1 here) via a normal mean-variance mixture of the following form:

$$x|\mu,\sigma^2,\theta,\lambda \sim N\left(\mu+\theta\lambda^{-1},\lambda^{-1}\sigma^2\right) \tag{6}$$

$$\lambda|\nu \sim Ga\left(\frac{\nu}{2},\frac{\nu}{2}\right) \tag{7}$$

where $N(a,b)$ is the normal distribution with mean a and variance b, $Ga(a,b)$ is the gamma distribution with mean a/b, μ is the location parameter, σ^2 the scale parameter, θ determines the skewness or asymmetry of the distribution, λ is a scale mixture variable and ν is the degrees of freedom of the skew t distribution. Alternatively, Branco and Dey (2001) proposed a different skew-t distribution (called the ST_2 here) via the following similar mean-variance mixture form:

$$x|\mu,\sigma^2,\theta,\lambda,V \sim N\left(\mu+\theta\lambda^{-1/2}V,\lambda^{-1}\sigma^2\right) \tag{8}$$

$$\lambda|\nu \sim Ga\left(\frac{\nu}{2},\frac{\nu}{2}\right) \tag{9}$$

where V is an additional scale mixture variable which follows a half normal distribution with probability density function (PDF) given by

$$f(v)=\sqrt{\frac{2}{\pi}}e^{-\frac{v^2}{2}}, \quad v>0 \tag{10}$$

In other words, $V=|W|$ where $W\sim N(0,1)$. Since the half normal distribution is a special case of the generalised gamma distribution, denoted by $GG(r,\mu,\beta)$ and proposed by Stacy (1962), having PDF

$$f_{GG}(v)=\frac{\beta}{\Gamma(r)}\mu^{\beta r}v^{\beta r-1}e^{-(\mu v)^{\beta}}, \quad v>0 \tag{11}$$

with $r = 2^{-1}$, $\mu = 2^{-1/2}$ and $\beta = 2$, statistical inference with half normal distribution can be easily implemented using WinBUGS (Bayesian analysis Using Gibbs Sampler) package (Spiegelhalter *et al.*, 2007). See Fung and Seneta (2010) for details.

The main difference between the two skew-t distributions is that the ST_2 distribution possesses asymptotic lower tail dependence (Dermata and McNeil, 2005) while the ST_1 distribution does not. In addition, the use of these distributions via their mean-variance mixture forms can simplify the MCMC algorithms and hence provide more efficient model estimation. See Choy and Chan (2008) for details.

4 Model Implementation and Estimation

Since the marginal likelihood function of the SV model does not have a closed form, likelihood approaches such as the quasi-maximum likelihood method (Ruiz, 1994) and simulated maximum likelihood method (Danielsson, 1994) are cumbersome and hence are not considered in this paper. Instead, we adopt the Bayesian MCMC approach which has been widely and successfully used for accurate inferences of complicated models since the early 1990s. The advantage of this approach is that exact finite sample inference can be drawn from the joint posterior distributions through simulation.

To estimate the SV models, we take the Bayesian MCMC approach using the Gibbs sampling algorithm of Jacquier, Polson and Rossi (1994). This approach iteratively simulates posterior samples from the univariate full conditional distribution of each model parameter conditional on the values of the other parameters and data and can be easily implemented using WinBUGS package. For comparison purpose, we also study the GARCH models in this paper. Due to the recursive nature of the GARCH models, the marginal likelihood function can be obtained. To increase computing efficiency, we implement the GARCH models using the adaptive MCMC algorithm of Gerlach and Chen (2008), which adopts a random walk Metropolis-Hastings algorithm for the burn-in period and an independent kernel Metropolis-Hastings algorithm for the estimation of parameters, in MATLAB programming language.

Using Bayesian approach, prior distributions of the model parameters must be specified in order to complete the Bayesian framework. To express ignorance about the parameter values before data collection and to get the results similar to those obtaining from the maximum likelihood approach, diffuse and non-informative prior distributions can be used. In this paper, we use as many diffuse or non-informative priors as possible and these prior distributions are assumed to be independent.

In the GARCH models, the priors for the coefficients in the conditional variance equation are $\alpha_0 \sim Ga(a_0, b_0)$, $\alpha_1 \sim Be(a_1, b_1)$ and $\alpha_2 \sim Be(a_2, b_2)$, where Be stands for the beta distribution. The priors for α_1 and α_2 suggest that $0 \leq \alpha_1 < 1$ and $0 \leq \alpha_2 < 1$ for stationarity condition. For the gamma and beta distributions, non-informative and diffuse priors can be obtained by setting $a_0 = b_0 = 0$ and $a_i = b_i = 1, i = 1, 2$, respectively. Significance test on $\alpha_2 = 0$ can be performed to assess

whether there is a persistence in the volatility. In the SV models, the priors are $\mu \sim N(\mu_0, \sigma_0^2)$, $\tau^{-2} \sim Ga(a_\tau, b_\tau)$, $2^{-1}(1+\phi) \sim Be(a_\phi, b_\phi)$ where the persistence parameter $-1 \leq \phi \leq 1$ as required. σ_0^2 can be set to a very large number to suggest a diffuse prior for μ and $a_\tau = b_\tau = 0$ for a non-informative prior for τ^{-2}. The priors used for the degrees of freedom of the return equation ν_1 and volatility equation ν_2 are truncated exponential distributions $\nu_i \sim Exp(\psi)I(a,b), i = 1,2$ where $\psi > 0$ is chosen such that the mean reflects an integer number around 5 or 10 and $I(a,b)$ restricts the range of possible values. When utilising a skew-*t* error distribution for the return equation, a non-informative prior is used for the skewness parameter γ to indicate our ignorance about the presence of skewness or the direction of it.

For all models, a single Markov chain is run for 110,000 iterations with the first 10,000 discarded as the burn-in period in the Gibbs sampling algorithm. To avoid highly correlated simulated realisations, we take every 100th iteration to mimic a sample of size 1,000 for posterior inferences. Convergence of the Markov chain is monitored through the trace plots of Markov chain iterations for each model parameter.

5 Results and Discussion

This section is divided into three subsections to facilitate the analysis of political cycles based on the various models used. In Section 5.1 we estimate the GARCH and GARCH-M models as in Worthington (2009) using updated regression variables. Section 5.2 analyses the data using the SV and SV-M models with various error distributions, and provides evidence that a stochastic volatility specification better fits the data. Model comparison is performed in Section 5.2 and the preferred model is used in Section 5.3 to interpret parameters which have an effect on market returns. Section 5.4 analyses the pre- and post-WWII data using the preferred model. Table 3 describes the model specification and error distributions considered throughout this Section.

Developed by Spiegelhalter *et al.* (2002) for complex hierarchical models where the number of parameters is not clearly defined, the Deviance Information Criterion (DIC) is used for model comparison. The DIC is defined as

$$\text{DIC} = E\left[D(\boldsymbol{\theta})\right] + \Delta D(\hat{\boldsymbol{\theta}})$$

Where $\hat{\boldsymbol{\theta}}$ is the posterior mean of the vector of model parameters $\boldsymbol{\theta}$, $E\left[D(\boldsymbol{\theta})\right]$, the expected value of the deviance, is a measure of the adequacy of model fitting and $\Delta D(\hat{\boldsymbol{\theta}}) = E\left[D(\boldsymbol{\theta})\right] - D(\hat{\boldsymbol{\theta}})$ estimates the effective number of parameters in the model. Amongst several models, the model having the smallest DIC is preferred. Despite there being other measures of model fit in a Bayesian context such as the Bayes factor and Bayesian Information Criterion (BIC), the DIC is used because it has been applied successfully in the past to a family of SV models in Berg, Meyer and Yu (2004) and it can be easily computed in WinBUGS and MATLAB.

Table 2 Monthly Return Comparisons

Statistic	*All*	*Non-Labor*	*Labor*
Returns (Jan 1901 - July 2011)			
Count	1327	892	435
Mean	0.4667	0.6135	0.1657
Median	0.6151	0.7174	0.3785
Maximum	20.08	13.24	20.08
Minimum	-39.08	-10.97	-39.08
Standard Deviation	3.590	2.960	4.611
Skewness	-1.301	-0.185	-1.665
Kurtosis	16.39	5.224	16.27
Jarque-Bera	10.2E+03***	185.7***	3313***
Returns (Jan 1901 - Dec 1945)			
Count	540	387	153
Mean	0.3874	0.6022	-0.1559
Median	0.4516	0.5977	0.1470
Maximum	20.0765	12.4045	20.0765
Minimum	-12.8517	-10.9261	-12.8517
Standard Deviation	2.6800	2.2702	3.4591
Skewness	0.1322	0.0328	0.4706
Kurtosis	12.0470	8.630	11.736
Jarque-Bera	1.80E+03***	495.3***	456.9***
Returns (Jan 1946 - July 2011)			
Count	787	505	282
Mean	0.5211	0.1512	0.3401
Median	0.8065	0.8635	0.7196
Maximum	13.2434	13.2434	11.1297
Minimum	-39.0799	-10.9745	-39.0799
Standard Deviation	4.1005	3.3973	5.1264
Skewness	-1.5230	-0.2290	-2.0105
Kurtosis	14.876	4.027	15.550
Jarque-Bera	4864***	1966***	25.65***

Notes: A two-sample t-test for equality of means (between Labor and non-Labor) fails to be rejected for returns for the full period (January 1901 to July 2011) (statistic = -1.85, p-value = 0.065) and for the latter half (January 1946 to July 2011) (statistic = -0.83, p-value = 0.41) but is rejected for the period before WWII (January 1901 to December 1945) (statistic = -2.51, p-value = 0.01). An F-test for equality of variances (between Labor and non-Labor) is rejected at the 1% level of significance for all three time periods. *** indicates 1% statistical significance.

5.1 GARCH Model Estimation and Comparison

Worthington (2009) used a GARCH-M specification to analyse the effects of political cycles on Australian stock market returns. In his return equation, Labor and

Table 3 Model Specification

	Specification	*Return Error Distribution*	*Volatility Error Distribution*
Model 1	GARCH, GARCH-M	Normal	Normal
Model 2	SV, SV-M	Normal	Normal
Model 3	SV, SV-M	Student-*t*	Student-*t*
Model 4	SV, SV-M	ST_1	Student-*t*
Model 5	SV, SV-M	ST_2	Student-*t*

Notes: ST_1 and ST_2 refer to the two forms of skew-*t* distributions discussed in Section 3.3. GARCH-M refers to the GARCH in mean specification and SV-M refers to the stochastic volatility in mean specification.

non-Labor were specified as two dummy variables and the model was fit without an intercept. However, his inferences might be subject to a form of misspecification in that the Labor and non-Labor variables contain the same information, and are in fact perfectly negatively correlated. We correct this misspecification by only defining one variable to indicate a Labor or non-Labor government and allowing for a constant as described in Section 3.

Table 4 provides parameter estimates for the GARCH(1,1) and GARCH-M(1,1) specifications, estimated via MCMC methods. With regard to the model specification, all volatility parameters are significant indicating that it is appropriate to model the volatility of returns as a time-varying variable. However, although the 'in-mean' specification is insignificant, the better model chosen according to DIC is the GARCH(1,1) model. In terms of political variables, the only significant variables are the Labor indicator variable and constant term, consistent with Worthington's findings. Contrary to his findings and perhaps an indication of model misspecification is that election indicator variable, β_3, is no longer marginally significant. Our constant term β_0 is highly significant, indicating strong evidence for a positive mean return for the time period considered. This is now separated from the effects of political party in power contained only in the Labor indicator variable.

5.2 SV Model Estimation

To increase the flexibility of volatility modelling, the deterministic volatility equation as in the GARCH models can be replaced by the stochastic volatility equation as in the SV models. Using Gibbs sampling algorithm, the SV models can incorporate different error distributions for the innovation terms of the return and volatility equations to account for heavy-tailed and skewed data and volatility. Table 5 outlines the parameter estimates for SV models with the normal, Student-*t* or skew-*t* distributions for the return and the normal or Student-*t* distributions for the log-volatility. Table 6 displays the estimates for the 'in-mean' versions of those models.

Table 4 Estimated GARCH Models

	GARCH(1,1)		*GARCH-M(1,1)*	
Parameter	*Coefficient*	*95% CI*	*Coefficient*	*95% CI*
Return equation				
β_0	0.6739***	(0.4195, 0.9197)	0.5509***	(0.2603, 0.8430)
	(0.1231)		(0.1483)	
β_1	-0.2924**	(-0.5834, -0.0102)	-0.4179***	(-0.7459, -0.0885)
	(0.1489)		(0.1716)	
β_2	-0.0001	(-0.0046, 0.0044)	-0.0002	(-0.0042, 0.0037)
	(0.0023)		(0.0021)	
β_3	0.2001	(-0.5174, 0.8753)	0.2063	(-0.5690, 0.9763)
	(0.3534)		(0.3964)	
β_4	0.0356	(-0.2692, 0.3702)	0.0841	(-0.2447, 0.4365)
	(0.1596)		(0.1731)	
γ_0			0.0002	(-0.0224, 0.0226)
			(0.0114)	
Volatility equation				
α_0	0.4378**	(0.2147, 0.7415)	0.4834***	(0.2605, 0.7911)
	(0.1409)		(0.1393)	
α_1	0.2466***	(0.1749, 0.3289)	0.2158***	(0.1539, 0.2943)
	(0.0400)		(0.0363)	
α_2	0.7431***	(0.6672, 0.8163)	0.7555***	(0.6761, 0.8259)
	(0.0392)		(0.0387)	
DIC	6691.3		6672.6	

Notes: The dependent variable is market returns for January 1901 to July 2011. The return equation includes an intercept term β_0, a dummy variable β_1 indicating Labor ministries, a continuous variable measuring time in office β_2, a dummy variable β_3 indicating an election, a dummy variable β_4 for data collected after the WWII and the in-mean parameter γ_0. The values in parentheses are the standard errors. *CI* stands for credible interval. *** indicates statistical significance at 1% level, ** at 5% level and * at 10% level. The values in parenthesis are standard errors. The volatility equation for the GARCH model includes a constant α_0, a first-order ARCH term α_1 and a first-order GARCH term α_2.

In comparing SV models with normal return and normal log-volatility (Model 2) to the GARCH models (Model 1), the smaller DIC value for the SV models reveals that the SV models are superior to the GARCH models in both original and in-mean specifications. This suggests that volatility is better modelled by a stochastic process than by a deterministic process.

The most important political variable, whether Labor is in power or not (β_1) is significant in all models. The coefficients for the other political variables, such as the time in power T_t and election event effects E_t are not useful in explaining market returns. We might improve upon these models by removing these variables and replacing them with macroeconomic factors which better describe market returns, such as the interest rate.

All coefficients in the volatility equation for all models are significant, suggesting that the SV model successfully captures the variation in market returns. Of note is also the degrees of freedom in models which use a Student-t or the skewed version in the return or volatility equations. The estimate for the degrees of freedom is small enough to confirm that the distribution of market returns and its volatility exhibit heavier tails than the normal distribution, as identified in Section 2.2. The degrees of freedom for the volatility equation are very small, ranging from 2.95 to 4.44 and have quite narrow confidence bands, suggesting that the distribution for the log-volatility have heavier tails than market returns. This is in contrast to the GARCH specifications, where the volatility is modelled by a deterministic process which is unable to capture random effects via a distribution. Although the asymmetry parameter θ is not significant in the skew-t models, there is the implication of slight positive skewness as evidenced by the Bayesian credible intervals.

Comparing the models by the choice of error distributions, the best model is in-mean version of Model 5, the SV-M model with a skew-t (ST_2) error distribution for the return equation and a Student-t error distribution for the volatility equation. While the two skew-t distributions are very similar, Model 5 may be preferred because of dependence in the tails of the distribution. We restrict our interpretation of the parameters to the in-mean version of Model 5 in the next section.

5.3 Parameter Interpretation

In the preferred model the Labor variable is significant, indicating that the presence of a Labor or non-Labor government has an effect on monthly market returns for the period between January 1901 and July 2011. In particular, the coefficient of this variable is negative which suggests that returns under Labor governments are less than market returns under non-Labor governments by about 0.26 percentage points. This result is very similar to what Worthington (2009) found under a GARCH-M(1,1) model and also follows the consensus that non-Labor (or right-of-centre) governments are more favourable to investors because of higher market returns. This finding also mirrors the market preference for right-of-centre presidents (Republicans) in the United States, as discussed in Section 1.

The other political variables, time in government T_t (β_2) and elections E_t (β_3) do not have any meaningful effect on market returns. This suggests that the market returns are independent of the time in which a political party has been in power. In addition, economic and social stimulus before an election as well as election surprises does not affect market returns. These results oppose the findings of similar analysis in the US, for example by Hensel and Ziemba (1995) who suggested that political re-election campaigns lead to higher stock market returns and Herbst and Slinkman (1984) who found support for a four-year political-economic cycle.

The persistence parameter ϕ in the volatility equation is also interpretable. The parameter is significantly different from zero and is very close to 1. This means that, in line with previous findings of volatility on market returns and the development of the GARCH methodology, volatility is persistent in that once high volatility of

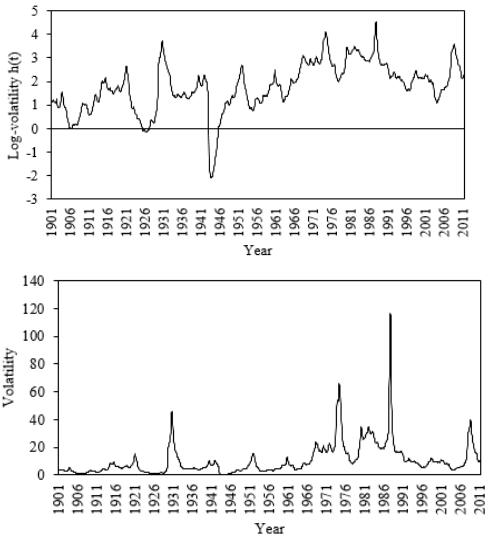


Fig. 2 Estimates of Log-Volatility and Volatility of Market Returns from 1901 - 2011

market prices exist, this persistently continues for some time. Similarly, there are clusters of low volatility as well. This phenomenon is displayed in Figure 2, which plots the estimates of unobserved volatility and log-volatility over time for the in-mean Model 5. An advantage of the SV model is that it can estimate the log-volatility and hence the volatility as a result of the parameters defined in the volatility equation. In particular, the model's estimation of high volatility corresponds to historic events such as the Great Depression in the early 1930's and recessions in Australia due to local and global factors in the 70's, 80's and early 1990's. It also captures the dip in volatility around 1942 during WWII, where there was very little trading in the markets. The average level of volatility has also increased in the late 20th century. It is clear from the graph that market volatility changes over time, hence SV models are appropriate.

In Model 5, the WWII indicator variable Wt is significant, indicating a positive change of about 0.39 percentage points in market returns after WWII compared to the period before. This makes sense as political endeavours and the macroeconomic environment has changed vastly since 100 years ago. The in-mean parameter γ_0 also suggests that market returns are dependent on the current value of their volatility. In particular, this parameter is negative which suggests that higher volatility leads to lower market returns.

Our interpretation was based on a model with a skewed distribution for the return equation, however the parameter associated with skewness θ is insensitive to the choice of a symmetric or asymmetric distribution. While the returns may resemble a symmetric distribution for the full period, they may follow different distributions

Table 5 Estimated Stochastic Volatility Models

	Model 2		*Model 3*		*Model 4*		*Model 5*	
Parameter	*Coefficient*	*95% CI*	*Coefficient*	*95% CI*	*Coefficient*	*95% CI*	*Coefficient*	*95% CI*
Return equation								
β_0	0.5788***	(0.3640, 0.7873)	0.5661***	(0.3579, 0.7802)	0.1293	(-0.9614, 1.219)	0.2246	(-0.3533, 0.9882)
	(0.1082)		(0.1091)		(0.5356)		(0.3475)	
β_1	-0.2873**	(-0.5418, -0.0360)	-0.2912**	(-0.5283, -0.0276)	-0.2992**	(-0.5124, -0.0577)	-0.2655**	(-0.5018, -0.0207)
	(0.1252)		(0.1266)		(0.1198)		(0.1226)	
β_2	-0.0003	(-0.0034, 0.0042)	-0.0001	(-0.0038, 0.0038)	0.0003	(-0.0033, 0.0047)	0.0001	(-0.0040, 0.0040)
	(0.0020)		(0.0019)		(0.0020)		(0.0020)	
β_3	0.1279	(-0.5193, 0.7648)	0.2329	(-0.4586, 0.8613)	0.2687	(-0.4605, 0.9495)	0.2386	(-0.4976, 0.8916)
	(0.3364)		(0.3408)		(0.3665)		(0.3404)	
β_4	0.1854	(-0.1089, 0.4683)	0.2299	(-0.0526, 0.5205)	0.2124	(-0.0621, 0.4965)	0.2073	(-0.0810, 0.4974)
	(0.1495)		(0.1458)		(0.1419)		(0.1477)	
Volatility equation								
μ	1.935***	(1.488, 2.395)	1.796***	(1.229, 2.398)	1.808***	(1.179, 2.452)	1.755***	(1.090, 2.323)
	(0.2378)		(0.3297)		(0.3290)		(0.3099)	
ϕ	0.9611***	(0.9390, 0.9801)	0.9751***	(0.9568, 0.9904)	0.9763***	(0.9567, 0.9930)	0.9750***	(0.9543, 0.9920)
	(0.0108)		(0.0086)		(0.0094)		(0.0099)	
τ	0.3082	(0.2512, 0.3806)	0.1519	(0.0787, 0.2502)	0.1296	(0.0378, 0.2396)	0.1335	(0.0548, 0.2468)
	(0.0339)		(0.0462)		(0.0524)		(0.0495)	
Model estimates								
ν_1			13.30	(7.43, 24.13)	14.27	(8.06, 25.67)	13.87	(7.32, 25.66)
			(4.35)		(4.66)		(4.63)	
ν_2			4.44	(1.82, 14.93)	3.14	(1.33, 9.45)	3.03	(1.49, 8.13)
			(3.49)		(2.11)		(1.73)	
θ					0.3803	(-0.5441, 1.413)	0.4126	(-0.4635, 1.078)
					(0.4725)		(0.3969)	
DIC	6479.06		6479.45		6441.93		6420.62	

Notes: The dependent variable is market returns for January 1901 to July 2011. The mean equation for both models include an intercept term β_0, a dummy variable β_1 indicating Labor ministries, a continuous variable measuring time in office β_2, a dummy variable β_3 indicating an election and a dummy variable β_4 for data collected after the WWII. The volatility equation includes a constant term μ and the persistence of the volatility ϕ. τ is the standard deviation of the log-volatility, which is always positive. ν_1 and ν_2 are the degrees of freedom for the Student-*t* error distributions of the mean and volatility equations respectively. θ is the degree of asymmetry in the skew-*t* return error distribution. The values in parentheses are the standard errors. *CI* stands for credible interval. *** indicates statistical significance at 1% level, ** at 5% level and * at 10% level which do not apply to τ and the degrees of freedom ν_1 and ν_2.

Table 6 Estimated Stochastic Volatility in-Mean Models

	Model 2 (in mean)		*Model 3 (in mean)*		*Model 4 (in mean)*		*Model 5 (in mean)*	
Parameter	*Coefficient*	*95% Credible Interval*	*Coefficient*	*95% Credible Interval*	*Coefficient*	*95% Credible Interval*	*Coefficient*	*95% Credible Interval*
Return equation								
β_0	0.6857***	(0.4607, 0.9182)	0.6476***	(0.4212, 0.8941)	0.0580	(-1.167, 1.422)	0.1322	(-0.4308, 0.9831)
	(0.3706)		(0.1226)		(0.6411)		(0.3706)	
β_1	-0.3074***	(-0.5593, -0.0542)	-0.3369**	(-0.5834, -0.0856)	-0.3205**	(-0.5777, -0.0689)	-0.2649*	(-0.5409, 0.0305)
	(0.1264)		(0.1272)		(0.1242)		(0.1418)	
β_2	-0.0006	(-0.0045, 0.0035)	-0.0007	(-0.0047, 0.0034)	-0.0004	(-0.0044, 0.0035)	-0.0004	(-0.0037, 0.0037)
	(0.0020)		(0.0021)		(0.0021)		(0.0020)	
β_3	0.1255	(-0.5250, 0.7911)	0.2592	(-0.4429, 0.8556)	0.2071	(-0.4689, 0.8555)	0.1666	(-0.6002, 0.7802)
	(0.3385)		(0.3442)		(0.3534)		(0.3483)	
β_4	0.3996**	(0.0543, 0.7292)	0.3916**	(0.0557, 0.7512)	0.4150**	(0.0398, 7610)	0.3855**	(0.0614, 0.7239)
	(0.1728)		(0.1774)		(0.1843)		(0.1710)	
γ_0	-0.0273**	(-0.0505, -0.0050)	-0.0242	(-0.0544, 0.0041)	-0.0280**	(-0.0593, 0.0005)	-0.0277*	(-0.0577, 0.0025)
	(0.0114)		(0.0149)		(0.0148)		(0.0152)	
Volatility equation								
μ	1.929***	(1.500, 2.442)	1.795***	(1.241, 2.346)	1.818***	(1.264, 2.389)	1.770***	(1.233, 2.336)
	(0.2318)		(0.2876)		(0.2842)		(0.3102)	
ϕ	0.9610***	(0.9369, 0.9817)	0.9754***	(0.9550, 0.9897)	0.9705***	(0.9503, 0.9884)	0.9735***	(0.9565, 0.9892)
	(0.0114)		(0.0088)		(0.0097)		(0.0087)	
τ	0.3044	(0.2442, 0.3812)	0.1292	(0.0510, 0.2339)	0.1624	(0.0800, 0.2845)	0.1325	(0.0602, 0.2452)
	(0.0361)		(0.0473)		(0.0544)		(0.0467)	
Model estimates								
ν_1			12.97	(7.73, 23.69)	16.39	(8.81, 27.28)	13.56	(7.77, 23.67)
			(4.11)		(4.90)		(4.10)	
ν_2			3.25	(1.49, 9.24)	4.02	(1.65, 13.59)	2.95	(1.50, 7.02)
			(2.23)		(3.07)		(1.88)	
θ					0.5385	(-0.6142, 1.713)	0.6234	(-0.3466, 1.235)
					(0.5727)		(0.4154)	
DIC	6475.22		6479.48		6433.78		6384.54	

Notes: The dependent variable is market returns for January 1901 to July 2011. The return equation for both models include an intercept term β_0, a dummy variable β_1 indicating Labor ministries, a continuous variable measuring time in office β_2 , a dummy variable β_3 indicating an election, a dummy variable β_4 for data collected after the WWII and the in-mean parameter γ_0. The volatility equation includes a constant term ν and the persistence of the volatility ϕ. τ is the standard deviation of the log-volatility, which is always positive. ν_1 and ν_2 are the degrees of freedom for the Student-t error distributions of the return and volatility equations respectively. θ is the degree of asymmetry in the skew-t return error distribution. The values in parentheses are the standard errors. *CI* stands for credible interval. *** indicates statistical significance at 1% level, ** at 5% level and * at 10% level which do not apply to τ and the degrees of freedom ν_1 and ν_2.

Table 7 Comparison before and after WWII

	Model 5 (before WWII)		*Model 5 (before in-mean)*		*Model 5 (after WWII)*		*Model 5 (after in-mean)*	
Parameter	*Coefficient*	*95% Credible Interval*	*Coefficient*	*95% Credible Interval*	*Coefficient*	*95% Credible Interval*	*Coefficient*	*95% Credible Interval*
Mean equation								
β_0	-0.1253	(-0.7003, 0.4800)	-0.2320	(-0.6695, 0.2370)	2.975***	(2.300, 3.628)	2.985***	(2.094, 3.752)
	(0.3031)		(0.2297)		(0.3364)		(0.4376)	
β_1	-0.2130	(-0.5025, 0.1235)	-0.2523	(-0.5754, 0.0725)	-0.1749	(-0.6514, 0.3360)	-0.1756	(-0.6682, 0.3225)
	(0.1544)		(0.1661)		(0.2472)		(0.0025)	
β_2	0.0042	(-0.0027, 0.0106)	0.0035	(-0.0033, 0.0102)	-0.0017	(-0.0065, 0.0033)	-0.0018	(-0.0065, 0.0034)
	(0.0034)		(0.0033)		(0.0025)		(0.0025)	
β_3	0.0721	(-0.8477, 0.9466)	0.0449	(-0.8698, 0.9755)	0.1743	(-1.016, 1.275)	0.1701	(-0.9394, 1.306)
	(0.4692)		(0.4705)		(0.5851)		(0.5708)	
γ_0			-0.0607	(-0.1443, 0.0151)			-0.0107	(-0.0451, 0.0253)
			(0.0409)				(0.0176)	
Volatility equation								
μ	0.7635*	(-0.2104, 1.668)	0.7944	(-0.3262, 1.994)	1.560*	(-0.8088, 3.150)	1.602*	(-0.0175, 3.102)
	(0.5540)		(0.5658)		(0.9387)		(1.303)	
ϕ	0.9754***	(0.9327, 0.9953)	0.9729***	(0.9137, 0.9956)	0.9824***	(0.9578, 0.9982)	0.9783***	(0.9508, 0.9972)
	(0.0155)		(0.0205)		(0.0108)		(0.0122)	
τ	0.1045	(0.0227, 0.3153)	0.1210	(0.0318, 0.4011)	0.1970	(0.1074, 0.3064)	0.2170	(0.1141, 0.3377)
	(0.0752)		(0.0911)		(0.0489)		(0.0591)	
Model estimates								
ν_1	6.226	(3.94, 9.94)	6.72	(4.11, 12.12)	16.36	(8.41, 27.56)	18.08	(8.87, 28.71)
	(1.577)		(2.05)		(5.22)		(5.31)	
ν_2	2.48	(1.10, 10.68)	2.23	(1.11, 6.69)	7.77	(2.42, 19.11)	8.08	(2.49, 21.92)
	(2.544)		(1.61)		(4.35)		(5.09)	
θ	0.6389*	(-0.0007, 1.221)	0.9569***	(0.5825, 1.451)	-2.658***	(-3.350, -1.980)	-2.560***	(-3.285, -1.607)
	(0.3280)		(0.2145)		(0.3618)		(0.4658)	
DIC	2218.57		2182.14		3979.16		3966.99	

Notes: The dependent variable is market returns for the dates specified in each Column. Model 5 is the SV model with ST_2 error distribution for the mean equation and Student-t error distribution for the volatility equation. The mean equation includes an intercept term β_0, a dummy variable β_1 indicating Labor ministries, a continuous variable measuring time in office β_2, a dummy variable β_3 indicating an election and the in-mean parameter γ_0. The volatility equation includes a constant term μ and the persistence parameter of the volatility ϕ. τ is the standard deviation of the log-volatility, which is always positive. ν_1 and ν_2 are the degrees of freedom for the error distributions of the return and volatility equations respectively.θ is the degree of asymmetry in the ST_2 error distribution. The values in parentheses are the standard errors. *CI* stands for credible interval. *** indicates statistical significance at 1% level, ** at 5% level and * at 10% level which do not apply to τ and the degrees of freedom ν_1 and ν_2.

if we divide the data as suggested by the positive level change associated with the WWII indicator variable. As such, in the next subsection we perform an additional analysis for Model 5 by physically splitting the data set into two periods and find a very different interpretation of political effects on market returns.

5.4 Before and after WWII Analysis

To see whether there is a difference in the effect of political variables on the market return before and after WWII, the data were split into two parts with 540 observations from January 1901 to December 1945 and 787 observations from January 1946 to July 2011. Table 7 summarises the results when the best model, SV-M model (Model 5), is fit to this data.

After separating the data, market returns are not affected by the political climate as no political variables are significant. This means that while there is a minor political effect in terms of differences between Labor and non-Labor when considering the whole time period, there is no difference between parties when only considering data before and after the war. In addition, the mean return of the AORD is around zero before WWII, but positive and significant around 3% as indicated by β_0 after the war. Of interest is also the estimate of μ in the volatility equation, which generally indicates the level at which the log-volatility deviates. When μ is absent from the volatility equation, the current log-volatility depends only on a fraction of the previous log-volatility, here given by ϕ which is close to 1, indicative of a nearly non-stationary process.

The most interesting of these results is that the asymmetry parameter θ is now significant. In particular, market return on the AORD index is slightly positively skewed before the war with parameter estimate of 0.6389 or 0.9569 (depending on preference of model) and negatively skewed after the war with parameter estimate around -2.6. This is consistent with previous studies such as in Ghysels, Plazzi and Valkanov (2011) who showed conditional and unconditional negative asymmetry for market returns in a number of developed global markets in the past 20 years, as well as from our descriptive analysis in Section 2.2.

6 Conclusion

This research built upon a previous study by Worthington (2009), by attempting to use more relevant models which capture market volatility in order to assess the effect of political cycles in the Australian stock market since January 1901. It was shown that modelling political variables using stochastic volatility techniques not only provides a better fit, but also allows for specification of more flexible error distributions for the return and volatility equations, as the market returns considered are not normally distributed. The best model used for analysis of political cycles was a SV model with a skew-*t* (ST_2) distribution for the return equation and Student-*t* distribution for the volatility equation.

The descriptive analysis indicated that there is only a significant difference of market returns between Labor and non-Labor ministries in the period before WWII, while there are significant differences between variances for all time periods. When using the best specification in which volatility varies stochastically, it was found that for the full period between January 1901 and July 2011 Labor governments have significantly lower mean return than non-Labor governments by about 0.28%. This is consistent with previous work mentioned in Section 1 in which there is evidence that the market prefers non-Labor governments. A reason for this may be because inflation is shown to be higher under Labor governments (Anderson, Malone and Marshall, 2008), and this analysis did not consider returns in excess of inflation. However, when splitting the data in a pre- and post-WWII analysis, the difference in returns between the governments disappears. Other political variables of note such as time in office and election surprises are also non-existent in Australia according to this analysis. Overall, there is very weak evidence to suggest political cycles have an effect and are able to predict accurately market returns in Australia.

This paper also indicated strong preference for non-deterministic modelling of the volatility of market returns. It was shown that market returns and volatility can be modelled appropriately by heavy tailed distributions, especially for market volatility which exhibits particularly heavy tails. It was also found that skewed distributions are preferred for modelling Australian market returns when considering the early 20th century compared to more recent data. If a robust distribution is not used in market analysis, possible outliers in return and volatility may affect parameter estimation.

Despite improving on the volatility model used in this basic analysis, we left modest scope for improving on the general methodology of choice and definition of political variables. Firstly, only political variables were used to explain market returns, ignoring other important macroeconomic and global effects. For example, Bohl and Gottschalk (2006) used macroeconomic variables in a similar analysis which capture the effects of the business cycle. Secondly, this analysis assumes the policies and workings behind each Labor and non-Labor government is constant over time, which is clearly not the case. A more substantial analysis should include the identification of government by phases in which policies on macroeconomic topics are consistent. Two important limitations, as also mentioned by Worthington (2009) is that we are unable to differentiate between large and small market capitalisation in Australian market return despite there being significant evidence between right and left wing parties in the US by Hensel and Ziemba (1995), as well as issues with the limited frequency of sampling.

As a final remark, this paper handles the model misspecification in Worthington (2009) and provides better techniques to study the effect of political cycles on Australian stock market return and volatility. Such techniques can also be used in many other areas other than political cycles. For the study of political cycles, the inclusion of additional variables such as government reforms, economic and financial market policies, returns of other international stock market indices, etc. can further improve the performance of the models.

References

1. Allvine, F., O'Neill, D.: Stock market returns and the presidential Election Cycle. Financial Analysts Journal 36, 49–56 (1980)
2. Anderson, H., Malone, C., Marshall, B.: Investment returns under right-and left-wing governments in Australasia. Pacifc-Basin Finance Journal 16, 252–267 (2008)
3. Andrieu, C., Doucet, A., Robert, C.: Computational advances for and from Bayesian analysis. Statistical Science 19, 118–127 (2004)
4. Beedles, W.: Asymmetry in Australian Equity Returns. Australian Journal of Management 11, 1–12 (1986)
5. Berg, A., Meyer, R., Yu, J.: Deviance information criterion for comparing stochastic volatility models. Journal of Business & Economic Statistics 22, 107–120 (2004)
6. Bialkowski, J., Gottschalk, K., Wisniewski, T.: Stock market volatility around national elections. Journal of Banking & Finance 32, 1941–1953 (2008)
7. Bohl, M., Gottschalk, K.: International evidence on the Democrat premium and the presidential cycle effect. The North American Journal of Economics and Finance 17, 107–120 (2006)
8. Bollerslev, T.: Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307–327 (1986)
9. Booth, J., Booth, L.: Is presidential cycle in security returns merely a reflection of business conditions? Review of Financial Economics 12, 131–159 (2003)
10. Branco, M., Dey, D.: A general class of multivariate skew-elliptical distributions. Journal of Multivariate Analysis 79, 99–113 (2001)
11. Cahan, J., Malone, C., Powell, J., Choti, U.: Stock market political cycles in a small, two-party democracy. Applied Economics Letters 12, 735–740 (2005)
12. Choy, S., Chan, J.: Scale mixtures distributions in statistical modelling. Australian and New Zealand Journal of Statistics 50, 135–146 (2008)
13. Danielsson, J.: Stochastic volatility in asset prices estimation with simulated maximum likelihood. Journal of Econometrics 64, 375–400 (1994)
14. Demarta, S., McNeil, A.: The t copula and related copulas. International Statistical Review 73, 111–129 (2005)
15. Fung, T., Seneta, E.: Modelling and estimation for bivariate financial returns. International Statistical Review 78, 117–133 (2010)
16. Gerlach, R., Chen, C.: Bayesian inference and model comparison for asymmetric smooth transition heteroskedastic models. Statistics and Computing 18, 391–408 (2008)
17. Ghysels, E., Harvey, A., Renault, E.: Stochastic Volatility. In: Maddala, G., Rao, C. (eds.) Handbook of Statistics: Statistical Methods in Finance, vol. 14. North-Holland, Butterworth Heinemann, Amsterdam (1996)
18. Ghysels, E., Plazzi, A., Valkanov, R.: Conditional skewness of stock market returns in developed and emerging markets and its economic fundamentals. Swiss Finance Institute Research Paper, 11-06 (2011)
19. Gilks, W., Richardson, S., Spiegelhalter, D.: Markov Chain Monte Carlo in Practice. Chapman and Hall, London (1996)
20. Gray, P., Kalotay, E.: Testing the multivariate normality of Australian stock returns. Australian Journal of Management 23, 135–150 (1998)
21. Hensel, C., Ziemba, W.: United States investment returns during Democratic and Republican Administrations. Financial Analysts Journal 51, 61–69 (1995)
22. Herbst, A., Slinkman, C.: Political-cconomic cycles in the U.S. stock market. Financial Analysts Journal 40, 38–44 (1984)

23. Hibbs, D.J.: Political parties and macroeconomic policy. The American Political Science Review 71, 1467–1487 (1977)
24. Huang, R.: Common Ssock returns and presidential elections. Financial Analysts Journal 41, 58–65 (1985)
25. Jacquier, E., Polson, N., Rossi, P.: Bayesian analysis of stochastic volatility models. Journal of Business & Economic Statistics 12, 371–389 (1994)
26. Koopman, S., Hol Uspensky, E.: The stochastic volatility in mean model: empirical evidence from international stock markets. Journal of Applied Econometrics 17, 667–689 (2002)
27. Niderhoffer, V., Gibbs, S., Bullock, J.: Presidential elections and the stock market. Financial Analysts Journal 26, 111–113 (1970)
28. Powell, J., Jing, S., Smith, T., Whaley, R.: The persistent presidential dummy. Journal of Portfolio Management 33, 133–143 (2007)
29. Praet, P., Wilson, E.: The distribution of stock market returns: 1958-1973. Australian Journal of Management 3, 79 (1978)
30. Riley, W.J., Luksetich, W.: The market prefers republicans: myth or reality. The Journal of Financial and Quantitative Analysis 15, 541–560 (1980)
31. Ruiz, E.: Quasi-maximum likelihood estimation of stochastic volatility models. Journal of Econometrics 63, 289–306 (1994)
32. Santa-Clara, P., Valkanov, R.: The presidential puzzle: political cycles and the stock market. The Journal of Finance 58, 1841–1872 (2003)
33. Smith, A., Roberts, G.: Bayesian computation via the Gibbs sampler and related Markov Chain Monte Carlo methods. Journal of the Royal Statistical Society, Series B 55, 3–23 (1993)
34. Snowberg, E., Wolfers, J., Zitzewitz, E.: Partisan impacts on the economy: Evidence from prediction markets and close elections. The Quarterly Journal of Economics 122, 807–829 (2007)
35. Spiegelhalter, D., Best, N., Carlin, B., Van Der Linde, A.: Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society, Series B 64, 583–639 (2002)
36. Spiegelhalter, D., Thomas, A., Best, N., Lunn, D.: Bayesian inference using Gibbs sampling for Windows (WinBUGS), Cambridge, UK (2007)
37. Stacy, E.: A Generalization of the Gamma Distribution. The Annals of Mathematical Statistics 33, 1187–1192 (1962)
38. Wisniewski, T.: Can political factors explain the behaviour of stock prices beyond the standard present value models? Applied Financial Economics 19, 1873–1884 (2009)
39. Worthington, A.: Political cycles in the Australian stock market since Federation. Australian Economic Review 42, 397–409 (2009)
40. Yu, J.: Forecasting volatility in the New Zealand stock market. Applied Financial Economics 12, 193–202 (2002)

Dependence Structure between Crude Oil, Soybeans, and Palm Oil in ASEAN Region: Energy and Food Security Context

Teera Kiatmanaroch and Songsak Sriboonchitta

Abstract. The increase in energy and food prices remains a challenge for the ASEAN Economic Community (AEC). Understanding the dependence between energy prices and food prices is imperative for the energy and food security for the people of the ASEAN countries. The C-vine copula model is a flexible tool to analyze the relationship between variables, in which the multivariate dependence modeling. It offers us to define the relationship structure between variables or we call pair-copula construction, according to the purpose of study. This study is interesting to examine an influence of crude oil price on palm oil price and soybeans price. The results can conclude that the change of crude oil price has influence on the prices of palm oil and soybeans. Moreover, the findings show that there exists the dependence between palm oil price and soybeans price, and crude oil price is one factor that has influence on relation of their prices. However, the dependence structure of the static copula for Crude oil–Palm oil (C,P), Crude oil–Soybeans (C,S), Palm oil–Soybeans (P,S), there exists a weak positive dependence in each pair-copula. This indicates that the price of each commodity is slightly related to the price of every other. In the case of soybeans, the ASEAN members should cooperate and incorporate their efforts to increase the capacity and performance in production to reduce relying on soybeans being imported from outside the region.

1 Introduction

By 2015, the nations in the Southeast Asian region consisting of Brunei, Cambodia, Indonesia, Laos, Malaysia, Myanmar, the Philippines, Singapore, Thailand, and Vietnam will agree on establishing an ASEAN Economic Community (AEC), which has a total population of approximately 600 million people. This regional integration shall lead to a single market and production that will induce free movement of goods, services, investment, capital, and skilled labor across the ASEAN

Teera Kiatmanaroch · Songsak Sriboonchitta
Faculty of Economics, Chiang Mai University, Chiang Mai 50200 Thailand
e-mail: pornan@kku.ac.th, songsakecon@gmail.com

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_21,

region [1, 2]. The ASEAN boundary is adjacent to the south of China and is linked to the east of India, both by sea and land. India and China are part of the BRIC countries, which are the newly industrialized countries, and are considered as two of the nations that have fast growing economies [3]. According to the information mentioned above, the premise is that the economic geography makes the AEC play an important role in the global economy. However, during the past several years, the AEC has remained restricted due to some challenging circumstances caused by the global financial crisis in 2008. In addition, the food and fuel crises have caused a huge burden on the people who are poor and near-poor in the ASEAN region, and created a negative impact with regard to their social and economic development [4]. The rise in food prices came about due to many factors, such as climate change which caused a decline in the agricultural production, a rise in fuel prices which led to a domino effect on the cost of production, and the increase in consumer demand [5].With regard to the rise in fuel prices, the incidence of such factors was due to an increasing demand in Asia, especially in the emerging markets of India and China [6]. The rise in food and energy prices is a real challenge for the ASEAN members while trying to find any crucial means to cooperate in the short- and long-term situations to solve the problems because food[1] and energy[2] security are fundamental for upholding the ASEAN economic and social development goals [7].

Palm oil and soybeans are food commodities that are related to food security in the ASEAN region because they are used as raw materials in food production and are converted to the necessary goods, and also used for other aspects of daily life. Palm oil can be modified as cooking oil, shortening, margarines, etc. Soybeans can be modified as cooking oil, soy milk, soy sauce, tempeh, tofu, etc. Moreover, palm oil and soybean oil can be used to produce alternative energy such as the biodiesel types, Palm Methyl Ester (PME) and Soy Methyl Ester (SME), respectively. In ASEAN, palm oil can be produced sufficiently for intra-regional demand and the remaining parts can be kept aside for exportation. In 2012/2013, Indonesia and Malaysia exported palm oil of an approximate volume of 37,300 thousand metric tons, or 89.66% of the total world exports, which was 41,603 thousand metric tons [8]. However, in the case of soybeans, it has to be imported from outside the region. In 2012/2013, Indonesia, Thailand, and Vietnam imported about 5,300 thousand metric tons or 5.66% of the total world imports, which was 93,587 thousand metric tons [9].

[1] FAO [14] definition: Food security exists when all people, at all times, have physical, social, and economic access to sufficient, safe, and nutritious food to meet their dietary needs and food preferences for an active and healthy life. The four pillars of food security are availability, access, utilization, and stability. The nutritional dimension is integral to the concept of food security.

[2] United Nations [15] definition: Energy security is a term that applies to the availability of energy at all times in various forms, in sufficient quantities, and at affordable prices, without unacceptable or irreversible impact on the environment. These conditions must prevail over the long term if energy is to contribute to sustainable development. Energy security has both a producer and a consumer side to it.

ASEAN has crude oil resources and oil production, but does not have a sufficient supply to meet the intra-regional demand. In 2011, ASEAN imported crude oil worth not less than 90,000 million US dollars [10]. ASEAN imports crude oil especially from the Middle East [11]. Although the crude oil benchmark prices of the international crude oil markets are from Brent, West Texas Intermediate (WTI), Dubai, and Maya, each of these markets is related to one another. It was found that in times of crude oil market stress, the crude oil price in each market tends to have co-movement with the same intensity [25]. In addition, we found that the crude oil markets are related to the food markets. As in the previous studies of the relationship between energy and agricultural prices, it can be concluded that the long-run agricultural prices can be driven by the energy prices and that volatility in the energy markets is transmitted to the food markets [13].

Over the past several years, there have been some evidences of significant volatility transmissions between the crude oil prices in each of these markets. Moreover, the volatility in the oil prices can be transmitted to the various food markets. Thus, it is interesting to analyze the relationship between the crude oil benchmark prices of the ASEAN and the prices of the two food commodities that can be used to generate alternative energy, which are the following: (1) palm oil, which can be produced and be sufficient for intra-regional demand and (2) soybeans, which rely on imports from outside the region. Since these commodities are related to the energy and food security for the people in the ASEAN region, and can also be substituted for each other, it would be quite interesting to learn about the dependence structure of these commodity prices. Furthermore, it will be useful for making decisions and plans for the economic and social development of the AEC. Therefore, the purposes of the study are as follows: (1) to analyze the dependence between crude oil prices (DME) and two food prices, namely, the prices of soybeans (CBOT) and palm oil (MDEX) and (2) to analyze the dependence between the soybeans and palm oil prices, with the crude oil prices as the conditioning variable. The GARCH model was applied to examine the volatility of the futures prices 1-Pos. of the three data series and the vine copula model was used to analyze the dependence structure between their marginal distributions. The data analyses were based on the daily observations from the period of June 2007 to March 2013.

The remainder of this work is organized as follows: part two is the methodology, and part three consists of the data and the empirical findings. Finally, part four comprises the conclusions.

2 Methodology

Over the past several years, there have been arguments about the relationship between the energy prices (e.g., crude oil, biodiesel, and ethanol) and the agricultural commodity prices (e.g., palm oil, soybeans, corn) as to whether they are related or not. The argument was always divided between a relation and an absence of relation. From the literature review, we come to know that relationships do exist between the energy prices and the agricultural commodity prices; what is more, there are

relationships between the prices of the different agricultural commodities themselves. The findings on these relationships depend on many factors such as the period of study, the data frequency, the statistical analysis, and the modeling. As for modeling, a number of different models were used in the studies prior to this study. Baffes [16] used the ordinary least squares (OLS) to analyze the relationship between the commodity prices and the crude oil price. Serra and Zilberman [13] mentioned about many econometrics and statistical models that the previous studies used to find the relationship between the energy prices and the agricultural commodity prices, and the relationships between the prices of the different commodities. A few of such applicable tools are cointegration, causality, vector error correction model (VECM), vector autoregressive (VAR), autoregressive distributed lag models (ARDL), vector auto regression moving-average (VARMA), stochastic volatility model with Merton jumps (SVMJ), panel data, minimal spanning and hierarchical trees, random parameter model, wavelet, GARCH modeling, and copula modeling. As mentioned above, we found that the statistics used for analyzing are both parametric and non-parametric, and that the relationship analysis between the variables is both linear and non-linear.

There were several models and each of the models was based on different assumptions in order to test the data. Sriboonchitta et al. [17] applied the copula based GARCH for modeling the volatility and dependency of the agricultural price and production indices of Thailand. Based on the study, the work mentioned that this approach provided more flexibility for finding out the joint distributions and the transformation of the invariant correlation, without the assumption of linear correlation. Therefore, in this study, we used the GARCH(1,1) model [18] to examine the volatility of the commodity daily prices which are generally non-normal distributions and applied the vine copula model to examine the relationship between each commodity.

The R-package fGarch by Wuertz and Chalabi [19] was used to estimate the GARCH(1,1) model with the skewed student T (SkT) residual distribution for the marginal distribution of the log-difference $\ln \frac{P_t}{P_{t-1}}$ or the growth rate of crude oil prices, palm oil prices, and soybeans prices. The standardized residuals with the skewed student T were transformed to copula data $(F_1(x_1), F_2(x_2), F_3(x_3))$ by using the empirical distribution function. After that, we used the R-package CDVine which was developed by Brechmann and Schepsmeier [20] to estimate the bivariate copula and C-vine copula.

This study used the C-vine copula modeling to analyze the dependence between the crude oil prices from the Dubai market (DME) and the two food prices consisting of palm oil prices from the Malaysia market (MDEX) and soybeans prices from the Chicago market (CBOT), which no one has studied before. The structure of the C-vine model is shown in Figure 1. This study selected crude oil which was the first root node, as Brechmann and Schepsmeier [20] hold the view that a vine structure can be chosen manually or through expert knowledge. Aas et al. [21] said that modeling C-vine might be advantageous when we know a main variable that governs the interactions in the data, or when it plays an important role in the dependence structure and when the others are linked to it. Therefore, our assumption in

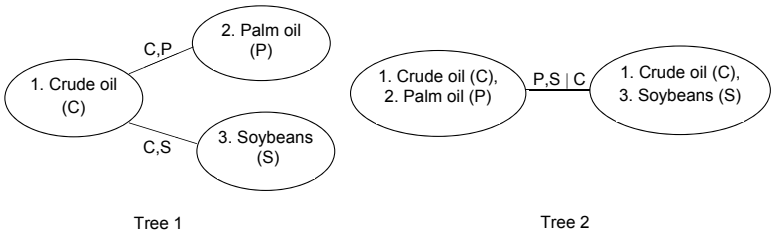


Fig. 1 The pair-copulas of three-dimensional C-vine trees

this study is that crude oil prices is a key variable as Serra and Zilberman [13] point out that energy prices can drive the long-run agricultural price levels.

3 Data and Empirical Findings

To analyze the relationship between crude oil prices and two food prices (palm oil and soybeans), we selected the commodity prices that are related to the AEC. The crude oil benchmark price for the Asian market is the Dubai (Oman) crude oil price [22] since the Middle East is the major source of crude oil for ASEAN [11]. Thus, the crude oil price of the Dubai Mercantile Exchange (DME) was used in this study. Palm oil prices were obtained from the Malaysia Derivatives Exchange (MDEX) because Malaysia is a major producer and a world exporter of palm oil [8]. In ASEAN, soybean production in the intra-region was insufficient for meeting the demand; most of the soybeans was imported from Brazil, Argentina, and America. Indonesia, Thailand, and Vietnam are the major importers of soybeans in the Asian region due to their demand for soybeans in the food industry, livestock industry, and so on [9, 23, 24, 25]. Therefore, we used the soybeans prices of the Chicago Board of Trade (CBOT) since it provides an updated data and it can be used as a reference price in the world market. The observations were based on the Futures 1-Pos of the daily close prices during the period from 1 June 2007 to 15 March 2013, from the EcoWin database. Each price data series was transformed into the log-difference $\ln \frac{P_t}{P_{t-1}}$, or the growth rates of the prices before were used to analyze by using the vine copula based GARCH model.

Table 1 presents a descriptive statistics of the growth rates of crude oil, palm oil, and soybeans. Crude oil and soybeans have positive average growth rates but palm oil has negative average growth rates. All of three data series exhibit negative skewness. If skewness is negative, the market has a downside risk or there is a substantial probability of a big negative return. The kurtosis of these data is greater than 3. Therefore, this kurtosis is called super Gaussian and leptokurtic. This means that the growth rates of the empirical data have a typically spiky probability distribution function with heavy tails. The null hypothesis of the normality of the Jarque-Bera tests are rejected in all the data series. The Dickey-Fuller test shows that these data series are stationary at p-value 0.01.

Table 1 Data Descriptive Statistics for Log-difference of Crude Oil, Palm Oil, and Soybeans Prices

	Crude oil	Palm oil	Soybeans
Mean	0.000354	−0.000107	0.000403
Median	0.000899	0.000000	0.001304
Maximum	0.133869	0.097638	0.203209
Minimum	−0.133661	−0.110391	−0.234109
Std. Dev.	0.023000	0.020276	0.020557
Skewness	−0.157438	−0.347154	−0.898968
Kurtosis	7.68	7.03	23.50
Jarque-Bera	1,265.75	961.06	24,341.97
(p-value)	(0.0000)	(0.0000)	(0.0000)
p-value of Dickey-Fuller test	0.01	0.01	0.01
Number of observations	1,379	1,379	1,379

From the data given in Table 1, it can be seen that the three data series are inappropriate with normal distribution, and exhibit negative skewness and excess kurtosis. Therefore, the GARCH(1,1) with the skewed student T residual distribution, $\varepsilon_t \sim SkT(\nu, \gamma)$, was modeled for examining the volatility and for estimating the marginal distributions.

Table 2 presents the result of GARCH(1,1) with skewed student T residual. The asymmetry parameters, γ, are significant and less than 1, exhibiting that all the data series are skewed to the left. For crude oil, palm oil, and soybeans, the $\alpha + \beta$ are 0.9980, 0.9901, and 0.9894, respectively; this implies that their volatilities have long-run persistence. For the short-run effect of the unexpected factors, we consider the event from the α parameter. Therefore, we can see that they have close values (0.0529, 0.0746 and 0.0483) and a small impaction for volatility.

Next, we transformed the standardized residuals from the GARCH(1,1) model into uniform [0,1] by using the empirical distribution function $F_n(x) = \frac{1}{n+1}\Sigma_{i=1}^{n} 1(X_i \leq x)$, where $X_i \leq x$ is the order statistics and 1 is the indicator function. The transformed data were used in the Kolmogorov-Smirnov (K-S) test for uniform [0,1] and the Box-Ljung test for serial correlation. More details are illustrated in Patton [26] and Manthos [27]. These tests are necessary to check for the marginal distribution models' misspecification before using the copula model.

The results of the K-S test show that these marginal distributions are uniform, by accepting the null hypothesis at p-values equal to 1 or nearly 1. The results of the Box-Ljung test provide that all of the four moments of all the marginal distributions are i.i.d. by accepting the null hypothesis that does not have a serial correlation at p-value greater than 0.05. Therefore, our marginal distributions were not misspecified and can be used for the copula model.

Table 2 Results of GARCH(1,1) with Skewed Student T Residual for Log-difference of Crude Oil, Palm Oil, and Soybeans Prices

	Crude oil	Std. error (p-value)	Palm oil	Std. error (p-value)	Soybeans	Std. error (p-value)
ω	2.325e-06	1.749e-06 (0.184)	3.903e-06	1.721e-06 (0.0233 *)	4.428e-06	1.674e-06 (0.00817 **)
α	0.0529	1.214e-02 (1.32e-05 ***)	0.0746	1.501e-02 (6.75e-07 ***)	0.0483	1.115e-02 (1.52e-05 ***)
β	0.9451	1.231e-02 (< 2e-16 ***)	0.9155	1.606e-02 (< 2e-16 ***)	0.9411	1.173e-02 (< 2e-16 ***)
ν (degree of freedom)	5.067	7.455e-01 (1.07e-11 ***)	7.681	1.485e+00 (2.31e-07 ***)	4.917	6.933e-01 (1.32e-12 ***)
γ (skewness)	9.418e-01	3.112e-02 (< 2e-16 ***)	9.685e-01	3.557e-02 (< 2e-16 ***)	8.795e-01	2.889e-02 (< 2e-16 ***)
Log likelihood	3,499.523	-	3,654.827	-	3,659.68	-
K-S test (p-value)	-	- (1)	-	- (0.9208)	-	- (1)
Box-Ljung test (p-value)	-	-	-	-	-	-
1st moment	-	- (0.5832)	-	- (0.2515)	-	- (0.9540)
2nd moment	-	- (0.7921)	-	- (0.8898)	-	- (0.4999)
3rd moment	-	- (0.7765)	-	- (0.0732)	-	- (0.4433)
4th moment	-	- (0.6423)	-	- (0.8803)	-	- (0.6692)

Note: Significant codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

3.1 Results of C-vine Copula Analysis

Figure 1, in the Part 2, presents each of the pair-copulas of the three-dimensional C-vine tree; there are two pair-copulas in Tree 1 and one pair-copula in Tree 2. The first and second pair-copulas in Tree 1 are Crude oil–Palm oil (C,P) and Crude oil–Soybeans (C,S), respectively. The third pair-copula in Tree 2 is a conditional pair-copula, Palm oil–Soybeans given Crude oil (P,S|C).

We use the Gaussian copula, Student's T copula, Clayton copula, Gumbel copula, Frank copula, Joe copula, rotated Clayton 180°, rotated Gumbel 180° copula, and rotated Joe 180° copula to fit the data. The AIC and the BIC are used to appraise as to which copula is the best fit. Kendall's tau correlation which was transformed from the copula parameter was used because each family of copula has a different range of copula parameters; hence we inverse a copula parameter into a Kendall's tau correlation, and it is bound on the interval $[-1,1]$. Kendall's tau is a measure of concordance which is a function of copula; thus, we can use it to assess the range of dependence covered by the families of copula. A goodness-of-fit test based on Kendall's tau provides the Cramér-von Mises (CvM) and Kolmogorov-Smirnov (KS) test statistics and the estimated p-values by bootstrapping [20] to test

the appropriateness of the copula model under the null hypothesis that the empirical copula C belongs to a parametric class C' of any of the copulas, $H_0 : C \in C'$.

The results of the pair-copulas Crude oil–Palm oil (C,P), Crude oil–Soybeans (C,S), and Palm oil–Soybeans given Crude oil (P,S|C) are presented in Table 3.

The first pair-copula, Crude oil–Palm oil, considering the values of the AIC and the BIC, the three most appropriate copulas in order are the Gaussian, Student's T, and rotated Gumbel 180°. But the second parameter (ν) of the Student's T copula is insignificant with p-values greater than 0.05. The CvM and KS tests of the Gaussian, Student's T, and rotated Gumbel 180° copula accept the null hypothesis with p-values greater than 0.05, which means that the dependence structure of the data series is appropriate for a chosen family. Therefore, the Gaussian copula is chosen to explain the dependence structure of this pair-copula with a copula parameter 0.2495 and a Kendall's tau correlation 0.16.

The second pair-copula, Crude oil–Soybeans, considering the values of the AIC and the BIC, the three most appropriate copulas in order are Student's T, Gaussian, and Frank. Although the Student's T copula is the best fit according to the AIC and the BIC, it does not give any results for the CvM and KS tests by estimation in the R-package CDVine. The Gaussian copula is a second order of the AIC and the BIC values, and shows that the CvM and KS tests accept the null hypothesis with p-values greater than 0.05. For the Frank copula, the CvM and KS tests reject the null hypothesis with p-values less than 0.05, which means that the Frank copula is not an appropriate model. Therefore, the Gaussian copula is chosen to explain the dependence structure of this pair-copula with a copula parameter of 0.3545 and a Kendall's tau correlation of 0.23.

The parameter of each pair-copula from an appropriate copula family in Tree 1 was used to construct the conditional pair-copula of Palm oil–Soybeans given Crude oil (P,S|C) in Tree 2 of the C-vine copula model, and the results are shown in Table 3.

For the conditional pair-copula, Palm oil–Soybeans given Crude oil, considering the values of the AIC and the BIC, the three most appropriate copulas in order are the Gaussian, Student's T, and Frank. Although the second parameter (ν) of the Student's T copula is insignificant with p-values greater than 0.05, it does not give any results for the CvM and KS tests by estimation in the R-package CDVine. The CvM and KS tests of the Gaussian and Frank copulas accept the null hypothesis with p-values greater than 0.05, which means that the dependence structure of the data series is appropriate for a chosen family. Therefore, the Gaussian copula is chosen to explain the dependence structure of this conditional pair-copula with a copula parameter 0.2303 and a Kendall's tau correlation 0.15.

In addition, the results of the bivariate copula analysis of Palm oil and Soybeans (P,S) are shown in Table 4. The Gaussian copula was chosen to explain the dependence structure between Palm oil and Soybeans by considering the AIC and the BIC values, and the CvM and KS tests accepted the null hypothesis with p-values greater than 0.05. The Gaussian copula gives a copula parameter of 0.2970 and a Kendall's tau correlation of 0.19.

Table 3 Results of C-vine Copula Model

Tree	Pair-copula	Copula family	Copula parameter	Std. error (p-value)	Kendall's tau	AIC	BIC	p-value CvM	p-value KS
1	C,P	Gaussian	0.2495	0.0245 (0.0000)	0.1600	−86.4655	−81.2364	0.24	0.19
		Student's T	0.2494	0.0250 (0.0000)	0.1605	−84.8997	−74.4415	0.59	0.61
			ν =53.4942	82.9504 (0.2596)					
		rotated Gumbel 180°	1.1675	0.0222 (0.0000)	0.1434	−77.9826	−72.7535	0.05	0.07
1	C,S	Gaussian	0.3545	0.0222 (0.0000)	0.2307	−182.9177	−177.6885	0.07	0.08
		Student's T	0.3606	0.0236 (0.0000)	0.2349	−190.5264	−180.0682	NA	NA
			ν =13.6722	5.1333 (0.0039)					
		Frank	2.2742	0.1692 (0.0000)	0.2407	−179.5645	−174.3354	0.01	0.01
2	P,S|C	Gaussian	0.2303	0.0249 (0.0000)	0.1480	−73.0587	−67.8296	0.98	0.99
		Student's T	0.2318	0.0258 (0.0000)	0.1489	−73.7226	−63.2643	NA	NA
			ν =26.1814	17.5220 (0.0677)					
		Frank	1.4004	0.1657 (0.0000)	0.1526	−69.5958	−64.3667	0.61	0.67

Table 4 Results of Palm Oil–Soybeans (P,S) of a Bivariate Copula Model

Pair-copula	Copula family	Copula parameter	Std. error (p-value)	Kendall's tau	AIC	BIC	p-value CvM	p-value KS
P,S	Gaussian	0.2970	0.0236 (0.0000)	0.1920	−125.1169	−119.8878	0.30	0.35
	Student's T	0.2990	0.0244 (0.0000)	0.1933	−125.5547	−115.0965	NA	NA
		ν =26.4778	18.4553 (0.0758)					
	Frank	1.8284	0.1666 (0.0000)	0.1967	−118.7307	−113.5015	0.02	0.16

By doing a comparison between a C-vine copula model, given in Table 3, and a bivariate copula model, given in Table 4, we found out that our results show that the copula parameters and the Kendall's tau correlations of a conditional pair-copula (P,S|C) in all the copula families are less than those that were obtained from

the bivariate pair-copula (P,S); for example, the Gaussian copula of the conditional pair-copula (P,S|C) offers the copula parameter and the Kendall's tau correlation as 0.2303 and 0.15, respectively. Further testing reveals that the Gaussian copula of the bivariate copula (P,S) offers the copula parameter and the Kendall's tau correlation as 0.2970 and 0.19, respectively.

This implies that crude oil price (C) has an influence on the relationship between palm oil price (P) and soybeans price (S). The crude oil price (C) is an important variable that governs the interactions in the dependence structure between the palm oil price (P) and the soybeans price (S).

4 Conclusions

The AEC plays an important role in the global economy. However, it remains in a state of challenge due to the many problems it faces, such as the global economic recession combined with food and fuel crises, which have an effect on the people who are poor and near-poor in the ASEAN region and can have a negative impact on the social and economic development. The rising prices of food and energy are the challenges for the ASEAN members to overcome. There exist evidences of significant price transmissions between the energy market and the food market. Thus, it is interesting to study the relationship between the crude oil benchmark prices of the ASEAN and the prices of the two food commodities that can be used to produce alternative energy, which are as follows: (1) palm oil, which can be produced and be sufficient for intra-regional demand and (2) soybeans, which relies on imports from outside the region. Gaining an understanding of the dependence structure of these commodity prices will be useful in making decisions and plans for the economic and social development of the AEC.

In this study, the data analyses were based on the daily observations from June 2007 to March 2013. The GARCH model was used to examine the volatility of the future prices 1-Pos. of the three data series and applied the C-vine copula model to examine the relationship between each commodity.

The empirical results of the GARCH(1,1) model with skewed student T residual show that the crude oil prices, palm oil prices, and soybeans prices have long-run persistence in volatility. The C-vine copula model was used to study the dependence structure between crude oil price, soybeans price, and palm oil price that related to ASEAN region. This study is interesting to examine an influence of crude oil price on palm oil price and soybeans price. The C-vine copula model is a flexible tool to analyze the relationship between variables, in which the multivariate dependence modeling. It offers us to define the relationship structure between variables according to the purpose of study, and it can describe the relationship between variables through the graphical model or are called pair-copulas, as shown in Figure 1. In this study, we assume crude oil price is a condition variable in C-vine structure. The finding results can conclude that the change of crude oil price has an influence on the prices of palm oil and soybeans. Moreover, the findings show that there exists

the dependence between palm oil price and soybeans price, and crude oil price is one factor that has an influence on relation of their prices.

The C-vine copula contains three pair-copulas: Crude oil–Palm oil (C,P) and Crude oil–Soybeans (C,S) in the first tree and a conditional pair-copula, Palm oil–Soybeans given Crude oil (P,S|C), in the second tree. For the pair-copula Crude oil–Palm oil (C,P), the Gaussian copula is chosen to explain its dependence structure with a copula parameter of 0.2495 and a Kendall's tau correlation of 0.16. Similarly, Crude oil–Soybeans (C,S) offers the Gaussian copula as the best fit with a copula parameter of 0.3545 and a Kendall's tau correlation of 0.23. For the last pair-copula, the conditional pair-copula, Palm oil–Soybeans given Crude oil (P,S|C), the Gaussian copula is chosen to explain its dependence structure with a copula parameter of 0.2303 and a Kendall's tau correlation of 0.15. Furthermore, considering to a bivariate pair-copula, Palm oil–Soybeans (P,S), we found that there exists a weak positive dependence and that the Gaussian copula is the best fit with a copula parameter 0.2970 and Kendall's tau correlation of 0.19. This indicates that the price of one commodity is slightly correlated with the prices of the other commodities.

Our results show that the dependence between the crude oil prices and the soybeans prices is stronger than the dependence between the crude oil prices and the palm oil prices due to the increase in biofuel demand and soybeans consumption [28]. Moreover, palm oil is produced on a large scale within the intra-ASEAN region, and the ASEAN nations do not have to rely on imports from the outside region. So the price of palm oil is slightly related to the change in crude oil prices. Thus, to reduce the price transmission and volatility spillover between crude oil prices and food prices, and to increase food security, the ASEAN members should get together and cooperate to incorporate innovative and effective plans to increase the capacity and performance in food production in order to reduce the reliance on food imports from outside the region, especially in the case of soybeans.

Acknowledgements. This work was granted support by the Energy Conservation Promotion Fund, the Energy Policy and Planning Office, the Ministry of Energy of Thailand. The first author is grateful for being granted a PhD scholarship to do his studies.

References

1. ASEAN Secretariat, ASEAN Economic Community Factbook. The ASEAN Secretariat (2011a), http://www.thaifta.com/ThaiFTA/Portals/0/ASEAN_AECFactBook.pdf (accessed May 20, 2013)
2. ASEAN Secretariat, ASEAN Community in Figures, ACIF, The ASEAN Secretariat (2012), http://www.asean.org/resources/publications/asean-publications/item/asean-community-in-figures-acif-2011-3 (accessed May 20, 2013)
3. Wikipedia, BRIC (2013), http://en.wikipedia.org/wiki/BRIC (accessed May 20, 2013)

4. ASEAN Secretariat, Regional and Country Reports of the ASEAN Assessment on the Social Impact of the Global Financial Crisis. The ASEAN Secretariat (2010), http://www.asean.org/archive/publications/ARCR/ASEANRegional&CountryReport.pdf (accessed May 20, 2013)
5. Asian Development Bank, Global food price inflation and developing Asia. Asian Development Bank (2011), http://www.adb.org/publications/global-food-price-inflation-and-developing-asia (accessed May 23, 2013)
6. Len, C.: Energy Security Cooperation in Asia: An ASEAN-SCO Energy Partnership (2007), http://www.silkroadstudies.org/new/docs/publications/2007/ENERGY1_015.pdf (accessed May 20, 2013)
7. ASEAN Secretariat, ASEAN Community in a Global Community of Nations.Co-Chairs' statement of the 4th ASEAN-UN summit Bali, Indonesia (November 19, 2011b), http://www.mofa.go.jp/region/asia-paci/eas/pdfs/declaration_1111_2.pdf (accessed May 20, 2013)
8. USDA, Table 11: Palm Oil: World Supply and Distribution. United States Department of Agriculture (2013a), http://www.fas.usda.gov/oilseeds/Current/ (accessed May 27, 2013)
9. USDA, Table 07: Soybeans: World Supply and Distribution. United States Department of Agriculture (2013b), http://www.fas.usda.gov/oilseeds/Current/ (accessed May 27, 2013)
10. ASEAN Secretariat, The ASEAN Economic Community (AEC) Chartbook, The ASEAN Secretariat (2013), http://www.asean.org/images/2013/resources/publication/2013%20-%20AEC%20Chartbook%202012.pdf (accessed May 25, 2013)
11. Speed, P.A.: ASEAN. The 45 Year Evolution of a Regional Institution. POLINARES working paper no. 61, University of Westminster (2012), http://www.polinares.eu/docs/d4-1/polinares_wp4_chapter11.pdf (accessed May 27, 2012)
12. Reboredo, J.C.: How do crude oil prices co-move? A copula approach. Energy Economics 33, 948–955 (2011)
13. Serra, T., Zilberman, D.: Biofuel-related price transmission literature: A review. Energy Economics 37, 141–151 (2013)
14. FAO, Declaration of the world summit on food security. World Summit on Food Security, Rome (November 16-18, 2009), ftp://ftp.fao.org/docrep/fao/Meeting/018/k6050e.pdf (accessed May 20, 2013)
15. United Nations, World Energy Assessment: Overview 2004 Update. United Nations Development Programme (2004), http://www.undp.org/content/dam/aplaws/publication/en/publications/environment-energy/www-ee-library/sustainable-energy/world-energy-assessment-overview-2004-update/World%20Energy%20Assessment%20Overview-2004%20Update.pdf (accessed May 23, 2013)
16. Baffes, J.: Oil spills on other commodities. Resources Policy 32, 126–134 (2007)
17. Sriboonchitta, S., et al.: Modeling volatility and dependency of agricultural price and production indices of Thailand: Static versus time-varying copulas. International Journal of Approximate Reasoning 54(6), 793–808 (2013)
18. Bollerslev, T.: Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics 31, 307–327 (1986)

19. Wuertz, D., Chalabi, Y.: Rmetrics-Autoregressive Conditional Heteroskedastic Modelling (2013), http://cran.r-project.org/web/packages/fGarch/index.html (accessed May 10, 2013)
20. Brechmann, E.C., Schepsmeier, U.: Modeling Dependence with C- and D-Vine Copulas: The R Package CDVine. Journal of Statistical Software 52(3), 1–27 (2013), http://www.jstatsoft.org/v52/i03/ (accessed February 20, 2013)
21. Aas, K., et al.: Pair-copula constructions of multiple dependence. Insurance: Mathematics and Economics 44, 182–198 (2009)
22. Koyama, K.: A Though on Crude Oil Pricing in Asia. The institute of energy economics, Japan (2011), https://eneken.ieej.or.jp/data/3711.pdf (accessed May 27, 2013)
23. USDA, Thailand Oilseeds and Products Annual. United States Department of Agriculture (2013c), http://gain.fas.usda.gov/Recent%20GAIN%20Publications/Oilseeds%20and%20Products%20Annual_Bangkok_Thailand_4-1-2013.pdf (accessed May 26, 2013)
24. USDA, Vietnam Oilseeds and Products Annual 2013. United States Department of Agriculture (2013d), http://gain.fas.usda.gov/Recent%20GAIN%20Publications/Oilseeds%20and%20Products%20Annual_Hanoi_Vietnam_4-5-2013.pdf (accessed May 26, 2013)
25. USDA, Indonesia Oilseeds and Products Update 2013. United States Department of Agriculture (2013e), http://usdaindonesia.org/wp-content/uploads/2013/02/Oilseeds-and-Products-Update_Jakarta_Indonesia_2-5-2013.pdf (accessed May 26, 2013)
26. Patton, A.J.: Modelling Asymmetric Exchange Rate Dependence. International Economic Review 47(2), 527–556 (2006)
27. Manthos, V.: Dynamic Copula Toolbox 3.0 (2010), http://www.mathworks.com/matlabcentral/fileexchange/29303-dynamic-copula-toolbox-3-0 (accessed December 15, 2012)
28. Abbott, P.C., et al.: Whats Driving Food Prices in 2011? Farm Foundation, NFP (2011), http://www.farmfoundation.org/news/articlefiles/1742-FoodPrices_web.pdf (accessed June 14, 2013)

Copula Based GARCH Dependence Model of Chinese and Korean Tourist Arrivals to Thailand: Implications for Risk Management

Ornanong Puarattanaarunkorn and Songsak Sriboonchitta

Abstract. China and Korea are two of the important tourist markets for Thailand. The growth rates of tourist arrivals from these two countries have volatility and also seem to have co-movement. Understanding the dependence between these tourists markets has importance for strategic planning and processes for decision-making. The purpose of this study is to find out the dependence between the growth rates of tourist arrivals from China and Korea to Thailand by using the copula based GARCH model. Copula provides a potential and flexible method to model the dependence between random variables. It is preferable to the conventional approach because the copula can cross over the restriction of normal distribution and linear assumption, according to the Pearson correlation. The results of the analysis can contribute to appropriate policy implications. The results show that there exists a weak positive dependence and that the rotated Joe 180° copula is the best fit, which provides an evidence of lower tail dependence. The growth rates of tourist arrivals from China and Korea have co-movement that is both upward and downward, but with a weak dependence. The rise or loss of tourism demand from China (Korea) is slightly correlated by the rise or loss of tourism demand from Korea (China). The time-varying rotated Joe 180° copula is the best fit and the most significant, which implies that the dependence parameter has varied over time. The policy implications for the risk management of the tourism demand should provide enough motivation for the marketing and promotion of the tourism demand by considering the time-varying dependency of China and Korea. Moreover, they should consider alternative target markets as substitutes when theres a loss of arrivals from these two markets in order to diversify the risk of tourism demand.

Ornanong Puarattanaarunkorn · Songsak Sriboonchitta
Faculty of Economics, Chiang Mai University, Chiang Mai 50200 Thailand
e-mail: pornan@kku.ac.th, songsakecon@gmail.com

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_22,

1 Introduction

The travel and tourism sector plays an important role in stimulating the economic growth of a country. In 2011, the total global contribution of travel and tourism sector to the GDP was 6,346 billion US dollars (9.1% of total GDP) [1]. Global tourism demand, which is measured by the number of international tourist arrivals, rose to 1,035 million and increased by 4% in 2012. Asia and Pacific, particularly, registered the highest growth among all the regions, with a 7% growth within that period [2]. Thailand is a major tourist destination in the Asia and Pacific region. In 2011, Thailand ranked fourth in the number of arrivals, with 19 million, and third in international receipts, with 26 million US dollars [3]. International tourist arrivals to Thailand have been growing over the past decades: 10 million in 2001 to 20 million in Nov 2012, which is a 6% yearly growth rate. However, the tourism sector continues to face challenges due to the prevailing social and economic uncertainties such as global economic recession, climate changes that influence travelers' behavior [4], terrorism, and natural disasters. These negative shocks and seasonal occurrences [5] have an effect on the volatility in tourism demand. Although the shock effects are not permanent, Lean and Smyth [6] found that the negative shocks made the growth of tourist arrivals slow down. Consequently, these can have some adverse effects on the businesses, employment, and economic growth. As shown in Table 1, the tourism demand for Thailand during the period 2002–2012 fluctuated greatly because of the many shock events. For example, in 2003, 2005, and 2009, the growth rate of international tourist arrivals to Thailand sharply declined because of the outbreak of SARS, tsunami, economic recession, and political disturbance.

Table 1 presents the top 10 ranking arrivals and the market share of the international tourist arrivals to Thailand. There is diversification in the countries of the various regions. This paper focuses on tourist arrivals to Thailand from the Asia region, particularly in the extra-ASEAN countries such as China and Korea, for the following reasons. First, the long distances of travel have increased the cost of travel and tourism. Second, the European economy has slowed down, thus causing an effect on the demand for long-haul travel. Therefore, we should lay the emphasis more on the intra-regional tourism demand. Third, when we consider the recent growth rate of arrivals during 2010–2012, China and Korea showed a rapid growth in the rate of arrivals. They ranked at the third and the fourth place of the market share, respectively. Fourth, China is a part of the BRIC countries; it is a newly industrialized country and a fast growing economy. It is the worlds largest exporter and manufacturer, and is the second largest economy [7] in the world. Korea is part of the OECD members, is a developed country, and has a high income level. These two countries have tourism potential and make for interesting studies regarding their interdependence as far as tourism demand to Thailand is concerned. In addition, when we look at the growth rates of the tourist arrivals from these two countries, they have volatility and seem to have co-movement, as is illustrated in Figure 1, Part 4. Thus, these interesting facts lead to our three research questions: (1) Is there dependence between the growth rates of tourist arrivals from China and Korea to Thailand? (2) If there is, then what is the magnitude of dependence? (3) What is the nature of

Table 1 International Tourist Arrivals to Thailand 2002-2012

	2002	2003	2004	2005	2006	2007	2008	2009	2012	2011	2012*
1. Malaysia	1.33	1.35	1.40	1.37	1.59	1.54	1.81	1.76	2.06	2.50	2.23
	(12%)	(2%)	(4%)	(-2%)	(15%)	(-3%)	(16%)	(-3%)	(16%)	(19%)	(-11%)
2. Japan	1.24	1.04	1.21	1.20	1.31	1.28	1.15	1.00	0.99	1.13	1.24
	(5%)	(-17%)	(15%)	(-1%)	(9%)	(-3%)	(-10%)	(-14%)	(-1%)	(13%)	(9%)
3. China	0.80	0.61	0.73	0.78	0.95	0.91	0.83	0.78	1.12	1.72	2.53
	(-0.4%)	(-27%)	(18%)	(6%)	(20%)	(-5%)	(-9%)	(-6%)	(37%)	(43%)	(38%)
4. Korea	0.70	0.70	0.90	0.82	1.09	1.08	0.89	0.62	0.81	1.01	1.05
	(25%)	(-1%)	(26%)	(-10%)	(29%)	(-1%)	(-20%)	(-36%)	(26%)	(22%)	(4%)
5. UK	0.70	0.74	0.76	0.77	0.85	0.86	0.83	0.84	0.81	0.84	0.77
	(6%)	(4%)	(3%)	(2%)	(9%)	(1%)	(-4%)	(2%)	(-4%)	(4%)	(-9%)
6. USA	0.56	0.51	0.63	0.64	0.69	0.68	0.67	0.63	0.61	0.68	0.68
	(5%)	(-8%)	(20%)	(2%)	(8%)	(-2%)	(-2%)	(-6%)	(-2%)	(11%)	(-1%)
7. Singapore	0.55	0.52	0.58	0.65	0.69	0.60	0.57	0.56	0.60	0.68	0.71
	(3%)	(-6%)	(11%)	(12%)	(5%)	(-13%)	(-6%)	(-1%)	(7%)	(12%)	(3%)
8. Australia	0.35	0.29	0.40	0.43	0.55	0.66	0.69	0.65	0.70	0.83	0.85
	(0.3%)	(-19%)	(31%)	(7%)	(25%)	(18%)	(5%)	(-7%)	(8%)	(17%)	(2%)
9. India	0.28	0.25	0.33	0.38	0.46	0.54	0.54	0.61	0.76	0.91	0.92
	(20%)	(-10%)	(27%)	(14%)	(19%)	(15%)	(0.1%)	(13%)	(21%)	(19%)	(1%)
10. Germany	0.41	0.39	0.46	0.44	0.52	0.54	0.54	0.57	0.61	0.62	0.59
	(2%)	(-6%)	(16%)	(-3%)	(16%)	(5%)	(-0.3%)	(6%)	(6%)	(2%)	(-4%)
11. Others	3.9	3.7	4.3	4.1	5.1	5.8	6.1	6.1	6.9	8.3	8.2
	(6%)	(-7%)	(16%)	(-6%)	(23%)	(12%)	(5%)	(1%)	(11%)	(19%)	(-1%)
Grand Total	10.9	10.1	11.7	11.6	13.8	14.5	14.6	14.1	15.9	19.2	19.8
	(7%)	(-8%)	(15%)	(-1%)	(18%)	(4%)	(1%)	(-3%)	(12%)	(19%)	(3%)

Note: *The total number of tourist arrivals from January to November.
The numbers in parenthesis are the growth rate of arrivals.
Source: Ecowin Database.

the dependence structure? If we know whether the growth rates of these two countries have dependence or not, then it becomes useful for policy makers and tourism businesses to plan for risk management of tourism demand and tourism supply. For example, if there is high positive dependence, then the shocks can have an effect that is either decreasing or increasing simultaneously on tourist arrivals from both China and Korea to Thailand. Conversely, if there is independence or low dependence, it would be beneficial in terms of risk diversification because the quantities of loss of arrivals from these two countries are not related.

In order to answer these research questions, it is the purpose of this study to find out the dependence between the growth rates of tourist arrivals from China and Korea to Thailand by using the copula based GARCH model. This model was chosen because GARCH can examine the volatility of the tourist arrivals and copula can model the dependence structure between the two marginal distributions that obtain from GARCH model. Copula can measure the dependence without making an assumption of normal distribution and linear relation as the Pearson correlation does. Another advantage of copula is that we can find out the dependence without actually knowing the real marginal distributions of the variables. The contributions

of this study are toward the policy implications in terms of risk management of the tourism demand for Thailand that are obtained from the findings in the analysis.

This paper is divided into seven parts. The next part is the literature review. The third part presents the methodology that describes the GARCH model and the copula model. The fourth part presents the data used. The fifth part shows the results of this study. The sixth part presents the policy implication that discusses in detail the risk management of the tourism demand for Thailand. The last part gives the conclusion and information on future research.

2 Literature Review

International tourism demand, which is measured as tourist arrivals, plays an important role in the economy of many countries. Therefore, analyses on the modeling of international tourism demands are vast. There is a vast array of literature that contains both the studies on the effects of the various determinants and the forecasting of future international tourism demands. In this paper, we review particularly the studies on international tourism demand forecasting by using various time series models. For example, Goh and Law [8] used the Box-Jenkin forecasting model along with a stochastic nonstationary seasonality (SARIMA) model and an intervention component (MARIMA) model for predicting tourist arrivals to Hong Kong. Similarly, Chang et al. [9] used the autoregressive integrated moving average (ARIMA) model and the seasonal ARIMA (SARIMA) model for forecasting tourist arrivals from East Asia to Thailand. Chu [10] used three autoregressive moving average (ARMA) based models to forecast tourist arrivals in nine Asia Pacific destinations. It's been known that international tourist arrivals undergo fluctuations due to many reasons, such as seasonality, economic changes, financial crisis, political instability, terrorism, diseases, and natural disasters. That is why a modeling of the international tourism demand was presented in the case of our study. The generalized autoregressive conditional heteroskedastic (GARCH) model has been widely used to investigate the volatility of tourism demand. For example, Chan et al. [11] used three multivariate GARCH models the constant condition correlation volatility model or the symmetric CCC-MGARCH, the symmetric vector ARMA-GARCH, and the asymmetric vector ARMA-AGARCH to investigate the volatility of international tourism demand to Australia. Shareef and McAleer [12] used ARMA-GARCH(1,1) and ARMA-GJR(1,1) to examine the international tourist arrivals to the Maldives. Coshall [13] used ARIMA for conditional mean, and GARCH and EGARCH for conditional variance to model the outbound UK tourism demand to international destinations, as well as to test the forecasting ability of these models.

There is no dearth of literature on correlation analysis across international tourism markets and tourism destinations. For example, Chan et al. [11] and Alvarez et al. [14] analyzed the conditional correlation-based GARCH model for monthly international tourist arrivals shocks. Hoti et al. [15] analyzed the conditional correlation-based GARCH model across two tourism destinations. Jang and Chen [16] and Chen et al. [17] analyzed the correlation across international tourist arrivals

for finding the optimal tourist market mixes by using a portfolio approach. All of the researches above measured the interdependence, or the correlation, using the conventional approach, namely, the Pearson correlation coefficient. But the drawback of the Pearson correlation is that it is restricted by the assumption based on normal distribution and linear relationship of the data series. However, many data series are not of normal distribution and have non-linear relationships. Therefore, to overcome that restriction, many studies used copulas to measure the dependency between the variables, especially in the financial field. Copulas can model the dependence between random variables without identifying the distribution of the individual variables [18]. Many studies used those copulas that have cooperated with the GARCH model, the copula based GARCH, to find the dependence structure of the marginal distribution of the conditional variance. The copula based GARCH model provides more flexibility for finding out the joint distributions and the transformation invariant correlation, without the assumption of linear correlation [19]. For example, Patton [20, 21] used the ARMA(p,q)-GARCH(1,1) model to estimate the marginal distributions of the Deutsche mark-US dollar and Japanese yen-US dollar exchange rates because the exchange rates had time variation in both the conditional mean and the conditional variance, and then used the copula to model the dependence structure of their marginals. Similarly, Goorbergh [22], Jondeau and Rockinger [23], and Wang and Cai [24] also used the copula based GARCH to model the dependence structure between stock markets. Reboredo [25] extended the copula based ARMA(p,q)-TGARCH(1,1) to find the co-movements between the world oil prices and the global prices for corn, soybean, and wheat. In the tourism field, a few pieces of literature are available on the application of the copula to model the dependence structure between variables. Zhang et al. [26] used a fully nested Archimedean copula function to find the dependence between three dependent variables: destination visits behavior, time use behavior, and expenditure behavior. Liu and Sriboonchitta [27] used a copula based GARCH to model the volatility and the dependence structure between tourist arrivals from China to two destination markets, Thailand and Singapore. The results of their study showed a strong dependence between two the destination markets, which led to a policy recommendation about cooperation in the policy planning of the two destination markets. Therefore, in this paper, we used the copula based GARCH to estimate the marginal distributions of the conditional variances of tourist arrivals from China and Korea to Thailand, and examined the dependence structure between these two marginal distributions. The contribution of this paper is toward the policy implication that points to the risk management of the variations in the tourism demand or tourist arrivals from these two countries to Thailand. This study is different from a previous study that focused on the cooperation between the two destination markets or the concept of multi-destinations.

3 Methodology

The copula based ARMA-GARCH model was used to answer the three research questions dealt with in this paper. The GARCH model [28] has been widely used

for modeling the volatility of the asset returns in the financial field. This is important as volatility is considered a measure of risk. Therefore, we applied the ARMA-GARCH model to estimate the marginal distributions since this model can capture the volatility of international tourism demand, as measured by the number of international tourist arrivals. The standardized residuals from ARMA-GARCH model were transformed to copula data $(F_1(x_1), F_2(x_2))$. After that, the copula approach was used to measure the dependence between the two marginal distributions of the growth rates of international tourist arrivals. This study used the R-package CDVine by Brechmann and Schepsmeier [29] to analyze the constant copula since it provides a range of tools for bivariate data analysis. And for the time-varying copula, we followed method endorsed by Patton [21].

We adopt the ARMA(1,0)-GARCH(1,1) with the skewed student T (SkT) distribution residual for the marginal distribution of the logarithm of the monthly growth rate of tourist arrivals to Thailand from China and Korea (y_t):

$$y_t = a_0 + a_1 y_{t-1} + \varepsilon_t \tag{1}$$

$$\varepsilon_t = z_t \sqrt{h_t}, z_t \sim SkT(\nu, \gamma) \tag{2}$$

$$h_t = \omega_t + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1} \tag{3}$$

In equation (1) presents ARMA(p,0) process where y_{t-1} is an autoregressive term of y_t and ε_t is an error term. Equation (2) then define this error term as the product between conditional variance h_t and a residual z_t. A residual z_t is assumed to follow the skewed student T (SkT) distribution with the degree of freedom parameter ν and the skewness parameter γ. Equation (3) presents GARCH(1,1) process where $\omega_t, \alpha \geq 0, \beta \geq 0$ are sufficient to ensure that the conditional variance $h_t > 0$. The ε_{t-1}^2 represent the ARCH term and α refers to the short run persistence of shocks, while βh_{t-1} represent the GARCH term and β refers to the contribution of shocks to long run persistence $(\alpha + \beta)$.

The property of the GARCH(1,1) model is that it requires the conditional variance, h_t, of the error term, ε_t, to be stationary and persistent. This paper used the second moment condition that was presented in the Bollerslev study [28] and the log moment condition that was presented by Nelson [30] and Lee and Hansen [31] to check these properties. The second moment condition: $\alpha + \beta < 1$ and the log moment condition: $\mathrm{E}[\ln(\alpha z_t^2 + \beta)] < 0$.

3.1 Copulas

One approach of modeling the multivariate dependence is the copula. The copula functions can offer us the flexibility of merging a univariate distribution to get a joint distribution with an appropriate dependence structure. The fundamental theorem of copula is Sklar's theorem by Sklar [32]. Nelson [33] has made a description of the copula theory, as follows:

Let H be a joint distribution function with marginal distributions F and G. Then there exists a copula C for all x, y in real line, with the following property:

$$H(x,y) = C(F(x), G(y)) \tag{4}$$

If F and G are continuous, C is unique. Conversely, if C is a copula and F and G are univariate distribution functions, then the above function H in (4) is a joint distribution function with marginal distributions F and G. If H is known, the copula is an equation (4) that one can get from the form,

$$C(u,v) = H(F^{-1}(u), G^{-1}(v)), \tag{5}$$

where F^{-1} and G^{-1} are the quantile functions of the marginal distributions.

Yan [34] provides a better understanding of the copula. If we have two or more continuous random variables that can be transformed to uniform [0,1] by probability integral transformation, then we can find the multivariate dependence structure by using the copula.

3.2 Characteristics of Copula Families

This paper uses constant copulas as well as time-varying copulas to describe the dependence of two marginal distributions. The copula types that were used to capture the dependence include the Gaussian copula, Student's T copula, Clayton copula, Gumbel copula, Frank copula, Joe copula, rotated Clayton copula, rotated Gumbel copula, and rotated Joe copula. Each copula type has its own distinct characteristics.

(*a*) Constant Copulas

Gaussian (Normal) Copula. Trivedi and Zimmer [35] mention in their work that the Gaussian copula allows for equal degrees of positive and negative dependence. The following copula function is offered by Lee [36].

$$C(u_1, u_2; \rho) = \Phi_G(\Phi^{-1}(u_1), \Phi^{-1}(u_2); \rho) \tag{6}$$

or

$$C(u_1, u_2; \rho) = \int_{-\infty}^{\phi^{-1}(u_1)} \int_{-\infty}^{\phi^{-1}(u_2)} \frac{1}{2\Pi\sqrt{(1-\rho^2)}} \times [\frac{-(s^2 - 2\rho st + t^2)}{2(1-\rho^2)}] dsdt \tag{7}$$

where Φ^{-1} is the inverse of the standard normal c.d.f. and $\Phi_G(u_1, u_2)$ is the standard bivariate normal distribution with the Pearson correlation parameter ρ restricted to the interval $(-1, 1)$. The Gaussian copula is the tail independence.

Student's T Copula. Trivedi and Zimmer [35] mention that the Student's T copula has two dependence parameters, ν degrees of freedom, and correlation ρ. Unlike

the Gaussian copulas, copulas extracted from T-distributions exhibit tail (upper and lower) dependence.

$$C^T(u_1,u_2;\rho,\nu)=\int_{-\infty}^{T_\nu^{-1}(u_1)}\int_{-\infty}^{T_\nu^{-1}(u_2)}\frac{1}{2\Pi\sqrt{(1-\rho^2)}}\times[1+\frac{(s^2-2\rho sT+T^2)}{\nu(1-\rho^2)}]^{-(\frac{\nu+2}{2})}dsdT \tag{8}$$

where $T_\nu^{-1}(u_1)$ is the inverse of the c.d.f. of the standard univariate T-distribution with ν degrees of freedom which is controlling the heaviness of the tails.

Clayton Copula. This family of copulas was discussed by Clayton [37].

$$C(u_1,u_2;\theta)=(u_1^{-\theta}+u_2^{-\theta}-1)^{-1/\theta} \tag{9}$$

for the dependence parameter θ is limited on the range (0,∞). Trivedi and Zimmer [35] state that the Clayton copula does not allow for negative dependence. Since the Clayton copula shows strong lower tail dependence and relatively weak upper tail dependence, it can be used to study the involved risks.

Gumbel Copula. Nelson [33] points out that the Gumbel copula was first discussed by Gumbel [38], and thus it was referred to as the Gumbel family.

$$C(u_1,u_2;\theta)=exp(-[(-\ln(u_1))^\theta+(-\ln(u_2))^\theta]^{\frac{1}{\theta}}) \tag{10}$$

for the dependence parameter θ is limited on the range [1,∞). The Gumbel copula cannot account for negative dependence. As the Gumbel copula shows strong upper tail dependence, we can say that it is in contrast to the Clayton copula [35]. Joe [39] is of the view that the Gumbel copula is an extreme value copula.

Frank Copula. Acorrding to Nelson [33], the Frank family was first presented in Frank [40].

$$C(u_1,u_2;\theta)=\frac{-1}{\theta}\ln(1+((e^{-\theta u_1}-1)(e^{-\theta u_2}-1))/(e^{-\theta}-1)) \tag{11}$$

for the dependence parameter $\theta\in(-\infty,\infty)/\{0\}$. Trivedi and Zimmer [35] speak of the Frank copula as their favorite because of many reasons. For example, the Frank copula allows for negative dependence. The dependence is tail symmetry, akin to the Gaussian and Student-t copulas. The Frank copula accounts for strong positive or negative dependence. Since the Frank copula can capture weak dependence in the tails better than the Gaussian, it is most appropriate for data that show weak tail dependence.

Joe Copula. Nelson [33] points out that this family was proposed in Joe's work [39].

$$C(u_1,u_2;\theta)=1-[(1-u_1)^\theta+(1-u_2)^\theta-(1-u_1)^\theta(1-u_2)^\theta]^{\frac{1}{\theta}} \tag{12}$$

for the dependence parameter θ restricted to [1,∞).The Joe copula also has upper tail dependence with $2-2^{1/\theta}$ as the limit and Family Gumbel is the extreme value limit of Family Joe [39].

Rotating the copula was made for the asymmetric dependence structures such as those of the Clayton, Gumbel, and Joe. Nelson [33] defined the rotation of the copulas by means of $C_R(u_1,u_2)=u_1+u_2-1+C(1-u_1,1-u_2)$. In practice, for rotated bivariate copulas, we can transform the input arguments u_1 and u_2 to $1-u_1$ and $1-u_2$ for 180 degrees. When rotating copulas by 180 degrees, we can also call the survival copulas of the corresponding family, for example, survival Clayton [29]. After we rotate copulas, such as with the rotated Gumbel, the copulas will show stronger dependency in the lower tail instead of the upper tail.

Rotated Clayton Copula. The rotated Clayton copula can capture the upper tail dependence, converse to the Clayton copula. Fisher [41] shows the functional form as

$$C(u_1,u_2;\theta)=u_1+u_2-1+[(1-u_1)^{-\theta}+(1-u_2)^{-\theta}-1]^{-\frac{1}{\theta}} \tag{13}$$

Rotated Gumbel Copula. The rotated Gumbel copula can capture the lower tail dependence, converse to the Gumbel copula. Fisher [41] shows the functional form as

$$C(u_1,u_2;\theta)=u_1+u_2-1+exp(-[(-\ln(1-u_1))^{\theta}+(-\ln(1-u_2))^{\theta}]^{\frac{1}{\theta}}) \tag{14}$$

Rotated Joe Copula. The rotated Joe copula can capture the lower tail dependence, converse to the Joe copula. Fisher [41] shows the functional form as

$$C(u_1,u_2;\theta)=u_1+u_2-(u_1^{\theta}+u_2^{\theta}-u_1^{\theta}u_2^{\theta})^{\frac{1}{\theta}} \tag{15}$$

Time-Varying Copulas. Since it is a fact that dependence between the marginal distributions of the time series variables are not constant through time, they should be considered as time-varying copulas. We used the functional form of the time-varying copulas by following an ARMA(1,10) process which was presented by Patton [21].

Time-Varying Gaussian Copula

$$\rho_t=\Lambda(\omega_\rho+\beta_\rho\rho_{t-1}+\alpha\frac{1}{10}\Sigma_{j=1}^{10}\Phi^{-1}(u_{1,t-j})\Phi^{-1}(u_{2,t-j})) \tag{16}$$

where Φ^{-1} is the inverse of the standard normal c.d.f., $\Lambda(x)\equiv(1-e^{-x})(1+e^{-x})^{-1}$ is the modified logistic transformation, used to hold ρ_t in the range $(-1,1)$ at all times, ρ_{t-1} is a regressor to measure any persistence in the dependence parameter, and 10 is used to average the transformed variables $\Phi^{-1}(u_{1,t-j})$ and $\Phi^{-1}(u_{2,t-j})$ over the previous 10 lags (ARMA(1,10) process), to capture the variation in the dependence [21].

Time-Varying for Non-Gaussian Copula. Patton [21] presented the modeling of the tail dependence parameters using the symmetrized Joe-Clayton (SJC) copula, where the upper tail T^U and the lower tail T^L were related to the parameters of the copulas. Thus, in addition to specifying the tail dependence parameters over the sample, the equation that follows also specifies the parameters of the copula. Manner and Reznikova [42] presented an equation which is based on Patton [21] as

$$\theta_t = \Lambda(\omega + \beta\Lambda^{-1}\theta_{t-1} + \alpha\frac{1}{10}\Sigma_{j=1}^{10}|u_{1,t-j} - u_{2,t-j}|) \tag{17}$$

where $\Lambda(x)$ is a transformation function to always keep the parameters in their intervals. $\Lambda(x) \equiv (1+e^{-x})^{-1}$ is the logistic transformation, used to hold the tail dependence in the range (0,1), $\Lambda(x) \equiv e^x$ for the Clayton copula, and $\Lambda(x) \equiv e^x + 1$ for the Gumbel copula. $\beta\Lambda^{-1}\theta_{t-1}$ is an autoregressive term, and the last term on the right-hand side of the equation is the mean absolute difference between u_1 and u_2 over the previous 10 observations. This is a forcing variable, and under perfect positive dependence it will be close to zero, in case of perfect negative dependence it will equal to 0.5, and in case of independence it will equal to 0.33 [21]. In addition, we used $\Lambda(x) \equiv e^x + 1$ for the Joe copula, same as that for the Gumbel copula.

Time-Varying for Rotated Non-Gaussian Copula. Patton [43] suggested that the rotated copulas can be formed thus: If (U_1, U_2) are distributed as the copula C, then $(1-U_1, 1-U_2)$ will be distributed as the rotated C copula. Thus, with regard to estimating time-varying for the rotated non-Gaussian copula, we will transform the input arguments and use the same function as in equation (17).

3.3 Tail Dependence

Tail dependence explains the degree of dependence in the upper and lower tails of a bivariate distribution. The distributions of the tail dependences in the case of financial risk are interesting because tail dependences can model the dependence of loss events across portfolio assets. Joe [44] explained the dependence of the tails of the bivariate copula.

Let X and Y be the random variables with marginal distribution functions F and G. The tail dependence of X and Y can be given as

$$T^U = \lim_{u\to 1}(P(X > F^{-1}(u)|Y > G^{-1}(u)) = \lim_{u\to 1}\frac{1-2u+C(u,u)}{1-u} \tag{18}$$

If $T^U \in (0,1]$ the joint distribution of X and Y shows upper tail, indicating that the probability of the joint occurrence of extreme values is positive; if $T^U = 0$, then there is no upper tail dependence. Similarly, in

$$T^L = \lim_{u\to 0}\frac{C(u,u)}{u} \tag{19}$$

Table 2 Function of Kendall's tau and Tail Dependence for Bivariate Copula

Copula family	Kendall's tau	Tail dependence (lower, upper)
Gaussian	$\frac{2}{\Pi}\arcsin\rho$	0
Student's T	$\frac{2}{\Pi}\arcsin\rho$	$T^L = T^U = 2T_{v+1}(-\sqrt{v+1}\sqrt{\frac{1-\rho}{1+\rho}})$
Clayton	$\frac{\theta}{\theta+2}$	$(2^{-1/\theta},0)$
Gumbel	$1-\frac{1}{\theta}$	$(0,2-2^{1/\theta})$
Frank	$1-\frac{4}{\theta}+4\frac{D_1(\theta)}{\theta}$	(0,0)
Joe	$1+\frac{4}{\theta^2}\int_0^1 t\log(t)(1-t)^{2(1-\theta)/\theta}dt$	$(0,2-2^{1/\theta})$
Rotated Clayton 180°	$\frac{\theta}{\theta+2}$	$(0,2^{-1/\theta})$
Rotated Gumbel 180°	$1-\frac{1}{\theta}$	$(2-2^{1/\theta},0)$
Rotate Joe 180°	$1+\frac{4}{\theta^2}\int_0^1 t\log(t)(1-t)^{2(1-\theta)/\theta}dt$	$(2-2^{1/\theta},0)$

Note: $D_1(\theta)=\int_0^\theta \frac{c/\theta}{exp(x)-1}$ is the Debye function. For the first six copula families, please refer to Brechmann and Schepsmeier [29]. For the next three rotated copulas, the functional forms of Kendall's tau are the same as the non-rotated copulas.

if $T^L \in (0,1]$, the joint distribution of X and Y shows lower tail dependence, indicating that the probability of the joint occurrence of extreme values is negative; if $T^L = 0$, then there is no lower tail dependence.

3.4 Maximum Likelihood Estimation

The method of maximum pseudo-log likelihood, studied by Genest et al. [45], was used for estimation since the marginal distribution functions F and G of the random vectors are unknown. Thus, we can construct the pseudo copula observations by using the empirical distribution functions to transform the standardized residual series into uniform [0, 1] as rank based.

Under the assumption that the marginal distributions F and G are continuous, the copula C_θ is a bivariate distribution with density c_θ and pseudo-observations $F_n(X_i)$ and $G_n(Y_i)$, $i = 1,2,...,n$. The pseudo-log likelihood function of θ can be given as

$$L(\theta) = \Sigma_{i=1}^n \log[c_\theta(F_n(X_i), G_n(Y_i))]. \tag{20}$$

Then, maximizing the pseudo-log likelihood yield as an estimator of θ,

$$\theta = \Sigma_{i=1}^n \frac{\partial}{\partial\theta}\log[c_\theta(F_n(X_i), G_n(Y_i))] = 0 \tag{21}$$

where $c_\theta = \frac{\partial^2 C_\theta(F_n(x),G_n(y))}{\partial x \partial y}$, $F_n(x) = \frac{1}{n+1}\Sigma_{i=1}^n 1(X_i \le x)$ and $G_n(x) = \frac{1}{n+1}\Sigma_{i=1}^n 1(Y_i \le y)$ are the empirical distributions.

3.5 Selection of Copulas

Selecting a family of copulas is based upon information criteria such as Akaike Information Criterion (AIC) by Akaike [46] and Bayesian Information Criterion (BIC) by Schwarz [47]. For examining whether the dependence structure of the data series is appropriate for a chosen family of copulas, we used a goodness-of-fit test based on a scoring approach by Vuong [48] and Clarke [49], and a second goodness-of-fit test based on Kendall's tau by Genest and Rivest [50], and Wang and Wells [51].

4 Data

The seasonal adjusted[1] data of tourist arrivals to Thailand from China and Korea that measure the tourism demands for these two countries were used. The monthly data of the two countries, during the period from January 1997 to November 2012, were taken into the logarithm of the monthly arrival rate (the growth rate of tourist arrivals).

$$\text{The growth rate of tourist arrivals} = \ln \frac{\text{Number arrivals}_t}{\text{Number arrivals}_{t-1}} \tag{22}$$

Table 3 Descriptive Statistics for Growth Rate of Tourist Arrivals

	China	Korea
Mean	0.011	0.005
Median	0.014	0.004
Maximum	0.739	0.826
Minimum	−1.099	−1.139
Std. Dev.	0.281	0.196
Skewness	−0.903	−1.164
Kurtosis	6.092	11.529
Jarque-Bera	101.538	618.798
(p-value)	(0.000)	(0.000)
Observations	190	190
Pearson's correlation	0.33	

Note: The null hypothesis of Jarque-Bera = data is taken as the normal distribution.

Table 3 presents a descriptive statistics of the growth rate of the tourist arrivals to Thailand from China and Korea. Both the countries have positive average growth rates during the period from January 1997 to November 2012. The negative skewness and the excess kurtosis are exhibited in the data of both China and Korea. This

[1] The X12-ARIMA monthly seasonal adjustment method by the U.S. Department of Commerce in the Eviews7 program was used.

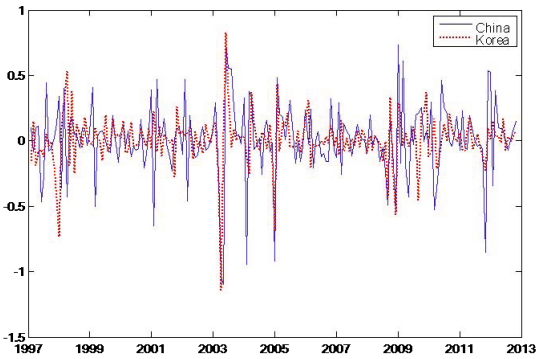


Fig. 1 The growth rate of the tourist arrivals to Thailand from China and Korea

means that the two data series have peakedness of distribution and heaviness of the tail. The null hypothesis of normality of the Jarque-Bera tests are rejected in both data series. Figure 1 presents the growth rate of the tourist arrivals to Thailand from China and Korea during this period. It can be seen that the growth rates of both the countries have considerable fluctuation and co-movement. The Pearson's correlation of 0.33 indicates that two data series have correlation.

5 Empirical Results

5.1 Results of ARMA-GARCH Model for Marginal Estimation

To test whether the data are stationary or not, we use the Augmented Dickey-Fuller test (ADF). The results show that the two data series are stationary at p-value 0.01. For examining the volatility of the growth rates of the tourist arrivals to Thailand from China and Korea, the ARMA(1,0)-GARCH(1,1) with skewed student T residual, $\varepsilon_t \sim \text{SkT}\ (\nu, \gamma)$, was modeled. The choice of skewed student T distribution is because of the fact that the two data series exhibit negative skewness and excess kurtosis. Identification of the optimal models was based on the Akaike information criterion (AIC). In Table 4, all the parameters of the model are significant at levels 0.01, 0.05, except that the parameter α of Korea is significant at level 0.1 and parameter ar1, β of China are insignificant at level 0.1. The degree of freedom and the skewness parameters are also significant at level 0.01. This indicates that the two data series are the skewed student T distributions. The results show that the volatility of the growth rates of the tourist arrivals from China has a short run persistence (α) and no long run persistence since the parameter β is insignificant. As for the volatility of the growth rates of the tourist arrivals from Korea, it has a weak long run persistence, although there is a value on the second moment which equals 1.08, which is more than one. But a value of the log moment equals -0.708, which is less

than zero; this is necessary and sufficient for it to be the strict stationarity, and the persistence of conditional variance is also satisfied.

Table 4 Result of ARMA(1,0)-GARCH(1,1) with Skewed Student T Residual for Growth Rate of Tourist Arrivals to Thailand from China and Korea

	China	Std. error	(p-value)	Korea	Std. error	(p-value)
ar1	−0.092	0.087	(0.292)	−0.219	0.078	(0.005)
ω	0.031	0.012	(0.011)	0.011	0.006	(0.044)
α	0.836	0.411	(0.042)	0.869	0.457	(0.057)
β	0.114	0.129	(0.377)	0.211	0.106	(0.045)
ν (degree of freedom)	3.249	0.873	(0.000)	2.980	0.808	(0.000)
γ (skewness)	0.870	0.076	(< 2e-16)	0.842	0.069	(< 2e-16)
Log likelihood	15.699	-	-	104.044	-	-
AIC	−19.398	-	-	−194.088	-	-
2^{nd} moment	0.950	-	-	1.080	-	-
Log moment	−0.978	-	-	−0.708	-	-

Table 5 P-values of K-S Test and Box-Ljung Test for Marginal Distributions

	Margin 1 (China)	Margin 2 (Korea)
K-S test	1.000	1.000
Box-Ljung test		
1^{st} moment	0.437	0.285
2^{nd} moment	0.419	0.894
3^{rd} moment	0.272	0.110
4^{th} moment	0.589	0.939

Note: The null hypothesis of the K-S test = data is uniform; the null hypothesis of the Box-Ljung test = no serial correlation.

Then we transformed the standardized residuals from these two models into uniform [0,1], u_1 and u_2, by using the empirical distribution function $F_n(x) = \frac{1}{n+1}\Sigma_{i=1}^{n} 1(X_i \leq x)$, where $X_i \leq x$ is the order statistics and 1 is the indicator function. For checking whether the marginal distributions that we transformed were correctly specified, which means that u_1 and u_2 are i.i.d. uniform [0,1]. We used the Kolmogorov-Smirnov (K-S) test for uniform [0,1] and the Box-Ljung test for the serial correlation. The results of the K-S test, as given in Table 5, show that both of these marginal distributions are uniform, by accepting the null hypothesis at p-values as equal to 1. The results of the Box-Ljung test show that all of the four moments of the marginal distributions are i.i.d. by accepting the null hypothesis that no serial correlation at p-value should be greater than 0.05. Therefore, our marginal distributions were not misspecified and can be used for the copula model.

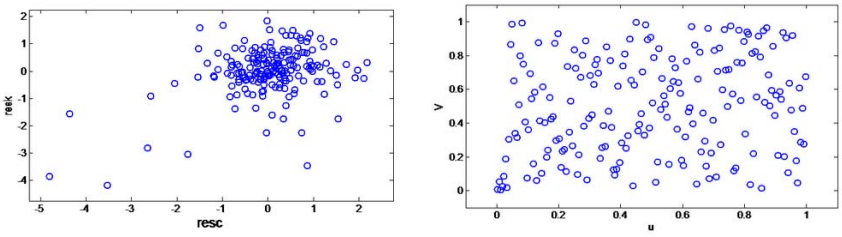


Fig. 2 The dependence of the standardized residuals (left) and the dependence of the transformed standardized residuals to uniform [0,1] (right)

Figure 2 shows the scatterplot of the standardized residuals from the ARMA-GARCH model for China and Korea, and the scatterplot of the transformed standardized residuals to uniform. This figure shows that there is weak dependence on the lower (left) tail between the marginal distributions of China and Korea. Therefore, to find the true dependence structure between the two marginal distributions of these two countries, various families of the copulas were used; the results are presented in the next section.

5.2 Results of Copula Estimations

We used various families of the copulas to examine the dependency and the dependence structure between the marginal distributions of the growth rates of the Chinese and Korean tourist arrivals to Thailand. For the static copula models, we used the Gaussian, Student's T, Clayton, Gumbel, Frank, Joe, rotated Clayton 180°, rotated Gumbel 180°, and rotated Joe 180°. Table 6 shows the results of each family of copula, including the copula parameter, standard error, and p-value. Kendall's tau correlation that was transformed from the copula parameter was used because each family of copula has a different range of copula parameters; so we inverse a copula parameter into a Kendall's tau correlation, and then it is bound on the interval $[-1,1]$. Kendall's tau is a measure of concordance which is a function of copula; thus, we can use it to assess the range of dependence covered by the families of the copula. The lower tail (T^L) and upper tail (T^U) dependences were used because, in the descriptive statistics, we found that China and Korea had heavy-tailed distributions, from the values of kurtosis. The tail dependences can explain the degree of dependence in the tails or it can model dependence of extreme events such as loss events. This is of interest as we can be aware of possible concurrent bad events in the tails. If there exists upper tail dependence, then it is an indication that the probability of the joint occurrence of extreme values is positive, or that the two variables rise together. But if there exists lower tail dependence, then it is an indication that the probability of the joint occurrence of extreme values is negative, or that the two variables crash together. The AIC and BIC were used for selection of copulas. The result shows that all of the estimated copula parameters from each family

are significant and that there are *weak positive dependences* between two marginal distributions. Moreover, most copulas present lower (left) tail dependence (T^L) and very weak upper (right) tail dependence (T^U). This implies that the growth rates of the Chinese and the Korean tourist arrivals to Thailand have co-movement. When the AIC and the BIC are looked at for selecting the copula model, the rotated Joe 180° copula is chosen for describing the dependence structure. The rotated Joe 180° can capture the lower (left) tail dependence. The estimated parameter of the rotated Joe 180° copula is 1.287, the Kendall's tau correlation is 0.140, and the lower (left) tail dependence (T^L) is 0.286, all of which provides evidence of lower tail dependence. This means that the growth rates of the tourist arrivals from China and Korea have co-movement that is both upward and downward, but with a weak dependence. The rise or loss of tourism demand from China (Korea) is slightly correlated by the rise or loss of tourism demand from Korea (China). Moreover, the lower (left) tail dependence indicates that there are chances that Thailand will have to face the probability of joint occurrences of large loss of tourist arrivals from China and Korea.

Table 6 Static Copula Models

Copula	parameter	Std. error (p-value)	Kendall's Tau	T^L (Lower tail)	T^U (Upper tail)	AIC	BIC
Gaussian	$\theta = 0.197$	0.069 (0.002)	0.127	0	0	−5.461	−2.214
Student's T	$\theta = 0.200$	0.0785 (0.006)	0.128	0.109	0.109	−10.322	−3.828
	$\nu = 4.487$	1.847 (0.008)					
Clayton	$\theta = 0.381$	0.109 (3.145e-04)	0.160	0.162	0	−15.659	−12.412
Gumbel	$\theta = 1.105$	0.060 (0.000)	0.095	0	0.127	−1.466	1.781
Frank	$\theta = 1.160$	0.451 (0.012)	0.127	0	0	−4.623	−1.376
Joe	$\theta = 1.048$	0.087 (0.000)	0.027	0	0.063	1.683	4.930
Rotated Clayton 180°	$\theta = 0.091$	0.099 (0.180)	0.044	0	0	1.078	4.325
Rotated Gumbel 180°	$\theta = 1.179$	0.059 (0.000)	0.152	0.199	0	−16.062	−12.815
Rotated Joe 180°	$\theta = 1.287$	0.092 (0.000)	0.140	0.286	0	−18.616	−15.369

Table 7 presents a goodness-of-fit test of copulas based on the Vuong and Clarke tests for bivariate copulas. The Vuong and Clarke method tests by comparing the copulas and also by taking into consideration the null hypothesis, which allows for statistically significant decision among the two models, and then gives a score for copulas. The copula with the highest score should be chosen. The results show that

the rotated 180° copula is the best fit which gives the highest score, for both the Vuong test and the Clake test, followed by the rotated Gumbel 180° copula. These results are consistent with the results obtained from the AIC and the BIC.

Table 7 Goodness-of-fit Test Scores Based on Vuong and Clarke Tests

	Gaussian	Student's T	Clayton	Gumbel	Frank	Joe	rotated Clayton 180°	rotated Gumbel 180°	rotated Joe 180°
Vuong	−2	0	0	−2	0	0	0	2	2
Clarke	−3	5	4	−1	−1	−6	−6	4	4

Note: The values in table are the scores at significance level = 0.05 under the null hypothesis that both copulas are statistically equivalent. The family with the highest score should be selected.

Table 8 Goodness-of-fit Test Based on Kendall's Process for Gaussian, Rotated Gumbel 180°, and Rotated Joe 180° Copulas

	Gaussian	rotated Gumbel 180°	rotated Joe 180°
p-value of CvM	0.22	0.94	0.49
p-value of KS	0.05	0.85	0.49

Note: Critical value α= 5%. If p-value > 0.05, it means that the dependence structure of the data series is appropriate for the chosen family of copulas.

Table 8 presents a second goodness-of-fit test of the copulas, based on Kendall's process, which offers the p-values of two statistical analyses, the Cramér-von Mises test (CvM) and the Kolmogorov-Smirnov test (KS). We selected two copulas that gave the highest score from the Vuong and Clarke tests and one family from the elliptical copula, the Gaussian. We used a second goodness-of-fit test to ensure that the dependence structure of the data series is appropriate for the chosen family of copulas. The results showed that the p-values of the CvM and KS tests of the four copula families are greater than 0.05, thus indicating the acceptance of the null hypothesis.

The dependence structures vary over time. Thus, we also used time-varying copula models to show the co-movement between the growth rates of the tourist arrivals from China and Korea to Thailand during this period. We used the time-varying copula models, as in Patton [21], such as the time-varying Gaussian copula, time-varying Gumbel copula, and time-varying rotated Gumbel 180° copula. Furthermore, we added the time-varying Joe copula and the time-varying rotated Joe 180° copula. As given in Table 9, the results showed that the time-varying rotated Joe 180° copula is the best fit from the AIC and the BIC, corresponding to the results from the static copula. All the three copula parameters of the time-varying rotated Joe 180° copula are significant at level 0.01. The parameter β represents the degree of persistence in the dependence. And the parameter α stands for significance,

Table 9 Time-varying Copula Models

Copula	parameters	Std. error	(p-value)	AIC	BIC
Time-varying Gaussian	$\omega = 0.560$	0.023	(0.000)	−3.287	6.454
	$\beta = -1.693$	0.040	(0.000)		
	$\alpha = 0.838$	0.044	(0.000)		
Time-varying Joe	$\omega = -0.610$	0.068	(1.110e-16)	4.914	14.655
	$\beta = 0.271$	0.054	(6.934e-07)		
	$\alpha = 1.814$	0.138	(0.000)		
Time-varying Gumbel	$\omega = -0.196$	0.079	(0.007)	2.297	12.039
	$\beta = 0.316$	0.069	(4.135e-06)		
	$\alpha = 0.564$	0.074	(6.984e-13)		
Time-varying rotated Gumbel 180°	$\omega = 2.204$	0.007	(0.000)	−14.355	−4.614
	$\beta = -1.217$	0.015	(0.000)		
	$\alpha = -1.195$	0.012	(0.000)		
Time-varying rotated Joe 180°	$\omega = 2.087$	0.034	(0.000)	−16.323	−6.582
	$\beta = -0.808$	0.022	(0.000)		
	$\alpha = -1.819$	0.085	(0.000)		

implying that there are variations over time in the dependences between the growth rates of the tourist arrivals from China and Korea. A comparison between the dependences of the static rotated Joe 180° copula and the time-varying rotated Joe 180° copula is shown in Figure 3. It can be observed that the dependences have fluctuated significantly through time.

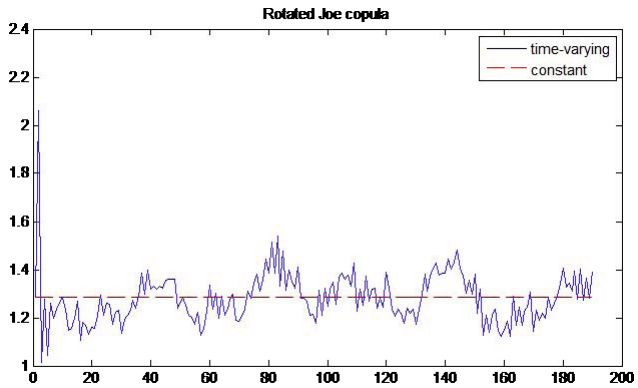


Fig. 3 The static and time-varying of the rotated Joe 180° copula

6 Policy Implications

Our results show that there exists a weak positive dependence between the growth rates of tourist arrivals from China and Korea to Thailand and that it keeps varying over time. The dependence structure is the rotated Joe 180° copula, and this provides evidence of lower tail dependence. This means that the growth rates of the tourist arrivals from China and Korea have co-movement that is both upward and downward, but with a weak dependence. The rise or loss of tourism demand from China (Korea) is slightly correlated by a rise or loss of tourism demand from Korea (China).

For risk management of the tourism demand, policy makers and tourism businesses should provide adequate and effective marketing and promotion to motivate the tourism demand by taking into consideration the time-varying dependency of China and Korea. Moreover, they should consider other target markets for substitution when there is loss of arrivals from these two markets. Especially in the low season, the simultaneous loss of arrivals will have more impact on the tourism industry and related businesses. Policy makers should examine the dependence structure of the tourist arrivals and the seasonal pattern of the other market countries for finding the movement across the countries and to ascertain which tourist market can be substituted when one or the other tourist markets have a loss from shock or low season so that they can diversify the risk of tourism demand.

7 Conclusion and Future Research

International tourism to Thailand has been growing over the past decade, but its successful management still remains a challenge due to the various uncertainties and events. China and Korea are important target markets which rank third and fourth, respectively. The growth rates of the tourist arrivals from these two countries have volatility and seem to have co-movement. This observation is interesting and raises certain questions, the answers to which we seek: Is there dependence between the growth rates of the tourist arrivals from China and Korea to Thailand? If so, what is the magnitude of the dependence? And, how is the dependence structure? In order to answer these research questions, this research work did a study to find out the dependence between the growth rates of the tourist arrivals from China and Korea to Thailand by applying the copula based GARCH model. The GARCH model can examine the volatility of the tourist arrivals and the copula model can find out the dependence structure between the two marginal distributions without having to assume normal distribution and linear correlation. The contribution of this study is based on the finding that it provides policy implications that are different from those given by Liu and Sriboonchitta [27], which was presented in the literature reviews. The findings are useful for policy makers and tourism businesses in terms of risk management of tourism demand.

For the empirical analysis, we first used the data descriptive statistics to analyze the two data series and found that they rejected the null hypothesis of normality

and exhibited skewness and excess kurtosis. Second, the ARMA(1,0)-GARCH(1,1) models with the skewed student-T residual were used to find out the marginal distributions of the growth rates of the arrivals of the two data series. We found that the ARMA(1,0)-GARCH(1,1) models could examine the volatility of the growth rates of the tourist arrivals from these two countries, that China has a short run persistence, and that Korea has a weak long run persistence. Third, the various copula families were used to measure the dependency and the dependence structure. The results show that there exists a weak positive dependence and that the rotated Joe 180° copula, which can capture the lower (left) tail, is the best fit. This implies that the growth rates of the tourist arrivals from China and Korea have co-movement that is both upward and downward, but with a weak dependence. The rise in or loss of tourism demand from China (Korea) is slightly correlated by the rise in or loss of tourism demand from Korea (China). Fourth, time-varying copula was used to show that the dependence parameter had varied over time. The results showed that the time-varying rotated Joe 180° copula is the best fit and that all the parameters are significant, corresponding to the results from the static copula. Our findings lead to policy implications on risk management of tourism demand, which has been discussed previously. For future research, we suggest that the copula based GARCH be applied to find out the dependence structure of the other tourist arrival countries for sketching the movement across the different countries.

References

1. World Travel & Tourism Council. Travel & Tourism Economic Impact 2012 (2012), http://www.wttc.org/site_media/uploads/downloads/world2012.pdf (accessed February 22, 2013)
2. World Tourism Organization. UNWTO World Tourism Barometer, vol. 11 (2013), http://dtxtq4w60xqpw.cloudfront.net/sites/all/files/pdf/unwto_barom13_01_jan_excerpt_0.pdf (accessed February 21, 2013)
3. World Tourism Organization. UNWTO Tourism Highlights 2012 Edition. World Tourism Organization (2012), http://dtxtq4w60xqpw.cloudfront.net/sites/all/files/docpdf/unwtohighlights12enlr_1.pdf (accessed February 21, 2013)
4. Goh, C.: Exploring impact of climate on tourism demand. Annals of Tourism Research 39(4), 1859–1883 (2012)
5. Vanhove, N.: The Economics of Tourism Destinations, 2nd edn. Elsevier, Ltd., London (2011)
6. Lean, H.H., Smyth, R.: Asian Financial Crisis, Avian Flu and Terrorist Threats: Are Shocks to Malaysian Tourist Arrivals Permanent or Transitory? Asia Pacific Journal of Tourism Research 14(3), 301–321 (2009)
7. World Bank. China 2030: Building a Modern, Harmonious, and Creative High-Income Society [pre-publication version]. World Bank, Washington, DC (2012), https://openknowledge.worldbank.org/handle/10986/6057 (accessed February 21, 2013)
8. Goh, C., Law, R.: Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention. Tourism Management 23, 499–510 (2002)

9. Chang, C.L., et al.: Modelling and Forecasting tourism from East Asia to Thailand under temporal and spatial aggregation. Mathematics and Computers in Simulation 79, 1730–1744 (2009)
10. Chu, F.L.: Forecasting tourism demand with ARMA-based methods. Tourism Management 30, 740–751 (2009)
11. Chan, F., et al.: Modelling multivariate international tourism demand and volatility. Tourism Management 26, 459–471 (2005)
12. Shareef, R., McAleer, M.: Modelling the uncertainty in monthly international tourist arrivals to the Maldives. Tourism Management 28, 23–45 (2007)
13. Coshall, J.T.: Combining volatility and smoothing forecasts of UK demand for international tourism. Tourism Management 30, 495–511 (2009)
14. Alvarez, G., et al.: Modeling Tourist Arrivals to Spain from the Top Five Source Markets. In: Ekasingh, B., Jintrawet, A., Pratummintra, S. (eds.) Proceedings of the 2nd International Conference on Asian Simulation and Modeling, Chiang Mai, Thailand, pp. 451–457 (2007)
15. Hoti, S., et al.: Modelling international tourism and country risk spillovers for Cyprus and Malta. Tourism Management 28, 1472–1484 (2007)
16. Jang, S.S., Chen, M.H.: Financial portfolio approach to optimal tourist market mixes. Tourism Management 29, 761–770 (2008)
17. Chen, M.H., et al.: Discovering Optimal Tourist Market Mixes. Journal of Travel Research 50(6), 602–614 (2011)
18. Chollets, L., et al.: International diversification: A copula approach. Journal of Banking & Finance 35, 403–417 (2011)
19. Sriboonchitta, S., et al.: Modeling volatility and dependency of agricultural price and production indices of Thailand: Static versus time-varying copulas. International Journal of Approximate Reasoning 54(6), 793–808 (2013)
20. Patton, A.J.: Modelling Asymmetric Exchange Rate Dependence Using the Conditional Copula. Unpublished Discussion paper. University of California (June 2001)
21. Patton, A.J.: Modelling asymmetric exchange rate dependence. International Economic Review 47(2), 527–556 (2006)
22. Goorbergh, R.: A Copula-Based Autoregressive Conditional Dependence Model of International Stock Markets. Unpublished DNB working paper No.22, Amsterdam, Netherlands (2004), http://www.dnb.nl/binaries/Working%20Paper%2022_tcm46-146679.pdf (accessed February 11, 2013)
23. Jondeau, E., Rockinger, M.: The Copula-GARCH model of conditional dependencies: An international stock market application. Journal of International Money and Finance 25, 827–853 (2006)
24. Wang, H., Cai, X.: A Copula Based GARCH Dependence Model of Shanghai and Shenzhen Stock Markets. Unpublished D-level Essay in Statistics, Dalarna University, Sweden (June 2011), http://www.statistics.du.se/essays/D11_HuilingWang_XinhuaCai.pdf (accessed February 10, 2013)
25. Reboredo, J.C.: Do food oil prices co-move? Energy Policy 49, 456–467 (2012)
26. Zhang, H., et al.: An integrated model of tourists? time use and expenditure behaviour with self-selection based on a fully nested Archimedean copula function. Tourism Management 33, 1562–1573 (2012)

27. Liu, J., Sriboonchitta, S.: Analysis of Volatility and Dependence between the Tourist Arrivals from China to Thailand and Singapore: A Copula-Based GARCH Approach. In: Huynh, V.N., Kreinovich, V., Sriboonchitta, S., Suriya, K. (eds.) Uncertainty Analysis in Econometrics with Applications. AISC, vol. 200, pp. 285–296. Springer, Heidelberg (2013)
28. Bollerslev, T.: Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics 31, 307–327 (1986)
29. Brechmann, E.C., Schepsmeier, U.: Modeling Dependence with C- and D-Vine Copulas: The R Package CDVine. Journal of Statistical Software 52(3), 1–27 (2013)
30. Nelson, D.B.: Stationarity and persistence in the GARCH(1,1) model. Econometric Theory 6, 318–334 (1990)
31. Lee, S.W., Hansen, B.E.: Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator. Econometric Theory 10, 29–52 (1994)
32. Sklar, A.: Fonctions de répartition á n dimensions et leurs marges. Publications de l'Institut de Statistique de L'Université de Paris 8, 229–231 (1959)
33. Nelson, R.B.: An Introduction to Copulas, 2nd edn. Springer, New York (2006)
34. Yan, J.: Enjoy the Joy of Copulas: With a Package copula. Journal of Statistical Software 21(4), 1–21 (2007)
35. Trivedi, P.K., Zimmer, D.M.: Copula Modeling: An Introduction for Practitioners. Foundations and Trends in Econometrics 1(1), 1–111 (2005)
36. Lee, L.: Generalized econometric models with selectivity. Econometrica 51, 507–512 (1983)
37. Clayton, D.G.: A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. Biometrika 65, 141–151 (1978)
38. Gumbel, E.J.: Distributions des Valeurs Extremes en Plusieurs Dimensions. Publications de l'Institute de Statistíque de l'Université de Paris 9, 171–173 (1960)
39. Joe, H.: Parametric Families of Multivariate Distributions with Given Margins. Journal of Multivariate Analysis 46(2), 262–282 (1993)
40. Frank, M.J.: On the simultaneous associativity of $F(x,y)$ and $x+y-F(x,y)$. Aequationes Math. 19, 194–226 (1979)
41. Fisher, M.: Tailoring copula-based multivariate generalized hyperbolic secant distributions to financial return data: An empirical investigation. Discussion papers, University of Erlangen-Nüremberg, Germany (2003), `http://www.statistik.wiso.uni-erlangen.de/forschung/d0047.pdf` (accessed January 25, 2013)
42. Manner, H., Reznikova, O.: A survey on time-varying copulas: specification, simulations and application. Econometric Reviews 31(6), 654–687 (2012)
43. Patton, A.J.: Applications of Copula Theory in Financial Econometrics. Dissertation, University of California (2002)
44. Joe, H.: Multivariate Models and Dependence Concepts. Chapman and Hall, London (1997)
45. Genest, C., et al.: A Semiparametric Estimation Procedure of Dependence Parameters in Multivariate Families of Distributions. Biometrika 82, 543–552 (1995)
46. Akaike, H.: Information Theory and an Extension of the Maximum Likelihood Principle. In: Petrov, B.N., Csaki, F. (eds.) Proceedings of the Second International Symposium on Information Theory, Budapest, pp. 267–281. Akademiai Kiado (1973)
47. Schwarz, G.: Estimating the Dimension of a Model. The Annals of Statistics 6(2), 461–464 (1978)
48. Vuong, Q.H.: Ratio Tests for Model Selection and Non-Nested Hypotheses. Econometrica 57(2), 307–333 (1989)

49. Clarke, K.A.: A Simple Distribution-Free Test for Nonnested Model Selection. Political Analysis 15(3), 347–363 (2007)
50. Genest, C., Rivest, L.P.: Statistical Inference Procedures for Bivariate Archimedean Copulas. Journal of the American Statistical Association 88(423), 1034–1043 (1993)
51. Wang, W., Wells, M.T.: Model selection and semiparametric inference for bivariate failure-time data. Journal of the American Statistical Association 95(449), 62–72 (2000)

Analyzing Relationship between Tourist Arrivals from China and India to Thailand Using Copula Based GARCH and Seasonal Pattern

Ornanong Puarattanaarunkorn and Songsak Sriboonchitta

Abstract. Chinese and Indian are the emerging tourist markets for Thailand. The two nations have tourism potential and make for interesting on doing a study about their tourism demand that was measure as the number of tourist arrivals. This study analyzed relationship between the tourist arrivals from China and India to Thailand by using the copula based GARCH model and the seasonal pattern. The findings by the copula based GARCH model show that there exists a weak positive dependence between the growth rates of tourist arrivals from China and India to Thailand and that this dependence keeps varying over time. The rotated Joe 180° copula, which can capture the lower (left) tail dependence, is chosen to describe the dependence structure. These mean that the growth rates of the tourist arrivals from China and India show a co-movement which is both upward and downward but with weak dependence. The rise or loss of tourism demand from China (India) is slightly correlated by a rise or loss of tourism demand from India (China). These results correspond to the seasonal patterns in which the seasonal pattern of China is in a direction opposite to the seasonal pattern of India in several periods, and the patterns showing a co-movement during some periods. Understanding the relationship between Chinese arrivals and Indian arrivals in each time period, it could contribute to policy implications such as developing the appropriate marketing and promotion strategies to attract other tourist markets as substitutes when we lose the regular tourist markets due to shock effects or low season.

1 Introduction

The travel and tourism sector has an important role to economic growth of Thailand. In 2011, Thailand's tourism receipts were about 26 million US dollars [1]. Thailand has a variety of the tourist arrivals markets. Strategic planning for all tourists

Ornanong Puarattanaarunkorn · Songsak Sriboonchitta
Faculty of Economics, Chiang Mai University, Chiang Mai 50200 Thailand
e-mail: pornan@kku.ac.th, songsakecon@gmail.com

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_23,

markets is difficult. Vanhove [2] said that it is difficult for any destination to make a choice among the potential markets by many reasons such as a destination cannot be operational on all these markets under limiting budget. Choice of markets is also one of the most important decisions in a strategic plan of a destination. Chinese and Indian are the interesting tourist markets for Thailand since the two countries are emerging markets (BRIC countries) with rapid economic growth and industrialization. These two countries have tourism potential and make for interesting on doing a study about their tourism demand that was measure as the number of tourist arrivals. The recent growth rates of Chinese arrivals to Thailand during 2010–2012 showed a rapid escalation, with over 30% yearly growth rate. The tourism receipt from the arrival of the Chinese to Thailand in 2012 was 3,409 million US dollars. The growth rate of Indian arrivals to Thailand was over 16%, with tourism receipt of 995 million US dollars in 2011 [3].

Seasonality is one characteristic of tourism demand [2] thus we assumed that the tourist arrivals from China and India to Thailand have seasonality. When we take into consideration the seasonal pattern by average seasonal index (SI)[1] that was calculated from the number of tourist arrivals during 2007–2012, we found that these two countries have the difference of seasonal pattern. Figure 1 shows a comparison of the seasonal patterns of the total tourist arrivals, both Chinese and Indian, to Thailand. The seasonal patterns of the Chinese and Indian tourists' arrivals move in different directions during several periods, and together in some periods. As evident from the data, the period from the months of January–February, and July–August were high season (SI>100) for the Chinese tourists but low season for the Indian tourists, while the stretch from May–June was low season (SI<100) for the Chinese tourists but high season for the Indian tourists. The difference in the arrival periods is, in fact, an advantage for Thailand in that one tourist market can act as a substitution market in case of loss of tourists from the other market.

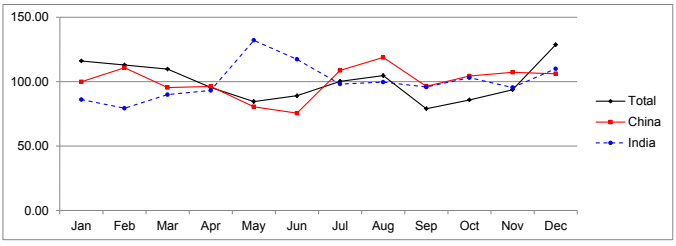


Fig. 1 Average Seasonal Index for international tourist arrivals to Thailand, 2007–2012

[1] A seasonal index (SI) is measured seasonal variation in terms of an index. It is an average that can be used to compare an actual observation relative to what it would be if there were no seasonal variation. This study used the method of simple average to calculate seasonal index, see more in Part 3.

However, the variation of tourism demand was not only due to the seasonal behavior but also because of irregular events, or due to the effect of a shock. Thus, when we consider the dependency, in addition to considering the seasonal pattern, we have to consider the dependence between the tourist arrivals due to irregular events. Understanding the dependence between the tourist arrivals due to irregular events and their seasonal patterns is important to Thailand's tourism industry because both of information can use to the decision making for policy maker. This study is interested in the dependence structure between the growth rates of tourist arrivals from China and India to Thailand. It is in the belief that if we can understand the dependence structures correctly and incorporate them with the seasonal pattern analysis, then these findings can become an important guideline for tourism promotion planning and management of risk in tourism demand.

The purpose of the study is to analyze the relationship between the tourist arrivals from China and India to Thailand by using the copula based GARCH model and seasonal index. The results from the copula based GARCH were analyzed incorporate with the seasonal pattern and contribute to policy recommendation in terms of the management of the tourism demand from these two emerging markets to Thailand. The copula based GARCH model was chosen because GARCH can examine the volatility of the tourist arrivals and copula can model the dependence structure between the two marginal distributions that obtain from GARCH model.

This paper is divided into five parts. The next part, which is the second part, is the literature review. The third part presents the methodology that describes the method is used to calculate seasonal index, the GARCH model, and the copula model. The fourth part presents the data used and the results of this study. The last part gives the conclusions and policy implications.

2 Literature Review

Modeling international tourism demand is vast. In this paper, we review, particularly, the studies of international tourism demand forecasting by using various time series models. For example, Chang et al. [4] used the autoregressive integrated moving average (ARIMA) model and the seasonal ARIMA (SARIMA) model for forecasting tourist arrivals from East Asia to Thailand. The generalized autoregressive conditional heteroskedastic (GARCH) model has been widely used to investigate the volatility of tourism demand. For example, Shareef and McAleer [5] used ARMA-GARCH(1,1) and ARMA-GJR(1,1) to examine the international tourist arrivals to the Maldives.

There is plenty of literature available on correlation analysis across international tourism markets and tourism destinations. For example, Chan et al. [6] and Alvarez et al. [7] analyzed the conditional correlation-based GARCH model of monthly international tourist arrival shocks. Jang and Chen [8] and Chen et al. [9] analyzed the correlation across international tourist arrivals for finding the optimal tourist market mixes by using a portfolio approach. All of the literature above measured the correlation, by the conventional approach, namely, the Pearson correlation coefficient.

The Pearson correlation is restricted within the assumption based on the normal distribution and the linear relationship of the data series. However, many data series are not normal distributions, and have non-linear relationships. To overcome that restriction, many studies used copulas to measure the dependency between the variables, especially in the financial field. Many studies used copulas that have cooperated with the GARCH model, that is, the copula based GARCH, to find the dependence structure of the marginal distribution of the conditional variance. The copula based GARCH model provides more flexibility for finding out the joint distributions and the transformation invariant correlation, without the assumption of linear correlation [10]. For example, Patton [11, 12] used the ARMA(p,q)-GARCH(1,1) model to estimate the marginal distributions of the Deutsche mark-US dollar and Japanese yen-US dollar exchange rates. Similarly, Jondeau and Rockinger [13] also used the copula based GARCH model to model the dependence structure between stock markets. In the tourism field, there is some literature on applying copula to model the dependence structure between variables. Zhang et al. [14] used a fully nested Archimedean copula function to find the dependence between three dependent variables: destination visits behavior, time use behavior, and expenditure behavior. Liu and Sriboonchitta [15] used a copula based GARCH model to model the volatility and the dependence structure between tourist arrivals from China to two destination markets, Thailand and Singapore.

3 Methodology

The copula based GARCH model and seasonal pattern of tourism demand by average seasonal index in Figure 1 were used to analyze the relationship between the tourist arrivals from China and India to Thailand. The GARCH model by Bollerslev [16] has been widely used for modeling volatility in asset returns and tourism demand. Therefore, we applied the ARMA-GARCH model to estimate the marginal distributions. The standardized residuals from ARMA-GARCH model were transformed to copula data $(F_1(x_1), F_2(x_2))$. After that, the copula approach was used to measure the dependence between the two marginal distributions. This study used the R-package CDVine by Brechmann and Schepsmeier [17] to analyze the constant copula since it provided a range of tools for bivariate data analysis. And, for the time-varying copula, we followed the method of Patton [12].

3.1 Average Seasonal Index

We used the method of simple average (see Sharma [18]) to calculate seasonal index of the tourist arrivals from China and India to Thailand during January 2007 to December 2012. The seasonal index (SI) is measured seasonal variation in terms of an index. It is an average that can be used to compare an actual observation relative to what it would be if there were no seasonal variations. The steps of this method are presented below,

(1) The number of tourist arrivals (x) were used to calculate the average of each month by $\bar{x}_i = \frac{\Sigma_{j=1}^{N} x_{i,j}}{N}$; i= Month 1,...,12, j= Year 1,...,6, N= Number of years. For example $\bar{x}_1 = \frac{x_{Jan,2007}+\ldots+x_{Jan,2012}}{6}$.

(2) The average 12 months denoted by $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_{12}$.

(3) We obtain an average of monthly averages by dividing the total of monthly averages by 12, $\bar{Y} = \frac{\bar{x}_1+\ldots+\bar{x}_{12}}{12}$.

(4) We compute the seasonal index (SI) for each month, $\mathrm{SI}_i = \frac{\bar{x}_i}{\bar{Y}} \times 100, i = 1,\ldots,12$, for SI > 100 means high season, SI < 100 means low season.

3.2 ARMA-GARCH Model

We adopt ARMA(1,0)-GARCH(1,1) and ARMA(2,0)-GARCH(1,1) with skewed student T (SkT) distribution residual for the marginal distribution of the logarithm of the monthly growth rates of tourist arrivals to Thailand from China and India (y_t), respectively:

$$y_t = a_0 + \Sigma_{i=1}^{p} a_i y_{t-i} + \varepsilon_t \tag{1}$$

$$\varepsilon_t = z_t \sqrt{h_t}, z_t \sim SkT(\nu, \gamma) \tag{2}$$

$$h_t = \omega_t + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1} \tag{3}$$

In equation (1) presents ARMA(p,0) process where y_{t-i} is an autoregressive term of y_t and ε_t is an error term. Equation (2) then define this error term as the product between conditional variance h_t and a residual z_t. A residual z_t is assumed to follow the skewed student T (SkT) distribution with the degree of freedom parameter ν and the skewness parameter γ. Equation (3) presents GARCH(1,1) process where $\omega_t > 0, \alpha \geq 0, \beta \geq 0$ are sufficient to ensure that the conditional variance $h_t > 0$. The $\alpha \varepsilon_{t-1}^2$ represent the ARCH term and α refers to the short run persistence of shocks, while βh_{t-1} represent the GARCH term and refers to the contribution of shocks to long run persistence $(\alpha + \beta)$. The second moment condition is $\alpha + \beta < 1$.

3.3 Copulas

One approach of modeling the multivariate dependence is the copula. The copula functions can offer us the flexibility of merging a univariate distribution to get a joint distribution with an appropriate dependence structure. The fundamental theorem of copula is Sklar's theorem by Sklar [19]. Nelson [20] has made a description of the copula theory, as follows:

Let H be a joint distribution function with marginal distributions F and G. Then there exists a copula C for all x, y in real line, with the following property:

$$H(x,y) = C(F(x), G(y)) \tag{4}$$

If F and G are continuous, C is unique. Conversely, if C is a copula and F and G are univariate distribution functions, then the above function H in (4) is a joint distribution function with marginal distributions F and G. If H is known, the copula is an equation (4) that one can get from the form,

$$C(u,v) = H(F^{-1}(u), G^{-1}(v)), \tag{5}$$

where F^{-1} and G^{-1} are the quantile functions of the marginal distributions.

3.4 Characteristics of Copula Families

This paper uses constant copulas and time-varying copulas to describe the dependence between two marginal distributions. Each copula family has different functions and characteristics; these copula families can be taken from Trivedi and Zimmer [21], Nelson [20], etc., and they are as follows.

(*a*) Constant Copulas

Gaussian (Normal) Copula. Trivedi and Zimmer [21] state that the Gaussian copula allows for equal degrees of positive and negative dependences. This copula function has been offered by Lee [22].

$$\begin{aligned} C(u_1,u_2;\rho) &= \Phi_G(\Phi^{-1}(u_1), \Phi^{-1}(u_2);\rho) \\ &= \int_{-\infty}^{\phi^{-1}(u_1)} \int_{-\infty}^{\phi^{-1}(u_2)} \frac{1}{2\Pi\sqrt{(1-\rho^2)}} \times [\frac{-(s^2-2\rho st+t^2)}{2(1-\rho^2)}] dsdt \end{aligned} \tag{6}$$

where Φ^{-1} is the inverse of the standard normal c.d.f. and $\Phi_G(u_1,u_2)$ is the standard bivariate normal distribution with the Pearson correlation parameter, $\rho \in (-1,1)$. Gaussian copula is tail independent.

Student's T Copula. Trivedi and Zimmer [21] point out that the Student's T copula has two dependence parameters, ν degrees of freedom, and correlation $\rho \in (-1,1)$. The student's T copula exhibits tail (upper and lower) dependence.

$$C^T(u_1,u_2;\rho,\nu) = \int_{-\infty}^{T_\nu^{-1}(u_1)} \int_{-\infty}^{T_\nu^{-1}(u_2)} \frac{1}{2\Pi\sqrt{(1-\rho^2)}} \times [1 + \frac{(s^2-2\rho sT+T^2)}{\nu(1-\rho^2)}]^{-(\frac{\nu+2}{2})} dsdT \tag{7}$$

where $T_\nu^{-1}(u_1)$ is the inverse of the c.d.f. of the standard univariate T-distribution with ν degrees of freedom which is controlling the heaviness of the tails.

Clayton Copula. This family of copulas was discussed by Clayton [23].

$$C(u_1,u_2;\theta) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-1/\theta} \tag{8}$$

for the dependence parameter $\theta \in (0,\infty)$. The Clayton copula shows strong lower tail dependence and relatively weak upper tail dependence, it can be used to study involved risks.

Gumbel Copula. The Gumbel copula was first discussed by Gumbel [24], and so it has been referred to as the Gumbel family.

$$C(u_1,u_2;\theta) = exp(-[(-\ln(u_1))^{\theta} + (-\ln(u_2))^{\theta}]^{\frac{1}{\theta}}) \tag{9}$$

for the dependence parameter $\theta \in [1,\infty)$. The Gumbel copula shows strong upper tail dependence, we can say that it contrasts with the Clayton copula. Joe [25] is of the view that the Gumbel copula is an extreme value copula.

Frank Copula. Nelson [20] points out that the Frank family was first presented in Frank [26].

$$C(u_1,u_2;\theta) = \frac{-1}{\theta}\ln(1+((e^{-\theta u_1}-1)(e^{-\theta u_2}-1))/(e^{-\theta}-1)) \tag{10}$$

for the dependence parameter $\theta \in (-\infty,\infty)/\{0\}$. The Frank copula allows for negative dependence. The dependences in the tail symmetry of the Frank copula are akin to those of the Gaussian and Student-t copulas. The Frank copula can capture weak dependence in the tails better than the Gaussian.

Joe Copula. Nelson [20] points out that this family was proposed in Joe [25].

$$C(u_1,u_2;\theta) = 1-[(1-u_1)^{\theta} + (1-u_2)^{\theta} - (1-u_1)^{\theta}(1-u_2)^{\theta}]^{\frac{1}{\theta}} \tag{11}$$

for the dependence parameter $\theta \in [1,\infty)$. The Joe copula also has upper tail dependence with $2-2^{1/\theta}$ as the limit.

Rotating the copula was made for the asymmetric dependence structures such as those of the Clayton, Gumbel, and Joe. Nelson [20] defined the rotation of the copulas by means of $C_R(u_1,u_2) = u_1+u_2-1+C(1-u_1,1-u_2)$. In practice, for rotated bivariate copulas, we can transform the input arguments u_1 and u_2 to $1-u_1$ and $1-u_2$ for 180 degrees. When rotating copulas by 180 degrees, we can also call the survival copulas of the corresponding family, for example, survival Clayton [17]. After we rotate copulas, such as with the rotated Gumbel, the copulas will show stronger dependency in the lower tail instead of the upper tail.

Rotated Clayton Copula. The rotated Clayton copula can capture the upper tail dependence conversely to the Clayton copula. Fisher [27] shows the functional form as

$$C(u_1,u_2;\theta) = u_1+u_2-1+[(1-u_1)^{-\theta} + (1-u_2)^{-\theta} - 1]^{-\frac{1}{\theta}} \tag{12}$$

Rotated Gumbel Copula. The rotated Gumbel copula can capture the lower tail dependence conversely to the Gumbel copula. Fisher [27] shows the functional form as

$$C(u_1,u_2;\theta) = u_1 + u_2 - 1 + exp(-[(-\ln(1-u_1))^{\theta} + (-\ln(1-u_2))^{\theta}]^{\frac{1}{\theta}}) \quad (13)$$

Rotated Joe Copula. The rotated Joe copula can capture the lower tail dependence conversely to the Joe copula. Fisher [27] shows the functional form as

$$C(u_1,u_2;\theta) = u_1 + u_2 - (u_1^{\theta} + u_2^{\theta} - u_1^{\theta}u_2^{\theta})^{\frac{1}{\theta}} \quad (14)$$

Time-Varying Copulas
Since it is a fact that dependence between the marginal distributions of the time series variables are not constant through time, they should be considered as time-varying copulas. We used the functional form of the time-varying copulas by following an ARMA(1,10) process which was presented by Patton [12].

3.5 Maximum Likelihood Estimation

The method of maximum pseudo-log likelihood, studied by Genest et al. [28], was used for estimation since the marginal distribution functions F and G of the random vectors are unknown. Thus, we can construct the pseudo copula observations by using the empirical distribution functions to transform the standardized residual series into uniform [0,1] as rank based.

Under the assumption that the marginal distributions F and G are continuous, the copula C_θ is a bivariate distribution with density c_θ and pseudo-observations $F_n(X_i)$ and $G_n(Y_i)$, $i = 1,2,...,n$. The pseudo-log likelihood function of θ can be given as

$$L(\theta) = \Sigma_{i=1}^{n} \log[c_\theta(F_n(X_i), G_n(Y_i))]. \quad (15)$$

Then, maximizing the pseudo-log likelihood yield as an estimator of θ,

$$\theta = \Sigma_{i=1}^{n} \frac{\partial}{\partial \theta} \log[c_\theta(F_n(X_i), G_n(Y_i))] = 0 \quad (16)$$

where $c_\theta = \frac{\partial^2 C_\theta(F_n(x),G_n(y))}{\partial x \partial y}$, $F_n(x) = \frac{1}{n+1}\Sigma_{i=1}^{n} 1(X_i \leq x)$ and $G_n(x) = \frac{1}{n+1}\Sigma_{i=1}^{n} 1(Y_i \leq y)$ are the empirical distributions.

3.6 Selection of Copulas

Selecting a family of copulas is based upon information criteria such as Akaike Information Criterion (AIC) by Akaike [29] and Bayesian Information Criterion (BIC) by Schwarz [30]. To examine whether the dependence structure of the data series is appropriate for a chosen family of copulas, we used a goodness-of-fit test in the R-package CDVine. A goodness-of-fit test based on a scoring approach by Vuong [31]

and Clarke [32]. A second goodness-of-fit test based on Kendall's tau by Genest and Rivest [33], and Wang and Wells [34] was conducted. It offered the Cramér-von Mises (CvM) and Kolmogorov-Smirnov (KS) test statistics and estimated the p-values by bootstrapping.

4 Data and Empirical Results

4.1 Seasonal Index

The seasonal index in Table 1 were used to construct the seasonal pattern, Figure 1, was shown in Part 1. The period from the months of January–February, and July–August were high season (SI>100) for the Chinese tourists but low season for the Indian tourists, while the stretch from May–June was low season (SI<100) for the Chinese tourists but high season for the Indian tourists.

Table 1 Seasonal Index of Tourist Arrivals to Thailand

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Total Arrivals	116	113	110	95	85	89	100	105	79	86	94	129
China	100	111	95	96	80	76	109	119	96	104	107	106
India	86	79	90	93	132	117	98	100	96	103	95	110

4.2 Copula Based GARCH

For modeling the volatility and the dependency of the tourist arrivals from irregular events, the seasonal adjusted (for removing the seasonal component of a time series data) of the monthly tourist arrivals to Thailand from China and India, during January 1997 to December 2012, data from the Ecowin database were used. The two data series are taken into the logarithm of the monthly arrival rates (the growth rates of tourist arrivals), the growth rate of tourist arrivals $=\ln \frac{\text{Number arrivals}_t}{\text{Number arrivals}_{t-1}}$.

Table 2 presents a descriptive statistics of the growth rates of tourist arrivals to Thailand from China and India. Both the countries have positive average growth rates during January 1997 to December 2012. Both the Chinese and the Indian data exhibit negative skewness and excess kurtosis. This implies that both the data series have peak of distribution and heaviness of tail. The null hypothesis of normality of the Jarque-Bera tests are rejected in both the data series. Figure 2 presents the growth rates of the tourist arrivals to Thailand from China and India along this period. It can be seen that the growth rates of both the countries have considerable fluctuation.

Table 2 Descriptive Statistics for Growth Rate of Tourist Arrivals

	China	India
Mean	0.011	0.010
Median	0.012	0.006
Maximum	0.739	0.526
Minimum	−1.099	−0.606
Std. Dev.	0.279	0.108
Skewness	−0.913	−0.341
Kurtosis	6.132	10.342
Jarque-Bera	104.580	432.757
(p-value)	(0.000)	(0.000)
Observations	191	191

Note: The null hypothesis of Jarque-Bera = data is taken as the normal distribution.

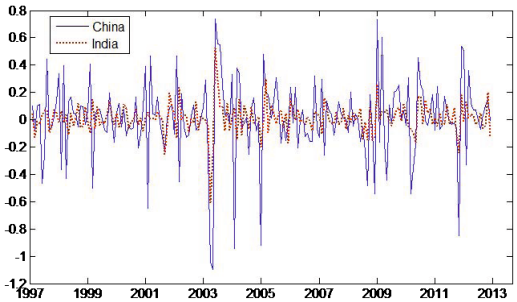


Fig. 2 The growth rates of tourist arrivals to Thailand from China and India

4.3 Results of ARMA-GARCH Model for Marginal Estimation

To test whether the data are stationary or not, we used the Augmented Dickey-Fuller test (ADF). The results showed that the two data series were stationary at p-value 0.01. The ARMA(1,0)-GARCH(1,1) and ARMA(2,0)-GARCH(1,1) with skewed student T residual $\sim SkT(\nu, \gamma)$ are modeled for estimating the marginal distributions of the Chinese and the Indian data, respectively. A choice of skewed student T distribution is based on the fact that the two data series exhibit negative skewness and excess kurtosis. The Akaike Information criterion (AIC) is used to identify the optimal models. Table 3 provides the estimation results from the ARMA-GARCH models, with all the parameters having significance at levels 0.01, 0.05, except for ar1, β of China and ar2 of India having insignificance at level 0.1. These observations imply that the volatility of the growth rates of tourist arrivals from China has a short run persistence, and India has a long run with a value of $\alpha + \beta = 0.96$.

Table 3 Results of ARMA(1,0)-GARCH(1,1), $\varepsilon_t \sim SkT(\nu, \gamma)$ for China and ARMA(2,0)-GARCH(1,1), $\varepsilon_t \sim SkT(\nu, \gamma)$ for India

	China	Std. error	(p-value)	India	Std. error	(p-value)
ar1	−0.093	0.087	(0.286)	−0.339	0.076	(8.72e-06)
ar2	-	-	-	−0.107	0.067	(0.107)
ω	0.032	0.013	(0.013)	0.004	0.001	(0.005)
α	0.846	0.423	(0.046)	0.479	0.201	(0.017)
β	0.114	0.129	(0.377)	0.337	0.144	(0.019)
ν (degree of freedom)	3.195	0.851	(0.000)	4.201	1.385	(0.002)
γ (skewness)	0.872	0.074	(< 2e-16)	0.724	0.078	(< 2e-16)
Log likelihood	16.759	-	-	196.325	-	-
AIC	−21.517	-	-	−378.651	-	-
2^{nd} moment	0.96	-	-	0.82	-	-

Then we transform the standardized residuals from these two models into uniform [0,1], u_1 and u_2, by using the empirical distribution function, $F_n(x) = \frac{1}{n+1}\Sigma_{i=1}^{n} 1(X_i \leq x)$, where $X_i \leq x$ is the order statistics, 1 is the indicator function. For checking whether the marginal distributions that we transformed are correctly specified, which means that u_1 and u_2 should be i.i.d. uniform [0,1], we use the Kolmogorov-Smirnov (K-S) test for Uniform [0,1] and the Box-Ljung test for serial correlation. In Table 4, the results of the K-S test are given, which show that both of these marginal distributions are uniform, by accepting the null hypothesis at p-values equal to 1. The results of the Box-Ljung test show that all of the four moments of the marginal distributions are i.i.d. by accepting the null hypothesis that no serial correlation at p-value is greater than 0.05. Therefore, our marginal distributions are not misspecified and can be used for the copula model.

Table 4 P-values of K-S Test and Box-Ljung Test for Marginal Distributions

	K-S test	Box-Ljung test			
		1^{st} moment	2^{nd} moment	3^{rd} moment	4^{th} moment
Margin 1 (China)	1.00	0.465	0.861	0.342	0.916
Margin 2 (India)	1.00	0.154	0.722	0.071	0.722

Note: The null hypothesis of the K-S test = data is uniform; the null hypothesis of the Box-Ljung test = no serial correlation.

4.4 Results of Copula Estimations

The various families of copulas were used to examine the dependence structure between the marginal distributions of China and India. The results show that the rotated Joe 180° copula, which can capture the lower (left) tail dependence, is the best fit by looking at the smallest values of the AIC and the BIC. Moreover, a goodness-of-fit test of the copulas based on the Vuong and Clarke tests and a goodness-of-fit

test based on the Kendall's process can be used to confirm the same. The process of Vuong and Clarke is tested by comparing the copulas and also by considering the null hypothesis which allows for a statistically significant decision, and then a score is given for the copulas. The copula with the highest score should be chosen. The results in Table 6 show that a rotated Joe 180° copula provides the highest scores. A second goodness-of-fit test based on the Kendall's process, which offers p-values of the two statistics, the Cramér-von Mises test (CvM) and the Kolmogorov-Smirnov test (KS). We selected three copulas, which give the highest score from Vuong and Clarke tests, and one family from the elliptical copula, the Gaussian, to assure that the dependence structure of the data series is appropriate as regards a chosen family of copulas. The results in Table 7 show that the p-values of the CvM and KS tests of the three copula families are greater than 0.05, thus proving that they accept the null hypothesis, with the exception of the Gaussian copula.

The estimated parameter of the rotated Joe 180° copula is 1.297, the Kendall's tau[2] is 0.144, and the lower (left) tail[3] (T^L) is 0.293. This means that the growth rates of the tourist arrivals from China and India have a co-movement which is both upward and downward but with weak dependence. The rise or loss of tourism demand from one country is slightly correlated by the rise or loss of tourism demand from the other. Moreover, the lower (left) tail dependence indicates that Thailand has chances of facing the probability of joint occurrences of large loss of tourist arrivals from China and India.

For the time-varying copula, the results show that the time-varying rotated Joe 180° copula is the best fit from the smallest AIC and the BIC, corresponding to the results from the constant copula. All the three copula parameters of the time-varying rotated Joe 180° copula, $\omega = -0.406$, $\beta = 0.675$, $\alpha = 0.258$, are significance at level 0.01. The parameter β represents the degree of persistence in the dependence and the parameter α stands for significance, implying that there are variations over time in the dependences between the growth rates of the tourist arrivals from China and India. A comparison between the dependences of the static rotated Joe 180° copula and the time-varying rotated Joe 180° copula is shown in Figure 3. It is evident that the dependences have fluctuated significantly over time.

[2] Kendall's tau correlation that was transformed from the copula parameter was used because each family of copula has a different range of copula parameter; so, we inverse a copula parameter into a Kendalls tau correlation, and then it is bounded on the interval $[-1, 1]$. Kendall's tau is a measure of concordance that is a function of copula; hence we can use it to assess the range of dependence covered by families of copula.

[3] The tail dependences can illustrate the degree of dependence in the tails or model the dependence of extreme events such as loss events. If there exists upper tail dependence, this indicates that the probability of the joint occurrences of extreme values is positive or that the two variables rise together. But if there exists lower tail dependence, then it is an indication that the probability of the joint occurrences of extreme values is negative or that the two variables to crash together.

Table 5 Constant Copula Models

Copula	parameter	Std. error (p-value)	Kendall's Tau	T^L (Lower tail)	T^U (Upper tail)	AIC	BIC
Gaussian	$\theta = 0.184$	0.069 (0.004)	0.118	0	0	−4.482	−1.230
Student's T	$\theta = 0.155$	0.082 (0.031)	0.099	0.123	0.123	−11.600	4.023
	$\nu = 3.777$	1.342 (0.003)					
Clayton	$\theta = 0.383$	0.107 (2.229e-04)	0.161	0.164	0	−16.956	−13.703
Gumbel	$\theta = 1.081$	0.058 (0.000)	0.075	0	0.101	−0.243	3.010
Frank	$\theta = 0.930$	0.454 (0.021)	0.102	0	0	−2.184	1.068
Joe	$\theta = 1.013$	0.078 (0.000)	0.007	0	0.018	1.972	5.224
Rotated Clayton 180°	$\theta = 0.056$	0.091 (0.271)	0.027	0	0	1.595	4.848
Rotated Gumbel 180°	$\theta = 1.175$	0.057 (0.000)	0.149	0.196	0	−16.196	−12.944
Rotated Joe 180°	$\theta = 1.297$	0.090 (0.000)	0.144	0.293	0	−20.402	−17.150

Table 6 Goodness-of-fit Test Scores Based on Vuong and Clarke Tests

	Gaussian	Student's T	Clayton	Gumbel	Frank	Joe	rotated Clayton 180°	rotated Gumbel 180°	rotated Joe 180°
Vuong	−3	0	2	−3	−3	−1	−2	3	7
Clarke	−2	7	4	−2	−2	−7	−7	3	6

Note: The values in table are the scores at significance level = 0.05 under the null hypothesis that both copulas are statistically equivalent. The family with the highest score should be selected.

Table 7 Goodness-of-fit Test Based on Kendall's Process for Gaussian, Clayton, Rotated Gumbel 180°, and Rotated Joe 180° Copulas

	Gaussian	Clayton	rotated Gumbel 180°	rotated Joe 180°
p-value of CvM	0.04	0.11	0.45	0.23
p-value of KS	0.05	0.21	0.55	0.29

Note: Critical value α= 5%. If p-value > 0.05, it means that the dependence structure of the data series is appropriate for the chosen family of copulas.

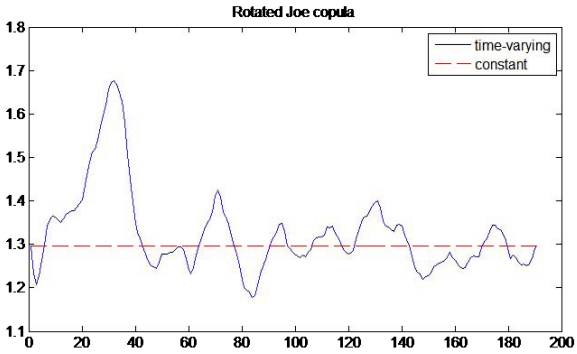


Fig. 3 The dependences of the constant and time-varying copulas of the rotated Joe 180° copula

5 Conclusions and Policy Implications

Analyzing relationship between tourist arrivals from China and India to Thailand, we used the dependency between the growth rates of tourist arrivals that was obtained by copula based GARCH model and the seasonal pattern of tourism demand.

Our empirical findings show that there exists a weak positive dependence between the growth rates of tourist arrivals from China and India to Thailand and that this dependence keeps varying over time. The results show that the rotated Joe 180° copula, which can capture the lower (left) tail dependence, is chosen to describe the dependence structure. This means that the growth rates of the tourist arrivals from China and India show a co-movement which is both upward and downward but with weak dependence. The rise or loss of tourism demand from China (India) is slightly correlated by a rise or loss of tourism demand from India (China). This is beneficial in the risk diversification of tourism demand – particularly, on those occasions where there is a loss of Chinese (Indian) tourists, we can promote tourism to Indian (Chinese) tourists as a substitute. On the other hand, if the two tourist markets had a high correlation, then the negative shock could lead to more loss of arrivals from both the countries, in addition to having a serious impact on the tourism industry and the related businesses.

The findings correspond to the seasonal patterns in which the seasonal pattern of China is in a direction opposite to the seasonal pattern of India in several periods, and the patterns showing a co-movement during some periods. In other words, low season for Chinese tourists is high season for Indian tourists in several periods, so Indian tourists can be a substitution market in the low season; this reduces the impact on the tourism industry.

Thus, it is evident that our findings have important implications for Thailand's tourism industry. For example, it can help policy makers examine, and make, the time-varying dependency cooperate with the seasonal patterns of the Chinese and the Indian tourists to effect some appropriate strategy plans for risk management.

For instance, when there is a loss of tourist market from shock effects or low season, policy makers can provide marketing and promotion strategy in order to attract other tourist markets as a substitute.

References

1. World Tourism Organization. UNWTO Tourism Highlights 2012 Edition. World Tourism Organization (2012), http://dtxtq4w60xqpw.cloudfront.net/sites/all/files/docpdf/unwtohighlights12enlr_1.pdf (accessed February 21, 2013)
2. Vanhove, N.: The Economics of Tourism Destinations, 2nd edn. Elsevier, Ltd., London (2011)
3. Department of Tourism of Thailand. Tourist Arrivals in Thailand 2012, Ministry of Tourism and Sports, Thailand (2013), http://www.tourism.go.th/tourism/th/home/tourism.php?id=11 (accessed April 20, 2013)
4. Chang, C.L., et al.: Modelling and Forecasting tourism from East Asia to Thailand under temporal and spatial aggregation. Mathematics and Computers in Simulation 79, 1730–1744 (2009)
5. Shareef, R., McAleer, M.: Modelling the uncertainty in monthly international tourist arrivals to the Maldives. Tourism Management 28, 23–45 (2007)
6. Chan, F., et al.: Modelling multivariate international tourism demand and volatility. Tourism Management 26, 459–471 (2005)
7. Alvarez, G., et al.: Modeling Tourist Arrivals to Spain from the Top Five Source Markets. In: Ekasingh, B., Jintrawet, A., Pratummintra, S. (eds.) Proceedings of the 2nd International Conference on Asian Simulation and Modeling, Chiang Mai, Thailand, pp. 451–457 (2007)
8. Jang, S.S., Chen, M.H.: Financial portfolio approach to optimal tourist market mixes. Tourism Management 29, 761–770 (2008)
9. Chen, M.H., et al.: Discovering Optimal Tourist Market Mixes. Journal of Travel Research 50(6), 602–614 (2011)
10. Sriboonchitta, S., et al.: Modeling volatility and dependency of agricultural price and production indices of Thailand: Static versus time-varying copulas. International Journal of Approximate Reasoning 54(6), 793–808 (2013)
11. Patton, A.J.: Modelling Asymmetric Exchange Rate Dependence Using the Conditional Copula. Unpublished Discussion paper, University of California (June 2001)
12. Patton, A.J.: Modelling Asymmetric Exchange Rate Dependence. International Economic Review 47(2), 527–556 (2006)
13. Jondeau, E., Rockinger, M.: The Copula-GARCH model of conditional dependencies: An international stock market application. Journal of International Money and Finance 25, 827–853 (2006)
14. Zhang, H., et al.: An integrated model of tourists' time use and expenditure behaviour with self-selection based on a fully nested Archimedean copula function. Tourism Management 33, 1562–1573 (2012)
15. Liu, J., Sriboonchitta, S.: Analysis of Volatility and Dependence between the Tourist Arrivals from China to Thailand and Singapore: A Copula-Based GARCH Approach. In: Huynh, V.N., Kreinovich, V., Sriboonchitta, S., Suriya, K. (eds.) Uncertainty Analysis in Econometrics with Applications. AISC, vol. 200, pp. 285–296. Springer, Heidelberg (2013)

16. Bollerslev, T.: Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics 31, 307–327 (1986)
17. Brechmann, E.C., Schepsmeier, U.: Modeling Dependence with C- and D-Vine Copulas: The R Package CDVine. Journal of Statistical Software 52(3), 1–27 (2013)
18. Sharma, J.K.: Business Statistics Problems and Solutions Dorling Kingdersley, p. 438. (India) Pvt. Ltd. (2010)
19. Sklar, A.: Fonctions de rpartition n dimensions et leurs marges. Publications de l'Institut de Statistique de L'Université de Paris 8, 229–231 (1959)
20. Nelson, R.B.: An Introduction to Copulas, 2nd edn. Springer, New York (2006)
21. Trivedi, P.K., Zimmer, D.M.: Copula Modeling: An Introduction for Practitioners. Foundations and Trends in Econometrics 1(1), 1–111 (2005)
22. Lee, L.: Generalized econometric models with selectivity. Econometrica 51, 507–512 (1983)
23. Clayton, D.G.: A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. Biometrika 65, 141–151 (1978)
24. Gumbel, E.J.: Distributions des Valeurs Extremes en Plusieurs Dimensions. Publications de l'Institute de Statistque de l'Université de Paris 9, 171–173 (1960)
25. Joe, H.: Parametric Families of Multivariate Distributions with Given Margins. Journal of Multivariate Analysis 46(2), 262–282 (1993)
26. Frank, M.J.: On the simultaneous associativity of $F(x,y)$ and $x+y-F(x,y)$. Aequationes Math. 19, 194–226 (1979)
27. Fisher, M.: Tailoring copula-based multivariate generalized hyperbolic secant distributions to financial return data: An empirical investigation. Discussion papers, University of Erlangen-Nüremberg, Germany (2003), `http://www.statistik.wiso.uni-erlangen.de/forschung/d0047.pdf` (accessed January 25, 2013)
28. Genest, C., et al.: A Semiparametric Estimation Procedure of Dependence Parameters in Multivariate Families of Distributions. Biometrika 82, 543–552 (1995)
29. Akaike, H.: Information Theory and an Extension of the Maximum Likelihood Principle. In: Petrov, B.N., Csaki, F. (eds.) Proceedings of the Second International Symposium on Information Theory, pp. 267–281. Akademiai Kiado, Budapest (1973)
30. Schwarz, G.: Estimating the Dimension of a Model. The Annals of Statistics 6(2), 461–464 (1978)
31. Vuong, Q.H.: Ratio Tests for Model Selection and Non-Nested Hypotheses. Econometrica 57(2), 307–333 (1989)
32. Clarke, K.A.: A Simple Distribution-Free Test for Nonnested Model Selection. Political Analysis 15(3), 347–363 (2007)
33. Genest, C., Rivest, L.P.: Statistical Inference Procedures for Bivariate Archimedean Copulas. Journal of the American Statistical Association 88(423), 1034–1043 (1993)
34. Wang, W., Wells, M.T.: Model selection and semiparametric inference for bivariate failure-time data. Journal of the American Statistical Association 95(449), 62–72 (2000)

Modeling Dependency in Tourist Arrivals to Thailand from China, Korea, and Japan Using Vine Copulas

Ornanong Puarattanaarunkorn and Songsak Sriboonchitta

Abstract. Market interdependence has always been an interesting topic in the study of tourism demand. China, Japan, and Korea are important tourist markets for Thailand tourism. Understanding how the arrivals relate to each other can help in tourism management, in a way that it prepares the tourism industry to plan for the risk management of the tourism demand and tourism supply. The vine copula model was used to analyze the multiple dependencies by decomposing the diversity of the pair-copulas which can be arranged and analyzed in a tree structure. For this study, both the C-vine copula and the D-vine copula were used to answer the research question. We give the same conditioning variable for both the C-vine and the D-vine copula models in order to find the answer to our question of whether these two models would give different results. The contributions of the study are obtained from the findings. The C-vine and D-vine copulas provided three pair-copulas, namely, China–Korea, China–Japan, and Korea–Japan given China and there exists a weak positive dependence in each pair. In addition, the results provide evidence that China has influence on the dependence between the tourist arrivals from Korea and Japan. Moreover, the three dimensions of the C-vine and D-vine copula models, which are given the same conditioning variable in the second tree, optimally provide the same estimates of the parameters of interest.

1 Introduction

International tourist arrivals to Thailand have been growing over the past decades: 10 million in 2001 to 20 million in Nov 2012, which is a 6% yearly growth rate. Thailand's international tourist market is diverse. This paper focuses on the short-haul tourist market from the Asia region. We are looking in particular at China, Japan, and Korea, the important tourist markets which rank 2^{nd} to 4^{th} in the market

Ornanong Puarattanaarunkorn · Songsak Sriboonchitta
Faculty of Economics, Chiang Mai University, Chiang Mai 50200 Thailand
e-mail: pornan@kku.ac.th, songsakecon@gmail.com

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_24,

share and are also three of the nations from the East Asia region which are in co-operation with ASEAN (ASEAN+3). The recent growth rates in the arrivals from these three countries have been on the rise. China was the highest and grew by 39% on a yearly average during 2010–2012 [1]. Moreover, China is a nation that has a fast-growing economy and has an impact on the global economy [2]. However, the management of tourism demand remains a challenge due to continuing social and economic uncertainty such as global economic recession, climate change that influences travelers' behavior [3], terrorism, and natural disasters. These negative shocks and seasonal effects [4] have an effect on the volatility in tourism demand. Although the shock effects are not permanent, Lean and Smyth [5] found that the negative shocks made the growth of tourist arrival slow down. Such occurrences and consequences can have some adverse effects on businesses, employment, and economic growth. For example, in 2003, 2005, and 2009, the growth rate of international tourist arrivals to Thailand declined sharply because of the outbreak of SARS, tsunami, economic recession, and political disturbance. Volatility in the international tourist arrivals can have an effect on the management decision making in the tourism industry. Because of the above-mentioned reasons, it is important to examine the following questions: How does the volatility of the tourist arrivals to Thailand from China, Korea, and Japan change over time? What is the nature of the dependence between the growth rates of tourist arrivals from these three countries? How do the tourist arrivals from China influence the dependence between the tourist arrivals from Korea and Japan? Understanding the dependence of their arrivals can help in tourism management, in that it will go a long way in planning for the risk management of the tourism demand and the tourism supply, for example, in formulating the promotion marketing strategies and in the decision-making of the budget allocation. In addition, there has been an increasing number of studies on the joint distribution of all the risk sources. Schirmacher and Schirmacher [6] stated that we should understand "how all risk sources relate to each other because their potential synergy can create catastrophic losses". Along with the risk management of international tourism demand, in this study, we take into consideration how the tourist arrivals to Thailand from these three countries, China, Korea, and Japan, are related to each other.

To answer the research questions that we mentioned above, we used the ARMA-GARCH model to examine the volatility of the growth rates of tourist arrivals from these three origin countries and the vine copula model to analyze the dependence between the three marginal distributions. Allen et al. [7] point out that the vine copula model is a method offering greater flexibility, which allows for the modeling of complex dependency patterns. Since vine copulas allow us to analyze multiple dependencies by decomposing the diversity of pair-copulas, which can then be arranged and analyzed in a tree structure. For this study, the C-vine copula and the D-vine copula were both used with a view to answering the question of whether these two models give different results. We gave the same conditioning variable for both the C-vine and the D-vine copula models.

The copula model is a popular tool in the financial field because this approach provided more flexibility for finding the joint distributions and the transformation of

the invariant correlation, without having to assume linear correlation [8], imposed by the conventional approach, namely, the Pearson correlation coefficient. The standard references of the copula theory were presented in Joe [9] and Nelson [10].

Vine copula modeling was introduced by Joe [11] and extended by Bedford and Cooke [12, 13]. It is a graphical model used for describing multivariate copulas, a graph composed of many bivariate copulas, and so it can be referred to as the pair-copula constructions (PCCs). In recent years, there has been an increase in the volume of literature on vine copula application, particularly, in the financial field. For example, Aas et al. [14] used the D-vine copula to examine the dependence of four variables of the financial data. Schirmacher and Schirmacher [6] used the canonical vine, or the C-vine, copula to model the dependence of three currency exchange rates. Allen et al. [7] used regular vine copula to analyze the dependence between stock indices. Zimmer [15] used both the C-vine and the D-vine copulas to analyze co-movement in housing prices. In the tourism field, a few pieces of literature are available on applying copula to model the dependence structure between variables. Zhang et al. [16] used a fully nested Archimedean copula function to find the dependence between three dependent variables, destination visit behavior, time use behavior, and expenditure behavior. Liu and Sriboonchitta [17] used a copula based GARCH to model the volatility and the dependence structure between tourist arrivals from China to two destination markets, Thailand and Singapore.

In this study, we applied the vine copula to examine the dependence of the growth rates of tourist arrivals to Thailand from the three major origin countries in Asia, namely, China, Korea, and Japan, by using the C-vine and D-vine models in the R-package CDVines. Moreover, we used the time-varying copula to model the dependence of the pair-copula of each of these three countries. The contributions of the study are obtained from the findings that can lead to policy recommendations in terms of risk management of the tourism demand for Thailand. The paper is divided into six parts. The next part is about the methodology used, which describes in detail the GARCH model and the vine copula model. The third part presents the data used. The fourth part shows the results of this study. The fifth part presents the policy implications that can help the risk management of the tourism demand for Thailand. The last part gives the conclusion remarks.

2 Methodology

The author of this study applied the marginal ARMA-GARCH models and the vine copula to answer the research questions in this paper. The GARCH model [18] has been widely used for modeling volatility. Volatility is a significant factor as it is considered a measure of risk. Therefore, we applied the ARMA-GARCH model to estimate the marginal distributions since this model can capture the volatility of international tourism demand, as measured by the number of international tourist arrivals. The standardized residuals from ARMA-GARCH model were transformed to copula data $(F_1(x_1), F_2(x_2), F_3(x_3))$. After that, the vine copula approach was used to measure the dependence of the marginal distributions. This study used the

R-package CDVine by Brechmann and Schepsmeier [19] to analyze the constant copula since it provided a range of tools for the bivariate data analysis. As for the time-varying copula, we followed the path laid by Patton [20].

2.1 ARMA-GARCH Model

We adopt the ARMA(1,0)-GARCH(1,1) with the skewed student T (SkT) distribution residual for the marginal distribution of the logarithm of the monthly growth rate of tourist arrivals to Thailand from China, Korea, and Japan (y_t) as

$$y_t = a_0 + a_1 y_{t-1} + \varepsilon_t \tag{1}$$

$$\varepsilon_t = z_t \sqrt{h_t}, z_t \sim SkT(\nu, \gamma) \tag{2}$$

$$h_t = \omega_t + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1} \tag{3}$$

In equation (1) presents ARMA(1,0) process where y_{t-1} is an autoregressive term of y_t and ε_t is an error term. Equation (2) then define this error term as the product between conditional variance h_t and a residual z_t. A residual z_t is assumed to follow the skewed student T (SkT) distribution with the degree of freedom parameter ν and the skewness parameter γ. Equation (3) presents GARCH(1,1) process where $\omega > 0$, $\alpha \geq 0$, $\beta \geq 0$ are sufficient to ensure that the conditional variance $h_t > 0$. The $\alpha \varepsilon_{t-1}^2$ represent the ARCH term and α refers to the short run persistence of shocks, while βh_{t-1} represent the GARCH term and β refers to the contribution of shocks to long run persistence $(\alpha + \beta)$. The property of the GARCH(1,1) model is that it requires the conditional variance, h_t, of the error term, ε_t, to be stationary and persistent. This paper used the second moment condition that was presented in the Bollerslev study [18] and the log moment condition that was presented by Nelson [21] and Lee and Hansen [22] to check these properties.

$$\text{The second moment condition} : \alpha + \beta < 1 \tag{4}$$

$$\text{The log moment condition: } \mathrm{E}[\ln(\alpha z_t^2 + \beta)] < 0 \tag{5}$$

2.2 Multivariate Copula

One approach of modeling the multivariate dependence is the copula. The copula functions can offer us to merge univariate distributions to get a joint distribution with an appropriate dependence structure. The fundamental theorem of copula was given by Sklar [23] as Sklar's theorem. The standard reference book of the copula theory was made by Nelson [10].

Let F be an n-dimensional distribution function with marginal distributions $F_1,...,F_n$. Then there exists a copula C for all $x = (x_1,...,x_n)' \in [-\infty,\infty]^n$,

$$F(x) = C(F_1(x_1),...F_n(x_n)) \tag{6}$$

If $F_1,...,F_n$ are continuous, then C is unique. Conversely, if C is a copula and $F_1,...,F_n$ are distribution functions, then the above function $F(x)$ in (6) is a joint distribution function with the marginal distribution $F_1,...,F_n$. C can be interpreted as the distribution function of an n dimensional random variable on $[0,1]^n$ with uniform margins [19].

2.3 Vine Copulas

Modeling copulas with high dimension is a difficult task because there are large numbers of variables. Vine copulas can cross over this restriction, vine copulas are a flexible tool for describing the multivariate copulas through the graphical model. The multivariate copulas are constructed from a cascade of bivariate copulas or are called pair-copulas. The principles of vine copulas propounded by Joe [11] and extended by Bedford and Cooke [12, 13]. For statistical inference techniques of two classes of C-vines and D-vines are described by Aas et al. [14].

In this study, we selected the order of the variables by choosing China (C) as the first root node of the C-vine copula model, with Korea (K) and Japan (J) was linked to it, as shown in Figure 1 (left). Brechmann and Schepsmeier [19] stated that a vine structure can be chosen manually or through expert knowledge, or be given by the data itself [19]. Aas et al. [14] pointed out that modeling C-vine might be advantageous when we know a main variable that governs the interactions. When we consider the recent growth rate of the tourist arrivals to Thailand during 2010–2012, China contributed the highest and had a rapid growth rate of 39% yearly on an average in this period [1]. China is a fast-growing economy, and the nation does have an impact on the global economy [2]. Therefore, we assumed and chose the tourist arrivals from China as the main variable of the C-vine copula model. Similarly, we also fitted the D-vine copula model with the order of the variables by giving China as the conditioning variable in order to find the answer to our question of whether these two models would give different results if we gave the same conditioning variable for both the C-vine and the D-vine copula models.

In Figure 1, we presented the three dimensions of C-vine and D-vine copulas, which were what we used in this paper. Let $X = (X_1,X_2,X_3) \sim F$ with marginal distribution functions F_1,F_2,F_3 and their density functions f_1,f_2,f_3, which was proposed as follows (see Aas et al. [14])

The density function of C-vine copula

$$\begin{aligned} f(x_1,x_2,x_3) = {} & f(x_1)\cdot f(x_2)\cdot f(x_3)\cdot c_{1,2}(F_1(x_1),F_2(x_2))\cdot c_{1,3}(F_1(x_1),F_3(x_3)) \\ & \cdot c_{2,3|1}(F_{2|1}(x_2 \mid x_1),F_{3|1}(x_3 \mid x_1)) \end{aligned} \tag{7}$$

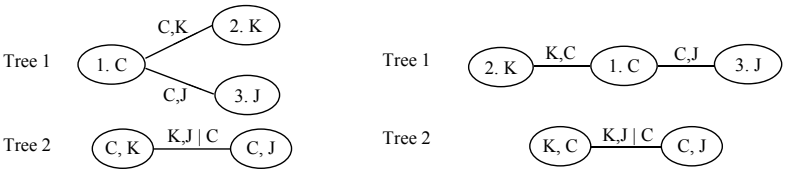


Fig. 1 The structures of the C-vine (left) and the D-vine (right) copulas of the tourist arrivals to Thailand from China (C), Korea (K), and Japan (J)

where $c_{1,2}, c_{1,3}$, and $c_{2,3|1}$ denote the densities of bivariate copulas $C_{1,2}, C_{1,3}$, and $C_{2,3|1}$, respectively. $F_{2|1}$ and $F_{3|1}$ are the marginal conditional distributions that can be derived from formula (9).

We also fitted the D-vine copula model with the order of the variables by giving China as the conditioning variable.

The density function of D-vine copula

$$\begin{aligned} f(x_2,x_1,x_3) = f(x_2)\cdot f(x_1)\cdot f(x_3)\cdot c_{2,1}(F_2(x_2),F_1(x_1))\cdot c_{1,3}(F_1(x_1),F_3(x_3)) \\ \cdot c_{2,3|1}(F_{2|1}(x_2 \mid x_1),F_{3|1}(x_3 \mid x_1)) \end{aligned} \tag{8}$$

where $c_{2,1}, c_{1,3}$, and $c_{2,3|1}$ denote the densities of bivariate copulas $C_{2,1}, C_{1,3}$, and $C_{2,3|1}$, respectively. $F_{2|1}$ and $F_{3|1}$ are the marginal conditional distributions that can be derived from formula (9).

The vine copulas involve marginal conditional distributions. The general form of a conditional distribution function is $F(x \mid v)$,

$$F(x \mid v) = \frac{\partial C_{x,v_j|v_{-j}}(F(x \mid v_{-j}),F(v_j \mid v_{-j}))}{\partial F(v_j \mid v_{-j})} \tag{9}$$

where v denotes all the conditional variables and $C_{x,v_j|v_{-j}}$ is a bivariate copula distribution function. For v is univariate, the marginal condition distribution, e.g. $F_{2|1}$ can be presented as

$$F_{2|1}(x_2 \mid x_1) = \frac{\partial C_{21}(F_2(x_2),F_1(x_1))}{\partial F_1(x_1)} \tag{10}$$

2.4 Vine Copula Estimation

In the R-package CDVine, the maximum likelihood was used to estimate the parameters of copulas. The log-likelihood of C-vine and D-vine copula with three dimensions in (7) and (8) can be written as

The log-likelihood of the C-vine copula is

$$\sum_{t=1}^{T} \log[c_{1,2}(F_1(x_{1,t}),F_2(x_{2,t}))\cdot c_{1,3}(F_1(x_{1,t}),F_3(x_{3,t}))\cdot c_{2,3|1}(F_{2|1}(x_{2,t}\mid x_{1,t}),F_{3|1}(x_{3,t}\mid x_{1,t}))]. \tag{11}$$

The log-likelihood of the D-vine copula is

$$\sum_{t=1}^{T} \log[c_{2,1}(F_2(x_{2,t}),F_1(x_{1,t}))\cdot c_{1,3}(F_1(x_{1,t}),F_3(x_{3,t}))\cdot c_{2,3|1}(F_{2|1}(x_{2,t}\mid x_{1,t}),F_{3|1}(x_{3,t}\mid x_{1,t}))]. \tag{12}$$

2.5 Copula Families

The R-package CDVine provide the various copula families to measure the values of dependence of the pair-copulas. The characteristics of the copula families that were used in this paper are shown in Table 1 and Table 2.

Table 1 Characteristics of Copula Families

Name	Pair-copula function	Parameter range
Gaussian	$C(u_1,u_2;\rho)=\Phi_G(\Phi^{-1}(u_1),\Phi^{-1}(u_2);\rho)$ $=\int_{-\infty}^{\phi^{-1}(u_1)}\int_{-\infty}^{\phi^{-1}(u_2)}\frac{1}{2\Pi\sqrt{(1-\rho^2)}}\times[\frac{-(s^2-2\rho st+t^2)}{2(1-\rho^2)}]dsdt$	$\rho\in(-1,1)$
Student's T	$C^T(u_1,u_2;\rho,\nu)=\int_{-\infty}^{T_\nu^{-1}(u_1)}\int_{-\infty}^{T_\nu^{-1}(u_2)}\frac{1}{2\Pi\sqrt{(1-\rho^2)}}\times$ $[1+\frac{(s^2-2\rho sT+T^2)}{\nu(1-\rho^2)}]^{-(\frac{\nu+2}{2})}dsdT$	$\rho\in(-1,1)$, $\nu>2$
Clayton	$C(u_1,u_2;\theta)=(u_1^{-\theta}+u_2^{-\theta}-1)^{-1/\theta}$	$\theta\in(0,\infty)$
Gumbel	$C(u_1,u_2;\theta)=exp(-[(-\ln(u_1))^\theta+(-\ln(u_2))^\theta]^{\frac{1}{\theta}})$	$\theta\in[1,\infty)$
Frank	$C(u_1,u_2;\theta)=-\frac{1}{\theta}\log(1+\frac{(e^{-\theta u_1}-1)(e^{-\theta u_2}-1)}{e^{-\theta}-1})$	$\theta\in(-\infty,\infty)\backslash\{0\}$
Joe	$C(u_1,u_2;\theta)=1-[(1-u_1)^\theta+(1-u_2)^\theta-(1-u_1)^\theta(1-u_2)^\theta]^{\frac{1}{\theta}}$	$\theta\in[1,\infty)$
Rotated Clayton 180°	$C(u_1,u_2;\theta)=u_1+u_2-1+[(1-u_1)^{-\theta}+(1-u_2)^{-\theta}-1]^{-\frac{1}{\theta}}$	$\theta\in(0,\infty)$
Rotated Gumbel 180°	$C(u_1,u_2;\theta)=u_1+u_2-1+exp(-[(-\ln(1-u_1))^\theta+$ $(-\ln(1-u_2))^\theta]^{\frac{1}{\theta}})$	$\theta\in[1,\infty)$
Rotated Joe 180°	$C(u_1,u_2;\theta)=u_1+u_2-(u_1^\theta+u_2^\theta-u_1^\theta u_2^\theta)^{\frac{1}{\theta}}$	$\theta\in[1,\infty)$

Source: The copula functions are given as presented in Trivedi and Zimmer [24], Nelson [10], and Fisher [25].

Table 2 Function of Kendall's tau and Tail Dependence for Bivariate Copula

Copula family	Kendall's tau	Tail dependence (lower, upper)
Gaussian	$\frac{2}{\Pi}\arcsin\rho$	0
Student's T	$\frac{2}{\Pi}\arcsin\rho$	$T^L = T^U = 2T_{v+1}(-\sqrt{v+1}\sqrt{\frac{1-\rho}{1+\rho}})$
Clayton	$\frac{\theta}{\theta+2}$	$(2^{-1/\theta}, 0)$
Gumbel	$1-\frac{1}{\theta}$	$(0, 2-2^{1/\theta})$
Frank	$1-\frac{4}{\theta}+4\frac{D_1(\theta)}{\theta}$	(0,0)
Joe	$1+\frac{4}{\theta^2}\int_0^1 t\log(t)(1-t)^{2(1-\theta)/\theta}dt$	$(0, 2-2^{1/\theta})$
Rotated Clayton 180°	$\frac{\theta}{\theta+2}$	$(0, 2^{-1/\theta})$
Rotated Gumbel 180°	$1-\frac{1}{\theta}$	$(2-2^{1/\theta}, 0)$
Rotate Joe 180°	$1+\frac{4}{\theta^2}\int_0^1 t\log(t)(1-t)^{2(1-\theta)/\theta}dt$	$(2-2^{1/\theta}, 0)$

Source: Kendall's tau and tail dependence are as presented in Brechmann and Schepsmeier [19]. Note: $D_1(\theta) = \int_0^\theta \frac{c/\theta}{exp(x)-1}$ is the Debye function.

3 Data

3.1 Descriptive Statistics

The seasonal adjusted[1] data of the tourist arrivals to Thailand from China, Korea, and Japan, which measures the tourism demands of these three countries, were used. The monthly data of the three countries, taken from the Ecowin database, during the period from January 1997 to November 2012, are taken into the logarithm of the monthly arrival rate, the growth rate of the tourist arrivals = $\ln \frac{\text{Number arrivals}_t}{\text{Number arrivals}_{t-1}}$.

Table 3 presents the descriptive statistics of the growth rate of the tourist arrivals to Thailand from China, Korea, and Japan. All the countries have positive average growth rates during the period from January 1997 to November 2012. Negative skewness and excess kurtosis are exhibited in the data of the three countries. This means that all the three data series have peakedness of distribution and heaviness of tail. The null hypothesis of normality of the Jarque-Bera tests are rejected in all the data series. Figure 2 presents the growth rate of the tourist arrivals to Thailand from China, Korea, and Japan along this period. It can be seen that the growth rates of all the three countries have considerable fluctuation and co-movement.

3.2 Marginal Distributions by ARMA-GARCH Model

To test whether the data are stationary or not, we use the Augmented Dickey-Fuller test (ADF). The results show that all the data series are stationary at p-value 0.01. For examining the volatility of the growth rates of the tourist arrivals to Thailand from China, Korea, and Japan, the ARMA(1,0)-GARCH(1,1) model with the

[1] The X12-ARIMA monthly seasonal adjustment method by the U.S. Department of Commerce in the Eviews7 program was used.

Table 3 Data Descriptive Statistics for Growth Rate of Tourist Arrivals from Three Countries

	China	Korea	Japan
Mean	0.011	0.005	0.003
Median	0.014	0.004	0.011
Maximum	0.739	0.826	0.415
Minimum	−1.099	−1.139	−0.366
Std. Dev.	0.281	0.196	0.093
Skewness	−0.903	−1.164	−0.501
Kurtosis	6.092	11.529	7.513
Jarque-Bera	101.538	618.792	169.171
(p-value)	(0.000)	(0.000)	(0.000)
Number of observations	190	190	190

Note: The null hypothesis of Jarque-Bera = data is taken as the normal distribution.

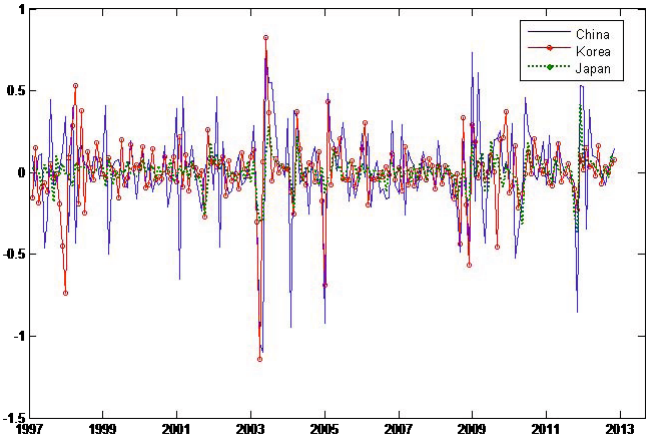


Fig. 2 The growth rate of the tourist arrivals to Thailand from China, Korea, and Japan

skewed student T residual $\varepsilon_t \sim SkT\ (\nu, \gamma)$, was modeled. The choice of the skewed student T distribution was based on the two data series exhibiting negative skewness and excess kurtosis. The identifying of the optimal models was based on the Akaike information criterion (AIC). In Table 4, it can be seen that all the parameters of the model have significance at levels 0.01 and 0.05, except the parameter α of Korea which has significance at level 0.1 and the parameters ar1 and β of China which have insignificance at level 0.1. These findings imply that the volatility of the growth rates of the tourist arrivals from China has a short run persistence (α) and no long run persistent since the parameter β is insignificance. The volatility of the growth rates of the tourist arrivals from both Korea and Japan has a weak long run persistence, although there are values of the second moment which are higher than one.

But the values of the log moment are less than zero; this is necessary and sufficient for the strict stationarity, and the persistence of the conditional variance is satisfied. The degree of freedom and the skewness parameters also have significance at level 0.01. This indicates that all the data series are skewed student T distributions.

Then we transformed the standardized residuals from these three models into uniform [0,1] by using the empirical distribution function $F_n(x) = \frac{1}{n+1}\Sigma_{i=1}^{n} 1(X_i \leq x)$ where $X_i \leq x$ is the order statistics and 1 is the indicator function. For checking whether the marginal distributions that we transformed were correctly specified, we checked them for being i.i.d. uniform [0,1] and confirmed that they were. We used the Kolmogorov-Smirnov (K-S) test for uniform [0,1] and the Box-Ljung test for serial correlation. It appeared that all of these marginal distributions were uniform because they accepted the null hypothesis at p-values equal to 1. The results of the Box-Ljung test demonstrated that all of the four moments of all the marginal distributions were i.i.d. by accepting the null hypothesis that no serial correlation at p-value is greater than 0.05. Therefore, our marginal distributions were not misspecified and can be used for the copula model.

Table 4 Results of ARMA (1,0)-GARCH(1,1) with Skewed Student T Residual for Growth Rates of Tourist Arrivals to Thailand from China , Korea, and Japan

	China	Std. error (p-value)	Korea	Std. error (p-value)	Japan	Std. error (p-value)
ar1	−0.092	0.087 (0.292)	−0.219	0.078 (0.005)	−0.280	0.070 (6.33e-05)
ω	0.031	0.012 (0.011)	0.011	0.006 (0.044)	0.002	0.001 (0.029)
α	0.836	0.411 (0.042)	0.869	0.457 (0.057)	0.768	0.379 (0.043)
β	0.114	0.129 (0.377)	0.211	0.106 (0.045)	0.294	0.099 (0.003)
ν (degree of freedom)	3.249	0.873 (0.000)	2.980	0.808 (0.000)	0.658	0.062 (<2e-16)
γ (skewness)	0.870	0.076 (<2e-16)	0.842	0.069 (<2e-16)	3.050	70.796 (0.000)
Log likelihood	15.699	-	104.044	-	241.188	-
AIC	−19.398	-	−194.088	-	−470.375	-
2^{nd} moment	0.950	-	1.080	-	1.062	-
Log moment	−0.978	-	−0.708	-	−0.530	-

4 Empirical Results

4.1 Results of C-vine and D-vine Copula Analysis

Table 5 and Figure 3 present the results of the pair-copula decomposition analysis for the C-vine copula model. The first tree consists of two pair-copulas. The first

pair is China–Korea (C,K), which has a weak positive dependence, and the rotated Joe 180° copula which can capture lower tail dependence was adjudged the best fit by using the AIC and BIC tests, and a scoring goodness-of-fit test based on the Vuong and Clarke tests and a goodness-of-fit test based on Kendall's process as the criteria. This pair-copula provides a copula parameter of 1.287 and a Kendall's tau correlation of 0.14. As for the second pair, which is China–Japan (C,J), it has a weak positive dependence, and the rotated Gumbel 180° copula which can capture lower tail dependence was adjudged the best fit, with a copula parameter of 1.265 and a Kendall's tau correlation of 0.21. Because the different families of the copulas have different ranges of copula parameters, we inverse the copula parameter into a Kendall's tau correlation, and it is bounded on the interval $[-1, 1]$. Kendall's tau is a measure of concordance which is a function of copula; thus, we can use it to assess the range of dependence covered by the families of the copula. A comparison using Kendall's tau correlation indicates that the pair-copula of China–Japan has a stronger correlation than the pair-copula of China–Korea.

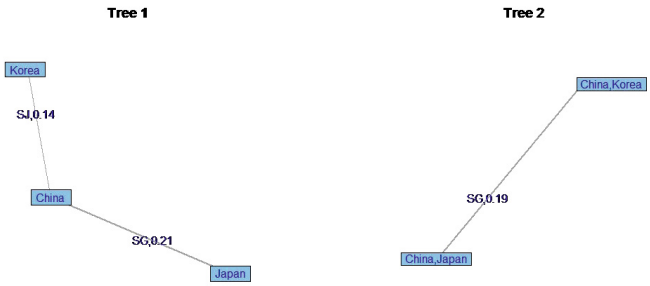


Fig. 3 The C-vine trees, tree 1 and tree 2, with the chosen pair-copula family and the Kendall's tau correlation

The second tree consists of the conditional pair-copula, Korea–Japan given China (K,J|C). There exists a weak positive dependence, and the rotated Gumbel 180° copula that can capture lower tail dependence was adjudged the best fit, with a copula parameter of 1.232 and a Kendall's tau correlation of 0.19. This Kendall's tau correlation is less than which was obtained from the bivariate pair-copula of Korea–Japan (0.24), which was obtained from a bivariate data analysis. This implies that China has an influence on the dependence between the tourist arrivals from Korea and Japan.

The results show that all of the estimated copula parameters from all the pairs have significance at p-value less than 0.01 and that all the pair-copulas provide evidence of lower tail dependence. This implies that the growth rates of the tourist arrivals in each of the pairs China–Korea, China–Japan, and Korea–Japan given China have co-movement that is both upward and downward, but with weak dependence. For example, for the first pair-copula in tree 1, the rise or loss of tourism

demands from China (Korea) is slightly correlated by the rise or loss of tourism demands from Korea (China); similar is the case with the second pair in tree 1. For the conditional pair-copula in tree 2, with the tourism demand from China as the conditional variable, the rise or loss of tourism demands from Korea (Japan) is slightly correlated by the rise or loss of tourism demands from Japan (Korea).

Table 6 presents the results of the pair-copula analysis for the D-vine copula model. When comparing the C-vine and the D-vine, the results show that both the models provide the same results for each of the pair-copulas, such as the appropriate copula family, the copula parameter, a Kendall's tau correlation value, an AIC value, and a BIC value. Moreover, we use the Voung test that is obtained in the R-package CDVines to compare both the models, as shown in Table 7. The test statistics are close to zero and the large p-values indicate that the three dimensions of the C-vine and D-vine copula models have to be given the same conditioning variable in the second tree; otherwise, it cannot be distinguished statistically in this study.

Table 5 Maximum Likelihood Parameter Estimates for C-vine Copula

Tree	Pair-copula	Copula family	Copula parameter (p-value)	Kendall's tau	AIC	BIC
1	C,K	rotated Joe 180°	$\theta = 1.287$ (0.000)	0.14	−18.617	−15.370
	C,J	rotated Gumbel 180°	$\theta = 1.265$ (0.000)	0.21	−21.848	−18.601
2	K,J|C	rotated Gumbel 180°	$\theta = 1.232$ (0.000)	0.19	−18.367	−15.122
			AIC and BIC of a model		−58.833	−49.093

Table 6 Maximum Likelihood Parameter Estimates for D-vine Copula

Tree	Pair-copula	Copula family	Copula parameter (p-value)	Kendall's tau	AIC	BIC
1	K,C	rotated Joe 180°	$\theta = 1.287$ (0.000)	0.14	−18.617	−15.370
	C,J	rotated Gumbel 180°	$\theta = 1.265$ (0.000)	0.21	−21.848	−18.601
2	K,J|C	rotated Gumbel 180°	$\theta = 1.232$ (0.000)	0.19	−18.367	−15.122
			AIC and BIC of a model		−58.833	−49.093

4.2 Time-Varying Copula

The dependence structures vary over time. Hence, we also used time-varying copula models to show the co-movement of each pair-copula during this period. We used

Table 7 Comparison of C-vine and D-vine Models by Voung Test

	Voung	Akaike	Schwarz
Statistic	0.023	0.023	0.023
p-value	0.982	0.982	0.982

Note: The null hypothesis = two model equivalent.

the time-varying copula models as given in Patton [20], such as the time-varying Gaussian copula, time-varying Gumbel copula, and time-varying rotated Gumbel 180° copula. Furthermore, we added the time-varying Joe copula and time-varying rotated Joe 180° copula. The smallest AIC and BIC values were used to select an appropriate copula family. The parameter β represents the degree of persistence in the dependences and the parameter α measures the variations over time in the dependences.

In tree 1, the pair-copula of China–Korea, the time-varying rotated Joe 180° copula is the best fit with the three copula parameters $\omega = 2.087$, $\beta = -0.808$, and $\alpha = -1.819$. Also, all the copula parameters have significance at level 0.01. For the pair-copula of China–Japan, the time-varying Gaussian is the best fit with the copula parameters $\omega = 0.872$, $\beta = -2.202$, and $\alpha = 1.201$. Also, all the copula parameters have significance at the same level 0.01.

In tree 2, the conditional pair-copula of Korea–Japan given China, the time-varying Gaussian is the best fit with the copula parameters $\omega = 0.800$, $\beta = 0.449$, and $\alpha = -1.042$. Also, all the copula parameters have significance at level 0.01.

The parameter α of all the pair-copulas have significance, indicating that the dependence between all the pairs of the tourist arrivals to Thailand, that is, between China–Korea, China–Japan and Korea–Japan given China keep varying over time. Moreover, the pair-copula China–Japan shows that it has the highest degree of persistence in the dependence, as indicated by the parameter β.

5 Policy Implications

Our results show that there exists a weak positive dependence between all the pairs in the growth rates of the tourist arrivals to Thailand from China, Korea, and Japan and that these values of dependence keep varying over time. This means that all of the pairs China–Korea, China–Japan, and Korea–Japan given China have co-movement that is both upward and downward, but with a weak dependence. In other words, the rise or loss of tourism demand — which is measured as tourist arrivals from each origin country — of each tourism origin country is slightly correlated to a rise in or loss of tourism demand of the other. This result has an important implication for Thailand's tourism management. In case there occurs a simultaneous rise or loss of arrivals, it will have considerable impact on the tourism industry and the related businesses. Therefore, policy makers should consider the time-varying dependency while planning the risk management of the tourism demand in each time period. This could be in the form of providing marketing and promotion strategies

to motivate the tourism demand when there is loss of arrivals from the effect of shocks or when it is low season. On the other hand, when there is a rise in tourism demand, the tourism industry can prepare and equip themselves with the appropriate resources, such as hotels, airline schedules, etc.

6 Concluding Remarks

This paper models the dependency in tourist arrivals to Thailand from the three major origin countries of the Asia region, namely, China, Korea, and Japan by using vine copulas. Volatility in international tourist arrival can have an effect on the management and decision making in the tourism industry. Therefore, we used the ARMA-GARCH model to examine the volatility of the growth rates of the tourist arrivals from these three countries during the period from January 1997 to November 2012. Thereafter, the marginal distributions from the ARMA-GARCH model were used to analyze the dependency by using the C-vine and D-vine copula models. Understanding the dependence of the tourists arrivals can help the tourism industry achieve better management, as well as in planning for the risk management of the tourism demand and tourism supply.

The empirical results provided evidence that, first, the three data series of the growth rates of the tourist arrivals from China, Korea, and Japan rejected the null hypothesis of normality and exhibited skewness and excess kurtosis through the use of data descriptive statistical analysis. Second, the ARMA(1,0)-GARCH(1,1) model with the skewed student T residual can examine the volatility of the growth rates of the tourist arrivals of each data series. The volatility of the growth rates of the arrivals from China has short run persistence and the volatility of the growth rates of the tourist arrivals from both Korea and Japan have weak long run persistence. Third, the C-vine and D-vine copula models can measure the dependency of the three marginal distributions. We identified the C-vine and D-vine structures that can decompose the multivariate copulas to many pair-copulas in a tree structure. We had two pair-copulas in tree 1: China–Korea and China–Japan, and one conditional pair-copula in tree 2: Korea–Japan given China. The results show that there exists weak positive dependence in all of the pairs and that all the pair-copulas provide evidence of lower tail dependence. China–Japan has the strongest dependence with a Kendall's tau correlation 0.21, followed by Korea–Japan given China, 0.19, and China–Korea, 0.14. In addition, the conditional pair-copula of Korea–Japan given China provided a Kendall's tau correlation of 0.19, which is less than that obtained from the bivariate pair-copula of Korea–Japan which is 0.24. This implies that China has an influence on the dependence between the tourist arrivals from Korea and Japan. Moreover, the three dimensions of the C-vine and D-vine copula models, which are given the same conditioning variable in the second tree, provided the same results; otherwise, it would not be possible to distinguish between these two models statistically in this study. Fourth, the time-varying copulas show that the dependence parameters of all the pair-copulas had varied over time. Our findings

have important implications and application in the risk management of the tourism demand and tourism supply for Thailand's tourism industry, as discussed above.

References

1. Department of Tourism, Thailand. Tourist Arrivals in Thailand (2012), http://www.tourism.go.th/tourism/th/home/tourism.php (accessed December 20, 2012)
2. World Bank. China 2030: Building a Modern, Harmonious, and Creative High-Income Society [pre-publication version]. World Bank, Washington, DC (2012), https://openknowledge.worldbank.org/handle/10986/6057 (accessed February 21, 2013)
3. Goh, C.: Exploring impact of climate on tourism demand. Annals of Tourism Research 39(4), 1859–1883 (2012)
4. Vanhove, N.: The Economics of Tourism Destinations, 2nd edn. Elsevier, Ltd., London (2011)
5. Lean, H.H., Smyth, R.: Asian Financial Crisis, Avian Flu and TerroristThreats: Are Shocks to Malaysian Tourist Arrivals Permanent or Transitory? Asia Pacific Journal of Tourism Research 14(3), 301–321 (2009)
6. Schirmacher, D., Schirmacher, E.: Multivariate Dependence Modeling using Pair- copulas. Technical report, Society of Acturaries: 2008 Enterprise Risk Management Symposium, Chicago, April 14-16 (2008), http://www.ermsymposium.org/2008/pdf/papers/Schirmacher.pdf (access March 10, 2013)
7. Allen, D.E., et al.: Financial Dependence Analysis: Applications of Vine Copulae. Unpublished Discussion paper, Tinbergen Institute Amsterdam Netherlands (2013), http://papers.tinbergen.nl/13022.pdf (accessed March 10, 2013)
8. Sriboonchitta, S., et al.: Modeling volatility and dependency of agricultural price and production indices of Thailand: Static versus time-varying copulas. International Journal of Approximate Reasoning 54(6), 793–808 (2013)
9. Joe, H.: Multivariate Models and Dependence Concepts. Chapman and Hall, London (1997)
10. Nelson, R.B.: An Introduction to Copulas, 2nd edn. Springer, New York (2006)
11. Joe, H.: Families of m-Variate Distributions with Given Margins and m(m1)/2 Bivariate Dependence Parameters. In: Rüschendorf, L., Schweizer, B., Taylor, M.D. (eds.) Distributions with Fixed Marginals and Related Topics, vol. 28, pp. 120–141 (1996)
12. Bedford, T., Cooke, R.M.: Probability Density Decomposition for Conditionally Dependent Random Variables Modeled by Vines. Annals of Mathematics and Artificial Intelligence 32, 245–268 (2001)
13. Bedford, T., Cooke, R.M.: Vines- A New Graphical Model for Dependent Random Variables. Annals of Statistics 30, 1031–1068 (2002)
14. Aas, K., et al.: Pair-copula constructions of multiple dependence. Insurance: Mathematics and Economics 44, 182–198 (2009)
15. Zimmer, D.M.: Analyzing Comovements in Housing Prices using Vine Copulas. Western Kentucky University (2013), http://people.wku.edu/david.zimmer/index_files/vine.pdf (access March 20, 2013)
16. Zhang, H., et al.: An integrated model of tourists' time use and expenditure behaviour with self-selection based on a fully nested Archimedean copula function. Tourism Management 33, 1562–1573 (2012)

17. Liu, J., Sriboonchitta, S.: Analysis of Volatility and Dependence between the Tourist Arrivals from China to Thailand and Singapore: A Copula-Based GARCH Approach. In: Huynh, V.N., Kreinovich, V., Sriboonchitta, S., Suriya, K. (eds.) Uncertainty Analysis in Econometrics with Applications. AISC, vol. 200, pp. 285–296. Springer, Heidelberg (2013)
18. Bollerslev, T.: Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics 31, 307–327 (1986)
19. Brechmann, E.C., Schepsmeier, U.: Modeling Dependence with C- and D-Vine Copulas: The R Package CDVine. Journal of Statistical Software 52(3), 1–27 (2013)
20. Patton, A.J.: Modelling Asymmetric Exchange Rate Dependence. International Economic Review 47(2), 527–556 (2006)
21. Nelson, D.B.: Stationarity and persistence in the GARCH(1,1) model. Econometric Theory 6, 318–334 (1990)
22. Lee, S.W., Hansen, B.E.: Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator. Econometric Theory 10, 29–52 (1994)
23. Sklar, A.: Fonctions de répartition á n dimensions et leurs marges. Publications de l'Institut de Statistique de L'Université de Paris 8, 229–231 (1959)
24. Trivedi, P.K., Zimmer, D.M.: Copula Modeling: An Introduction for Practitioners. Foundations and Trends in Econometrics 1(1), 1–111 (2005)
25. Fisher, M.: Tailoring copula-based multivariate generalized hyperbolic secant distributions to financial return data: An empirical investigation. Discussion papers, University of Erlangen-Nüremberg, Germany (2003), `http://www.statistik.wiso.uni-erlangen.de/forschung/d0047.pdf` (accessed January 25, 2013)

Relationship between Exchange Rates, Palm Oil Prices, and Crude Oil Prices: A Vine Copula Based GARCH Approach

Teera Kiatmanaroch and Songsak Sriboonchitta

Abstract. The dollar is the leading international currency, and it is used widely in the majority of international financial transactions. The various food products that comprise agricultural commodities, as also crude oil, have been using the dollar exchange rate for international trade. Over the past several years, the changes in the dollar exchange rate have shown more volatility in addition to a depreciation trend, which has had an influence on the prices of those commodities. We analyzed the relationship between the dollar exchange rates and the prices of two commodities, palm oil and crude oil, by using the GARCH(1,1) model to examine the volatility of the exchange rates and the future prices 1-Pos. of the prices of both the commodities. The vine copula model is used to analyze the dependence structure between their marginal distributions. The data analyses were based on the daily observations from June 2007 to March 2013. The empirical results of GARCH(1,1) show that the exchange rates, palm oil prices, and crude oil prices have a long-run persistence in volatility. The C-vine copula model reveals that there exists a weak negative dependence for each pair-copula, that is, Exchange rate–Palm oil (E,P) and Exchange rate–Crude oil (E,C) in tree 1. Also, a conditional pair-copula of Palm oil–Crude oil given Exchange rate (P,C|E) in tree 2 offers a weak positive dependence. Moreover, the findings of this study provide evidence that the exchange rate (E) is an important variable that governs the interactions in the dependence structure between palm oil price (P) and crude oil price (C).

1 Introduction

At present, there are continuous changes in food prices due to the complex interactions between the several factors. The demand-side factors include population growth, and rising food consumption of emerging economies. The supply-side

Teera Kiatmanaroch · Songsak Sriboonchitta
Faculty of Economics, Chiang Mai University, Chiang Mai 50200 Thailand
e-mail: pornan@kku.ac.th, songsakecon@gmail.com

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_25,

factors include the simultaneous use of food grains to produce biofuel, in which gives rise to a yield of low crops. The others are cyclical factors, such as the depreciation of the U.S. dollar, speculative activities, rising oil prices, input costs in food production, and trading policies of nations [1]. Therefore, it is difficult to take all the factors into account to estimate the percentage of food price changes. To elaborate on the increase of food prices, there exist pieces of literature that have a mention of energy prices and exchange rates which have an effect on food or agricultural commodity prices. There has also been a wide study on energy prices that are related to many types of agricultural commodities, from which it can be concluded that the long-run agricultural prices are driven by the energy prices and that the volatility in the energy markets is transmitted to the food markets [2]. A study on the exchange rate's effect was conducted by Abbott et al. [3] who reviewed and analyzed the twenty five studies and arrived at the conclusion that there were three board factors that drive up the food price. The first is the global changes in production and consumption of key commodity goods. The second is the changing rate of the U.S. dollar. The final factor is the increase in the production of biofuels. The other studies have been those pertaining to econometric modeling, which is based on utilization, to explain the relationship between the exchange rate, energy prices, agricultural commodity prices, and other variables. Their empirical results showed that the depreciation of the U.S. dollar can have an influence on the energy price and/or commodity price [4, 5, 6, 7, 8, 9, 10, 11]. Moreover, Anzuini et al. [12] also found that the expansionary U.S. monetary policy was the cause of the increase in the crude oil price, food price, and other components of the broad commodity price index.

The rise in food and energy prices have caused a burden on the people who are poor and near-poor in the ASEAN region as well as created a negative impact on social and economic development [13]. Under these circumstances, these are major challenges for all the ASEAN members, and they have to find any crucial means to cooperate in the short- and long-term situations in these solving problems because food[1] and energy[2] security are fundamental for upholding the ASEAN economic and social development goals [16]. It also well known that the dollar is a leading currency that is widely used in international financial transactions. The dollar is also used in the international trade of food, agricultural commodities, and crude oil; thus, it is clear that it has been used constantly in the market. Over the past several years, it has been found that the changes in the dollar exchange rate have more

[1] FAO [14] definition: Food security exists when all people, at all times, have physical, social, and economic access to sufficient, safe, and nutritious food to meet their dietary needs and food preferences for an active and healthy life. The four pillars of food security are availability, access, utilization, and stability. The nutritional dimension is integral to the concept of food security.

[2] United Nations [15] definition: Energy security is a term that applies to the availability of energy at all times in various forms, in sufficient quantities, and at affordable prices, without unacceptable or irreversible impact on the environment. These conditions must prevail over the long term if energy is to contribute to sustainable development. Energy security has both a producer and a consumer side to it.

volatility and show a depreciation trend [3, 17]. Thus, it is interesting to analyze the manner in which the volatility of the dollar exchange rates influence the relationship between palm oil prices (MDEX) and crude oil prices (DME), These two commodity prices spark an interest in this study whose results are relevant for ASEAN. In the ASEAN region, palm oil can be produced sufficiently intra-regional demand and the remaining parts can be kept aside for exportation [18]. Moreover, it can be used for producing alternative energy in the form of biodiesel to reduce the effects from the crude oil price crisis. ASEAN has relied on imported crude oil from the Middle East [19]: its price is related to the crude oil prices of other regions such as West Texas Intermediate (WTI) [20].

The purpose of this study are the following: (1) to analyze the dependence between the exchange rates (the strength of the U.S. dollar) and two commodity prices: palm oil (MDEX) and crude oil prices (DME); (2) to analyze the dependence between palm oil and crude oil prices by considering the exchange rate as a conditioning variable.

The copula based GARCH model provides more flexibility for finding out the joint distribution and the transformation of the invariant correlation, without having to assume linear correlation [21]. Therefore, in this study, we used the GARCH(1,1) model [22] to examine the volatility of the exchange rate and the commodity daily prices, which are generally non-normal distributions, and applied the vine copula model to examine the relationship between each commodity, by using the R-package CDVine which was developed by Brechmann and Schepsmeier [23].

The remainder of this paper is organized as follows: part two is the methodology, and part three consists of the data and the empirical findings. Finally, part four provides the conclusions and the policy implications.

2 Methodology

2.1 Marginal Distribution Model

We adopt the GARCH(1,1) model [22] with an appropriate distribution (D), residual distribution, for the marginal distribution of the log-difference $\ln \frac{P_t}{P_{t-1}}$ of the three data series: palm oil prices, crude oil prices, and exchange rates.

$$y_t = \mu_t + \varepsilon_t \tag{1}$$

$$\varepsilon_t = z_t \sqrt{h_t}, z_t \sim (D) \tag{2}$$

$$h_t = \omega_t + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1} \tag{3}$$

In equation (1), we decompose the log-difference y_t into a mean μ_t and an error term ε_t. Equation (2) define the error term ε_t as the product between conditional variance h_t and a residual z_t. The residual z_t will be assumed to follow an appropriate distribution. Equation (3) presents GARCH(1,1) process where $\omega_t > 0, \alpha \geq 0, \beta \geq 0$

are sufficient to ensure that the conditional variance $h_t > 0$. The $\alpha\varepsilon_{t-1}^2$ represent the ARCH term and α refers to the short run persistence of shocks, while βh_{t-1} represent the GARCH term and β refers to the contribution of shocks to long run persistence $(\alpha+\beta)$. The properties of the GARCH(1,1) model require stationary and persistence of the conditional variance, h_t, of the error term, ε_t. This paper used the second moment condition that was $\alpha+\beta<1$ to check for these properties. In this study, the R-package fGarch by Wuertz and Chalabi [24] was used to estimate the parameters of GARCH(1,1) model.

For the next analysis by copula functions, the standardized residuals from GARCH(1,1) model were transformed to copula data $(F_1(x_1), F_2(x_2), F_3(x_3))$.

2.2 Copula Function

One approach of modeling the multivariate dependence is the copula. The copula functions can offer us to merge univariate distributions to get a joint distribution with an appropriate dependence structure. The fundamental theorem of copula was given by Sklar [25] as Sklar's theorem. The standard reference book of the copula theory was made by Nelson [26].

Let F be an n-dimensional distribution function with marginal distributions $F_1,...,F_n$. Then there exists a copula C for all $x=(x_1,...,x_n)' \in [-\infty,\infty]^n$,

$$F(x) = C(F_1(x_1),...F_n(x_n)) \tag{4}$$

If $F_1,...,F_n$ are continuous, then C is unique. Conversely, if C is a copula and $F_1,...,F_n$ are distribution functions, then the above function $F(x)$ in (4) is a joint distribution function with the marginal distribution $F_1,...,F_n$. C can be interpreted as the distribution function of an n dimensional random variable on $[0,1]^n$ with uniform margins [23].

We used various copula families contained in the R-package CDVine to measure the dependence of the pair-copula. Table 1 presents the characteristics of the copula families that were used in this study. Table 2 presents the function of Kendall's tau.

2.3 Vine Copula Modeling

Modeling copulas with high dimension is a difficult task because there are large numbers of variables. Vine copulas can cross over this restriction, vine copulas are a flexible tool for describing the multivariate copulas through the graphical model. The multivariate copulas are constructed from a cascade of bivariate copulas or are called pair-copulas. The principles of vine copulas propounded by Joe [29] and extended by Bedford and Cooke [30, 31]. For statistical inference techniques of two classes of C-vines and D-vines are described by Aas et al. [32]. Brechmann and Schepsmeier [23] said that a vine structure can be chosen manually or through expert knowledge, or be given by the data itself. Aas et al. [32] was of the opinion that modeling C-vine might be an advantage when we know that the main variable

Table 1 Characteristics of Copula Families

Name	Pair-copula function	Parameter range
Gaussian	$C(u_1,u_2;\rho)=\Phi_G(\Phi^{-1}(u_1),\Phi^{-1}(u_2);\rho)$ $=\int_{-\infty}^{\phi^{-1}(u_1)}\int_{-\infty}^{\phi^{-1}(u_2)}\frac{1}{2\Pi\sqrt{(1-\rho^2)}}\times[\frac{-(s^2-2\rho st+t^2)}{2(1-\rho^2)}]dsdt$	$\rho\in(-1,1)$
Student's T	$C^T(u_1,u_2;\rho,\nu)=\int_{-\infty}^{T_\nu^{-1}(u_1)}\int_{-\infty}^{T_\nu^{-1}(u_2)}\frac{1}{2\Pi\sqrt{(1-\rho^2)}}\times$ $[1+\frac{(s^2-2\rho sT+T^2)}{\nu(1-\rho^2)}]^{-(\frac{\nu+2}{2})}dsdT$	$\rho\in(-1,1)$, $\nu>2$
Clayton	$C(u_1,u_2;\theta)=(u_1^{-\theta}+u_2^{-\theta}-1)^{-1/\theta}$	$\theta\in(0,\infty)$
Gumbel	$C(u_1,u_2;\theta)=exp(-[(-\ln(u_1))^\theta+(-\ln(u_2))^\theta]^{\frac{1}{\theta}})$	$\theta\in[1,\infty)$
Frank	$C(u_1,u_2;\theta)=-\frac{1}{\theta}\log(1+\frac{(e^{-\theta u_1}-1)(e^{-\theta u_2}-1)}{e^{-\theta}-1})$	$\theta\in(-\infty,\infty)\backslash\{0\}$
Joe	$C(u_1,u_2;\theta)=1-[(1-u_1)^\theta+(1-u_2)^\theta-(1-u_1)^\theta(1-u_2)^\theta]^{\frac{1}{\theta}}$	$\theta\in[1,\infty)$
Rotated Clayton 90°	$C(u_1,u_2;\theta)=u_2-[(1-u_1)^{-\theta}+u_2^{-\theta}-1]^{-\frac{1}{\theta}}$	$\theta\in(-\infty,0)$
Rotated Gumbel 90°	$C(u_1,u_2;\theta)=u_2-exp(-[(-\ln(1-u_1))^\theta+(-\ln(u_2))^\theta]^{\frac{1}{\theta}})$	$\theta\in(-\infty,-1]$
Rotated Joe 90°	$C(u_1,u_2;\theta)=u_2-1-[u_1^\theta+(1-u_2)^\theta-u_1^\theta(1-u_2)^\theta]^{\frac{1}{\theta}}$	$\theta\in(-\infty,-1)$
Rotated Clayton 180°	$C(u_1,u_2;\theta)=u_1+u_2-1+[(1-u_1)^{-\theta}+(1-u_2)^{-\theta}-1]^{-\frac{1}{\theta}}$	$\theta\in(0,\infty)$
Rotated Gumbel 180°	$C(u_1,u_2;\theta)=u_1+u_2-1+exp(-[(-\ln(1-u_1))^\theta+$ $(-\ln(1-u_2))^\theta]^{\frac{1}{\theta}})$	$\theta\in[1,\infty)$
Rotated Joe 180°	$C(u_1,u_2;\theta)=u_1+u_2-(u_1^\theta+u_2^\theta-u_1^\theta u_2^\theta)^{\frac{1}{\theta}}$	$\theta\in[1,\infty)$

Source: The copula functions are given as presented in Trivedi and Zimmer [27], Nelson [26], and Fisher [28].

governs interactions in the data or plays an important role in the dependence structure, and that the others are linked to it. So, the C-vine copula model offers us to define the relationship structure between variables according to the purpose of study, and it can describe the relationship between variables through the graphical model or are called pair-copulas.

This study used C-vine copula modeling to analyze the dependence between the palm oil prices, crude oil prices, and exchange rates, a kind of analysis which no one has attempted before with a view to exploring it in depth. The structure of the C-vine model is shown in Figure 1. This study selected the exchange rate which was the first root node. Therefore, our assumption in this study is that the exchange rate is a key variable that plays a role in the linkage between palm oil prices and crude oil prices; this is based on the available literature, and includes the currency, the dollar, which is widely used in international financial transactions. Moreover, the

Table 2 Function of Kendall's tau and Tail Dependence for Bivariate Copula

Copula family	Kendall's tau
Gaussian	$\frac{2}{\Pi}\arcsin\rho$
Student's T	$\frac{2}{\Pi}\arcsin\rho$
Clayton	$\frac{\theta}{\theta+2}$
Gumbel	$1-\frac{1}{\theta}$
Frank	$1-\frac{4}{\theta}+4\frac{D_1(\theta)}{\theta}$
Joe	$1+\frac{4}{\theta^2}\int_0^1 t\log(t)(1-t)^{2(1-\theta)/\theta}dt$
Rotated Clayton 90°	$\frac{\theta}{\theta-2}$
Rotated Gumbel 90°	$-1-\frac{1}{\theta}$
Rotate Joe 90°	$-1-\frac{4}{\theta^2}\int_0^1 t\log(t)(1-t)^{-2(1+\theta)/\theta}dt$
Rotated Clayton 180°	$\frac{\theta}{\theta+2}$
Rotated Gumbel 180°	$1-\frac{1}{\theta}$
Rotate Joe 180°	$1+\frac{4}{\theta^2}\int_0^1 t\log(t)(1-t)^{2(1-\theta)/\theta}dt$

Source: Kendall's tau is as presented in Brechmann and Schepsmeier [23].
Note: $D_1(\theta)=\int_0^\theta \frac{c/\theta}{exp(x)-1}$ is the Debye function.

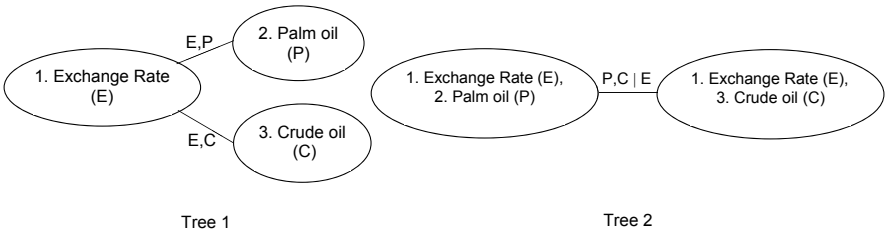


Fig. 1 The pair-copulas of three-dimensional C-vine trees

international trading of food, agricultural commodities, and crude oil is done using the dollar in their respective markets [3, 17].

We presented the three dimensions, which was what we used in this paper. Let $X=(X_1,X_2,X_3)\sim F$ with marginal distribution functions F_1,F_2,F_3 and their density functions f_1,f_2,f_3, which was proposed as follows (see Aas et al. [32]).

$$F(x_1,x_2,x_3)=C(F_1(x_1),F_2(x_2),F_3(x_3)) \tag{5}$$

$$\begin{aligned} f(x_1,x_2,x_3)=f(x_1)\cdot f(x_2)\cdot f(x_3)\cdot c_{1,2}(F_1(x_1),F_2(x_2))\cdot c_{1,3}(F_1(x_1),F_3(x_3)) \\ \cdot c_{2,3|1}(F_{2|1}(x_2\mid x_1),F_{3|1}(x_3\mid x_1)) \end{aligned} \tag{6}$$

where $c_{1,2}$, $c_{1,3}$, and $c_{2,3|1}$ denote the densities of bivariate copulas $C_{1,2}$, $C_{1,3}$, and $C_{2,3|1}$, respectively. $F_{2|1}$ and $F_{3|1}$ are the marginal conditional distributions that can be derived from formula (7).

The vine copulas involve marginal conditional distributions. The general form of a conditional distribution function is $F(x \mid v)$,

$$F(x \mid v) = \frac{\partial C_{x,v_j|v_{-j}}(F(x \mid v_{-j}), F(v_j \mid v_{-j}))}{\partial F(v_j \mid v_{-j})} \tag{7}$$

where v denotes all the conditional variables and $C_{x,v_j|v_{-j}}$ is a bivariate copula distribution function. For v is univariate, the marginal condition distribution, e.g. $F_{3|1}$ can be presented as

$$F_{3|1}(x_3 \mid x_1) = \frac{\partial C_{31}(F_3(x_3), F_1(x_1))}{\partial F_1(x_1)} \tag{8}$$

2.4 Vine Copula Estimation

In the R-package CDVine, the maximum likelihood was used to estimate the parameters of copulas. The log-likelihood of C-vine copula with three dimensions in (6) can be written as

$$\sum_{t=1}^{T} \log[c_{1,2}(F_1(x_{1,t}), F_2(x_{2,t})) \cdot c_{1,3}(F_1(x_{1,t}), F_3(x_{3,t})) \cdot c_{2,3|1}(F_{2|1}(x_{2,t} \mid x_{1,t}), F_{3|1}(x_{3,t} \mid x_{1,t}))]. \tag{9}$$

3 Data and Empirical Findings

To analyze the relationship between the exchange rate and the two commodity prices (palm oil and crude oil), we selected the commodity prices that are related to the AEC. Palm oil prices were obtained from the Malaysia Derivatives Exchange (MDEX) because Malaysia is a major producer and world exporter of palm oil [18]. The crude oil benchmark price for the Asian market is the Dubai (Oman) crude oil price [33] since the Middle East is the major source of crude oil for ASEAN [19]. Hence, the crude oil price of Dubai Mercantile Exchange (DME) was used in this study. The exchange rate data, or the broad dollar index, was measured as a weighted average of the foreign exchange values of the U.S. dollar against the currencies of a large group of major U.S. trading partners (definition from the EcoWin database).

The observations of the three data series were obtained from the EcoWin database during the period from 1 June 2007 to 15 March 2013. For the prices of palm oil and crude oil, we used the Futures 1-Pos of daily close prices. Each data series was transformed into the log-difference, $\ln \frac{P_t}{P_{t-1}}$, before it was used to analyze using the vine copula based GARCH model.

Table 3 presents the descriptive statistics of the log-difference of exchange rate, palm oil price, and crude oil price. Palm oil has a negative average growth rate but crude oil has a positive average growth rate. All of the three data series exhibited negative skewness. If skewness is negative, the market has a downside risk, or there is substantial probability of a big negative return. The kurtosis of these data is greater

Table 3 Data Descriptive Statistics for Log-difference of Exchange Rate, Palm Oil Price, and Crude Oil Price

	Exchange rate	Palm oil	Crude oil
Mean	0.0000	−0.0001	0.0004
Median	−0.0001	0.0000	0.0009
Maximum	0.0174	0.0976	0.1339
Minimum	−0.0230	−0.1104	−0.1337
Std. Dev.	0.0039	0.0203	0.0230
Skewness	−0.1172	−0.3472	−0.1574
Kurtosis	6.5570	7.0304	7.6829
Jarque-Bera	730.13	961.06	1265.75
(p-value)	(0.0002)	(0.0000)	(0.0001)
p-value of Dickey-Fuller test	0.01	0.01	0.01
Number of observations	1,379	1,379	1,379

than 3. Hence, this kurtosis can be said to be super Gaussian and leptokurtic. This means that the growth rates of the empirical data have a typically spiky probability distribution function with heavy tails. The null hypothesis of normality of the Jarque-Bera tests are rejected in all the data series. The Dickey-Fuller test shows that these data series are stationary at p-value 0.01.

Table 4 Results of GARCH(1,1) Test with Normal Residual for Exchange Rate Data, and of Skewed Student T Residual for Palm Oil and Crude Oil Data

	Exchange rate	Std. error (p-value)	Palm oil	Std. error (p-value)	Crude oil	Std. error (p-value)
ω	1.007e-07	5.176e-08 (0.0518*)	3.903e-06	1.721e-06 (0.0233*)	2.325e-06	1.749e-06 (0.184)
α	0.0636	1.049e-02 (1.34e-09***)	0.0746	1.501e-02 (6.75e-07***)	0.0529	1.214e-02 (1.32e-05***)
β	0.9304	1.107e-02 (<2e-16***)	0.9155	1.606e-02 (<2e-16***)	0.9451	1.231e-02 (<2e-16***)
ν (degree of freedom)	-	-	7.6810	1.485e+00 (2.31e-07***)	5.0670	7.455e-01 (1.07e-11***)
γ (skewness)	-	-	0.9685	3.557e-02 (<2e-16 ***)	0.9418	3.112e-02 (<2e-16 ***)
Log likelihood	5,869.953	-	3,654.827	-	3,499.523	-
K-S test (p-value)	-	- (1)	-	- (0.9208)	-	- (1)
Box-Ljung test (p-value)	-	-	-	-	-	-
1st moment	-	(0.4301)	-	(0.2515)	-	(0.5832)
2nd moment	-	(0.9363)	-	(0.8898)	-	(0.7921)
3rd moment	-	(0.6521)	-	(0.0732)	-	(0.7765)
4th moment	-	(0.8513)	-	(0.8803)	-	(0.6423)

Note: Significant codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Table 4 presents the results of GARCH(1,1) with normal residual for the exchange rate data and skewed student T residual for the palm oil and crude oil data. The asymmetry parameters, γ, are significant and less than 1, thus indicating that the palm oil and crude oil data series are skewed to the left.

For the exchange rate, palm oil, and crude oil, the $\alpha + \beta$ are 0.9940, 0.9901, and 0.9980, respectively; this implies that their volatilities have a long-run persistence. For the short-run effect of the unexpected factors, we considered the event from the α parameter. Therefore, we can see that they nearly have the values 0.0636, 0.0746, and 0.0529, and a small impaction for volatility.

Next, we transformed the standardized residuals from the GARCH(1,1) model into the uniform [0,1] by using the empirical distribution function $F_n(x) = \frac{1}{n+1}\Sigma_{i=1}^{n} 1(X_i \leq x)$, where $X_i \leq x$ the order statistics and 1 is the indicator function. The transformed data were used in the Kolmogorov-Smirnov (K-S) test for uniform [0,1] and the Box-Ljung test for serial correlation. More details are available in Patton [34] and Manthos [35]. These tests are necessary to check for the marginal distribution models' misspecification before using the copula model.

The results of the K-S test show that these marginal distributions are uniform, by accepting the null hypothesis at p-values equal to 1 or nearly 1. The results of the Box-Ljung test provide that all of the four moments of all the marginal distributions are i.i.d., by accepting the null hypothesis that there is no serial correlation at p-values greater than 0.05. Therefore, our marginal distributions were not misspecified and can be used for the copula model.

3.1 Results of C-vine Copula

Figure 1 presents each pair-copula of the three-dimensional C-vine tree; there are two pair-copulas in tree 1 and one pair-copula in tree 2. The first and second pair-copulas in tree 1 are Exchange rate–Palm oil (E,P) and Exchange rate–Crude oil (E,C), respectively. The third pair-copula in tree 2 is a conditional pair-copula, Palm oil–Crude oil given Exchange rate (P,C|E).

We used the Gaussian copula, Student's T copula, Clayton copula, Gumbel copula, Frank copula, Joe copula, rotated Clayton 90° and 180°, rotated Gumbel 90° and 180°, and rotated Joe 90° and 180° copula to fit the data.

The AIC and the BIC are used to appraise which copula is the best fit. The Kendall's tau correlation which was transformed from the copula parameter was used because each family of copula has a different range of copula parameters; hence we inverse a copula parameter into a Kendall's tau correlation, and it is bound on the interval $[-1, 1]$. Kendall's tau is a measure of concordance and is a function of copula; hence, we can use it to assess the range of dependence covered by the families of copula. A goodness-of-fit test based on Kendall's tau provides the Cramér-von Mises (CvM) and Kolmogorov-Smirnov (K-S) test statistics and the estimated p-values by bootstrapping [23] in order to test the appropriateness of the copula model under the null hypothesis that the empirical copula C belongs to a parametric class C' of any copulas, $H_0 : C \in C'$.

The results of the pair-copulas, the Exchange rate–Palm oil (E,P), the Exchange rate–Crude oil (E,C), and the Palm oil–Crude oil given Exchange rate (P,C|E), are presented in Table 5.

Table 5 Results of C-vine Copula Model

Tree	Pair-copula	Copula family	Copula parameter	Std. error (p-value)	Kendall's tau	AIC	BIC	p-value CvM	p-value KS
1	E,P	Gaussian	−0.2438	0.0246 (0.0000)	−0.1568	−82.3568	−77.1276	0.38	0.47
	E,C	Gaussian	−0.4260	0.0203 (0.0000)	−0.2802	−273.8174	−268.5883	0.76	0.74
2	P,C|E	Gaussian	0.1660	0.0259 (0.0000)	0.1062	−36.4692	−31.2401	0.61	0.38

Table 5 presents the results of C-vine copula model. The first pair is the Exchange rate–Palm oil (E,P), the Gaussian copula provided the smallest AIC and BIC, and the CvM and K-S tests accepted the null hypothesis with p-values greater than 0.05, which means that the dependence structure of the data series is appropriate for a chosen family. Therefore, the Gaussian copula is the best fit copula, with a copula parameter of -0.2438 and a Kendall's tau correlation of -0.16. This implies that when the exchange rate increases (i.e., when the U.S. dollar is stronger), the palm oil price decreases, and vice versa. However, there exists a weak negative dependence in this pair-copula, thus indicating that a change in palm oil price is slightly related to a change in exchange rate.

For the second pair, the Exchange rate–Crude oil (E,C), the Gaussian copula is chosen to explain the dependence structure of this pair-copula with a copula parameter of -0.4260 and a Kendall's tau correlation of -0.28. This means that when the exchange rate increases (i.e., when the U.S. dollar is stronger), the crude oil price decreases, and vice versa. However, this pair-copula has a weak negative dependence, thereby indicating that a change in crude oil price is slightly related to a change in exchange rate, which is similar to the result of the first pair-copula.

The parameter of each pair-copula from an appropriate copula family in tree 1 was used to construct a conditional pair-copula of Palm oil–Crude oil given Exchange rate (P,C|E) in tree 2. This pair-copula provides the Gaussian copula is the best fit with the copula parameter of the Gaussian copula is 0.1660 and the Kendall's tau correlation is 0.11. Therefore, whether it is an upward or a downward trend, both the commodity prices tend to move together. However, this pair-copula has a weak positive dependence; this means that a change in palm oil price is slightly related to a change in crude oil price.

According to our results, the copula parameters and the Kendall's tau correlations of a conditional pair-copula (P,C|E), 0.1660 and 0.11, are less than those that are obtained for the bivariate pair-copula Palm oil–Crude oil (P,C). Further testing reveals that the Gaussian copula of a bivariate copula (P,C) offers a copula parameter and a Kendall's tau correlation of 0.2495 and 0.16, respectively. This implies

that the exchange rate (E) has an influence in the relationship between the palm oil price (P) and the crude oil price (C). The exchange rate (E) is an important variable that governs the interactions in the dependence structure between the palm oil price (P) and the crude oil price (C).

4 Conclusions and Policy Implications

We analyzed the relationship between the dollar exchange rates and two commodity prices, palm oil price and crude oil price. The analysis was done by using the GARCH(1,1) model to examine the volatility of the exchange rates and the future prices 1-Pos. of both the commodity prices. The vine copula model was used to analyze the dependence structure between their marginal distributions. The data analyses were based on the daily observations during the period from June 2007 to March 2013. The empirical results of GARCH(1,1) showed that the exchange rates, palm oil prices, and crude oil prices have a long-run persistence in volatility. The C-vine copula consisted of three pair-copulas (Figure 2), which are Exchange rate–Palm oil (E,P), Exchange rate–Crude oil (E,C), and Palm oil–Crude oil given Exchange rate (P,C|E). Of the three, the Exchange rate–Palm oil (E,P) and Exchange rate–Crude oil (E,C) are in the first tree, and the conditional pair-copula, Palm oil–Crude oil given Exchange rate (P,C|E), is in the second tree.

The Gaussian copula was chosen to explain the dependence structure of the Exchange rate–Palm oil (E,P) pair-copula with a copula parameter of -0.2438 and a Kendall's tau correlation of -0.16. Similarly, the Exchange rate–Crude oil (E,C) indicates the Gaussian copula as the best fit with a copula parameter of -0.4260 and a Kendall's tau correlation of -0.28.

As for the last pair-copula, the conditional pair-copula Palm oil–Crude oil given Exchange rate (P,C|E), the Gaussian copula was chosen to explain the dependence structure with a copula parameter of 0.1660 and a Kendall's tau correlation of 0.11. Furthermore, the findings of this research provide evidence that the exchange rate (E) is an important variable that governs the interactions in the dependence structure between the palm oil price (P) and the crude oil price (C).

Our results showed that the volatility of the exchange rate, palm oil price and crude oil price are interrelated (see Figure 2). Considering the Kendall's tau correlation, the pair-copulas Exchange rate–Palm oil (E,P) and Exchange rate–Crude oil (E,C) have a weak negative correlation. The conditional pair-copula Palm oil–Crude oil given Exchange rate (P,C|E) has a weak positive correlation.

This study found some evidences of excess kurtosis, skewness, and non-normal distribution in each data series, exchange rate, palm oil price and crude oil price. That is why the copula model is an appropriate tool to measure the relationship between each variable considered in this study. The copula can model the dependence between random variables without an assumption of linear correlation.

From the empirical results of this study, it can be concluded that a depreciating exchange rate has a relation with an increase in palm oil price in the Malaysian market (MDEX) and crude oil price in the Dubai market (DME). As far as the palm

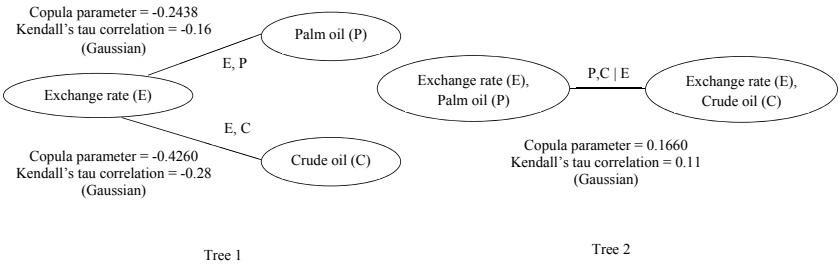


Fig. 2 The C-vine copula for the exchange rate, and palm oil and crude oil data with the pair-copula families and the Kendall's tau values

oil exports of ASEAN are concerned, a depreciating dollar exchange rate would prove advantageous to ASEAN because that would generate more income for the region. But, on the other hand, the incentive in world market price and the increased profitability in international trade will cause an increase in the volume of palm oil that is exported from the region. The consequences can be negative in that it can lead to a rise in the local price, or a shortfall for consumers in some areas of the ASEAN region. For example, if this were to occur in Indonesia, it would tend to make palm oil producers increase exports when the world market price increases, and this would create a shortfall for domestic consumers [18]. Thus, the palm oil producers and exporters of ASEAN who are from Indonesia, Malaysia, and Thailand should endeavor to keep the balance between the intra-regional demand and the exportation demand. As for the crude oil imports of ASEAN, a rise in crude oil price will increase the cost of living of the people living in the ASEAN region.

So, a depreciation in the exchange rate is related to an increase in the palm oil price and the crude oil price. The dollar exchange rate is an important variable in that the ASEAN nations have to monitor and manage its impact in terms of food security and energy security. As for the investors, they should take into consideration the risk that could arise from a change in the exchange rate, which, again, is related to the palm oil price and the crude oil price.

ASEAN can produce enough quantity of palm oil and is the world largest exporter of this food commodity [18]. However, ASEAN has to rely on import crude oil from the Middle East [19]: the fact is that the crude oil price of the Dubai Mercantile Exchange (DME) is related to the West Texas Intermediate (WTI) as well as the other markets in the world [20]. Thus, the dollar exchange rate should have more influence on the crude oil price from the Middle East (DME) than the palm oil price from Malaysia (MDEX). This corresponds to the empirical results of this research: the negative dependence between the dollar exchange rate and the crude oil price (DME) is greater than the negative dependence between the dollar exchange rate and the palm oil price (MDEX).

For the bivariate copula analysis of palm oil price and crude oil price, there exists a weak positive dependence (a copula parameter 0.2495 and Kendall'stau

correlation 0.16); this means that the intensity of the co-movement of their prices is less. It can be explained by the fact that ASEAN has a high production capacity of palm oil, and so it can reduce the direct and indirect effects of fluctuations in crude oil prices in the world market. This is the reason why the dependence between the two commodity prices is weak.

From our findings, it is evident that the dependence between the palm oil price and the crude oil price is still weak. Also, there exists a high production capacity of palm oil in ASEAN. Therefore, using biodiesel as alternative energy is one of the choices that should be considered. Palm biodiesel can reduce energy cost for consumers when they are faced with a continuous rise in crude oil prices. For example, in Malaysia, where there are a lot of cultivated areas of oil palm trees and there is a high potential for producing palm biodiesel, if the production of biodiesel were implemented very effectively, it would have a positive impact on the economy in many ways [36]. However, while deciding to use the produce from the oil palm tree for biodiesel production, the policy makers should take into consideration the suitability as regards food security, environment, and critical social needs [18, 37].

Acknowledgements. This work was granted support by the Energy Conservation Promotion Fund, the Energy Policy and Planning Office, the Ministry of Energy of Thailand. The first author is grateful for being granted a PhD scholarship to do his studies.

References

1. Asian Development Bank. Global food price inflation and developing Asia. Asian Development Bank (2011), http://www.adb.org/publications/global-food-price-inflation-and-developing-asia (accessed May 23, 2013)
2. Serra, T., Zilberman, D.: Biofuel-related price transmission literature: A review. Energy Economics 37, 141–151 (2013)
3. Abbott, P.C., et al.: What's Driving Food Prices? Farm Foundation Issue Report (July 2008)
4. Harri, A., et al.: The relationship between oil, exchange rates, and commodity prices. Journal of Agricultural and Applied Economics 41, 501–510 (2009)
5. Harri, A., Hudson, D.: Mean and variance dynamics between agricultural commodity prices and crude oil prices. Paper presented at the Economics of Alternative Energy Sources and Globalization, The Road Ahead Meeting, Orlando, FL, November 15-17 (2009)
6. Kwon, D., Koo, W.W.: Price transmission channels of energy and exchange rate on food sector: a disaggregated approach based on stage of process. Selected Paper prepared for presentation at the Agricultural & Applied Economics Association 2009 AAEA & ACCI Joint Annual Meeting, Milwaukee, Wisconsin, July 26-29 (2009)
7. Akram, Q.F.: Commodity Prices, Interest Rates and the Dollar. Energy Economics 31, 838–851 (2009)
8. Cooke, B., Robles, M.: Recent Food Prices Movements. A Time Series Analysis. International Food Policy Research Institute (IFPRI) Discussion Paper No. 00942. IFPRI, Washington DC (2009)

9. Gilbert, C.L.: How to understand high food prices. Journal of Agricultural Economics 61(2), 398–425 (2010)
10. Balcombe, K.: The nature and determinants of volatility in agricultural prices: an empirical study. In: Prakash, A. (ed.) Safeguarding Food Security in Volatile Global Markets, pp. 85–106. FAO, Rome (2011)
11. Nazlioglu, S., Soytas, U.: Oil Price, agricultural commodity prices, and the dollar: a panel cointegration and causality analysis. Energy Economics 34, 1098–1104 (2012)
12. Anzuini, et al.: The impact of monetary policy shocks on commodity prices. Working paper No. 851, Bank of Italy (2012)
13. ASEAN Secretariat. Regional and Country Reports of the ASEAN Assessment on the Social Impact of the Global Financial Crisis. The ASEAN Secretariat (2010), http://www.asean.org/archive/publications/ARCR/ASEANRegional&CountryReport.pdf (accessed May 20, 2013)
14. FAO. Declaration of the world summit on food security. World Summit on Food Security, Rome, November 16-18 (2009), ftp://ftp.fao.org/docrep/fao/Meeting/018/k6050e.pdf (accessed May 20, 2013)
15. United Nations. World Energy Assessment: Overview 2004 Update. United Nations Development Programme (2004), http://www.undp.org/content/dam/aplaws/publication/en/
16. ASEAN Secretariat. ASEAN Community in a Global Community of Nations.Co-Chairs' statement of the 4th ASEAN-UN summit Bali, Indonesia (November 19, 2011), http://www.mofa.go.jp/region/asia-paci/eas/pdfs/declaration_1111_2.pdf (accessed May 20, 2013)
17. Abbott, P.C., et al.: What's Driving Food Prices in 2011? Farm Foundation Issue Report (July 2011)
18. Sheil, D., et al.: The impacts and opportunities of oil palm in Southeast Asia: What do we know and what do we need to know? Occasional paper no. 51. CIFOR, Bogor, Indonesia (2009)
19. Speed, P.A.: ASEAN. The 45 Year Evolution of a Regional Institution. POLINARES working paper n. 61, University of Westminster (2012), http://www.polinares.eu/docs/d4-1/polinares_wp4_chapter11.pdf (accessed May 27, 2003)
20. Reboredo, J.C.: How do crude oil prices co-move? A copula approach. Energy Economics 33, 948–955 (2011)
21. Sriboonchitta, S., et al.: Modeling volatility and dependency of agricultural price and production indices of Thailand: Static versus time-varying copulas. International Journal of Approximate Reasoning 54(6), 793–808 (2013)
22. Bollerslev, T.: Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics 31, 307–327 (1986)
23. Brechmann, E.C., Schepsmeier, U.: Modeling Dependence with C- and D-Vine Copulas: The R Package CDVine. Journal of Statistical Software 52(3), 1–27 (2013), http://www.jstatsoft.org/v52/i03/ (accessed February 20, 2013)
24. Wuertz, D., Chalabi, Y.: Rmetrics-Autoregressive Conditional Heteroskedastic Modelling (2013), http://cran.r-project.org/web/packages/fGarch/index.html (accessed May 10, 2013)
25. Sklar, A.: Fonctions de répartition á n dimensions etleursmarges. Publications de l'Institut de Statistique de L'Université de Paris 8, 229–231 (1959)
26. Nelson, R.B.: An Introduction to Copulas, 2nd edn. Springer, New York (2006)

27. Trivedi, P.K., Zimmer, D.M.: Copula Modeling: An Introduction for Practitioners. Foundations and Trends in Econometrics 1(1), 1–111 (2005)
28. Fisher, M.: Tailoring copula-based multivariate generalized hyperbolic secant distributions to financial return data: An empirical investigation. Discussion papers, University of Erlangen-Nüremberg, Germany (2003), http://www.statistik.wiso.uni-erlangen.de/forschung/d0047.pdf (accessed January 25, 2013)
29. Joe, H.: Families of *m*-Variate Distributions with Given Margins and m(m1)/2 Bivariate Dependence Parameters. In: Rüschendorf, L., Schweizer, B., Taylor, M.D. (eds.) Distributions with Fixed Marginals and Related Topics, vol. 28, pp. 120–141 (1996)
30. Bedford, T., Cooke, R.M.: Probability Density Decomposition for Conditionally Dependent Random Variables Modeled by Vines. Annals of Mathematics and Artificial Intelligence 32, 245–268 (2001)
31. Bedford, T., Cooke, R.M.: Vines- A New Graphical Model for Dependent Random Variables. Annals of Statistics 30, 1031–1068 (2002)
32. Aas, K., et al.: Pair-copula constructions of multiple dependence. Insurance: Mathematics and Economics 44, 182–198 (2009)
33. Koyama, K.: A Though on Crude Oil Pricing in Asia. The institute of energy economics, Japan (2011), https://eneken.ieej.or.jp/data/3711.pdf (accessed May 27, 2013)
34. Patton, A.J.: Modelling Asymmetric Exchange Rate Dependence. International Economic Review 47(2), 527–556 (2006)
35. Manthos, V.: Dynamic Copula Toolbox 3.0 (2010), http://www.mathworks.com/matlabcentral/fileexchange/29303-dynamic-copula-toolbox-3-0 (accessed December 15, 2012)
36. Lim, S., Teong, L.K.: Recent trends, opportunities and challenges of biodiesel in Malaysia: An overview. Renewable and Sustainable Energy Reviews 14, 938–954 (2010)
37. Gasparatos, A., et al.: Sustainability impacts of first-generation biofuels. Animal Frontiers 3(2), 1–15 (2013), doi:10.2527/af.2013-0011

An Analysis of Interdependencies among Energy, Biofuel, and Agricultural Markets Using Vine Copula Model

Phattanan Boonyanuphong and Songsak Sriboonchitta

Abstract. This paper aims to study the structure of interdependencies between the energy, biofuel and agricultural commodity markets. The work concentrates on the dependence between ethanol and agricultural futures returns conditional to crude oil returns, and interdependence among agricultural commodities conditional to crude oil and ethanol futures returns. The C-vine copula based ARMA-GARCH model was used to explain the dependence structure of crude oil and the four related variables, and applied to investigate the risk of energy-agricultural commodity futures portfolio. We generally found symmetry in the tail dependence between the energy, biofuel, and agricultural commodities, and also found a greater significant variability in dependence, specifically, the dependence between the ethanol and agricultural commodity futures returns conditional to crude oil as well as interdependence between corn and soybean conditional to crude oil and ethanol return. This indicates that there is a rise in ethanol productions and that higher crude oil prices have caused a price increase in agricultural commodities such as corn and soybean. Moreover, the higher dynamic dependence and symmetric tail dependences indicate that opportunities for portfolio diversification are reduced, particularly during a downturn in the markets. Finally, our result suggests that the time-varying copula model captures the portfolio risk better than the static copula models.

Phattanan Boonyanuphong
Faculty of Economics, Chiang Mai University, Chiang Mai 50200, Thailand
Department of Social Sciences, Prince of Songkla University, Pattani 94000, Thailand
e-mail: bphattanan@bauga.pn.psu.ac.th

Songsak Sriboonchitta
Faculty of Economics, Chiang Mai University, Chiang Mai 50200, Thailand
e-mail: songsak@econ.cmu.ac.th, songsakecon@gmail.com

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_26,

1 Introduction

The rising trend in food commodity prices has shown a significant increase in the last few years; this has caused substantial impact on the global economic activity, especially in developing countries. In such a situation, the volatility of food prices, particularly grain prices, has increased concerns about the world food supply and security.

One key factor is growth in demand of grain production in the biofuels industry which is the culprit behind the rising trend in agricultural prices. The growth in global biofuels production has reached phenomenal proportions, especially in the case of ethanol and biodiesel production, which were, roughly, 30 billion gallons and 9 million tons in 2012[1, 2]. The following two main agricultural productions were used to produce ethanol from their coarse grains: corn and sugarcane amounted to 50% of the total global ethanol production via feedstock during 2008-2010. Biodiesel production is mostly carried out from vegetable oils, particularly, soybean oil and rapeseed oil[3]. During 2008-2011, the global productions of coarse grains and sugar were around 11% and 21%, respectively; they were used to produce ethanol, and 11% of the global production of vegetable oil was used to produce biodiesel[4].

This dramatic rise in biofuel production has increased the adverse effect on the prices of the main agricultural commodities that are used for producing ethanol and biodiesel. The increased production of biofuel has been blamed for being one of the causes of the 2007/08 and 2010/11 global food crises[3]. Moreover, the emergence of the new trend of large-scale production of biofuels used for transportation has reshaped the relationship between agricultural commodities and energy markets. Traditionally, an energy-food linkage has been connected through the input costs channels, especially through fuel, fertilizer, and transportation.

The strong connection and increasing volatility of the agricultural and energy markets have attracted a growing interest in the academic world, especially among policy makers and researchers, with the latter having closely examined the evolution of the relationship between energy and the agricultural prices. A considerable body of research has been devoted to investigating the links between energy and food prices through the input costs channels. Hanson et al.[5] showed that soaring crude oil prices drive higher costs of production which, in turn, cause the agricultural commodity prices to increase. Baffes[6] also pointed out that crude oil price should be included in the aggregate production function that passes through input functions such as fertilizer, fuel, and transportation. Similarly, the European Commission recognized that the rise in agricultural commodity prices has been effected by the energy prices through the input channel for a rising cost i n fertilizers, chemical materials, and transportation[7].

Opposed to this point of view is a piece of literature which deals with new links that are focused only on crude oil and agricultural commodities. There is an overwhelming amount of study that analyzes the impact of biofuels on food and energy prices, which can be categorized on the basis of the data used in the empirical study as relying on biofuel prices or not. Chang and Su[8], Ciaian and Kancs[9],

Natanelov et al.[10], Nazlioglu and Soytas[11], Du et al.[12], and Boonyanuphong et al.[13] did a study on the dependency between crude oil and food prices by ignoring biofuel prices. Chang and Su[8] found evidence that the substitutive effect can be represented in the period of the high crude oil price due to the significant price spillover effects from crude oil futures to corn and soybean futures. Ciaian and Kancs[9] also discovered a cointegration between crude oil and food commodities, and found that the interdependencies keep rising over time. Similarly, Nazlioglu and Soytas[11], and Boonyanuphong et al.[13] discovered that crude oil is in cointegration with agricultural commodities, especially in the recent years. Results derived from the work of Boonyanuphong et al.[13] also show that there exists symmetric tail dependence between crude oil and agricultural commodity prices, and that the dependences are very volatile over time.

Furthermore, some research papers deal with new linkages that make an analysis of the interdependency between the energy, biofuel, and food markets that rely on the biofuel prices[14, 15, 16, 17]. Serra et al.[17] provided evidence of two cointegration relationships: crude oil-gasoline and ethanol-corn-gasoline. The results show the existence of long-running relationships between ethanol, corn, and gasoline, thus indicating that the energy-agricultural price relationship is in linkage with the biofuel market. Du and McPhail[14] also found that the ethanol, gasoline, and corn prices are more closely linked to a strengthened biofuel relationship. This could be examined using new developments in the biofuel industry and bioenergy policy instruments. While the results of the correlation between the oil, ethanol, and agricultural commodities are relatively strong, the evidence for a causal link from oil to the commodity prices is still mixed [15].

Literature considering the price link and price volatility interactions among the biofuel-related markets are extensively analyzed using different econometric techniques. Most of the common methodological approach applied in the study price level link consists of cointegration analysis and/or estimation of a VECM, while the prominent methodological approaches employed in investigating the price volatility interactions consist of the VECM, BEKK, and GARCH-type models[3]. Although the cointegration analysis methods, the VECM, BEKK, and GARCH-type models, are still good for measurement and enough for analyzing interdependence and volatility between random variables, they are based on some strong assumptions that were not conforming to the data in the empirical studies. Given the drawbacks of the conventional methodology, researchers are motivated to utilize copulas since they are more flexible in modeling the volatility and dependence structures. There are several collections o f bivariate copulas with distinct features that can fit with the various forms of dependence. However, in a multivariate case, there is only a standard multivariate copula, such as the Gaussian or the Student-t, as well as the Archimedean copulas that are lacking in flexibility due to the imposition of strong restrictions on equal dependence with all pairs of variables. Vine copulas, first proposed by Joe[18] and developed further by Bedford and Cooke[19], are very flexible for use in multivariate variables. The vine copulas can be constructed using a cascade of bivariate copulas; they produce large collections of bivariate copulas available in multivariate cases.

In this study, we attempt to fill the gaps and handle the drawbacks of the traditional models via the vine copula based ARMA-GARCH model by investigating the interdependencies between the energy, biofuel, and agricultural markets. The main purpose of this paper is to analyze the volatilities and dependencies among crude oil, ethanol, and agricultural commodities future prices including corn, soybeans, and sugar prices. Moreover, we are interested in the co-movement between the agricultural commodity and ethanol prices conditional to crude oil prices, and the co-movement between the agricultural commodity prices conditional to crude oil and ethanol prices. Finally, our paper calculates the value at risks and expected shortfalls using the results obtained from the use of copulas and the Monte Carlo simulation method, which can give some revelations for risk management. By answering these questions, we hope to enhance the understanding of the interdependent among the prices of crude oil , ethanol, and agricultural commodities.

The remainder of the paper is organized as follows. The second section will provide a brief review of the copula model, vine copula, the marginal models used for estimating the volatility and dependence structures of the energy and the other four related variables. Section 3 provides details of the data set and the empirical results of this study. Section 4 provides the applications for portfolio management in the technical field, as well as the empirical results. Finally, the conclusions are presented in section 5.

2 Econometrics Models

2.1 Copula Models

Of late, copula models have been widely applied for use in measuring and analyzing the dependence structures of joint probability distributions. The copula concept was first developed by Sklar. For a random vector $X = (X_1, X_2, ..., X_d) \sim F$ with a univariate margins $F_i, i = 1, ..., d$, there exists a unique function C called copula for which:

$$F(x_1, ...x_d) = C(F_1(x_1), ..., F_d(x_d)) \tag{1}$$

If F is an absolutely continuous function and $F_i, i = 1, ..., d$ are strictly increasing, we have the density function as

$$f(x_1, ...x_d) = \prod_{i=1}^{d} f_i(x_i) \times c(F(x_1, ..., F(x_d))) \tag{2}$$

where c is copula density function. In other words, copulas can be separately decomposed in the modeling of the marginal densities and the dependency part in terms of the copula density.

A large collection of copula families find application in empirical studies, especially in the finance markets. Two of the most commonly used in this field are the Gaussian copula and the t-copula. The Gaussian copula has zero tail

dependence, whereas the tail dependence for the t-copula is symmetric and non-zero with the same probability of occurrence[20]. The Clayton and Gumbel copulas are non-symmetric and commonly used to investigate asymmetric tail dependence[20, 21]. The Clayton copula provides strong lower tail dependence and the Gumbel copula exhibits strong upper tail dependence. Likewise, the Joe copula has higher dependence in the upper tail than in the lower tail, where it is zero. Moreover, Joe[22] employed two bivariate copula families, namely BB1 and BB7, that provide non-zero upper and lower tail dependences. Also, the rotation of the bivariate copula families is utilized to analyze the dependence structure in our empirical study.

2.2 Vine Copulas

Although there exists a large collection of bivariate copula families, the multivariate distributions carry many restrictions on the dependence relationships between the random variables. Vine copula is helpful in constructing multivariate distributions by incorporating the bivariate copula into the dependence structure under the specified marginal conditional distributions. For the d-dimensional density corresponding to a canonical vine (C-vine) is given by[23]

$$f(x_1,...,x_d) = \prod_{k=1}^{d} f(x_k) \cdot \prod_{j=1}^{d-1} \prod_{i=1}^{d-j} F(x_j|x_1,..,x_{j-1}), F(x_{j+i}|x_1,..,x_{j-1}). \tag{3}$$

For example, the 5-dimensional version of C-vine density (3) in our case can be written as

$$\begin{aligned} f(x_1,...,x_5) = \prod_{i=1}^{5} f(x_i) &\cdot c_{12}(F(x_1),F(x_2)) \cdot c_{13}(F(x_1),F(x_3)) \\ &\cdot c_{14}(F(x_1),F(x_4)) \cdot c_{15}(F(x_1),F(x_5)) \cdot c_{23|1}(F(x_2|x_1),F(x_3|x_1)) \\ &\cdot c_{24|1}(F(x_2|x_1),F(x_4|x_1)) \cdot c_{25|1}(F(x_2|x_1),F(x_5|x_1)) \\ &\cdot c_{34|12}(F(x_3|x_1,x_2),F(x_4|x_1,x_2)) \cdot c_{35|12}(F(x_3|x_1,x_2),F(x_5|x_1,x_2)) \\ &\cdot c_{45|123}(F(x_4|x_1,x_2,x_3),F(x_5|x_1,x_2,x_3)). \end{aligned} \tag{4}$$

The vine-copula requires marginal conditional distributions of the form $F(x|v)$. Joe[18] showed that for every v_j in the vector v, we can write $F(x|v)$ as

$$F(x|v) = \frac{\partial C_{x,v_j|v_{-j}}\{F(x|v_{-j}),F(v_j|v_{-j})\}}{\partial F(v_j|v_{-j})} \tag{5}$$

where $C_{x,v_j|v_{-j}}$ is an arbitrary bivariate copula distribution function. As an example, the C-vine specification in equation (4) requires $F(x_2|x_1)$, which can be written as

$$F(x_2|x_1) = \frac{\partial C_{1,2}\{F(x_2), F(x_1)\}}{\partial F(x_1)} \tag{6}$$

or equation (4) also requires $F(x_3|x_1, x_2)$, which can be calculated as

$$F(x_3|x_1, x_2) = \frac{\partial C_{3,2|1}\{F(x_3|x_1), F(x_2|x_1)\}}{\partial F(x_2|x_1)} \tag{7}$$

Hence, each of the marginal conditional distributions can be calculated from bivariate copulas and marginal distributions.

2.3 Dynamic C-vine Model

Although the C-vine copula used in our study is flexible in a multivariate setting, the dependence parameters are still constant over time. By following Heinen and Valdesogo[24] and Patton[25], we proceeded to introduce the time-varying aspect into the multivariate dependence model. This method consists of the two-step setting: first, we used the C-vine copula to construct multivariate structures and, second, in each building block, the bivariate copula allowed the dependence parameters to be time-varying. As for the time-varying aspect of the bivariate copula, we will allow the dependence parameter of the copula to vary according to the ARMA(1,10)-type process. In accordance with Patton[25], we proposed some time-varying copula candidates in the following manner.

The time-varying Gaussian copula can be defined as

$$\rho_t = \Lambda\left(\psi_0 + \psi_1 \rho_{t-1} + \psi_2 \frac{1}{10} \sum_{j=1}^{10} \Phi^{-1}(u_{t-j}) \Phi^{-1}(v_{t-j})\right) \tag{8}$$

where $\Lambda = (1 - e^{-x})(1 + e^{-x})^{-1}$ is the modified logistic transformation used to maintain the correlation coefficient, ρ_t, belonging to $(-1, 1)$ at all times. For the t copula with the timevarying aspect, $\Phi^{-1}(x)$ is replaced by $t_v^{-1}(x)$.

The time-varying Clayton copula and the time-varying Gumbel copula also assumed the tail dependence parameters to follow the ARMA(1,10) process. We proposed that the time-varying Clayton and Gumbel copulas could be given as follows:

$$\tau_t = \Lambda\left(\psi_0 + \psi_1 \tau_{t-1} + \psi_2 \frac{1}{10} \sum_{j=1}^{10} |(1 - u_{t-j}) - (1 - v_{t-j})|\right), \tag{9}$$

$$\delta_t = \Lambda\left(\psi_0 + \psi_1 \delta_{t-1} + \psi_2 \frac{1}{10} \sum_{j=1}^{10} |(1 - u_{t-j}) - (1 - v_{t-j})|\right) \tag{10}$$

where $\Lambda = (1 + e^{-x})^{-1}$ is the logistical transformation which guarantees that τ_t and δ_t will be between the interval (0,1) at all times.

2.4 Marginal Models

In order to consider the characteristics of the conditional mean and the conditional variance, we constructed the marginal distributions of each returns series by using an ARMA-GARCH model where the type of innovations satisfied a skewed-t distribution. The models employed for the marginal distributions followed Hansen[26], and we denote the log-difference of the crude oil future price or the agricultural commodity future prices as the variable r_t. Thus, the univariate skewed-t GARCH model can be formed as

$$r_{i,t} = \mu_i + \sum_{j=1}^{p} \phi_{i,j} r_{i,t-j} + \sum_{k=1}^{q} \psi_{i,j} \varepsilon_{i,t-k} + \varepsilon_{i,t}, \tag{11}$$

$$\varepsilon_t = \eta_{i,t} \cdot \sqrt{h_{i,t}} \; and \; h_t = \omega_i + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i h_{i,t-1}, \tag{12}$$

where $\varepsilon_{i,t} \sim skewed - t(\nu_i, \lambda_i)$. The skewed-t distribution captured the characteristics of asymmetric heavy tail dependence, in which the model has two parameters λ and ν; they measure the asymmetry and kurtosis behavior of the return series that we expect to observe in our work.

3 The Data and Empirical Results

3.1 Data

In this paper, we used the daily time series data on the one-month futures prices of the five closely linked energy and agricultural commodities, namely, crude oil, ethanol, corn, soybean, and sugar. All that data regarding the futures prices were collected from the Datastream, where corn and soybean commodities and ethanol are traded on the Chicago Board of Trade. Sugar futures is sugar no. 11 futures traded at the ICE market. For crude oil futures prices, we used the Brent crude one-month futures traded at the ICE Market. Our sample covers the period from March 23, 2005, to January 1, 2013. The length of the sample data depended on the ethanol futures available on the Chicago Board of Trade.

The descriptive summaries of all futures returns are demonstrated in Table1. The sample means of all the daily returns series are lower relative to their standard deviations. There is negative skewness in almost all of the returns series, except for corn, thus revealing the fact that the unconditional distributions have longer left tails than right tails. With respect to the excess kurtosis statistics, the data show that all the return series are highly leptokurtic with respect to the normal distribution, indicating that there is a higher probability for extreme movement occurring in these futures markets. Furthermore, the Jarque-Bera test results strongly reject the null hypothesis of normality in all returns series, and that is the reason for the inappropriateness in using the multivariate normal distribution to explain the financial data.

Table 1 Data Description of Returns

	Crude oil	Ethanol	Corn	Soybean	Sugar
Mean	0.0004	0.0003	0.0006	0.0004	0.0004
Max.	0.1271	0.1603	0.1276	0.2032	0.1306
Min.	-0.1095	-0.1365	-0.1041	-0.2341	-0.1237
Std. Dev.	0.0218	0.0198	0.0213	0.0186	0.0238
Skewness	-0.1628	-0.5211	0.0552	-0.7754	-0.2510
Kurtosis	6.5129	8.9927	4.8428	24.0945	5.6761
Jarque-Bera	1038.2380	3086.2750	284.2866	37319.0400	618.4206
Prob.	0.0000	0.0000	0.0000	0.0000	0.0000

3.2 Results of Marginal Models

The findings given in Table2 represent the result of each of the univariate skewed-t GARCH models. The coefficients of the GARCH term, β are strong as they are in the range of 0.91 to 0.95. The sums of the ARCH term and the GARCH term are close to one for all of the series, thus indicating that the conditional variance converges to the long-run variance which takes a longer time. Therefore, we assumed that the selected time series models are adequate to construct the conditional mean and variance of the five return series in our study. For all the asymmetry coefficients of the conditional distribution, λ are positive and significant, implying that the return series of crude oil, ethanol, and the three agricultural commodities, namely, corn, soybean, and sugar, are skewed to the right. The degree of freedom parameters range from 4.02 to 10.00, suggesting that the error terms were not normal.

The correct specifications of the marginal distributions are necessary in the joint copula models. We checked that the marginal models are well-specified by introducing the Box-Ljung test to assess the serial correlation for the first four moments of each return series as well as the Kolmogorov-Smirnov (K-S) test to check the density specification of the marginal distribution assumption. The p-values presented in Table3 suggest that for all series the null hypothesis of no serial correlation could be rejected at the 5% significance level; also, the p-values from the K-S test show that all marginal distribution series can pass at the 5% significance level. Hence, the results imply that all marginal distribution models were in correct specification.

3.3 Results of Copula Models

We selected the structure of the C-vine with the empirical linkage between the energy and agricultural commodities. In fact, the crude oil is a dominant variable that has an influence on the dependencies with the all other related variables. The findings correspond to the objective of our study, which focuses on the dependencies between the crude oil prices and the prices of the other four related variables, which consist of ethanol, corn, soybean, and sugar, and the co-movement between the

Table 2 Parameter Estimates for Marginal Distribution Models

	ω	α	β	λ	ν	LogL
Crude oil	2.38e-06	0.0488	0.9467	0.9169	10.0000	5097.7240
	(1.52e-06)	(9.86e-03)	(1.09e-02)	(2.93e-02)	(1.7670)	
Ethanol	1.25e-05	0.0672	0.9067	0.8826	4.0200	5263.3540
	(6.07e-06)	(2.02e-02)	(3.07e-02)	(2.37e-02)	(3.91e-01)	
Corn	4.66e-06	0.0636	0.9289	1.0190	6.6660	5006.1870
	(2.35e-06)	(1.18e-02)	(1.30e-02)	(3.11e-02)	(9.33e-01)	
Soybean	4.09e-06	0.0495	0.9378	0.9095	5.6510	5482.0720
	(1.36e-06)	(9.74e-03)	(1.11e-02)	(2.52e-02)	(7.47e-01)	
Sugar	3.58e-06	0.0515	0.9455	1.0180	5.2330	4830.8570
	(1.93e-06)	(1.02e-02)	(1.06e-02)	(2.76e-02)	(6.50e-01)	

Note: The numbers in the parentheses are the standard errors.

agricultural commodity and the ethanol prices conditional to the crude oil prices. Also, we investigated the interdependence among the agricultural commodity prices conditional to the crude oil and ethanol prices. By following this method, we could do an ordering with regard to the sequential arrangement of the variables for the C-vine structure, as follows: crude oil, ethanol, corn, soybean, and sugar futures returns.

Table4 reports the estimate of the bivariate copula parameters that were selected according to the AIC and the BIC criteria for each of the building blocks in the appropriate C-vine structures. The sequential procedure is used to select an appropriate C-vine copula for the crude oil and the related copula data, and then used those parameters are used as starting values to calculate the corresponding maximum likelihood estimation (MLE) parameters. The correlation coefficient parameters from Table4 are statistically significant, but they have relatively low dependence.

For the first tree, the optimal choices of the copula are the symmetry t copula, except for the pair of crude oil and soybean that is fitted with the Gaussian copula;

Table 3 Goodness-of-fit Test for Marginal Distributions

	Crude oil	Ethanol	Corn	Soybean	Sugar
Box-Ljung test					
first moment	0.973	0.151	0.395	0.978	0.129
second moment	0.268	0.114	0.876	0.957	0.316
third moment	0.983	0.120	0.865	0.622	0.385
fourth moment	0.640	0.082	0.992	0.991	0.581
K-S test	1.000	1.000	1.000	1.000	1.000

Note: The table presents p-values from the Box-Ljung tests and the K-S tests, respectively.

the degree of freedom of the t copula ranges from 13.52 to 23.27. These facts imply that the co-movements and tail dependence between crude oil and ethanol, corn, and sugar are not strong, especially during the extreme market events. In the second tree, we show the relationship between ethanol and corn, soybean, and sugar, respectively, with conditional prices on crude oil. The best-performing dependence models for ethanol and corn conditional to crude oil, and ethanol and soybean conditional to crude oil are the t copula. The dependence parameters of the t copula are slightly above 0.3 for both the pairs of the return series, especially, the pair of ethanol and corn conditional to crude oil, which has a very low degree of freedom.

Table 4 Structure and Parameter Estimate Results of C-vine Copula for Static Cases

	Copula	par1	par2	λ_U	λ_L	τ	AIC	BIC
$C_{1,2}$	t	0.3293	13.5178	0.0166	0.0166	0.2136	-226.7206	-215.5168
		(0.0203)	(5.0295)					
$C_{1,3}$	t	0.2843	23.2735	0.0012	0.0012	0.1835	-169.9709	-158.7671
		(0.0205)	(10.5845)					
$C_{1,4}$	N	0.3473				0.2258	-268.0105	-262.4086
		(0.0184)						
$C_{1,5}$	t	0.2306	20.5515	0.0014	0.0014	0.1482	-107.3756	-96.1718
		(0.0215)	(10.0642)					
$C_{2,3\|1}$	t	0.4801	3.8110	0.2524	0.2524	0.3188	-596.1210	-584.9172
		(0.0195)	(0.2905)					
$C_{2,4\|1}$	t	0.2920	11.9419	0.0196	0.0196	0.1887	-170.5151	-159.3112
		(0.0223)	(1.6695)					
$C_{2,5\|1}$	N	0.1266				0.0808	-29.2218	-23.6199
		(0.0218)						
$C_{3,4\|1,2}$	t	0.5019	6.8857	0.1450	0.1450	0.3347	-632.2034	-620.9996
		(0.0174)	(1.1653)					
$C_{3,5\|1,2}$	N	0.1396				0.0892	-37.0299	-31.4280
		(0.0213)						
$C_{4,5\|1,2,3}$	R-180G	1.0460			0.0600	0.0439	-12.4627	-6.8608
		(0.0143)						

Note: 1 = crude oil, 2 = ethanol, 3 = corn, 4 = soybean, 5 = sugar. The numbers in the parentheses are the standard errors.

Likewise, the dependence among corn and soybean prices conditional to ethanol and crude oil prices is relatively high, 0.502, and has a very low degree of freedom of the t copula models. The empirical evidence shows that a rise in ethanol production and higher crude oil prices are causing an increase in the prices of related agricultural commodities like corn and soybean, as well as an increase in the probability of extremely symmetric co-movement among ethanol price and the related agricultural prices under the unidirectional prices of crude oil.

The time-varying copula was applied in all the trees within our study by following the ARMA (1,10) process of Patton[25]. The data given in Table5 reveal that all

Table 5 Parameter Estimate Results of C-vine copula for Time-varying Cases

	copula	ψ_0	ψ_1	ψ_2	AIC	BIC
$C_{1,2}$	t	0.0416 (0.0005)	0.1149 (0.0009)	1.8466 (0.0014)	-261.8275	-261.8191
$C_{1,3}$	t	-0.0189 (5.07e-05)	0.0280 (0.0002)	2.1051 (0.0003)	-191.7716	-191.7632
$C_{1,4}$	N	1.3053 (0.0039)	0.3659 (0.0026)	-1.9204 (0.0077)	-274.2329	-274.2245
$C_{1,5}$	t	0.1833 (0.0009)	0.2523 (0.0009)	0.9907 (0.0018)	-126.8346	-126.8262
$C_{2,3\|1}$	t	0.0303 (0.0007)	0.1027 (0.0008)	1.9904 (0.0005)	-749.6977	-749.6893
$C_{2,4\|1}$	t	0.0071 (4.5e-05)	0.0658 (0.0004)	1.9851 (0.0006)	-239.5575	-239.5491
$C_{2,5\|1}$	N	0.5147 (0.0008)	0.0735 (0.0017)	-2.0429 (0.0005)	-35.5384	-35.5300
$C_{3,4\|1,2}$	t	-0.0969 (0.0007)	0.0278 (0.0003)	2.3555 (0.0018)	-639.7591	-639.7507
$C_{3,5\|1,2}$	N	0.3343 (0.0047)	0.0278 (0.0037)	-0.4827 (0.0034)	-38.4581	-38.4497
$C_{4,5\|1,2,3}$	R-180G	-0.9927 (0.8418)	1.2425 (0.6479)	-0.3147 (0.7155)	-17.9531	-17.9447

Note: 1 = crude oil, 2 = ethanol, 3 = corn, 4 = soybean, 5 = sugar. The numbers in the parentheses are the standard errors.

the time-varying copulas were able to improve the performance of the entire static copulas in each tree, consistent with the AIC and the BIC.

The parameter ψ_1 indicates that the persistence effect is relatively low, which implies that the related interdependence structure among the crude oil, ethanol, corn, soybean, and sugar futures returns are generally weak. Meanwhile, the variability of the dependence parameters (ψ_2) is significant and displays greater variability over time on the dependence between each pair of the crude oil, ethanol, corn, soybean, and sugar futures returns.

4 Applications for Portfolio Management

In this section, we want to forecast the Value-at-Risk (VaR) and the Expected Shortfall (ES) of an equally weighted portfolio composed of five assets, crude oil, ethanol, corn, soybean, and sugar futures return. Moreover, we applied the above results to compute the optimal weights of each asset, which is one of the major concerns in the field of portfolio risk management. In order to estimate the VaR and ES of the portfolio, we need to investigate the joint distribution of the return series of the

portfolio. Therefore, in this paper, we used the Monte Carlo simulation and the estimation results of the copula based ARMA-GARCH to measure the VaR and ES of equally weighted portfolio. The procedures to forecast the VaR and ES one day in advance, based on the copulas at 90%, 95%, and 99% confidence levels, were the following:

(1) Using the estimated bivariate copula parameters (C12, C13, C14, and C15) and the inverse distribution of the estimated marginal to simulate a sample of standardized residuals $(\hat{\eta}_{T+1,1},...,\hat{\eta}_{T+1,M})'$ of the assets $1,...,M$.

(2) Computing the ex-ante returns using the estimated ARMA-GARCH parameters for the assets returns $j, j = 1,...,M$,

$$\hat{r}_{T+1,j} = \hat{\mu}_i + \Sigma_{k=1}^{p} \hat{\phi}_{k,j} r_{T+1-k,j} + \Sigma_{k=1}^{q} \hat{\psi}_{k,j} \hat{h}_{T+1-k,j} \hat{\eta}_{T+1-k,j} + \hat{h}_{T+1,j} \hat{\eta}_{T+1,j} \quad (13)$$

(3) The portfolio returns forecast is then given by $\hat{r}_{T+1,p} = \Sigma_{j=1}^{M} w_j \hat{r_{T+1,j}}$, where w_j, $j = 1,...,M$ with $\Sigma_{j=1}^{M} w_j = 1$.

(4) We compute the VaR and ES by taking the 10%, 5% and 1% quantiles of the portfolio returns forecasts.

(5) Loop the steps (1)-(4) for 1000 times, 2000 times, and 5000 times, and compute the averages of the VaR and ES.

Furthermore, to deal with the optimal portfolio allocation, we assumed that the weight in each asset within a portfolio is $w_i = (w_1,...,w_M)\prime$, for M assets namely, crude oil, ethanol, corn, soybean, and sugar futures returns. As for any returns of the portfolio that obtain a minimum ES and correspond to each given investment weight vector w, it becomes the optimal portfolio.

Then the optimal portfolio weights are the solution for the following optimization problem:

$$\begin{aligned} &minES \\ &s.t.\ r = w_j \times r_{j,t+1} \end{aligned} \quad (14)$$

where $0 \leq w_j \leq 1$, $\sum_{j=1}^{M} w_j = 1$, and $j = 1,...,M$. However, the choice of minimum risk measure depends on the behavior of the investor.

Applying the Monte Carlo simulation and the results of the copula based ARMA-GARCH model, as discussed in the previous section, we can simulate the returns at time of the five assets (crude oil, ethanol, corn, soybean, and sugar futures), and calculate the VaR and ES of the portfolio with equal weights, as shown in Table6. Table7 presents the results of the optimal portfolio weight analysis and the one with a minimum portfolio risk.

The data in Table6 reveal that the VaR and ES which are calculated from the time-varying copula are slightly smaller than those calculated from a static copula at the same confidence level. This is because the time-varying copulas consider the dynamic variation of the dependence between each pair of return series that varies over time corresponding to the situation in the financial markets. Moreover, the results are consistent with the data given in section 3.3, in which it is pointed out that the time-varying copulas usually perform better than the constant copulas.

Table 6 VaR and ES of Portfolio under Equally Weighted Criterion

Panel A : static copula						
	0.99		0.95		0.90	
	VaR	ES	VaR	ES	VaR	ES
1000 times	0.0148	0.0178	0.0096	0.0128	0.0072	0.0105
2000 times	0.0147	0.0178	0.0096	0.0127	0.0072	0.0105
5000 times	0.0148	0.0178	0.0096	0.0128	0.0072	0.0105
Panel B : time-varying copula						
	0.99		0.95		0.90	
	VaR	ES	VaR	ES	VaR	ES
1000 times	0.0147	0.0175	0.0095	0.0127	0.0071	0.0104
2000 times	0.0146	0.0176	0.0094	0.0127	0.0072	0.0104
5000 times	0.0146	0.0176	0.0095	0.0127	0.0071	0.0104

Table 7 shows the estimate results of the optimal portfolio weights under minimum ES with different levels of significance. We found that the proportions of investment focused on three assets, namely, crude oil, corn, and soybean, and that they rise along with a gradual increase in risk, for a static copula. This means that investors are willing to take more risks to achieve higher expected returns. However, in the time-varying copula, the weights of corn and soybean become smaller with higher levels of significance. The proportions of investment in the crude oil, corn, and soybean become obviously reduced with higher confidence levels. This gives an indication that the investor changes the weight of each asset in a portfolio with changing market dynamics in order to minimize portfolio risk, which is in correspondence with the ES at each level of significance. Hence, we reached the same conclusion that calculating the VaR and ES of the portfolio by using equally

Table 7 Optimal Weights of Portfolio with Minimum Expected Shortfall

Panel A : static copula							
Alpha	Crude oil	Ethanol	Corn	Soybean	Sugar	ES	VaR
0.10	0.2620	0.1103	0.2068	0.2361	0.1848	0.0104	0.0070
0.05	0.2775	0.0920	0.2067	0.2470	0.1768	0.0125	0.0099
0.01	0.2811	0.0643	0.2169	0.2776	0.1602	0.0158	0.0141
Panel B : time-varying copula							
Alpha	Crude oil	Ethanol	Corn	Soybean	Sugar	ES	VaR
0.10	0.2529	0.1247	0.2236	0.2186	0.1802	0.0103	0.0074
0.05	0.2463	0.1310	0.2235	0.2089	0.1903	0.0123	0.0093
0.01	0.2992	0.1319	0.1976	0.1679	0.2034	0.0153	0.0138

weighted as sets with a time-varying copula has greater advantage and accuracy in forecasting the VaR and ES than with a static copula.

5 Conclusion

In this paper we introduced the C-vine copula based ARMA-GARCH model to construct the interdependence among the energy, biofuel, and agricultural futures returns, especially the dependence between the ethanol and agricultural futures returns conditional to crude oil futures returns, as well as the co-movement among the agricultural futures returns conditional to crude oil and ethanol futures returns. We also used the advantage of a new approach to estimate the VaR and ES of an equally weighted portfolio and calculated the optimal portfolio weights with minimum portfolio risk. The C-vine copula is a very flexible multivariate copula, which can measure symmetry and time variation in the dependence structures of multivariate series of financial returns.

We found evidence that the dependencies among the energy, biofuel, and agricultural futures returns have a significant variability over time and a higher variation of dependence, especially in the dependence between the ethanol and agricultural futures returns conditional to crude oil futures returns, and in the interdependence between corn and soybean futures returns conditional to crude oil and ethanol returns. Furthermore, we found that there was symmetrical tail dependence among the crude oil, ethanol, and agricultural commodity futures returns. This provides an implication that the rise in ethanol production and higher crude oil prices are the cause of the increase in the prices of related agricultural commodities such as corn and soybean. The higher time-varying dependence and symmetric tail dependences between the energy, biofuel, and agricultural commodity returns are an indication that the opportunities for portfolio diversification do become reduced, particularly during the downturn in markets. Finally, we found that calculating the VaR and ES of the portfolio under an equally weighted criterion and estimating the optimal portfolio under minimum portfolio risk using the time-varying copula is better than computing the process using the static copula. These findings can help investors better manage their energy-agricultural commodities portfolio risk.

Acknowledgements. We would like to thank the Prince of Songkla University for providing the PhD fellowship funding for Mr. Phattanan Boonynuphong.

References

1. European Biodiesel Board (EBB), Statistics the EU biodiesel industry (2012), `http://www.ebb-eu.org/stats.php` (cited May 8, 2013)
2. Renewable Fuels Association (RFA), Ethanol industry statistics (2012), `http://www.ethanolrfa.org/pages/statistics` (cited May 8, 2013)
3. Serra, T., Zilberman: Biofuel-related price transmission literature: A review. Energy Economics 37(0), 141–151 (2013)
4. OECD-FAO. Agricultural Outlook 2011-2020. OECD, Paris (2011)

5. Hanson, K.A., Sherman, R., Schluter, G.E.: Sectoral Effects of a World Oil Price Shock: Economywide Linkages to the Agricultural Sector. Journal of Agricultural and Resource Economics 18(1), 96–116 (1993)
6. Baffes, J.: More on the energy/nonenergy price link. Applied Economics Letters 17(16), 1555–1558 (2010)
7. European Biodiesel Board (EBB), Statistics the EU biodiesel industry (2012), http://www.ebb-eu.org/stats.php (cited May 8, 2013)
8. Chang, T.H., Su, H.M.: The substitutive effect of biofuels on fossil fuels in the lower and higher crude oil price periods. Energy 35(7), 2807–2813 (2010)
9. Ciaian, P.: Kancs dA Interdependencies in the energy? bioenergy? food price systems: A cointegration analysis. Resource and Energy Economics 33(1), 326–348 (2011)
10. Natanelov, V., Alam, M.J., McKenzie, A.M., et al.: Is there co-movement of agricultural commodities futures prices and crude oil? Energy Policy 39(9), 4971–4984 (2011)
11. Nazlioglu, S., Soytas, U.: Oil price, agricultural commodity prices, and the dollar: A panel cointegration and causality analysis. Energy Economics 34(4), 1098–1104 (2012)
12. Du, X., Yu, C.L., Hayes, D.J.: Speculation and volatility spillover in the crude oil and agricultural commodity markets: A Bayesian analysis. Energy Economics 33(3), 497–503 (2011)
13. Boonyanuphong, P., Sriboonchitta, S., Chaiboonsri, C.: Modeling Dependency of Crude oil Price and Agricultural Commodity Prices: A Pairwise Copulas Approach. In: Huynh, V.N., Kreinovich, V., Sriboonchitta, S., Suriya, K., et al. (eds.) Uncertainty Analysis in Econometrics with Applications. AISC, vol. 200, pp. 259–270. Springer, Heidelberg (2013)
14. Du, X., McPhail, L.L.: Inside the Black Box: the Price Linkage and Transmission between Energy and Agricultural Markets. The Energy Journal 33(2), 171–194 (2012)
15. Saghaian, S.H.: The Impact of the Oil Sector on Commodity Prices: Correlation or Causation? Journal of Agricultural and Applied Economics 42(3), 477–485 (2010)
16. Serra, T.: Volatility spillovers between food and energy markets: A semiparametric approach. Energy Economics 33(6), 1155–1164 (2011)
17. Serra, T., Zilberman, D., Gil, J.M., et al.: Nonlinearities in the U.S. corn-ethanol-oil-gasoline price system. Agricultural Economics 42(1), 35–45 (2011)
18. Joe, H.: Families of m-Variate Distributions with Given Margins and m(m-1)/2 Bivariate Dependence Parameters. Lecture Notes-Monograph Series 28, 120–141 (1996)
19. Bedford, T., Cooke, R.M.: Vines: A New Graphical Model for Dependent Random Variables. The Annals of Statistics 30(4), 1031–1068 (2002)
20. Embrechts, P., Lindskog, F., McNeil, A.: Modelling Dependence with Copulas and Applications to Risk Management. In: Rachev, S. (ed.) Handbook of Heavy Tailed Distributions in Finance. Elsevier (2003)
21. Nelson, R.: An Introduction to Copulas. Springer, New York (2006)
22. Joe, H.: Multivariate Models and Dependence Concepts. Chapman & Hall, London (1997)
23. Aas, K., Czado, C., Frigessi, A., et al.: Pair-copula constructions of multiple dependence. Insurance: Mathematics and Economics 44, 182–198 (2009)
24. Heinen, A., Valdesogo, A.: Asymmetric CAPM Dependence for Large Dimensions: The Canonical Vine Autoregressive Copula Model (2009), SSRN: http://ssrn.com/abstract=1297506
25. Patton, A.J.: Modelling asymmetric exchange rate dependence. International Economic Review 47(2), 527–556 (2006)
26. Hansen, B.E.: Autoregressive conditional density estimation. International Economic Review 35(3), 705–730 (1994)

An Analysis of Volatility and Dependence between Rubber Spot and Futures Prices Using Copula-Extreme Value Theory

Phattanan Boonyanuphong and Songsak Sriboonchitta

Abstract. This paper aims to estimate the dependency between spot rubber price and futures prices using the copula-extreme value theory based on semi-parametric approaches, which combine copula functions with the conditional extreme value theory to construct the dependence models. The C-EVT model is used to estimate the marginal distributions of the returns of rubber spot price and futures prices that enable the model's flexibility for the tail behavior. Both static and time-varying copulas are applied to construct the dependence structure between the returns of the rubber spot price and the futures prices. The empirical results showed weak spot-futures dependence between the spot rubber price and the futures prices of Thai markets, implying that we could not accept the efficient market hypothesis. However, we found symmetric tail dependence between the spot rubber price and the futures prices of the Singapore, Tokyo, and Shanghai markets. This means that cash rubber price is do minated by the futures prices of the Singapore, Tokyo, and Shanghai markets. The best-fitting dependence models are the time-varying t-copulas, but the tail dependence for all pairs is relatively low. This result means that the futures prices are weak in explaining the changes in spot prices under extreme events.

1 Introduction

In recent years, agricultural commodity prices have experienced strong fluctuations as a consequence of high demand in the emerging market, demand for biofuels, global climate changes, and financial issues. Situations of abnormal oscillation during such periods can be attributed to the highly important role played by the financial

Phattanan Boonyanuphong
Department of Social Sciences, Prince of Songkla University, Pattani 94000, Thailand
e-mail: bphattanan@bauga.pn.psu.ac.th

Songsak Sriboonchitta
Faculty of Economics, Chiang Mai University, Chiang Mai 50200, Thailand
e-mail: songsak@econ.cmu.ac.th, songsakecon@gmail.com

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_27,

instruments. However, financial derivatives such as futures, options, and swaps provide economic benefits: for example, price discovery, information dissemination and efficient allocation of resources, and high trading activities with financial derivatives exacerbated the volatility of the agricultural commodity prices.

During the years from 2005 to 2010, the exchange-traded agricultural derivatives were on a growing trend of up to 29 percent per year, and in 2010, the number of contracts traded in the exchange-traded agricultural derivatives went up to 1,436 million[1]. Under these circumstances, many studies tend to investigate the relationship between the spot and futures prices in the agricultural commodity markets, which concerns the role of futures and arbitrage, and results in an increase in the spot prices of agricultural commodities and market inefficiency.

Natural rubber is one of the agricultural commodities that had been traded in high volumes in the futures market. Therefore, price volatility in the futures rubber market might affect the price fluctuation in the countries of production. Thailand is the worlds largest producer of rubber, which in the period from 2005 to 2010 produced on an average around 31% of the worlds total output. However, rubber prices are not determined only by the Thai market because of the importance of the other markets in the Asian region, such as the Tokyo Commodity Exchange $(TOCOM)$, Singapore Commodity Exchange and Agriculture Futures Exchange $(SICOM)$, and Shanghai Futures Exchange $(SHFE)$ which play a significant role in exchange-trading in the world rubber markets. It is important, therefore, to understand the relationship between the three major rubber futures markets and the Thai spot markets, as well as the efficiency of the Thai rubber futures markets.

The presence of co-movement between the spot and futures prices, mostly in practice, has had many studies done to explain this relationship using different means and techniques. Yang et al.[2] analyzed the price discovery for corn, oats, soybeans, and three major types of wheat, and found that the futures markets play a dominant role in the spot market for storable commodities. Kaur and Rao[3] also found that no significant volatility had been observed in the spot and futures prices of the chosen agricultural commodities, which implies that the co-movement in the spot and futures prices is because of their close relation to each other. Hernandez and Torero[4] confirmed that the spot prices are generally discovered in futures markets and also found that the causal linkages are stronger than the reverse. Similarly, Chang et al.[5] analyzed the relationship between the spot and futures rubber prices to find evidence that there were spillover effects between most pairs of spot and futures rubber prices. In addition, they found asymmetric effects of market shocks on conditional volatility. Other studies, however, discovered contrasting results. Kuiper et al.[6] tested the futures and spot corn price relationship in the CBOT and showed that there is weak exogenous change for both long- and short-run parameters. This implies that the spot price is not just driven by the futures price. Mohan and Love[7] also found that changes in spot prices do not depend on changes in lagged futures prices for the coffee futures market. Analogous to this, Wang and Ke[8] investigated the efficiency of China futures markets for agricultural commodities, namely, wheat and soybean. They found a weak relationship between the cash and futures prices for soybean futures markets and also found that the futures market for wheat was

inefficient. They suggest that speculation and government intervention are the causes of market ine fficiency.

Some interesting studies have previously examined the relationship between spot and futures prices by using econometric techniques, such as the co-integration theory, Granger causality test, and multivariate GARCH models, with the aim of explaining the market efficiency, or the volatility transmission across the markets. However, very few of them have focused on the spot-future relationship of the rubber markets. For example, Siamwalla et al.[21] used cointegration models to test the efficiency of AFET futures markets. In another study, Chang et al.[5] analyzed the relationship between the spot and futures rubber prices by employed the multivariate GARCH models, which is based on some strong assumptions was often used in order to obtain a desirable variance-covariance matrix. Furthermore, multivariate GARCH models were assumed to have a linear relationship with multivariate normal distribution or student-t [9]. These assumptions may be considered as strong assumptions in empirical studies.

Our study attempts to fill this gap by re-examining the spot-future dependence structure of the rubber markets by employing the conditional extreme value theory (C-EVT) and the copulas. The C-EVT method provides better flexibility to the models in the conditional returns distribution due to their stochastic volatility and fat-tailed behavior; in addition to this, the concept of copula offers a more simple and flexible method to model the multivariate dependence. More specifically, we attempt to answer three questions: (1) What is the dependence structure between the spot and futures prices?, (2) Is the dependence symmetric or asymmetric?, and (3) Is there an existence of extreme tail dependence between the spot and futures markets? By answering these questions, we hope to enhance the understanding of the dependence structure between the spot and futures rubber markets, and the market efficiency hypothesis.

2 Econometrics Models

2.1 Copula Models

We considered several static copulas and time-varying copula models that capture the different patterns of dependence. The static copula models employed in our work are the Gaussian, t, (rotated) Clayton, (rotated) Gumbel, and (rotated) Joe copulas.

The bivariate Gaussian is described by $C_N(u,v;\rho) = \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v))$, where the variables u and v are CDFs of the standardized residuals from the marginal models, where $0 \leq u,v \leq 1$ and where $\Phi^{-1}(u)$ and $\Phi^{-1}(v)$ are the inverses of the univariate normal distribution function. The Gaussian copula has zero tail dependence.

Similarly, the t copula is defined by $C_T(u,v;\rho,\nu) = T_{\nu,\rho}(t_\nu^{-1}(u), t_\nu^{-1}(v))$, where $T_{\nu,\rho}$ is the bivariate student-t distribution with degrees of freedom ν and correlation ρ, and $t_\nu^{-1}(u)$ and $t_\nu^{-1}(v)$ are the inverse student-t distribution functions. The t-copula provides the symmetric structure non-zero tail dependence with the same probability of occurrence, λ_U, $\lambda_L > 0$ on both the positive and negative sides.

To deal with the asymmetric tail dependence, we used the (rotated) Clayton, (rotated) Gumbel, and (rotated) Joe copulas which consider lower (upper) tail dependence and negative tail dependence. The Clayton copula is defined as follows $C_{CL}(u,v;\theta) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}$, where $\theta \in (0,\infty)$. The Clayton copula exhibits strong lower tail dependence, where there is a high probability that the trend of the two variables is to go down together at the same time. In contrast, the Gumbel copula exhibits strong upper tail dependence. If the dependences of the two random variables follow the Gumbel copula, then there is a high probability that the trend of the two variables would be to go up together at the same time. The Gumbel copula is given by $C_G(u,v;\theta) = exp(-((-\ln(u))^{1/\theta} + (-\ln(v))^{1/\theta})^{\theta})$, where $\theta \in [1,\infty)$. While the Joe copula is defined by $C_J(u,v;\theta) = 1 - [(1-u)^{\theta} + (1-v)^{\theta} - (1-u)^{\theta}(1-v)^{\theta}]^{-1/\theta}$, where $\theta \geq 0$. The Joe copula has higher dependence in the upper tail than in the lower tail, where it is zero.

All of the above Archimedean copulas display only positive tail dependence, whereas the bivariate random variable has negative dependence and it cannot fit these structures. The rotated copulas are the bivariate copulas that deal with the negative dependence problems by using $(1-u)$ and $(1-v)$, respectively, instead of u and v. According to Nguyen and Bhatti[10], a copula is a rotated Clayton copula if $C_{RC}(u,v;\theta) = u + v - 1 + [(1-u)^{-\theta} + (1-v)^{-\theta}]^{-1/\theta}$. Also, a copula is a rotated Gumbel copula if $C_{RG}(u,v;\theta) = u + v - 1 + exp\{-[(-\ln u)^{\theta} - (-\ln v)^{\theta}]^{-1/\theta}\}$. The rotated Clayton and rotated Gumbel copulas can capture upper (lower) tail dependence instead of the lower (upper) tail dependence as compared to the Clayton and Gumbel copulas, respectively. The property and dependence range of the rotated copulas are also the same as the original copula functions.

In order to allow for time-varying dependence, we will allow the dependence parameter of the copula to vary in accordance with the ARMA(1,10)-type process. Consistent with Patton [11], we propose some time-varying copula candidates as follows.

The time-varying Gaussian copula can be defined as

$$\rho_t = \Lambda\left(\psi_0 + \psi_1 \rho_{t-1} + \psi_2 \frac{1}{10} \sum_{j=1}^{10} \Phi^{-1}(u_{t-j}) \Phi^{-1}(v_{t-j})\right) \tag{1}$$

where $\Lambda = (1-e^{-x})(1+e^{-x})^{-1}$ is the modified logistic transformation used to maintain the correlation coefficient ρ_t belonging to $(-1,1)$ at all times. For the t-copula with the time-varying aspect, $\Phi^{-1}(x)$ is replaced by $t_v^{-1}(x)$.

The time-varying Clayton copula and the time-varying Gumbel copula also assumed the tail dependence parameters to follow the ARMA(1,10) process. We propose that the time-varying Clayton and Gumbel copulas are as follows:

$$\tau_t = \Lambda\left(\psi_0 + \psi_1 \tau_{t-1} + \psi_2 \frac{1}{10} \sum_{j=1}^{10} |(1-u_{t-j}) - (1-v_{t-j})|\right), \tag{2}$$

$$\delta_t = \Lambda(\psi_0 + \psi_1 \delta_{t-1} + \psi_2 \frac{1}{10} \sum_{j=1}^{10} |(1-u_{t-j}) - (1-v_{t-j})|) \tag{3}$$

where $\Lambda = (1+e^{-x})^{-1}$ is the logistic transformation to keep τ_t and δ_t within the interval $(0,1)$ at all times.

2.2 Marginal Models

The extreme value theory (EVT) models provide better estimations in extremely volatile markets than the standard approach, which assumes normal distribution. However, the data series applied in EVT should be independent and identically distributed (i.i.d.) random variables. To deal with these problems, integration is performed of the EVT with the time series model, which develops into the conditional extreme value theory (C-EVT), which is helpful in filtering the data.

(*a*) GARCH Application

To prepare a series of i.i.d. random variables, we propose the use of an $ARMA(p,q) - GJR-GARCH(1,1)$ model by following the method suggested by Glosten et al.[12] to filter the returns series of the rubber prices. The model is defined as

$$r_t = \omega + \sum_{i=1}^{p} \phi_i r_{t-1} + \sum_{i=1}^{q} \psi_i \varepsilon_{t-i} + \varepsilon_t \tag{4}$$

$$\varepsilon_t = G_t \sqrt{h_t} \text{ and } h_t = \alpha_0 + \sum_{i=1}^{r} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{q} \beta_i h_{t-i} \gamma \varepsilon_{t-i}^2 I_{t-i} \tag{5}$$

where $I_{t-i} = 0$ if $\varepsilon_{t-1} > 0$ and $I_{t-i} = 1$ if $\varepsilon_{t-1} < 0$. Also, $\alpha_0 > 0, \alpha_i \geq 0, \beta_i \geq 0$ and $\sum_{i=1}^{k} \alpha_i + \sum_{i=1}^{l} \beta_i < 1$. G_t are the standardized residuals, which satisfies the student-t distribution to compensate for the fat tails often associated with financial returns.

(*b*) Extreme Value Theory

The standardized residuals from the GARCH process, if we only estimate the empirical CDF with a Gaussian kernel, do not catch price jumps caused by extreme events. In our study, we apply EVT to the residuals of each series, which provides better estimation for the tails of distribution than the Gaussian kernel. There are two methods for the extreme value, which can be constructed by the block maxima or the peaks-over-threshold (POT) method[13]. Therefore, we select the peaks-over-threshold models because they model all the large observations which exceed a high threshold and are more suitable for use for the data of the extreme outcomes.

(*c*) **Generalized Pareto Distribution**

Let $G(g_1, g_2, ..., g_n)$ be a sequence of independent and identically distributed (iid) random variables which represents the standardized residuals of the returns series. The excess distributions $F(g)$ that explain the probability of G exceeds a fixed threshold u, which can be estimated using a generalized Pareto distribution (GPD) fitted using the maximum likelihood method[14]. The tail estimator is defined by

$$F(\hat{g}) = 1 - \frac{k}{n}\left[1 + \hat{\xi}\frac{g-u}{\hat{\beta}}\right]^{-1/\hat{\xi}}, \; for \; g > u \tag{6}$$

where β and ξ are the scale parameter and the shape parameter, respectively, n is the number of observation, and k is the number of observations exceeding the threshold u. The value of shape parameter (ξ) represents the tailed behavior of the distributions. When $\xi = 0$, the distribution belongs to the Gumbel type, as with normal, log normal, and exponential distributions. When $\xi < 0$, the distribution belongs to the Weibull family, as with the beta and uniform distributions. And, when $\xi > 0$, the distribution belongs to the heavy-tailed family, including the Pareto, log-gamma, and t distributions.

2.3 Estimation and Testing

We estimate the parameters in two stages. In the first stage, the scale and sharp parameters of the marginal distributions are estimated via maximum log-likelihood, as follows:

$$\hat{\theta}_i = arg\max \sum_{t=1}^{N} log f_i(x_{i,t}, \theta_i) \tag{7}$$

The GPD (for the tail of the distribution) and the Gaussian kernel (for the interior of the distribution) distributions are employed to construct the marginal distributions u_i and v_i, which are used to estimate the copula parameters. In the second stage, the parameters of the copula are estimated via maximum log-likelihood, as follows:

$$\hat{\delta}_c = arg\max \sum_{t=1}^{N} logc(\hat{u}_{i,t}, \hat{v}_{i,t}, \delta_c) \tag{8}$$

The performances of the different copula models are important after fitting the copulas with the marginal distributions, and were evaluated as follows: (1) using the Akaike Information Criterion (AIC) and the Bayesian Information Criteria (BIC) in agreement with Brechmann[15] and Bhatti and Nguyen[16] to evaluate the goodness of fit.The copula model with the smallest AIC and/or BIC should be considered as the best fit; and (2) using two tests based on the Kendalls transform to perform

the goodness of fit of the copulas, which calculate the Cramer-von Mises (CvM) and Kolmogorov-Smirnov (K-S) statistics as well as the corresponding p-values using bootstrapping[17].

3 Data

In our analysis, we focused on the ribbed smoked sheet no.3 (RSS3) by using the daily closing prices of the spot and futures returns from the period 1 June 2004 to 14 September 2012, making a total of 2,164 observations. For the spot data set, we used the averages of the FOB Bangkok and FOB Songkla prices (FOB) as a proxy for the daily spot prices, whereas the futures data set comprised the four daily futures prices from different futures markets including the Agricultural Futures Exchange of Thailand $(AFET)$, Singapore Commodity Exchange and Agriculture Futures Exchange $(SICOM)$, Tokyo Commodity Exchange $(TOCOM)$ and, Shanghai Futures Exchange $(SHFE)$. Almost all of the data were obtained from DataStream, except AFET which was collected from the website of the Agricultural Futures Exchange of Thailand. The daily prices were calculated as returns of market i at time t, as follows $r_{i,t} = log(P_{i,t+1}) - log(P_{i,t})$ wh ere $P_{i,t}$ and $P_{i,t+1}$ are the closing prices of spot or futures for days t and $t+1$, respectively.

Table 1 Descriptive Statistics of Returns

	FOB	AFET	SICOM	TOCOM	SHFE
Mean	0.0002	0.0003	0.0004	0.0002	0.0002
Min.	-0.0825	-0.0859	-0.1340	-0.2573	-0.1448
Max.	0.0464	0.0680	0.1086	0.1058	0.0849
Std. Dev.	0.0115	0.0157	0.0170	0.0220	0.0156
Skew.	-0.9468	-0.4142	-0.6086	-1.1466	-1.0319
Kurt.	10.6150	6.7141	10.7314	14.2414	11.6586
JB stat.	0.0010	0.0010	0.0010	0.0010	0.0010

The descriptive summaries of the spot and futures returns are demonstrated in Table1. The sample means of all the daily returns series are lower relative to their standard deviations, and the standard deviations range from 0.012 to 0.022, showing relatively weak volatility in all the data series. The negative skewness in all the returns series reveals that the distributions have long tails to the left. As regards the excess kurtosis statistics, the data show that all the returns series are highly leptokurtic with respect to the normal distribution. Furthermore, the Jarque-Bera test results strongly reject the null hypothesis of the normality in all the returns series, and that is the reason for the inappropriateness in using multivariate normal distribution to explain financial data.

4 Empirical Results

4.1 Results for Marginal Models

In order to deal with the stochastic volatility and fat-tailed behavior of the conditional returns distribution, we used the C-EVT which was suggested by McNeil and Frey[18] and Fernandez[19]. Therefore, we employed the ARMA-GJR-GARCH to model for volatility, in which the TOCOM and SHFE returns series are fitted with the ARMA(4,2)-GJR-GARCH(1,1), whereas the AFET, FOB, and SICOM returns series are fitted with the ARMA(2,1)-GJR-GARCH(1,1), the ARMA(4,3)-GJR-GARCH(1,1), and the ARMA(4,1)-GJR-GARCH(1,1), respectively. Then, the GPD is used to estimate the tails of the distributions in each returns series.

Table 2 Estimated Tails from GPD

	FOB	AFET	SICOM	TOCOM	SHFE
Upper tail					
ξ	0.1027	-0.0561	-0.0543	-0.1726	0.0275
	(0.0696)	(0.0112)	(0.0085)	(0.0599)	(0.0061)
β	0.7074	0.6874	0.6353	0.5811	0.5974
	(0.0696)	(0.0051)	(0.0066)	(0.0492)	(0.0153)
Lower tail					
ξ	0.0322	-0.1004	0.0323	0.0778	0.2003
	(0.0930)	(0.0109)	(0.0304)	(0.0573)	(0.0744)
β	0.7954	0.7881	0.7070	0.6381	0.5771
	(0.1104)	(0.0698)	(0.0411)	(0.0567)	(0.0523)

Note: The numbers in the parentheses are the standard errors.

Table2 summarizes the EVT estimations of the tails from the GPD of each returns series. Of the upper tails, the negative shape parameters (ξ) of the futures returns in AFET, SICOM, and TOCOM are significant, which reveals that the upper tails are finite. In contrast, the shape parameters of the upper tail of the returns distributions in FOB and SHFE are statistically significantly fat-tailed. Of the lower tails, the TOCOM and SHFE have significant fat tails on the left side of the returns distribution, while AFET has significant finite tails, but the shapes of the lower tail of the returns distributions in FOB and SICOM are not significant. The correct specifications of the marginal distributions are necessary in the joint copula models. If the marginal distributions are correctly specified, then the probability transformations will be i.i.d. uniform (0,1)[20], and hence the copula model will also be correctly specified. In accordance with the technique promulgated by Patton[11], we used the Lagrange Multiplier (LM) test for serial in dependence of the probability transforms and the Kolmogorov-Smirnov (K-S) test for uniform (0,1) to test the specification of the marginal distribution assumption. First, the LM independence tests are used

Table 3 Goodness-of-fit Test for Marginal Distributions

	FOB	AFET	SICOM	TOCOM	SHFE
LM test					
first moment	0.16	0.11	0.40	0.87	0.24
second moment	0.13	0.20	0.31	0.64	0.19
third moment	0.20	0.12	0.02	0.77	0.26
fourth moment	0.77	0.12	0.31	0.59	0.33
K-S test	0.75	0.07	0.73	0.30	0.40

Note: The table presents p-values from the LM tests and K-S tests, respectively.

to explain the independence of the first four moments of the variables u_t and v_t. For this, we regress $(\hat{u}_t - \bar{u}_t)^k$ and $(\hat{v}_t - \bar{v}_t)^k$ on 20 lags of both variables for k=1, 2, 3, 4. The LM test statistic $(T-20)R^2$ for each regression follows the asymptotic χ^2_{20} distributions under the null hypothesis, no se rial correlation. Table3 summarizes all the marginal distribution models that could not reject the null hypothesis of no serial correlation at the 5% significance level for all series. Second, we employed the K-S tests to test the null hypothesis that $\hat{u}_t$ and $\hat{v}_t$ are uniform (0,1) according to the specifications of the marginal assumptions, and also that the p-values from the K-S tests, given in Table3, present the fact that all the marginal distribution series can pass at the 5% significance level. The results provide significant evidence that the marginal distribution models are correctly specified. Therefore, it is evident that the copula model can correctly measure the dependence structures of two returns series.

4.2 Results for Copula Models

Table4 presents a report of the parameter estimates for the constant dependence copulas. For all pairs, the parameter dependences of the Gaussian and t-copulas are positive and strongly significant. The correlation coefficient, ρ, of the Gaussian and t-copulas range from 0.262 to 0.366 for the pairs of FOB-SICOM, FOB-SHFE, and FOB-TOCOM, whereas the correlation coefficient of FOB-AFET is close to 0.08. This means that the spot returns are positive and generally strong in relation with the futures returns, except for the interdependence between FOB and AFET, which is relatively low. The degrees of freedom of the t-copula range from 10.96 t0 28.031, indicating intermediate extreme co-movements and tail dependence in each pair. In fact, all the tail dependence values of the t-copula were relatively low, and the tail dependence values between FOB and AFET, FOB and SICOM, FOB and TOCOM, and FOB and SHFE were 0.0002, 0.0022, 0.0324, and 0.0249, respectively.

In considering asymmetric tail dependence, the parameter estimates for (rotated) Clayton, (rotated) Gumbel, and (rotated) Joe are positively significant. This shows that the dependence between the spot returns and the futures returns vary during the different states of an economy. The Clayton, rotated Gumbel, and rotated Joe

Table 4 Static Copula Estimates of Spot PriceFutures Prices

		N	t	C	G	J	RC	RG	RJ
AFET	ρ	0.083	0.078	0.101	1.039	1.042	0.064	1.054	1.073
		(0.021)	(0.022)	(0.025)	(0.013)	(0.018)	(0.024)	(0.013)	(0.019)
	ν		28.030						
			(19.346)						
	λ_L		0.0002	0.001				0.069	0.092
	λ_U		0.0002		0.052	0.055	0.0001		
	AIC	-12.189	-12.436	-17.164	-7.997	-4.114	-5.659	-20.131	-19.904
	BIC	-6.511	-1.080	-11.485	-2.319	1.563	0.018	-14.452	-19.904
SICOM	ρ	0.262	0.265	0.294	1.173	1.201	0.281	1.178	1.217
		(0.019)	(0.021)	(0.031)	(0.018)	(0.026)	(0.031)	(0.018)	(0.027)
	ν		19.885						
			(9.0375)						
	λ_L		0.002	0.095				0.199	0.233
	λ_U		0.002		0.195	0.219	0.084		
	AIC	-143.134	-146.664	-112.302	-120.978	-80.545	-100.499	-125.956	-90.024
	BIC	-137.456	-137.656	-137.456	-115.301	-74.868	-94.821	-120.278	-84.347
TOCOM	ρ	0.362	0.366	0.437	1.277	1.339	0.440	1.277	1.339
		(0.018)	(0.019)	(0.033)	(0.021)	(0.031)	(0.033)	(0.021)	(0.031)
	ν		11.485						
			(3.036)						
	λ_L		0.032	0.205				0.278	0.322
	λ_U		0.032		0.279	0.322	0.207		
	AIC	-283.858	-300.627	-217.726	-261.625	-189.415	-218.587	-259.243	-186.523
	BIC	-278.181	-289.272	-212.049	-255.948	-183.737	-212.910	-253.566	-180.846
SHFE	ρ	0.283	0.292	0.344	1.198	1.229	0.310	1.210	1.264
		(0.019)	(0.021)	(0.032)	(0.019)	(0.027)	(0.032)	(0.019)	(0.028)
	ν		10.969						
			(2.941)						
	λ_L		0.025	0.133				0.227	0.270
	λ_U		0.025		0.217	0.242	0.107		
	AIC	-169.136	-183.556	-142.212	-151.429	-101.349	-118.790	-168.947	-128.4962
	BIC	-163.459	-172.201	-136.534	-145.752	-95.672	-113.112	-163.269	-122.818

Note: Note: The numbers in the parentheses are the standard errors. λ_L and λ_U present the lower and upper tail dependence values. N = "Gaussian", C = "Clayton", RC ="Rotated Clayton", G = "Gumbel", RG = "Rotated Gumbel"and J = "Rotated Joe", RJ = "Rotated Joe"

copulas represent the lower tail dependence. In contrast, in the rotated Clayton, Gumbel, and Joe copulas, the upper tail dependence is represented. In general, the lower tail dependence and the upper tail dependence in each pair are likely to be the same, signifying that there is a high possibility of the values of spot price and futures prices crashing (booming) together at the same time.

However, the AIC and the BIC demonstrated that the rotated Gumbel displayed the best performance among the static copulas for the pair of FOB and AFET, whereas the t-copula is best suited for the pairs of FOB and SICOM, and TOCOM and SHFE, respectively. Table5 presents the goodness of fit of the copula models according to the Kendalls transform process[17]. The results also showed that the rotated Gumbel copula is acceptable for the pair of FOB and AFET, while the t-copula is acceptable for the pairs of FOB and SICOM, and TOCOM and SHFE. As a result, we could not reject the null hypothesis that the rotated Gumbel copula and the t-copula are the suitable ones among the constant copula models.

Finally, the time-varying copula, as reported in Table6, reveals that the time-varying rotated Gumbel copula can improve the performance of all the other copula specifications for the pair of FOB and AFET. Likewise, the time-varying t-copula would be able to improve the performance of all the other copula specifications for the pairs of FOB and SICOM, and TOCOM and SHFE, which are consistent with the AIC and the BIC, respectively.

Consequently, the results from our analysis can be concluded as follows:

(1) The dependence of the spotfutures prices is generally strong, except for the dependence parameters between FOB and AFET, which are relatively low, which implies that the efficient market hypothesis could not be accepted. This finding is associated with the fact that the AFET market had low liquidity and low total trading volume besides having to deal with high policy intervention in the rubber market by the Thailand government[6, 8, 21]. However, our empirical results show that FOB co-moves with SICOM, TOCOM, and SHFE. This indicates that the spot rubber price in Thailand is dominated by the futures prices of the SICOM, TOCOM, and SHFE markets.

(2) In the dependence structure between the spot prices and the futures prices, there exists extreme dependence for the pairs of FOB-SICOM, FOB-TOCOM, and

Table 5 Goodness of Fit of Cramer-von Mises and K-S Statistics

	AFET		SICOM		TOCOM		SHFE	
	CvM	KS	CvM	KS	CvM	KS	CvM	KS
Gaussian	0.07	0.07	0.60	0.60	1.00	1.00	0.80	0.60
Clayton	0.00	0.10	0.00	0.00	0.00	0.00	0.00	0.00
Gumbel	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Joe	0.60	0.30	1.00	1.00	1.00	1.00	1.00	1.00
Rotated Clayton	0.30	0.30	0.00	0.00	0.00	0.00	0.00	0.00
Rotated Gumbel	0.07	0.05	0.00	0.00	0.00	0.10	0.10	0.10
Rotated Joe	0.70	0.70	1.00	1.00	1.00	1.00	1.00	1.00
t	0.30	0.50	0.50	0.70	1.00	1.00	0.80	0.70

Note: This table shows the p-values of the Cramer-von Mises and K-S statistics.

Table 6 Copula-Time Varying Estimates of Spot PriceFutures Prices

		N	t	RC	RG
AFET	ψ_0	0.220	0.212	-3.623	-1.271
		(0.004)	(0.005)	(0.012)	(0.009)
	ψ_1	-0.116	-0.123	-1.627	1.467
		(0.004)	(0.003)	(0.006)	(0.008)
	ψ_2	-0.514	-0.542	-0.297	-0.144
		(0.046)	(0.038)	(0.003)	(0.003)
	AIC	-14.659	-17.009	-9.039	-23.807
	BIC	-14.651	-17.001	-9.031	-23.800
SICOM	ψ_0	1.108	1.127	0.048	1.930
		(0.001)	(0.001)	(0.000)	(0.002)
	ψ_1	-0.262	-0.249	-0.842	-1.082
		(0.001)	(0.003)	(0.002)	(0.011)
	ψ_2	-1.878	-1.897	0.889	-0.842
		(0.002)	(0.002)	(0.000)	(0.006)
	AIC	-147.235	-152.921	-114.085	-131.551
	BIC	-147.227	-152.913	-114.077	-131.543
TOCOM	ψ_0	0.409	0.288	0.052	0.502
		(0.002)	(0.001)	(0.023)	(0.007)
	ψ_1	0.256	0.163	-0.289	0.222
		(0.005)	(0.002)	(0.137)	(0.005)
	ψ_2	0.739	1.141	0.982	-1.032
		(0.007)	(0.001)	(0.010)	(0.006)
	AIC	-289.126	-307.965	-265.157	-279.806
	BIC	-272.091	-290.930	-265.149	-262.771
SHFE	ψ_0	0.005	0.007	0.024	-0.392
		(0.000)	(0.000)	(0.000)	(0.003)
	ψ_1	0.058	0.049	-0.462	0.762
		(0.000)	(0.000)	(0.001)	(0.002)
	ψ_2	1.995	1.995	0.936	-0.277
		(0.001)	(0.001)	(0.000)	(0.003)
	AIC	-182.476	-195.703	-123.335	-174.343
	BIC	-165.441	-178.668	-123.327	-157.308

Note: The numbers in the parentheses are the standard errors. N = "Gaussian", RC ="Rotated Clayton",RG = "Rotated Gumbel"

FOB-SHFE, and the best-fitting dependence model is the time-varying t-copula. However, the tail dependence for all the pairs is generally weak. This evidence reflects the fact that the futures prices are weak in explaining the changes in the spot prices under extreme situations. According to the discussion given in Garcia et al.[22], the grain futures markets fail to explain the convergence of the spot and futures prices due to the fact that the futures prices at expiration are up to 35% above the cash grain price.

5 Conclusion

In this study, we examined the co-movement between the cash rubber market and the futures markets, including Agricultural Futures Exchange of Thailand ($AFET$), Singapore Commodity Exchange and Agriculture Futures Exchange($SICOM$), Tokyo Commodity Exchange ($TOCOM$), and Shanghai Futures Exchange ($SHFE$) by using copula-EVT models with time invariant and time varying. In order to choose the correct specification copula function, we used C-EVT, that is, GPD, to model the tails of the marginal distributions because the GPDs are appropriate in being able to explain the tail behaviors of financial data. The results revealed that the interdependence between the spot rubber price and the futures price of the AFET market is relatively low, indicating that we could not accept the efficient market hypothesis. However, we found symmetric tail dependence between the spot rubber price and the futures prices of the SICOM, TOCOM, and SHFE markets, res pectively. This means that the cash rubber price is dominated by the futures prices of the SICOM, TOCOM, and SHFE markets. The best-fitting dependence models are the time-varying t-copulas, but the tail dependence for all the pairs is relatively low. This means that the futures prices are weak in being able to explain the change in the spot prices under extreme events; also, the dependence parameters are very volatile over time and deviate from their constant levels.

The important implication from these results is the need of a good price transmission system in which the futures market has to be closely related to the fundamentals and be good indicators for the spot market. Therefore, the government should consider the policy of promoting the market efficiency of the Agricultural Futures Exchange of Thailand and reduce policy intervention in the rubber market which could be unfavorable to market development. Hedgers and investors can benefit from this information by hedging in the futures market. However, they should be aware in the case of an extreme event of the fact that futures prices are weak in explaining the reason for the changes in the spot prices.

References

1. WFE, IOMA. IOMA/IOCA Derivatives Market Survey 2010. World Federation of Exchanges (2011)
2. Yang, J., Bessler, D.A., Leatham, D.J.: Asset storability and price discovery in commodity futures markets: A new look. Journal of Futures Markets 21(3), 279–300 (2001)
3. Kaur, G., Rao, D.N.: Do the Spot Prices Influence the Pricing of Future Contracts? An Empirical Study of Price Volatility of Future Contracts of Select Agricultural Commodities Traded on NCDEX, India (2009), http://dx.doi.org/10.2139/ssrn.1469700
4. Hernandez, M., Torero, M.: Examining the dynamic relation between spot and futures prices of agricultural commodities. Commodity Market Review. FAO, Rome (2010)
5. Chang, C.L., Khamkaew, T., McAleer, M., et al.: Modelling conditional correlations in the volatility of Asian rubber spot and futures returns. Mathematics and Computers in Simulation 81(7), 1482–1490 (2011)

6. Kuiper, W.E., Pennings, J.M.E., Meulenberg, M.T.G.: Identification by full adjustment: evidence from the relationship between futures and spots prices. European Review of Agricultural Economics 29, 67–84 (2002)
7. Mohan, S., Love, J.: Coffee futures: role in reducing coffee producers' price risk. Journal of International Development 16(7), 983–1002 (2002)
8. Wang, H.H., Ke, B.: Efficiency tests of agricultural commodity futures markets in China. The Australian Journal of Agricultural and Resource Economics 49, 125–141 (2005)
9. Sriboonchitta, S., Nguyen, H.T., Wiboonpongse, A., et al.: Modeling volatility and dependency of agricultural price and production indices of Thailand: Static versus time-varying copulas. International Journal of Approximate Reasoning 54(6), 793–808 (2013)
10. Nguyen, C.C., Bhatti, M.I.: Copula model dependency between oil prices and stock markets: Evidence from China and Vietnam. Journal of International Financial Markets, Institutions and Money 22(4), 758–773 (2012)
11. Patton, A.J.: Modelling asymmetric exchange rate dependence. International Economic Review 47(2), 527–556 (2006)
12. Glosten, L.R., Jagannathan, R., Runkle, D.E.: On the Relation between Expected Value and the Volatility of the Nominal Excess Return on Stocks. The Journal of Finance 48, 1779–1801 (1993)
13. McNeil, A.J., Frey, R., Embrechts, P.: Quantitative Risk Management: Concepts, Techniques and Tools. Princeton University Press (2005)
14. Hsu, C.P., Huang, C.W., Chiou, W.J.: Effectiveness of copula-extreme value theory in estimating value-at-risk: empirical evidence from Asian emerging markets. Rev. Quant. Finan. Acc. 39(4), 447–468 (2012)
15. Brechmann, E.C.: Truncated and simplied regular vines and their applications. Technische Universitaet Muenchen (2010)
16. Bhatti, M.I., Nguyen, C.C.: Diversification evidence from international equity markets using extreme values and stochastic copulas. Journal of International Financial Markets, Institutions and Money 22(3), 622–646 (2012)
17. Genest, C., Rémillard, B., Beaudoin, D.: Goodness-of-fit tests for copulas: A review and a power study. Insurance: Mathematics and Economics 44(2), 199–213 (2009)
18. McNeil, A.J., Frey, R.: Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. Journal of Empirical Finance 7(3-4), 271–300 (2000)
19. Fernandez, V.: Extreme value theory and value at risk. Revista de Analisis Economico 18(1), 86–102 (2003)
20. Diebold, F.X., Gunther, T.A., Tay, A.S.: Evaluating Density Forecasts with Applications to Financial Risk Management. International Economic Review 39(4), 863–883 (1998)
21. Siamwalla, A., Pomlaktong, N., Panpiemras, J., et al.: Evaluation of the success in the establishment of the Agricultural Futures Exchange of Thailand. Agricultural Futures Trading Commission, Thailand (2007)
22. Garcia, P., Irwin, S.H., Smith, A.: Futures Market Failure (2011), http://dx.doi.org/10.2139/ssrn.1950262

Effect of Markets Temperature on Stock-Price: Monte Carlo Simulation on Spin Model

Arjaree Thongon, Songsak Sriboonchitta, and Yongyut Laosiritaworn

Abstract. In this study, we used the Monte Carlo simulations to investigate the phenomena in the stock-price market which we considered as a function of temperature and external field which reflect the effects of the environment (e.g., access to external information). The Monte Carlo simulation was used to simulate the Ising model with heat-bath algorithm. The results show that the average orientation of the agents varies with the external field at constant temperature. In other words, the agents always buy when they get good news. And at high temperature, with constant positive external field, the average orientation of the agents is decreased to near zero.

1 Introduction

Bubbles and crashes always take place in the stock market. When they occur, they impact the economy in a tremendous manner. So, their in-depth study is extremely important. There are many models that are used to explain these phenomena. For example, the random matrix theory (RMT) [1],[2] finds application in finance. The RMT which is a matrix-valued random variable has recently been applied to noise filtering in financial time series. The advantage of the RMT method is that it eliminates random properties from the financial time series. But the value of the eigenvalue in this method increases in proportion to the number of stocks. So, the RMT method is difficult to use in cases where there are large amounts of stocks. Another method is the minimum spanning tree (MST) method [3]. This method constructs the asset tree using the correlations between the stock prices [4]. It is very difficult to form links between the various stocks.

Arjaree Thongon · Yongyut Laosiritaworn
Faculty of Science, Chiang Mai University, Chiang Mai 50200 Thailand
e-mail: arjare@hotmail.com

Songsak Sriboonchitta
Faculty of Economics, Chiang Mai University, Chiang Mai 50200 Thailand
e-mail: songsakecon@gmail.com

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_28,

So, a simple spin model such as the Ising model is widely used to study stock markets for explaining the bubbles and crashes that occur in them. To improve this model, the interacting-agent model is used. The interacting-agent model, motivated by the dynamic of traders in a market, has two main influences: The first is the strength of others trading on the interested trader, and the second is the strength of the reaction of the interested trader. There are many methods to solve using model. One example would be the use of mean field approximation which is a deterministic algorithm. From this method, some parameters that can describe phase transition (Bull phase to Bear phase, or Bear phase to Bull phase) can be calculated [5]. After that, more parameters would be added into the term, the strength of the other trader, such as fluctuating interaction network and fluctuating environment, to improve the model [6]. But the results are not effective because this method works only for a stable state. So, to describe phase transition, Monte Carlo simulation, which is a stochastics algorithm, is used. In this model, the probability for updating agents or spins can be measured by using the heat-bath dynamics [7]. Like in the mean field approximation, the first interacting-agent model has two main influences. The result from this method is compared to the real result from the Dow Jones daily changes. It is found that this method closely corresponds to the trading volume in the stock market [8]. Thereafter, to improve this method, more parameters, as done in the previous method, are added [9]. But the actual behavior of the investors is not limited to just buying and selling. Many investors choose to wait and watch at the stock market without making a purchase or a sale. Hence, the state that describes this situation is added in the value of 0 [10]. From the above simulation, it was found that the prices of assets could be predicted. If the demand is high (average spin is more than 0), the assets will be more expensive. On the other hand, if the supply is high (average spin is less than 0), the assets will be cheaper.

In the previous works, the temperature and the external field were considered constant in each step of time. But in reality, the money in the stock market does not remain constant. The investors often invest in markets that have better returns. This means that the temperature in the stock markets is not constant either. So, in this work, we will improve the model by changing the temperature in each step of time. And, in future work, we will change the external field in each step of time.

2 Ising Model

The Ising model is a mathematical model that is commonly used in ferromagnetism in statistical mechanics. The model consists of discrete variables that represent the magnetic dipole moments of atomic spins which can be in one of the two states (spin up or spin down). In the study of stock market, the spin is used to describe the investments behaviors: +1 when the agent is buying and -1 when the agent is selling.

In physics, the factor energy is used to find the probability of the changing state of the spin. For this model, the energy can be found from the Hamiltonian function

$$H = -\sum_{<ij>} J_{ij}\sigma_i\sigma_j - \sum_i h_i\sigma_i$$

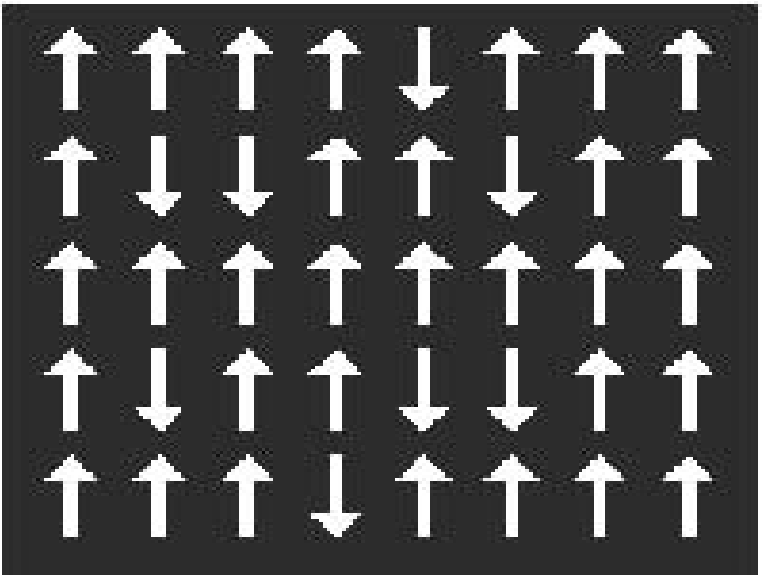


Fig. 1 The alignment of the spins of the two-dimensional square lattice as an Ising model. The spin up is configured as +1 and the spin down is configured as -1.

where the first sum is over pairs of spins (every pair is counted once). J_{ij} is the exchange interaction between sites i and j, and σ is the spin. The second term is for the external field. This can also be written as

$$H = -\sum_i I_i \sigma_i$$

where $I_i = \sum_j J_{ij}\sigma_j + h_i$ is the local field.

So, in this work, we used the local field to find the probability of changing the orientation of the agent (σ) in the Monte Carlo simulation.

3 Monte Carlo Simulation

In this work, the orientation of the agent $i(\sigma_i)$ at time $t+1$ depends on the local field [11]

$$I_i(t) = \frac{1}{z}\sum_{j=1}^{N} A_{ij}\sigma_j(t) + h_i$$

where z is the mean coordination number (the mean number of the non-zero connections between the agents), A_{ij} is the interaction strengths per agent, or the influence from the others, and h_i stands for the external fields reflecting the effect of the environment (e.g., access to external information). In the general situation, the values of A_{ij} and $h_i(t)$ are not static. But in this work, we assume that A_{ij} is constant and

equal to 1. It's mean that the influence from the others is equal. And h_i in the general situation should be a function of some parameters such as time, volume etc. which we will evaluate the relationship of function with other parameters in next work. So, in this work, we will assume h_i is constant.

The configuration probability which is used in the decision of the agent is given by the heat-bath algorithm with inverse temperature ($\beta = 1/T$):

$$\text{probability}(p) = \frac{1}{\{1+\exp[-2\beta I_i(t)]\}}$$

For updating the orientations,

$$\sigma_i(t+1) = \begin{cases} 1 & \text{with probability } p \\ -1 & \text{with probability } 1-p \end{cases}$$

This section shows all the parameters that affect the decisions of investors. The parameter which we were interested in this work is the temperature which is described in the next section.

4 Temperature

For finding the temperature of the stock-price market, we consider the Boltzmann-Gibbs distribution [12]

$$P(m) = Ce^{-m/T}$$

where m is money and T is temperature. To find the equation between C and T, the normalization conditions are used.

$$\int_0^\infty P(m)\,dm = 1 \quad \text{and} \quad \int_0^\infty mP(m)\,dm = M/N$$

where M is the total money (assume that M become form the total money of this situation) and N stands for the total agents. By solving the normalization conditions, we find that

$$C = 1/T \quad \text{and} \quad T = M/N$$

The price [13] of an asset can be found out as

$$\Pr(t_n) = \Pr(t_{n-1})\, e^{S(t_{n-1})}$$

where $\Pr(t)$is price, t is time, n is a time index, and $S(t)$ is the average orientation of the agents:

$$S(t_n) = \frac{1}{N}\sum_{i=1}^{N} \sigma_i(t_n)$$

The total money in this model will be calculated from the total number of successful trade which come from the number of buyers and sellers match fit and the price at that time.

$$T(t_n) \propto \left(\left(\frac{N - (N \times |S(t_{n-1})|)}{2} \right) \times \Pr(t_{n-1}) \right) / N$$

Or $T(t_n) \propto \left(\frac{1-|S(t_{n-1})|}{2} \right) \times \Pr(t_{n-1})$

The term $\left(\frac{N-(N\times|S(t_{n-1})|)}{2} \right)$ is the number of successful trade.

5 Result and Discussion

In this work, the orientation of the agent, or the total number of agents (N), is 10,000, the interaction strength per agent is constant ($A_{ij} = A = 1$), and the time-unit is defined in terms of Monte Carlo step (mcs) which is widely used in Monte Carlo Simulation. In each step of time, all agents will determine to buy or sell based on probability in each step. From the simulations, the effect of the temperature on the average orientation of the agents at a fixed external field has been revealed. As can be seen in Figure 2, the average orientation of the agents shows a decrease, and the error bar shows an increase when the temperature is high. This result is relative to probability. When the local field is positive and the temperature is very close to zero, the probability will be very close to 1. Also, when the temperature is very high, the probability will be 1/2. So, when the temperatures rise, the orientation of the agents shows the tendency to fluctuate more.

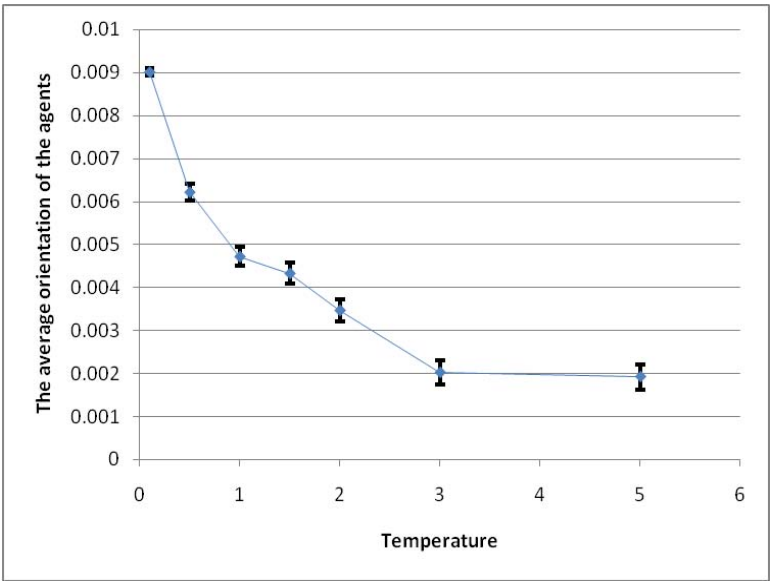


Fig. 2 The average orientation of the agents as a function of temperature (T) at fixed external field (h) = 0.01

On considering the effect of the external field only, Figure 3 shows that the average orientation of the agents varies with the external field, and also that it is a linear function. This result implies that the external field influences the investor behavior — or, in other words, that the investors will buy if there is good news but they will sell if there is bad news.

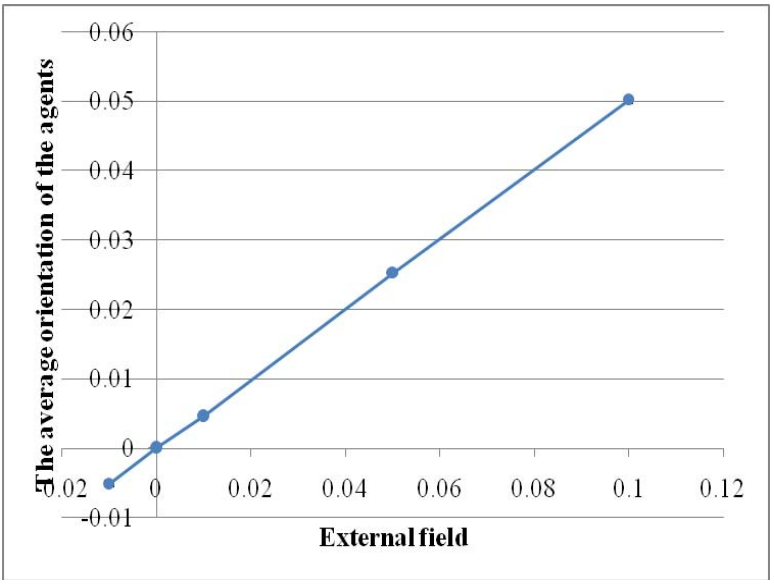


Fig. 3 The average orientation of the agents as a function of the external field (h) at fixed temperature (T) = 1

As for the effect of changing the temperature in each step of time, as illustrated in Figure 4, although the external field is positive, the average orientation of the agents is both positive and negative. This is due to the higher temperature, or price (see Figure 5 and Figure 6). When the temperature is very high, the average orientation of the agents is allowed to be negative so that the local field will become negative too. When the local field is negative, the probability of the low temperature will be of opposite value. It will be very close to 0. In such a situation, the price may be reduced so that the temperature will get reduced too.

This paper presents the spin model that was used to simulate the stock market. The study evaluates the temperature via the Monte Carlo simulation. The results show the effect of temperature on the investments behavior. Furthermore, we found that the external field, too, has an effect on the investments behavior. So, in the work done in future, we will change the external field in each step of time. In addition, we will enable the parameters used in this work to represent the real parameters and compare the results from this model to the real data in the stock market.

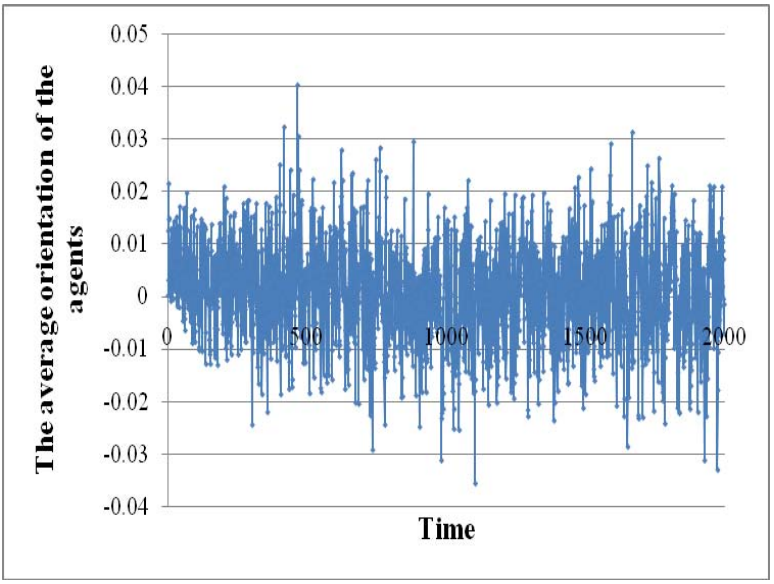


Fig. 4 The average orientation of the agents as a function of time at fixed external field (h) = 0.01 and initial temperature (T_0) =1

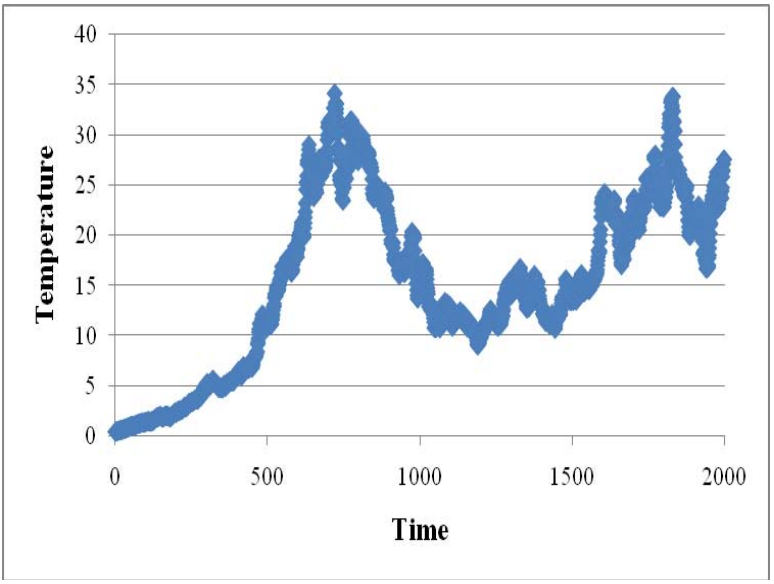


Fig. 5 The temperature as a function of time

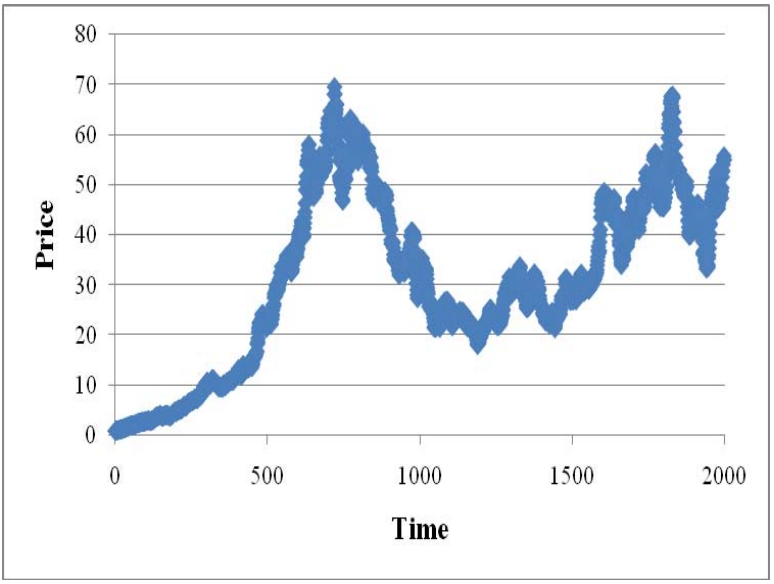


Fig. 6 Price at fixed external field (h) = 0.01 and initial temperature (T_0) =1

Acknowledgements. The authors gratefully acknowledge financial supports from the Graduate School Chiang Mai University.

References

1. Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L.A.N., Guhr, T., Stanley, H.E.: Random matrix approach to cross correlations in nancial data. Phys. Rev. E 65, 066126 (2002)
2. Eom, C., Jung, W., Kaizoji, T., Kim, S.: Effect of changing data size on eigenvalues in the Korean and Japanese stock markets. Physica A 388, 4780–4786 (2009)
3. Jung, W.-S., Kwon, O., Wang, F., Kaizoji, T., Moon, H.T., Stanley, H.E.: Group dynamics of the Japanese market. Physica A 387, 537–542 (2008)
4. Mantegna, R.N.: Hierarchical structure in financial markets. Eur. Phys. J. B 11, 193–197 (1999)
5. Kaizoji, T.: Speculative bubbles and crashes in stock markets: an interacting-agent model of speculative activity. Physica A 287, 493–506 (2000)
6. Krawiecki, A., Holyst, J.A., Helbing, D.: Volatility clustering and scaling for financial time series due to attractor bubbling. Physical Review Letters 89, 158701-1–158071-4 (2002)
7. Bornholdt, S.: Expectation bubbles in a spin model of markets: intermittency from frustration across scales. Int. J. Mod. Phys. C12, 667–670 (2001)
8. Kaizoji, T., Bornholdt, S., Fujiwara, Y.: Dynamics of price and trading volume in a spin model of stock markets with heterogeneous agents. Physica A 316, 441–452 (2002)

9. Krawiecki, A.: Microscopic spin model for the stock market with attractor bubbling on scale-free networks. J. Econ. Interact. Coord. 4, 213–220 (2009)
10. Sieczkaand, P., Hoyst, J.A.: A Threshold Model of Financial Markets. ACTA Physica Polonica A 114, 525–530 (2008)
11. Krawiecki, A., Holyst, J.A.: Stochastic resonance as amodelforfinancial market crashes and bubbles. Physica A 317, 597–608 (2003)
12. Dragulescu, A., Yakovenko, V.M.: Statistical mechanics of money. Eur. Phys. J. B 17, 723–729 (2000)
13. Sieczkaand, P., Holyst, J.A.: A Threshold Model of Financial Markets. Acta Physica Polonica A 114, 525–530 (2008)

An Analysis of Relationship between Gold Price and U.S. Dollar Index by Using Bivariate Extreme Value Copulas

Mutita Kaewkheaw, Pisit Leeahtam, and Chukiat Chaiboosri

Abstract. In this study, we analyse the behaviour of the gold price and U.S. dollar index by using bivariate extreme value and extreme value copulas. For measuring the dependence structure between the returns on gold price and U.S. dollar index, the paper uses the extreme value copula theory. This study presents the result that the returns on gold price and the U.S. dollar index are independence in the extreme.

1 Introduction

There are many methods to estimate co-movement of two variables. The copula analysis is one important technique used in dependence study. Of late, the copula method has been finding application in various fields in the study of correlation. The copula method can define and examine the dependence structure between two variables more than the classical dependence measures, such as linear correlation, with their limitations. As for analysing non-linear dependence, the copula can measure dependence for heavy-tail distributions and is flexible in the cases of parametric, semi-parametric, or non-parametric models. The study about the asymptotic properties of the dependence structures is conducted using the copula method (Lei, 2009). Moreover, the most general margin-free description of the dependence structure of a multivariate distribution can use the copula approach (Segers, 2005). The major benefit of using the copula method is avoidance of the impact of the marginal from the structure when used in the joint probability distribution (Chinnakum et al, 2013). In the case of extreme events, multivariate distribution - which the copulas can adapt themselves to - was employed to measure the parameter dependence. Extreme value theory (EVT) is a branch of statistics which deals with extreme deviations from the mean of the probability distribution (Lu et al, 2008). The EVT explains the analysis

Mutita Kaewkheaw · Pisit Leeahtam · Chukiat Chaiboosri
Faculty of Economics, Chiang Mai University, Chiang Mai 50200 Thailand
e-mail: {muknajah,chukiat1973}@hotmail.co.th,
pisitleeahtam@gmail.co.th

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_29,

and modeling of the extreme maxima and minima observation (Chuangchid et al, 2012). However, there are many distribution functions, and the bivariate extreme value distribution can handle the problem via multivariate extension (Rakonzai and Tajvidi, 2010). Many of the latest research papers have presented the topic of extreme value copulas. Lu, Tien, and Zhang (2008) analyzed the risk of the foreign exchange data dependence by using the extreme value copulas. They found that three members of the copula family could be applied to measure the joint tail risk and tail dependence for the data. Chuangchid et al (2012) arrived at the conclusions about the dependence measure of palm oil price from the futures prices of Malaysia, Singapore, and Dalian commodities using the extreme value copula, especially by using the generalized extreme value as well as the extreme value copulas of HuslerRiess and Gumble for estimation. The result shows that the extreme value copulas can explain the dependence structure for palm oil futures prices. Chinnakum et al (2013) focused on the effect of the economic output in the developed countries. They tested the panel data that have determinants of economic output in 22 developed countries. First, they applied the maximum likelihood method and the method from Heckman (1979) so that the assumption is a multivariate normal distribution. However, the result showed that the copula approach, especially the Archimedean copula, fitted the sample selection model and was successful in the identification of the significant factors affecting the economic output. Sriboonchitta et al (2013) focused on estimating the dependency between the percentage changes of the agricultural prices and the production indices of Thailand, and their conditional volatilities using copula-based GARCH models. The results showed that the skewed-t distributions were appropriate in marginal density for the growth rates of the agricultural production and the price indices. The time-varying rotated Joe copula was the best among the various copula candidates. Liu and Sriboonchitta (2013) found that the time-varying Gaussian copula has the highest explanatory power of all the dependence structures to estimate the dependency between the growth rates of tourist arrivals in Thailand and Singapore from China, as well as to estimate the conditional volatilities.

For centuries, investors have been found to protect their capital by investing in assets that offer safer stores of value (World Gold Council, 2008). Gold is an important asset that provides stability to international money markets and international currency reserves (Chang et al, 2013). The World Gold Council has published the gold demand trends of the first quarter of 2013 (Q1). The gold demand trends of the first quarter show strong inclination toward demand for gold jewelry, bars, and coins, but the overall demand showed a decrease as compared to the first quarter of 2012. Also, the price of gold in most currencies weakened in Q1. However, gold prices in Q1 of 2013 were higher than Q1 of 2012 which offered the rise in demand was not only gold price-driven. Moreover, a lot of investment requires the efficiency exchange currencies that cannot accept the U.S dollar is the one important selection of investor. Additionally, the prices of gold are determined in U.S. dollars and influenced by changes in the exchange rate of the U.S. dollar (Wang, 2012). The alternative tool such as dollar index is a significant tool that can tell the direction of investment or investment flow (Investor Chart Co., Ltd., 2010). The U.S. Dollar Index (USDX) is a geometrically-averaged calculation of major currencies weighted

against the U.S. dollar. The USDX was created by the U.S. Federal Reserve in 1973. Following the ending of the 1944 Bretton Woods agreement, which had established a system of fixed exchange rates, the U.S. Federal Reserve Bank began calculation of the U.S. Dollar Index to provide an external bilateral trade-weighted average of the U.S. dollar as it freely floated against global currencies. USDX is calculated by six currencies weight namely; Euro (EUR) 57.6%, Japanese yen (JPY) 13.6%, Pound sterling (GBP) 11.9%, Canadian dollar (CAD) 9.1%, Swedish krona (SEK) 4.2% and Swiss franc (CHF) 3.6%. All weight rate of USDX is Euro has most affect the dollar because weight is 57.6%, and followed by the Japanese yen and the pound respectively (Intercontinental Exchange Inc, 2012). The paper focuses on the tail behavior of gold and USDX also, dependence structure between them by using the bivariate extreme values copulas.

The components of this paper are organized as follow. Section 2 presents bivariate extreme value distribution (BEVD) with bivariate block maxima model. Section 3 reviews the concept of copulas and extreme value copulas. Section 4 explains the data uses in the empirical analysis. Section 5 discusses the results and last section is conclusion.

2 Bivariate Extreme Value

The Extreme Value Theory (EVT) is an approach used for modelling and measuring extreme events which occur with a very small probability (Alves & Neves, 2010). There are two approaches to find the extremes in data. The first is Block Maxima (BM) and the second is Peaks-Over Threshold (POT). BM and POT are the statistical analyses of maxima or minima, and exceeds over a higher or a lower threshold (Lai and Wu, 2007). This paper uses bivariate BM models to examine the relationship between the gold prices and the U.S. dollar index.

2.1 Bivariate Block Maxima

The bivariate block maxima model involves both parametric and non-parametric cases. The study chooses the parametric models that can summarize the bivariate BM, as follows:

Let (X, Y) denote a bivariate random vector representing the component-wise maxima of an i.i.d. sequence over a given period of time. Under the appropriate conditions, the distribution of (X, Y) can be approximated by a bivariate extreme value distribution (BEVD) with the cumulative distribution function (cdf) G. The BEVD does the examination using its two univariate margins G1 and G2, which are necessarily EVD, and also by using its Pickands dependence function A (Rakonczai and Tajvidi, 2010).

$$G(x_1,x_2) = \exp\left\{\log(G_1(x_1)G_2(x_2))A\left(\frac{\log(G_2(x_2))}{\log(G_1(x_1)G_2(x_2))}\right)\right\} \tag{1}$$

Under these provisions, A(t) is responsible for capturing the dependence structure between the margins. A is the Pickands dependence function which is convex and is inside the triangle that is plotted from the points (0, 1), (1, 1), (1/2, 1/2) tying upper left and right corners. In addition, A(t) has three properties: 1) A(t) is convex, 2) $max\{(1-t),t\} \leq A(t) \leq t$, and 3) A(0) = A(1) = 1. In the second property of A, the lower bound corresponds to complete dependence $G(x,y) = min\{G_1(x), G_2(y)\}$, whereas the upper bound corresponds to (complete) independence $G(x,y) = \{G_1(x), G_2(y)\}$.

For the bivariate block maxima model, this paper identified one parametric model that has minimum Akaike Information Criterion (AIC) from a total of nine models (Chuangchid et al, 2012). So, the paper concentrated on the bivariate negative logistic distribution function of A(t) with parameter dependence = r, which is

$$G(x,y) = \exp\{-x-y+[x^{-r}+y^{-r}]^{-1/r}\} \tag{2}$$

where $r > 0$. Upon the implementation of this, independence is obtained in the limit as r closes to zero. Complete dependence is obtained when r tends to infinity.

The estimation Parameter of the bivariate block maxima applied generalized extreme value (GEV) distribution (Chuangchid et al, 2012):

Let $Z_i (i = 1,\ldots,n)$ denote maximum observation in each block. Z_n is normalized to obtain a non-degenerated limiting distribution. The bivariate block maxima distribution associate with the use of GEV distribution with cdf:

$$H(z) = exp\left\{-\left[1+\xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} \tag{3}$$

where μ, σ and ξ are location, scale and shape parameter respectively. Note that ,¿0. The generalized extreme value has 3 types depending on ξ (shape parameter); $\xi = 0$ is Gumbel or double-exponential distribution, $\xi > 0$ is Frchet distribution, $\xi < 0$ is Fisher-Tippet or Weibull distribution. Under the assumption: $Z_1,\ldots,Z_n$ are independent variables having the GEV distribution, the log-likelihood for the GEV parameters when $\xi \neq 0$ is given by:

$$l(\xi,\mu,\sigma) = -n\log\sigma\sum_{i=1}^{n}\log\left[1+\xi\left(\frac{Z_i-\mu}{\sigma}\right)\right] - \sum_{i=1}^{n}\left[1+\xi\left(\frac{Z_i-\mu}{\sigma}\right)\right]^{-1/\xi} \tag{4}$$

and $1+\xi\left(\frac{Z_i-\mu}{\sigma}\right) > 0$ for $i = 1,\ldots,n$. The case $\xi = 0$ requires separate treatment using the Gumbel limit of the GEV distribution. The log-likelihood in that case is;

$$l(\mu,\sigma) = -n\log\sigma - \sum_{i=1}^{n}\left(\frac{Z_i-\mu}{\sigma}\right) - \sum_{i=1}^{n}\exp\left[-\left(\frac{Z_i-\mu}{\sigma}\right)\right] \tag{5}$$

The maximization of this equation with respect to the parameter vector (ξ,μ,σ) leads to the maximum likelihood estimate with respect to the entire GEV family.

3 Copulas and Extreme Value Copulas

The copula approach was introduced by Sklar (1959), and it involves an n-dimensional distribution into two parts: the marginal distribution functions and the copulas (Ruschendorf, 2013). This study uses the extreme value distribution for the extreme value copulas. When determining the bivariate case, let (U,V) be a pair of random variables, and let both U and V be uniformly distributed on the interval [0,1]. Then the joint distribution function of (U,V) (Segers, 2005) would be as follows:

$$C(u,v) = Pr[U \leq u, V \leq v] \tag{6}$$

Then, let (X, Y) be the stochastic behaviour of the two random variables, with the joint distribution function (Chuangchid et al, 2012)

$$H(x,y) = Pr[X \leq x, Y \leq y] \tag{7}$$

and marginal distribution functions

$$F(x) = P(X \leq x)\, and\, G(y) = P(Y \leq y) \tag{8}$$

Since F(x) and G(y) are uniformly distributed between 0 and 1, the joint distribution function C on $[0,1]^2$ for all (x.y) $\in R^2$ is such that

$$H(x,y) = C(F(x), G(y)) \tag{9}$$

where C is called the copula related with X and Y which couples the joint distribution H with its margins. Set equation (8) is equalled to $H(F^{-1}(u), G^{-1}(v)) = C(u,v)$ on account of the Sklars theorem, where u = F(x) and v = G(y) are the marginal distributions of X, Y. The implication of the Sklars theorem is after standardizing the effects of the margins, and the dependence between X and Y is fully explained by the copula. This paper combines the copula construction with the extreme value theory. The extreme value copula family is used to represent the Multivariate Extreme Value Distribution (MEVD). Consider a bivariate sample $(X_i, Y_i), i = 1, \ldots, n$. Denote component-wise maxima by $M_n = \max(X_1, \ldots, X_n)$ and $N_n = \max(Y_1, \ldots, Y_n)$. The object of interest is the vector of the component-wise block maxima: $M_c = (M_n, N_n)$ The bivariate extreme distribution H can be connected by an extreme value copula (EV copula) C_0 :

$$H(x,y) = C_0(F(x; \mu_1, \sigma_1, \xi_1), G(y; \mu_2, \sigma_2, \xi_2)) \tag{10}$$

where μ_i, σ_i, ξ_i are the location, scale and shape parameters of the gold price return and USDX return. F(x) and G(y) are the BEV margins. By Sklars theorem, the unique copula C_0 of H is given by

$$C_0(u^t, v^t) = C_0^t(u,v), t > 0 \tag{11}$$

where u^t is marginal of x and v^t is marginal of y There are many families that belong to the extreme value copula. The copula HuslerReiss is applied in this paper.

HuslerReiss copula is given as

$$C(u,v) = \exp\left\{-\tilde{u}\Phi\left(\frac{1}{r}+\frac{1}{2}r\ln\left(\frac{\tilde{u}}{\tilde{v}}\right)\right) - \tilde{v}\Phi\left(\frac{1}{r}+\frac{1}{2}r\ln\left(\frac{\tilde{v}}{\tilde{u}}\right)\right)\right\} \qquad (12)$$

where $u = -\ln u$, $v = -\ln v$, and Φ is the standardized normal distribution. The independence copula is obtained in the limit as the value of r becomes 0, and the complete dependence copula is obtained in the limit as the value of r becomes ∞. This paper used the Exact Maximum Likelihood (EML) method for the estimation of the copula parameters.

4 Data

This paper used the time series data was obtained from the World Gold Council and Reuters. We worked with gold price and U.S. dollar index. This study applied the gold price and the U.S. dollar index in terms of the U.S. dollar. The first method is of taking the daily prices of gold and the U.S. dollar index, converted to return series. Daily prices are computed as return of price i at time t relatively; $R = \ln(p_{i,t}/p_{i,t-1})$ where R is return of price, $p_{i,t}$ and p_{t-1} are daily price for day t and day t-1, respectively. The study period was from January 2008 till June 7, 2013.

5 Empirical Result

The test results from using the bivariate BM are shown in Table 1. This table reveals the distribution function parameter (r), and estimates for the location (μ), shape

Table 1 Bivariate Block Maxima Gold Prices and U.S. Dollar Index

Variables	Parameter estimation	bivariate BM
Gold Price	μ_1	-0.010788 (0.00587)
	σ_1	0.043609 (0.00396)
	ξ_1	-0.207766 (0.05932)
USDX	μ_2	-0.008616 (0.00300)
	σ_2	0.22267 (0.0020)
	ξ_2	-0.145127 (0.061058)

Note: The terms in the parentheses are the standard errors of the parameter estimates.

(ξ), and scale (σ) parameters. The negative logistic model between the log return rates of gold and USDX estimates r as equal to 0.0576, which implies that the log return rates of gold and USDX have independence in the extremes.

For the calculation of the copula parameter (r), the result can be estimated using the HuslerReiss copula analysis. The result gives the value of the copula parameter as 0.0037, which reveals the independence measure structure between the gold price and the USDX. The standard error for this parameter is too large to calculate because the two variables, the log return gold price and the U.S. dollar index, have independent structures.

Table 2 Parameter Estimation

Market	HuslerReiss copula
Gold price and USDX	0.0037

6 Conclusion

This paper aims to study the extreme behavior of daily global gold price and the U.S. dollar index by using bivariate extreme value distribution. The study explains the dependence measure structure of both the variables using the extreme value copula HuslerReiss. The result presents the independence structure extreme event of return in gold and USDX. The consequence of this study corresponded with result from the paper of Forecasting Gold Prices Using Multiple Linear Regression Method by Ismail Z. et al (2009). Ismail Z. studied eight independent variables to forecast gold price and USDX is one of the variables. When they computed model to gold price prediction, they found insignificant for USDX in multiple linear regression and Commodity Research Bureau, USD/Euro Foreign Exchange rate, inflation rate, Money Supply affected gold price change. Additionally, the result could be constructive for future research about the U.S. dollar index and beneficial to many investors.

References

1. Chang, C.-L., Chang, D.J.-C., Huang, Y.-W.: Dynamic price integration in the global gold market. North American Journal of Economics and Finance (2013)
2. Chaithep, K.: Value at Risk analysis of Gold price returns using Extreme Value Theory. Master's Thesis of Economics Chiang Mai University (2012)
3. Chinnakum, W., Sriboonchitta, S., Pastpipatkul, P.: Factors affecting economic output in developed countries: A copula approach to sample selection with panel data. International Journal of Approximate Reasoning 54, 809–824 (2013)
4. Chuangchid, K., et al.: Application of Extreme value Copulas to palm oil prices analysis. Business Management Dynamics 2, 25–31 (2012)
5. Chuangchid, K., et al.: Factors Affecting Palm Oil Price Based on Extremes Value Approach. International Journal of Marketing Studies 4(6) (2012)

6. Chuangchid, K., et al.: Predicting Malaysian palm oil price using Extreme Value Theory. International Journal of Agricultural Management 2(2), 91–99 (2013)
7. Fraga, A.M.I., Neves, C.: Extreme Value Distributions. International Encyclopedia of Statistical Science 3, 493–496 (2010)
8. Hua, L.: A Brief Introduction to Copulas. Department of Statistics University of British Columbia (2009)
9. Hurlimann, W.: Hutchinson-Lais conjecture for bivariate extreme value copulas. Statistic & Probability Letters 61, 191–198 (2001)
10. Intercontinental Exchange Inc. The ICE U.S Dollar Index and US Dollar index Futures Contracts. FAQ (June 2012), www.theice.com (July 11, 2013)
11. Lai, L., Wu, P.: An Extreme Value Analysis of Taiwan's Agriculture Natural Disaster loss data. In: International Conference on Business and Information (BAI), Tokyo, Japan (2007)
12. Liu, J., Sriboonchitta, S.: Analysis of Volatility and Dependence between the Tourist Arrivals from China to Thailand and Singapore: A Copula-Based GARCH Approach. In: Huynh, V.N., Kreinovich, V., Sriboonchitta, S., Suriya, K. (eds.) Uncertainty Analysis in Econometrics with Applications. AISC, vol. 200, pp. 285–296. Springer, Heidelberg (2013)
13. Lu, J., Tian, W.-J., Zhang, P.: The Extreme Value Copulas Analysis of the Risk Dependence for the Foreign Exchange Data. Wireless Communications, Networking and Mobile Computing, 1–6 (2008)
14. Pukthuanthong, K., Roll, R.: Gold and the Dollar (and the Euro, Pound, and Yen). Journal of Banking & Finance 35, 2070–2083 (2011)
15. Rakonczai, P., Tajvidi, N.: On Prediction of Bivariate Extreme. International Journal of Intelligent Technologies and Applied Statistics 3(2), 115–139 (2010)
16. Segers, J.: Extreme-Value Copulas. Medium Econometrische Toepassingen 13(1), 9–11 (2005)
17. Sriboonchitta, S., et al.: Modelling volatility and dependency of agricultural price and production indices of Thailand: Static versus time-varying copulas. International Journal of Approximate Reasoning 54, 793–808 (2013)
18. Stephenson, A.: Functions for extreme value distributions. Package 'evd', Version 2.2-4
19. The World gold council. 2008. About a gold: Story of Gold, Heritage (2011), http://www.gold.org/about_gold/story_of_gold/heritage/ (July 11, 2013)
20. The World gold council. Gold Demand Trends (First Quarter, 2013) www.gold.org (July 11, 2013)
21. Wang, M.-L., Wang, C.-P., Huang, T.-Y.: Relationships among Oil Price, Gold Price, Exchange Rate and International Stock Markets. International Research Journal of Finance and Economics 47, 83–93 (2010)
22. Wang, Z.: The Relationships between Silver Price, Gold Price and U.S. Dollar Index Before and After the Sub-prime Crisis. A research project submitted in partial fulfilment of the requirements for Master of Finance Saint Mary's University (2012)

An Integration of Eco-Health One-Health Transdisciplinary Approach and Bayesian Belief Network

Chalisa Kallayanamitra, Pisit Leeahtam, Manoj Potapohn,
Bruce A. Wilcox, and Songsak Sriboonchitta

Abstract. Animal health economics is becoming increasingly important as the assistance for decision making on animal health intervention at all levels in attempting to optimize animal health management. Economic analysis of the optimal control of zoonoses associated with livestock production is complex as it depends on the nature of occurrence, transmission, and circulation of the diseases. Recent studies show that the emphasis of most of the veterinary economists is usually on the practical field of the economic evaluation of animal diseases based on a detailed knowledge of the production system. However, the field had not yet begun to address the more complex and real-world problems such as cause of emerging diseases. This empirical research employs a more holistic approach such as that advocated by the Eco-Health One-Health approach, together with the transdisciplinary analytical framework and Bayesian Belief Network analysis that integrates uncertainties into consideration to explain Trichinellosis risk. This fundamental research found that the Bayesian Belief Network modeling for the analysis of zoonoses risk and a combined human and animal health framework can be used to guide decision making for interventions to solve the Eco-Health One-Health problem of Trichinellosis risk. However, the scoring rule results from Netica, an easy to use software for working with Bayesian Belief Network, provide only symmetric loss values based on the assumption that the loss from misestimating is the same in any direction. Nonetheless, this assumption may not be valid in some practical situations such as what we are interested in this research, Trichinellosis risk. The research suggests an approach that takes the idea of decision theory combining the cost of collecting a sample to

Chalisa Kallayanamitra · Pisit Leeahtam · Manoj Potapohn · Songsak Sriboonchitta
Faculty of Economics, Chiang Mai University, Chiang Mai 50200 Thailand
e-mail: {c.kallayanmitra,pisitleeah-tam,potapohnm,
songsake-con}@gmail.com

Bruce A. Wilcox
Integrative Research and Education Program, Faculty of Public Health, Mahidol University, Bangkok 73170 Thailand, Tropical Disease Research Laboratory, Khon Kaen University, Khon Kaen 40002 Thailand
e-mail: wilcox.bruce@gmail.com

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_30,

minimize the pre-posterior expected cost. If the sampling cost of collecting data is very high, or if there is strong prior information about the risk, it is not worth sampling. Also, if the loss of illness is very high, a thorough protection strategy would be more efficient.

1 Introduction

Animal health economics is becoming increasingly important as the assistance for decision making on animal health intervention at all levels [20] in attempting to optimize animal health management [17]. Economic analysis of the optimal control of zoonoses associated with livestock production is complex as it depends on the nature of occurrence, transmission, and circulation of the diseases. Economic approaches to infectious diseases are embedded in many areas of work nowadays and cannot be ignored [21]. Even though this was unpopular with the pure economists, it is a novel way of utilizing economics in explaining a complex and real-world problem.

Recent studies show that the emphasis of most of the veterinary economists is usually on the cost-benefit analysis [8, 11, 12, 26, 29] in the evaluation of zoonoses intervention and control efforts [24]. Many of the veterinary economists, such as Ramsay, Tisdell, and Harrison (1997), concentrated on how better information communication in the field of animal health could enhance decision making. Some veterinary economists have been working on the development of economic analysis techniques in the study of diseases and their control. Tim Carpenter was the first to examine the use of various economic analysis techniques such as decision tree analysis [2, 3, 22, 23], microeconomics analysis of diseases [1], simulation models to assess animal diseases [4], dynamic programming [5], dual estimation approach to derive shadow prices for diseases [28], estimation of consumer surplus [18], willingness to pay for vaccination [25], linear programming [1, 5, 6], and use of economic analysis to review subsidies to veterinary support institutions [5]. On the theoretical side, some veterinary economists such as McInerney and Howe began researching the economics of livestock diseases through the development of conceptual models of farmer behavior toward disease [13, 14].

All the above approaches are in the practical field of the economic evaluation of animal diseases based on a detailed knowledge of the production system. However, the field had not yet begun to address the more complex and real-world problem of cause of emerging diseases. This research employs a more holistic approach such as that advocated by the Eco-Health One-Health approach[1], together with the trans-

[1] A systematic and participatory approach to understanding and promoting sustainable health and well-being of humans, animals, and the environment thought of as all part of one ecosystem, as well as making decisions, taking action, and evaluating outcomes. The Eco-Health One-Health approach is an emerging field of study and practice that examines the biological, social, and economic dynamics of an ecosystem, and it relates these changes to human and animal health, holistically. It brings together people from various disciplines such as veterinarians, ecologists, economists, social scientists, policy makers, and others to explore and understand how the above dynamics affect human and animal health.

disciplinary analytical framework and Bayesian Belief Network analysis[2] that integrates uncertainties into the consideration to explain Trichinellosis risk, a historic problem worldwide in humans consuming raw or undercooked pork.

This research explains the population, the sampling design and methodology used to collect the data, how the data is analyzed, purpose of the modeling, how to develop the model, model structure and parameters, and how to test the modeling. Finally, we conclude with a presentation of the constraints, limitations, and benefits of the application of the Bayesian Belief Network and the Eco-Health One-Health approach.

2 Objectives

The focus of this empirical research is to illustrate how the Bayesian Belief Network analysis and the Eco-Health One-Health approach are integrated to explain the complex and real-world problem of Trichinellosis risk, and to find out the constraints, limitations, and benefits of this methodology.

3 Methodology

3.1 Population and Sampling Design

The target population of this study includes ethnic minority groups residing in two selected highlanders villages in Mae Ai district, Chiang Mai Province. There are a total of 84 households in the Huai Chan Si village and 118 households in the Huai Ma Fueang village. Twenty-six households from the Huai Chan Si village and 28 households from the Huai Ma Fueang village were selected using simple random selection.

3.2 Data Collection

A questionnaire was developed for the household survey, based on the Trichinellosis risk factors derived from experts opinions. Twelve enumerators including eight students from the faculty of Veterinary Medicine and four students from the faculty of Economics, Chiang Mai University, were trained on how to conduct the questionnaire session in the selected villages; at the same time, the questionnaire was also tested.

[2] A statistical method invented in the 1940s and 1950s to take into account the effects of uncertainty in management systems in the decision-making processes. It is a graphical description of the conceptual model that captures the analyst beliefs in the causal relationships of significant variables in the system of interest.

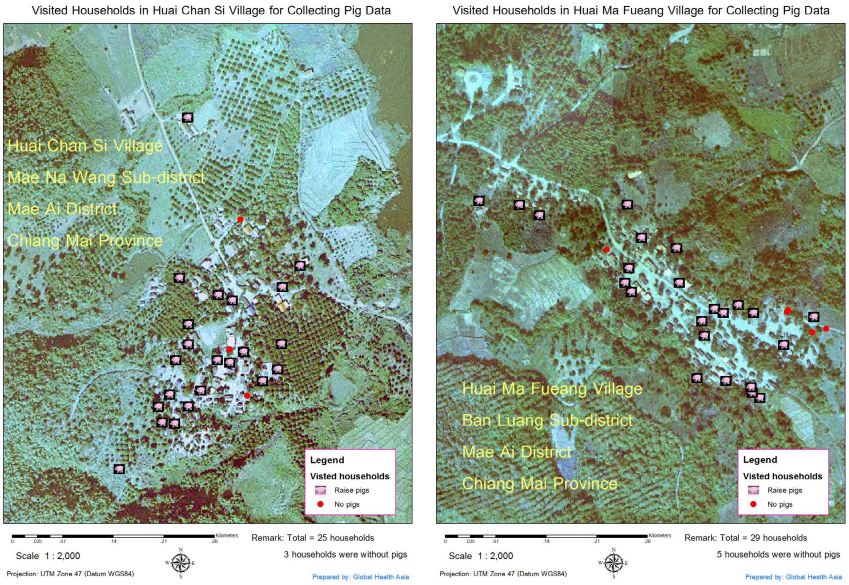


Fig. 1 Location of the random households in the two selected villages in Mae Ai District

After conducting the household survey, we developed a set of data by preparing it suitably for the experts to evaluate the Trichinellosis risk circumstance in the selected villages at the experts meeting. Seven experts were invited to join the focus groups.[3]

Data collected about the components of the Trichinellosis risk factors are transformed into the variables used in a Bayesian Belief Network model.

[3] Animal Health Experts:
Assist. Prof. Panuwat Yamsakul, Faculty of Veterinary Medicine, Chiang Mai University
Dr. Veerasak Punyapornwithaya, Faculty of Veterinary Medicine, Chiang Mai University
Ms. Pornpen Tablerk, Department of Livestock Development, Nan Province
Disease Ecologist:
Prof. Bruce A. Wilcox, Integrative Research & Education Program, Faculty of Public Health, Mahidol University and Tropical Disease Research Laboratory, KhonKaen University
Human Health Experts:
Assoc. Prof. Dr. Pichart Uparanukraw, Faculty of Medicine, Chiang Mai University
Assoc. Prof. Dr. Nimit Morakote, Faculty of Medicine, Chiang Mai University
Mr. Adulsak Wijit, Office of Diseases Prevention and Control 10

3.3 Data Analysis

(*a*) Modeling

The complexity of the circumstances to cope with Trichinellosis has led to model-based approaches for investigating the interconnections and for predicting the management outcomes [16]. A probabilistic graphical model for qualitative instrument called Bayesian Belief Network is applied for this analysis. The conceptual transdisciplinary framework of the Trichinellosis risk is developed by experts based on the existing knowledge and the experience from the field study to explain the interconnection between animal health, environment, and human health.

(*b*) Purpose of Modeling

The purpose of this modeling is to find the posterior probability of the Trichinellosis risk, given the evidence of the related risk factors.

(*c*) Developing the Model

The Trichinellosis risk framework was developed based on the Eco-Health One-Health concept by making note of the opinions of veterinarians, disease ecologists, medical doctors, and public health officers (see Fig. 2). There are a total of 13 variables to be studied. There are two kinds of variables in this study, including discrete data and continuous data. These variables are associated with probabilistic functions. There are two sources of information to be fed into the model, including the data from the field study and the data from the experts opinions. Netica, a powerful and easy-to-use program for working with the Bayesian Belief Network, and influence diagrams are applied to analyze this set of data.

(*d*) Model Structure and Parameters

This part explains how the Bayesian Belief Network determines the posterior probability of the risk of contracting Trichinellosis, based on the associated risk factors.

Posterior Probability Equation of Risk of Getting Trichinellosis

Based on the Bayesian statistics, the posterior probability equation for this model is defined as

$$p(RSKHUM_i|X) = \frac{p(X|RSKHUM_i}{p(X)} \cdot (RSKHUM_i)$$

where i is the levels of the risk that humans will be infected by Trichinellosis (high, medium and low)

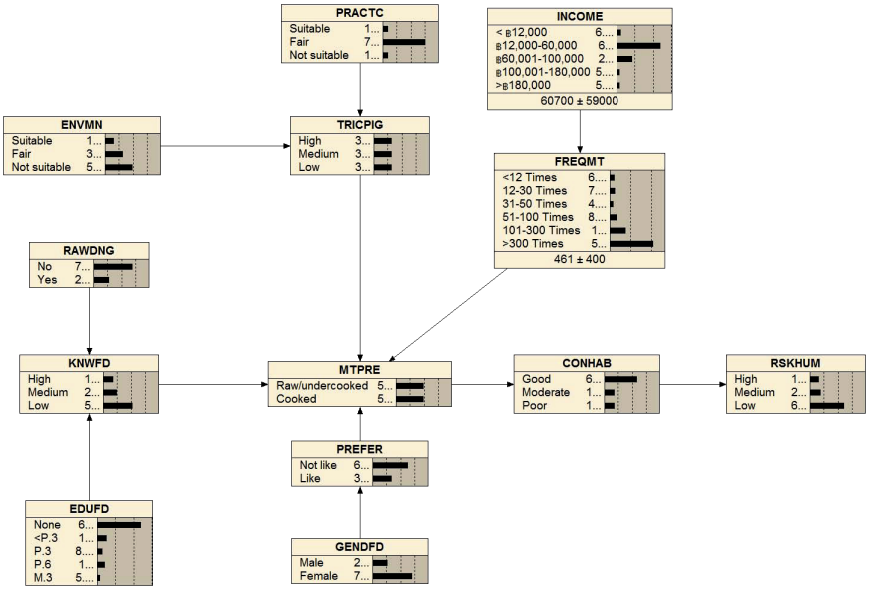


Fig. 2 The Trichinellosis risk framework

X is the set of all the risk factors associated with the Trichinella infection in pigs. That is, $\{TRICPIG, ENVMN, PRACTC, GENDFD, INCOME, EDUFD, MTPRE, PREFER, RAWDNG, FREQMT, CONHAB, KNWFD\}$.

$p(RSKHUM_i|X)$ is posterior probabilities (or the probabilities of the parameters $RSKHUM_i$),given evidence.

$p(X|RSKHUM_i)$ is likelihood functions (or the probabilities of the evidence X),given the parameters $RSKHUM_i$.

$p(RSKHUM_i)$ is prior probability probabilities (or the probabilities of the risk that humans will be infected by Trichinellosis based on the subjective assessment of experienced experts).

$p(X)$ is probabilities of all the evidence in set X, regardless of any other information.

(*e*) Testing Modeling

The objective of this test is to evaluate the quality of the Bayesian Belief Network using Netica to process a set of real cases. This test illustrates how well the models match the actual cases by considering the actual belief levels of the states in determining how well they agree with the value of the case file. We first incorporate 60 percent of the cases into the model. Then, the nodes in which we wish to find their inferences which is the *RSKHUM* node is selected. We use the rest 40 percent of the samples to verify the validity of the model. When the Netica was done, it printed a

report called scoring rule results of each of the selected nodes. The reports included error rate, logarithmic loss score, quadratic (Brier score), and spherical payoff score.

The error rate determines how many times the classifier misclassifies a case divided by the number of classifications. It is only with respect to the probability distribution of the test cases [19].

Bayes Risk Minimization Using Loss Function

In decision theory, a Bayes estimator is an estimator that minimizes the posterior expected value of a loss function [9]. Suppose parameter $RSKHUM_i$ is known to have a prior distribution $p(RSKHUM_i)$. Let $\widehat{RSKHUM_i}$ be an estimator of $RSKHUM_i$ and $L(RSKHUM_i, \widehat{RSKHUM_i})$, a loss function.

The Bayes risk of $RSKHUM_i$ is;

$$r(\widehat{RSKHUM_i}, p(RSKHUM_i|X))$$

$$= \sum_{RSKHUM_i} L(RSKHUM_i, \widehat{RSKHUM_i}) p(RSKHUM_i|X))$$

In a Bayesian framework, the Bayes risk is based on the data set X and a prior $p(RSKHUM_i|X)$ representing the beliefs about $RSKHUM_i$ with density $p(RSKHUM_i)$. An estimator $\widehat{RSKHUM_i}$ is considered to be a Bayes estimator if it minimizes the Bayes risk.

$$r^*(\widehat{RSKHUM_i}, p(RSKHUM_i|X))$$

$$= min \sum_{RSKHUM_i} L(RSKHUM_i, \widehat{RSKHUM_i}) p(RSKHUM_i|X))$$

There are many different types of loss functions. In this analysis, we consider only the logarithmic loss function and the quadratic loss function.

Logarithmic loss function is one of the symmetric loss functions given by Brown (1968). Logarithmic loss values are calculated using the natural log;

$$L_{ll}(\widehat{RSKHUM_i}, RSKHUM_i) = \left| ln \frac{\widehat{RSKHUM_i}}{RSKHUM_i} \right|$$

Therefore, the Bayes estimator of $RSKHUM_i$ under the logarithmic loss function is;

$$\widehat{RSKHUM_i}^{ll^*} = exp \sum_{RSKHUM_i} ln(RSKHUM_i) p(RSKHUM_i|X)$$

The logarithmic loss values are between zero and infinity. Zero indicates the best performance [15, 19].

The logarithmic score only considers the estimated probability for the actual value of $RSKHUM_i$, whereas the quadratic loss function (the squared error loss

function) which was first introduced by Brier (1950) also takes into account how the estimated probabilities are distributed on the false states. Under the quadratic rule, the forecaster is penalized in proportion to the mean squared difference between the parameter and the estimator.

$$L_{ql}(\widehat{RSKHUM}_i, RSKHUM_i) = (\widehat{RSKHUM}_i - RSKHUM_i)^2$$

Therefore, the Bayes estimator of $RSKHUM_i$ under the quadratic loss function is;

$$\widehat{RSKHUM}_i^{ql^*} = \sum_{RSKHUM_i} (RSKHUM_i) p(RSKHUM_i|X)$$

The quadratic loss values are between zero and two. The lower the quadratic score is for a set of predictions, the better the predictions are calibrated [10, 15].

Another scoring rule is the spherical payoff value which is also used to assess the quality of the probabilistic forecaster between zero and one. One represents the best performance.

$$Spherical payoff = MOAC \left| \frac{p_c}{\sqrt{\sum_i^n p_i^2}} \right|$$

where p_c is probability predicted for the correct state.

p_i is probability predicted for state i, where n is the number of states.

$MOAC$ is mean (average) over all cases [19].

Table 1 Scoring Rule Results of Trichinellosis Risk in Humans

Scoring rule results	Values
Logarithmic loss	0.3628
Quadratic loss	0.1804
Spherical payoff	0.8963
Error rate	15%

All scoring rule results show that the prediction of the model is well calibrated.

Another way to verify the validity of the models is to use Netica to pass through the case file by processing the cases one by one. For each case, the software reads the case except the nodes whose inferences we wish to find. After that, the software will revise the actual values for those nodes and compare them with the beliefs the model generated. Netica accumulates all the comparisons, as illustrated in Table 2.

Table 2 Probability Table of Trichinellosis Risk in Humans

Household	p($RSKHUM$=High)		p($RSKHUM$=Medium)		p($RSKHUM$=Low)	
	$RSKHUM_i$	$\widehat{RSKHUM_i}$	$RSKHUM_i$	$\widehat{RSKHUM_i}$	$RSKHUM_i$	$\widehat{RSKHUM_i}$
1	0	0	0.33	0.07	0.67	0.93
10	0	0	0	0	1.00	1.00
13	0	0	0.67	0.86	0.33	0.14
15	0	0	0.67	0.75	0.33	0.25
17	0	0	0.33	0.05	0.67	0.95
19	0.33	0.24	0.33	0.51	0.34	0.25
20	0.33	0.06	0	0	0.67	0.94
21	0.33	0.08	0	0	0.67	0.92
22	0	0	0	0	1.00	1.00
23	0	0	0	0	1.00	1.00
25	0.33	0.14	0.67	0.86	0	0
26	0	0	0	0	1.00	1.00
27	0	0	0	0	1.00	1.00
31	0	0	1.00	1.00	0	0
36	0.67	0.49	0.33	0.51	0	0
37	1.00	1.00	0	0	0	0
41	0	0	0	0	1.00	1.00
42	1.00	1.00	0	0	0	0
44	0.67	0.48	0	0	0.33	0.52
46	0.67	0.40	0.33	0.60	0	0
47	0	0	0.33	0.11	0.67	0.89
51	0.33	0.11	0.33	0.19	0.34	0.70

4 Conclusion

4.1 Benefit of Integration of Eco-HealthOne-Health Transdisciplinary Approach and Bayesian Belief Network

A Bayesian Belief Network is a complex mathematical model incorporating the qualitative and quantitative aspects of a problem, and has been used for decision making in human management systems. This research combined the Bayesian Belief Network modeling for the analysis of zoonoses risk and for creating a combined human and animal health framework that can be used to guide decision making for interventions to solve the Eco-Health One-Health problem of Trichinellosis risk.

4.2 Constraint of Bayes Risk Minimization Using Symmetric Loss Function

The scoring rule results from Netica provide only symmetric loss values based on the assumption that the loss is the same in any direction. Nonetheless, this

assumption may not be valid in some practical situations such as the one that is of interest to us in this research, Trichinellosis risk. The false classification of the risk in the less severe direction (underestimation) would be more serious than a classification in the more severe direction (overestimation); in such a case, the policy maker would have to provide protection intervention [15]. Let $p(RSKHUM_i|X)$ be the posterior distribution of a discrete random variable $RSKHUM_i$. The objective is to find the estimator $\widehat{RSKHUM_i}$ that minimizes the Bayes risk;

$$r^*(\widehat{RSKHUM_i}, p(RSKHUM_i|X))$$

$$= min \sum_{RSKHUM_i} L(RSKHUM_i, \widehat{RSKHUM_i}) p(RSKHUM_i|X))$$

$(\widehat{RSKHUM_i} - RSKHUM_i)$ is the error from the estimation. Therefore, the loss function $L(RSKHUM_i, \widehat{RSKHUM_i})$ can be written in the form;

$$L(RSKHUM_i, \widehat{RSKHUM_i})$$

$$= \begin{cases} L_1(RSKHUM_i, \widehat{RSKHUM_i});\ \widehat{RSKHUM_i} \geq RSKHUM_i \\ L_2(RSKHUM_i, \widehat{RSKHUM_i});\ \widehat{RSKHUM_i} < RSKHUM_i \end{cases}$$

In the case of overestimation, $\widehat{RSKHUM_i} \geq RSKHUM_i$, the society bears only the cost of protection, such as field visits, to educate people to stop consuming raw or undercooked pork while the underestimation bears a lot more if the outbreak of the disease takes place.

4.3 Lack of Optimal Sample Size Determination

Due to the time and financial constraints, this research does not fulfill the optimal number of the sample size. The optimal sample size of the symmetric and asymmetric loss functions can be determined differently. Also, there are two main approaches in determining the size of the sample in the Bayesian study. The first approach is based on the probability coverage and the length of the interval containing the parameter. The second approach takes the idea of the decision theory that combines the costs of collecting a sample, which minimizes the pre-posterior expected cost. The later approach suggests that if the sampling cost of collecting data is very high, or if there is strong prior information about the risk, it is not worth sampling [15]. Also, if the loss of illness is very high, a thorough protection strategy would be more efficient.

Acknowledgements. We gratefully acknowledge the administrative assistance of Eco-HealthOne-Health Research Center, Faculty of Veterinary Medicine, Chiang Mai University; research funds from the EcoZD program of the International Livestock Research

Institute (ILRI); guidance and advice provided by Fred Unger (ILRI) and Jeff Gilbert (ILRI); veterinary technical advice from Jenny Steele (Tufts University), Karin Hamilton (University of Minnesota), Veerasak Punyapornwithaya (CMU), Khwanchai Kreausukon (CMU), Warangkhana Chaisowwong (CMU), and Pranee Rodtian (DLD Thailand); and modeling technical advice from Chalermpol Samranpong. We are especially indebted to the headmen of our two study villages, Mr. Abhinan Taotao and Mr. Lisor Jalor. Field assistance was provided by the enumerators from Faculty of Economics and Faculty of Veterinary, Chiang Mai University. GIS and mapping support was provided by Kongchak Jaidee, Global Health Asia, Faculty of Public Health, Mahidol University.

References

1. Carpenter, T.E.: A Microeconomic Evaluation of the Impact of Mycoplasma Meleagridis Infection in Turkey Production. Preventive Veterinary Medicine 1(4), 289–301 (1983)
2. Carpenter, T.E., Berry, S.L., Glenn, J.S.: Economics of Brucella Ovis Control in Sheep: Computerized Decision-tree Analysis. Journal of the American Veterinary Medical Association 190(8), 983–987 (1987)
3. Carpenter, T.E., Norman, B.B.: An Economic Evaluation of Metabolic and Cellular Profile Testing in Calves to be Raised in a Feedlot. Journal of the American Veterinary Medical Association 183(1), 72–75 (1983)
4. Carpenter, T.E., Thieme, A.: A Simulation Approach to Measuring the Economic Effects of Foot-and-mouth Disease in Beef and Dairy Cattle. In: Proceeding of the Second International Symposium on Veterinary Epidemiology and Economics, ISVEE, Canberra, Australia, pp. 511–516 (1980)
5. Carpenter, T.E., Howitt, R.E.: A Model to Evaluate the Subsidization of Governmental Animal Disease Control Programs. Preventive Veterinary Medicine 1(1), 17–25 (1988)
6. Christiansen, K.H., Carpenter, T.E.: Linear Programming as a Planning Tool in the New Zealand Brucellosis Eradication Scheme. In: Third International Symposium on Veterinary Epidemiology and Economics, pp. 369–376. Veterinary Medicine Publishing, Edwardsville (1983)
7. Dambacher, J.M., Shenton, W., Hayes, K.R., Hart, B.T., Barry, S.: Qualitative Modelling and Bayesian Network Analysis for Risk-based Biosecurity Decision Making in Complex Systems (2007), http://www.acera.unimelb.edu.au/materials/endorsed/0601.pdf
8. Ellis, P.R.: An Economic Evaluation of the Swine Fever Eradication Programme in Great Britain Using Cost Benefit Analysis Techniques. Department of Agriculture, University of Reading, Reading, U.K. (1972)
9. Figueiredo, M.A.T.: Lecture Notes on Bayesian Estimation and Classification. Technical University of Lisbon, Lisbon, Portugal (2004)
10. Gertheiss, J., Tutz, G.: Feature Selection and Weighting by Nearest Neighbor Ensembles. Technical Report Number 033, Department of Statistics, University of Munich, Munich, Germany (2008), http://epub.ub.uni-muenchen.de/4479/1/tr033.pdf
11. Roy, K.C., Blomqvist, H.C., Hossein, I.: Development that Lasts, pp. 201–214. New Age International, New Delhi (2007)

12. Harrison, S.R.: Cost-Benefit Analysis with Applications to Animal Health Programmes. Research Papers and Reports in Animal Health Economics, No.18-23. The University of Queensland, Brisbane, Australia (1996)
13. Howe, K.: The Economics of Disease Control. Veterinary Record 117(15), 375 (1985)
14. Howe, K., Christinsen, K.H.: The State of Animal health Economics: A Review. In: Proceedings of the Society for Veterinary Epidemiology and Preventive Medicine, pp. 153–165 (2004)
15. Islam, S.: Loss Functions, Utility Functions and Bauesian Sample Size Determination. Doctor of Philosophys thesis. University of London, London, U.K. (2011)
16. Jakeman, A.J., Letcher, R.A., Norton, J.P.: Ten Iterative Steps in Development and Evaluation of Environmental Models. Environmental Modelling & Software, 1–13 (2006), http://www.iemss.org/iemss2006/papers/w4/TenSteps.pdf
17. Marsh, W.: The Economics of Animal Health in Farmed Livestock at the Herd Level. Rev. Sci. Tech. Off. Int. Epiz. 18(2), 357–366 (1999)
18. Mohammed, H.O., Carpenter, T.E., Yamamoto, R.: Economic Impact of Mycoplasma Gallisepticum and M. Synoviae in Commercial Layer Flocks. Avian Diseases 31(3), 477–482 (1987)
19. Norsys Software Corp. Decision-Making Nets (2013), http://www.norsys.com/WebHelp/NETICA/X_Quick_Tour_Decision_Problems.htm
20. Otte, M.J., Chilonda, P.: Animal Health economics: An Introduction. Livestock Information, Sector Analysis and Policy Branch, Animal Production and Health Division, FAO, Rome (2001)
21. Roberts, T.: Chapter 13: Risk Assessment Models, Economic Analysis and Food Safety Policy. The Economics of Infectious Disease. Oxford University Press Inc., New York (2006)
22. Rodrigues, C.A., Gardner, I.A., Carpenter, T.E.: Financial Analysis of Pseudorabies Control and Eradication in Swine. Journal of the American Veterinary Medical Association 197(10), 1316–1323 (1990)
23. Ruegg, P.L., Carpenter, T.E.: Decision-tree analysis of Treatment Alternatives for Left Displaced Abomasum. Journal of the American Veterinary Medical Association 195(4), 464–467 (1989)
24. Rushton, J.: The Economics of Animal Health and Production. CABI, Wallingford (2009)
25. Thorburn, M.A., Carpenter, T.E., Plant, R.E.: Perceived Vibriosis Risk by Swedish Rainbow Trout Net-pen Farmers: Its Effect on Purchasing Patterns and Willingness-to-pay for Vaccination. Preventive Veterinary Medicine 4(5-6), 419–434 (1987)
26. Tisdell, C.A.: Economics of Controlling Livestock Diseases: Basic Theory. Economics, Ecology and the Environment. Working Paper No. 134. Brisbane: University of Queensland, School of Economics (2006)
27. UNBC (n.d.) UNBC-Ecohealth 2011 (2011), http://www.unbc.ca/qrrc/photo_albumeco.html
28. Vagsholm, I., Carpenter, T.E., Howitt, R.E.: Shadow Costs of Disease and Its Impact on Output Supply and Input Demand: The Dual Estimation Approach. Preventive Veterinary Medicine 10(3), 195–212 (1991)
29. Zessin, K.H., Carpenter, T.E.: Benefit-Cost Analysis of an Epidemiological Approach to Provision of Veterinary Services in the Sudan. Preventive Veterinary Medicine 3, 323–337 (1985)

Appendix

Table 3 Abbreviations of Variables Used in Bayesian Belief Network

Abbrev.	Topic	States	Descriptions
TRICPIG	Trichinella infection in pig	High	High risk that pigs are infected by Trichinella
		Medium	Medium risk that pig are not infected by Trichinella
		Low	Low risk that pig are not infected by Trichinella
RSKHUM	Risk of getting Trichinellosis in human	High	High risk of getting Trichinellosis in human
		Medium	Medium risk of getting Trichinellosis in human
		Low	Low risk of getting Trichinellosis in human
ENVMN	Environment suitability of Trichinella circulation	Suitable	The environment of this household is suitable for Trichinella circulation
		Fair	The environment of this household is fair for Trichinella circulation
		Not suitable	The environment of this household is not suitable for Trichinella circulation
PRACTC	Rearing practice	Suitable	Rearing practice is suitable
		Fair	Rearing practice is fair
		Not suitable	Rearing practice is not suitable
GENDFD	Gender of food-preparing person	Male	Food-preparing person is a man
		Female	Food-preparing person is a woman

Table 3 (*continued*)

Abbrev.	Topic	States	Descriptions
INCOME	Income level	$< 12,000$	People receive less than 12,000 Baht annually
		$12,000 - 60,000$	People receive $12,000 - 60,000$ Baht annually
		$60,001 - 100,000$	People receive $12,000 - 60,000$ Baht annually
		$100,001 - 180,000$	People receive $12,000 - 60,000$ Baht annually
		$> 180,000$	People receive greater than 180,000 Baht annually
EDUFD	Formal education of food-preparing person	None	Food-preparing person does not go to school
		< P.3	Food-preparing person does not finish Prathom 3 (Grade 3)
		P.3	Food-preparing person finishes Prathom 3 (Grade 3)
		P.6	Food-preparing person finishes Prathom 6 (Grade 6)
		M.3	Food-preparing person finishes Mathayom 3 (Grade 9)
MTPRE	Meat preparation	Raw or undercooked meat	People usually consume raw or undercooked meat
		Cooked meat	People usually consume cooked meat
PREFER	Raw/undercooked meat consumption preference	Not like	People do not like consuming raw/undercooked meat
		Like	People like consuming raw/undercooked meat
RAWDNG	Recognition of the danger of consuming raw/undercooked meat	No	People do not know the danger of consuming raw/undercooked meat

Table 3 (*continued*)

Abbrev.	Topic	States	Descriptions
		Yes	People know the danger of consuming raw/undercooked meat
FREQMT	Frequency of consuming meat in a year	< 12 times	People consume meat less than 12 times in a year
		$12 - 30$ times	People consume meat around $12 - 30$ times in a year
		$31 - 50$ times	People consume meat around $31 - 50$ times in a year
		$51 - 100$ times	People consume meat around $51 - 100$ times in a year
		$101 - 300$ times	People consume meat around $101 - 300$ times in a year
		> 300 times	People consume meat more than 300 times in a year
CONHAB	Consumption habits	Good	The consumption habits are good
		Fair	The consumption habits are fair
		Poor	The consumption habits are poor
KNWFD	Knowledge of food-preparing person	High	Food-preparing person has high knowledge about how to prepare good food
		Medium	Food-preparing person has medium knowledge about how to prepare good food
		Low	Food-preparing person has low knowledge about how to prepare good food

Factors Affecting Hospital Stay Involving Drunk Driving and Non-Drunk Driving in Phuket, Thailand

Jirakom Sirisrisakulchai and Songsak Sriboonchitta

Abstract. The purpose of this paper is to investigate the factors affecting hospital stay involving drunk driving and non-drunk driving of accident victims in Phuket, Thailand. However, the decisions to drink and drive are made by the road users before the accidents occurred. Self-selection bias may arise when some component of the drink and drive decision is relevant to the length of stay in hospital of accident victims. We discuss a new approach to specifying and estimating zero-inflated negative binomial models with endogenous switching, based on copula function. These models provide a framework of analysis for the effect of self-selection in drunk-driving behaviors on the length of stay in the hospitals. We use the concept of pair-copula constructions for discrete margins to derive the likelihood function of our model system. The results suggest that drunk-driving accident victims are positively selected.

1 Introduction

Drunk driving (or driving under the influence of alcohol) increases the risk of crashes and violation of traffic laws, which leads to higher severity of injury. In developed countries about 20% of fatally injured drivers have alcohol in their blood more than the legal limit. For developing countries, this proportion may be up to 69% (WHO, 2007). Injury surveillance (IS) data of Thailand from 1999 to 2004 showed that 35% to 48% of injury victims aged older than 15 years had illegal blood alcohol concentration (BAC) above 50 mg/100ml (Ditsuwan et al., 2013). Moreover, there is probably and underestimate of true proportion since the victims who died at scene had no BAC test administered (Aungkasuvapala, 2003).

It is often argued that drunk driving in comparison with non drunk driving should have the higher probability of encountering a serious injury under the same

Jirakom Sirisrisakulchai · Songsak Sriboonchitta
Faculty of Economics, Chiang Mai University, Chiang Mai 50200 Thailand
e-mail: sirisrisakulchai@hotmail.com, songsakecon@gmail.com

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_31,

circumstances. Eluru and Bhat (2007) studied the influence of seat belt usage on the crash related injury severity. They found that safety conscious drivers are more likely to wear seat belts, and this defensive habit leads to less severe injury. Krull, Khattak, and Council (2000) studied the factors affecting the probability of fatal and incapacitating injuries. They found that the factors contributing to fatal and severe injury are rollover involvement, seat belt usage, alcohol consumption, rural roads and violating speed limits. Kasantikul et al. (2005) studied the role of alcohol in Thailand motorcycle crashes and found that drinking riders were more likely to be hospitalized, stay longer in the hospital and more likely to be killed. Santolino et al. (2012) showed that the factors affecting the hospital stay are age of victim, gender, and nature of injuries.

The purpose of this paper is to investigate the factor affecting hospital stay involving drunk driving and non drunk driving of accident victims in Phuket, Thailand. The analysis is based on IS data from Vachira Phuket Hospital. A central issue here is the debate whether any effect of the self selection in alcohol consumption of accident victims on the length of stay in the hospital is causal or associative. The conventional analysis assumes that, once the explanatory variables that affect the outcome variable are taken into account, the process by which individual are sorted into positions is independent of factors influencing the outcome variable itself (Mare and Winship, 1988).

However, sample selection problem arises when sampling observations are generated from the population by rules other than simple random sampling (Lung-Fei Lee, 2003). This make the sample representation of a true population is distorted. One possible reason in practice is called self-selection bias in which distorted sample generation result from self-selection decisions by the agents being studied. The possibility of sample selection bias arises when there are unobservable characteristics that influence both the observed outcomes and the decision process. In our situation, the decisions to drink and drive are made by the road users before the accidents occurred. Self-selection bias may arise when some component of the drink and drive decision is relevant to the length of stay in hospital of accident victims.

If this self-selection is not accounted for, the parameter estimates using the conventional method will give inconsistent estimates of the effect of the individual decisions (the decisions to drink and drive) and of the other explanatory variables in the model. In this paper, we apply the copula approach, which allows for the joint determination of the discrete variables and the outcome that they affect, to examine the effect of the alcohol consumption of the accident victim on the length of the stay in the hospital.

The rest of this paper is organized as follows. The next section provides a description of the data used. In section 3 and 4, we provide a brief discussion of switching regression model for hospital stay and the copula approach for this model. Section 5 presents and discusses the modeling results, followed by conclusion in section 6.

2 Switching Regression Model for Hospital Stay

Consider two cases; drunk driving and non drunk driving; that are self selected by the individual rather than randomly assigned. We want to estimate the difference in length of stay in the hospitals between alcohol related injuries and non-alcohol related injuries. However, we can observe two outcomes from the hospital data, namely, injuries (including slightly injured and seriously injured) and fatalities. Therefore, only people injured can observe the length of stay in the hospitals (either zero or positive numbers) and people who died from accident cannot observe the length of stay.

The situation discussed above can be modeled by applying Roy's (1951) endogenous switching model system (See Cameron and Trivedi, 2005 for the detailed discussion) with additional stage three on the censored outcome (length of stay). The first stage distinguishes drunk driving from non drunk driving using binary outcome model. In the second stage, similar to the first stage, we use separated binary outcome models to distinguish the injury cases from the fatality cases. Finally, in the third stage, the determinants of the length of stay are identified in separated count regressions.

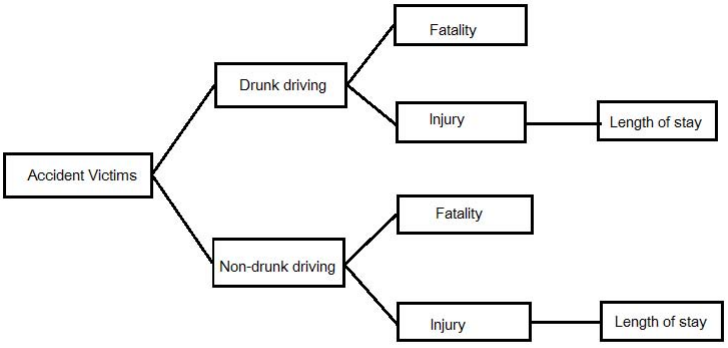


Fig. 1 Graphical Illustration of the three stages model for length of stay in the hospitals

To derive a likelihood function of the above model system, we start in the first stage where accident victims are identified according to whether they are drunk driving or not using binary outcome model. Let $Y_1 = 0, 1$ be the binary outcomes; where 1 is drunk driving and 0 is non drunk driving; X_1 the vector of all explanatory variables thought to explain self selected drunk driving behavior, and β_1 a vector of parameters to be estimated. We can think of this random variable Y_1 is generated from the binomial distribution:

$f(k;n,p) = \Pr(Y_1 = k) = \binom{n}{k} p^k (1-p)^{n-k}$, for $k = 0, 1, 2, \ldots, n$, where

$\binom{n}{k} = \frac{n!}{k!(n-k)!}$, n is the number of trials, and p is the probability of success.

In our case, $n = 1$[1] and if $p$is assumed to be a standard normal distribution (Φ). We can derive the probability distribution as follows:

$$\Pr(Y_1 = 1 \,|X_1, \beta_1\,) = \Phi\,(X_1\beta_1)$$

$$\Pr(Y_1 = 0\,|X_1, \beta_1\,) = 1 - \Phi\,(X_1\beta_1).$$

Focusing on the second stage, we define $Y_{2,i} = 0, 1$ as the binary outcomes; where 1 is an accident victim who dies in the hospital and 0 is an injured accident victim; $X_{2,i}$as the vector of all variables explaining the characteristics of the accident victims, $\beta_{2,i}$as a vector of parameters to be estimated, and i as the indicator for drunk driving or non drunk driving as defined previously for selection variable Y_1 . However, this is not a standard binary outcome model since the probability of observing fatality or injury depends on the outcome of the selection variableY_1, and $Y_1$and $Y_{2,i}$are not necessarily independent. Thus we have to specify the joint distribution of these two variables, which will be discussed in the next section.

Finally, in the third stage, let $Y_{3,i}$be defined as the number of day that accident victim stay in the hospital, $X_{3,i}$the vector of all determinants for the length of stay in the hospital, $\beta_{3,i}$a vector of parameters to be estimated, and i is the indicator for drunk driving or non drunk driving as defined previously for selection variable Y_1. Each of these variables$Y_{3,i}$ is assumed to be discrete distribution for count variables such as Poisson or Negative Binomial distribution, for instance. However, as discussed previously in the second stage, $Y_{3,i}$ are dependent with $Y_{2,i}$andY_1. Thus the joint distribution for these three variables is needed. Now, we can derive the likelihood function of our model system as follows:

For $Y_1 = 0$ and $Y_{2,0} = 0$,
$L_1 = \prod \{ \Pr(Y_{3,0}\,|Y_{2,0} = 0, Y_1 = 0\,) \times \Pr(Y_{2,0} = 0\,|Y_1 = 0\,) \times \Pr(Y_1 = 0)\}$
For $Y_1 = 0$ and $Y_{2,0} = 1$,
$L_2 = \prod \{ \Pr(Y_{2,0} = 0\,|Y_1 = 0\,) \times \Pr(Y_1 = 0)\}$
For $Y_1 = 1$ and $Y_{2,1} = 0$,
$L_3 = \prod \{ \Pr(Y_{3,1}\,|Y_{2,1} = 0, Y_1 = 0\,) \times \Pr(Y_{2,1} = 0\,|Y_1 = 0\,) \times \Pr(Y_1 = 0)\}$
For $Y_1 = 1$ and $Y_{2,1} = 1$,
$L_4 = \prod \{ \Pr(Y_{2,1} = 0\,|Y_1 = 0\,) \times \Pr(Y_1 = 0)\}$

For our model system above, it is natural to model the interdependence between outcome equations and selection equation using copula functions which will be discussed in the next section.

3 Copula Approach for Modeling Switching Regression

The notion of a copula was introduced by Sklar (1959). A copula is a function that links multivariate joint distribution with pre-specified univariate distribution

[1] This is the special case of binomial distribution which is called Bernoulli distribution.

function. Introduction of copula can be found in Joe (1996) and Nelsen (2006). An introduction for empirical researcher is provided by Trivedi and Zimmer (2007).

The foundation of copula is based on the theorem of Sklar (1959), which states that given a joint distribution function$F$, and respective marginal distribution functions, there exists a copula C such that the copula binds the margins to give the joint distribution, that is

$$F(y_1, y_2, ..., y_m) = C(F_1(y_1), F_2(y_2), ..., F_m(y_m)) \tag{1}$$

Where $y = (y_1, y_2, ..., y_m)$ is the realization of an m-dimensional random vector$Y = (Y_1, Y_2, ..., Y_m)$. $F_j(y_j)$ is the marginal distribution function of thej^{th}margin for$j = 1, 2, ..., m$.

The fact that copula function can be used to build new multivariate distribution for given univariate marginal distribution is useful for econometric modeling. The key is that copula can introduce dependence between the two or more random variables. The degree and type of dependence depends on the choice of copula. In our analysis, it is essential that copula allows for positive and negative correlation. We want to learn from the data whether drunk driving has more, less, or equally length of stay in the hospital in comparison to non drunk driving. Therefore, we consider three copula functions in this paper, namely, the Normal copula, the Frank copula, and the Independence copula. Both Normal and Frank copula can reach the Frechet upper bound and lower bound (See Trivedi and Zimmer, 2007, for a detailed discussion). They can span the full range of dependence.

Some authors proposed the copula approach for modeling endogenous switching regressions and sample selection models. Ophem (2000) and Zimmer and Trivedi (2006) used copula approach for event count outcomes. Prieger (2002) and Smith (2003) offer copula-based selectivity models. Chinakum et al. (2013) applied the copula approach to a sample selection modeling of panel data. Bhat and Eluru (2009) showed the application of endogenous switching regressions for continuous outcome in transportation research fields. Spissu et al. (2009) applied copula approach to the case of sample selection with a multinomial treatment effect. Luechinger et al. (2010) proposed the copula model of ordered outcome with endogenous self-selection. It should be noted that these papers re-formulate discrete random variables as continuous latent variables and then construct a joint distribution of the latent variables by a continuous copula. In contrast, we use a pair copula construction for discrete margins as proposed by Panagiotelis (2012).

Our model system can be viewed as a three dimensions copula with discrete margins. The dependent variables for the first and the second stage are binomial distribution margins. The dependent variable for the third stage is zero-inflated negative binomial distribution, which has the following form:

$$\Pr(Y_{3,i} \leq n) = \varphi_{3,i} + (1 - \varphi_{3,i}) \sum_{k=0}^{n} f(k; \mu_{3,i}, \theta_{3,i}),\ n = 0, 1, 2, ...;\ i = 0, 1 \tag{2}$$

Where $f(k;\mu_{3,i})$represents the probability mass function (pmf) of the univariate negative binomial distribution with the dispersion parameter$\theta_{3,i}$, and$\mu_{3,i}$and $\phi_{3,i}$are given by

$$\mu_{3,i} = \exp(X_{3,i}\beta_{3,i}), \log\left(\frac{\varphi_{3,i}}{1-\varphi_{3,i}}\right) = Z_{3,i}\gamma_{3,i}$$

Where $Z_{3,i}$are the vectors of explanatory variables for zero outcome equations and $\gamma_{3,i}$ are the corresponding vectors of parameter.

For each state,$i = 0, 1$, in our analysis we need three bivariate copula functions, i.e., $C_{12,i}, C_{13,i}$,and $C_{23|1,i}$. We perform the derivation of three-dimensional joint pmf by following Panagiotelis (2012);

$$\begin{aligned}&\Pr(Y_1 = y_1, Y_{2,i} = y_2, Y_{3,i} = n)\\ &\quad = \Pr(Y_{3,i} = n\,|Y_{2,i} = y_2, Y_1 = y_1\,) \times \Pr(Y_{2,i} = y_2\,|Y_1 = y_1\,) \times \Pr(Y_1 = y_1)\end{aligned} \tag{3}$$

where

$$\begin{aligned}&\Pr(Y_{3,i} = n\,|Y_{2,i} = y_2, Y_1 = y_1)\\ &\quad = \frac{\left\{\Sigma_{i_1=0,1}\Sigma_{i_2=0,1}(-1)^{i_1+i_2}C_{23|1,i}\,(F_{2|1,i}\,(y_2-i_2|y_1\,),F_{3|1,i}\,(y_1-i_1|y_1\,))\right\}}{\Pr(Y_{2,i}=y_2|Y_1=y_1)}\end{aligned} \tag{4}$$

and the arguments in the copula function are

$$F_{2|1,i}\,(y_2 - i_2\,|y_1\,) = \frac{C_{12,i}(F_1(y_1), F_{2,i}(y_2 - i_2)) - C_{12,i}(F_1(y_1 - 1), F_{2,i}(y_2 - i_2))}{\Pr(Y_1 = y_1)},$$

and

$$F_{3|1,i}\,(n - i_3\,|y_1\,) = \frac{C_{13,i}(F_1(y_1), F_{3,i}(y_3 - i_3)) - C_{13,i}(F_1(y_1 - 1), F_{3,i}(n - i_3))}{\Pr(Y_1 = y_1)},$$

$$\begin{aligned}\Pr(Y_{2,i} = y_2\,|Y_1 = y_1\,) = {}& C_{12,i}(F_1(y_1), F_{2,i}(y_2)) - C_{12,i}(F_1(y_1 - 1), F_{2,i}(y_2))\\ & - C_{12,i}(F_1(y_1), F_{2,i}(y_2 - 1)) + C_{12,i}(F_1(y_1 - 1), F_{2,i}(y_2 - 1))\end{aligned}, \tag{5}$$

and

$$\Pr(Y_1 = y_1) = F_1(y_1) - F_1(y_1 - 1). \tag{6}$$

By substituting equation (4) – (6) into the likelihood function, we can maximize the log-likelihood function to get the parameter estimates for our model system.

4 Data

The data used in this paper is an individual data of accident victim who was sent to Vachira Phuket hospital, Thailand. The data consist of a sample of 10,218 accident victims involved in traffic crashes during the period of 2008 to 2012. The data on fatality are the death occurred at the scene and at the emergency department. Since there is no regulation required for BAC testing for all traffic injury cases, drunk-driving accident victims were subjectively evaluated for alcohol involvement by

the hospital officers. The accident victims are simply divided into two groups: had been drinking (drunk driving) or had not been drinking (non-drunk driving). The duration of hospital stay is observed for all sample victims. Victims have to be fully recovered and discharged from hospital. About 30% of the victims are classified as drunk driving and 70% are classified as non-drunk driving. The average length of stay is about 3.95 days for non-drunk driving, and 4.98 days for drunk driving. The description of variables used in this paper and main statistics are shown in Table 1. We found the existence of excessive zero for duration of hospital stay in the dataset. Zero represent 28.5% of the data.

Table 1 Description of variables and statistics

Variable	Label	Description	N	Mean	SD	Min.	Max.
Dependent variables							
Y1	Had been drinking or not	1 if victim had been drinking; 0 otherwise	10218	0.298	0.458	0	1
Y2	Fatality or injury	1 if victim died; 0 otherwise	10218	0.011	0.104	0	1
Y3	Hospital stay	Length of hospital stay (number of days)	10218	4.26	6.641	0	50
Explanatory variables							
x1	Age	Age of the victim	10218	30.182	14.773	0	89
x2	Male	1 if the victim is male; 0 otherwise	10218	0.654	0.476	0	1
x3	Driver	1 if the victim was the driver; 0 otherwise	10218	0.771	0.42	0	1
x4	Motorcycle	1 if the victim used motorcycle; 0 otherwise	10218	0.853	0.354	0	1
x5	Car	1 if the victim used passenger car or pickup or van; 0 otherwise	10218	0.035	0.185	0	1
x6	Night	1 if the accident occurred during 6.01-24.00; 0 otherwise	10218	0.322	0.467	0	1
x7	Severe	The severity index 0-6; 6 is the highest level and 0 is the lowest level	10218	1.751	1.096	0	6
x8	Head	1 if injury located in head; 0 otherwise	10218	0.216	0.412	0	1
x9	Thorax	1 if injury located inthorax; 0 otherwise	10218	0.019	0.138	0	1
x10	Pelvic	1 if injury located in pelvic; 0 otherwise	10218	0.35	0.477	0	1
x11	Abdomen	1 if injury located in abdomen; 0 otherwise	10218	0.032	0.175	0	1

5 Empirical Results

The explanatory variables considered in our analysis consist of several categories, including individual characteristics, mode of travels, time of accident, severity index, and the injured body regions. We selected the individual characteristics, mode of travels and time of accident as the explanatory variables for selection equation and binary outcomes for fatality or injury equations. For hospital stay equations, we added severity index and the injured body region variables for the determinants of the length of stay in the hospitals. We encounter the convergence problems for Normal copula. Therefore, the results provided here are only Independence and Frank copula for the discussions. The Frank copula model rejects the independence assumption based on likelihood ratio test, indicating the significant existence of self-selection effects. In the following discussions, we focus on the results of the Frank copula model specification.

5.1 Binary Choice Equation for Alcohol Consumption

Table 2 gives the results of selection equation. The results of the binary outcome equation of self-selected alcohol consumption provide the effects of variable on the

propensity to drink driving relative to non-drink driving. All parameter estimates were statistically significant at standard level. The parameter estimate shows that males are more likely to drink and drive. Both car and motorcycle users are more likely to consume alcohol before travel when compare with other modes such as bicycle, truck and taxi. There is a linear relationship between age and propensity to drink driving as expected. Older victims are more likely to drink and drive when compare with the younger.

Table 2 Estimation results of selection equation

	Had been drinking or not	
	Independence	Frank
selection equation		
(Intercept)	**-1.865***	**-1.597***
male	**0.893***	**0.787***
driver	**0.165***	**0.109***
age	**0.008***	**0.007***
motorcycle	**0.406***	**0.302***
car	**0.385***	**0.354***
night	**-0.143***	**-0.148***

*indicate statistical significance at the 5% level.

5.2 Binary Outcome for Fatality or Injury

Table 3 provides the estimation results of binary outcome equation for fatality or injury. Only the dummy variable for motorcycle user was found statistically significant. The result indicates that motorcycle users are more likely to injure from the accidents rather than die at the scene. The parameter estimates of non-drunk driving regime are higher than drunk driving regime, indicating that drunk driving has higher probability to die from the accident when compare with non-drunk driving. The dependence parameters $\theta_{12,0}$and$\theta_{12,1}$translate to a Kendall's Tau value of 0.11 and 0.13, respectively.

5.3 Zero-Inflated Negative Binomial Models for the Length of Stay in the Hospital

The results of hospital stay outcomes are shown in Table 4. The length of stay in the hospitals can be explained by the injury location variables and severity index. However, the factor affecting the length of stay in the hospitals for drunk-driving and non-drunk driving regimes are different. For instance, head injuries, thorax injuries, and abdomen injuries are associated with longer stay in the hospitals for non-drunk driving but injury located at pelvic and abdomen are associated with longer stay in

Table 3 Estimation results of binary outcome equation for fatality or injury

	Independence		Frank	
	Non-drunk driving	Drunk driving	Non-drunk driving	Drunk driving
Outcome equation 1: fatality or injury				
(Intercept)	**-1.936***	**-1.546***	**-1.983***	**-2.411***
male	-0.052	-0.023	0.597	0.123
motorcycle	**-0.778***	**-0.317***	**-0.547***	**-0.291***
age	-0.002	-0.004	0.005	-0.003
night	-0.041	-0.04	-0.157	-0.071
$\theta_{12,0}$			1.008	
$\theta_{12,1}$				1.198

*indicate statistical significance at the 5% level.

Table 4 Estimation results of hospital stay outcomes

	Independence				Frank			
	Non-drunk driving		Drunk driving		Non-drunk driving		Drunk driving	
Variable	zero	outcome	zero	outcome	zero	outcome	zero	outcome
Outcome equation: Hospital stay								
intercept	**1.957***	**0.634***	1.251	**0.583***	-8.99	**-0.461***	**2.641***	**-1.15***
male	**-0.282***	0.036	0.243	**0.152***	**-0.615***	**0.458***	**-1.558***	**0.689***
age	**-0.009***	**0.008***	-0.002	**0.008***	-0.109	**0.0117***	**-0.012***	**0.013***
night	-0.245	0.009	**-0.993***	0.008	0.2	-0.007	-0.205	**-0.121**
motorcycle	**1.011***	**-0.186***	0.671	0.075	0.275	**-0.17***	**-0.627***	**0.276***
car	-0.293	0.037	1.099	0.126	0.611	**0.233***	-0.399	0.227
severe	**-3.178***	**0.416***	**-3.118***	**0.316***	**-1.778***	**0.814***	**0.864***	**0.379***
head	**-1.817***	**-0.165***	**-1.605***	-0.017	**-0.682***	**0.100***	0.120	0.048
thorax	-2.350	0.056	0.365	-0.129	-1.509	**0.315***	-0.756	-0.008
pelvic	**-1.143***	-0.080	-2.635	0.101	**-1.608***	0.061	-0.361	**0.139***
abdomen	-3.071	**0.343***	-15.956	**0.430***	-3.8606	**0.508***	-1.044	**0.274***
dispersion		**0.811***		**0.905***		**0.956***		**0.835***
$\theta_{13,0}$					**6.082***			
$\theta_{13,1}$							**21.939***	
$\theta_{23,0}$						**5.614***		
$\theta_{23,1}$								**10.322***
LL			29880			29420		

*indicate statistical significance at the 5% level.

the hospital for drunk driving. The surprising result for non-drunk driving regime is that motorcycle users are less likely to stay longer in the hospital. One explanation might be non-drunk motorcycle users ride more careful than drunk driving motorcycle users. For the time of accident occurred, we found that when accident occurred between 0:00 to 6:00 a.m. the accident victims are more likely to stay longer in

the hospitals. Finally, as expected the male accident victims are more likely to have serious injury when compare with the female victims.

The dependency parameters for the Frank copula model are highly statistically significant and positive. The dependence parameters $\theta_{13,0}, \theta_{13,1}, \theta_{23|1,0}$, and $\theta_{23|1,1}$ translate to a Kendall's Tau value of 0.52, 0.83,0.49, and 0.67 respectively. The positive dependency indicates that the probability of reporting longer stay in the hospital in drunk-driving regime is higher for persons who actually chose that regime, relative to the others.

6 Conclusions

In this paper, we apply a copula based approach to model zero-inflated negative binomial regression with endogenous switching for the length of stay in the hospitals of accident victims using injury surveillance (IS) data from Vachira Phuket Hospital. Pair copula constructions for discrete margin are used to generate joint distributions for multivariate discrete random variables. We encountered the convergence problem for normal copula. The models presented in the text are Frank copula and Independence models. We found statistical evidence for positive self-selection on alcohol consumption. Failure to accommodate these self-selection effects can lead to an inconsistent estimate of parameters and mis-estimation of true effects. Although our model systems in the paper were motivated by a substantive issued related to health economic research, it is clear that they can be applied to other areas of empirical economics.

References

1. Aungkasuvapala, N., Santikarn, C., Chadbunchachai, W., et al.: Road traffic safety, a long journey of health promotion. The War Veterans Organization of Thailand, Bangkok (2003)
2. Bhat, C.R., Eluru, N.: A Copula-Based Approach to Accommodate Residential Self-Selection Effects in Travel Behavior Modeling. Transportation Research Part B 43(7), 749–765 (2009)
3. Cameron, A.C., Trivedi, P.K.: Microeconometrics: Methods and Applications. Cambridge University Press, New York (2005)
4. Chinnakum, W., Sriboonchitta, S., Pastpipatkul, P.: Factors affecting economic output in developed countries: A copula approach to sample selection with panel data. International Journal of Approximate Reasoning (2013)
5. Ditsuwan, V., Veerman, J.L., Bertram, M., Vos, T.: Cost-Effectiveness of Interventions for Reducing Road Traffic Injuries Related to Driving under the Influence of Alcohol. Value in Health 16, 23–30 (2013)
6. Eluru, N., Bhat, C.R.: A Joint Econometric Analysis of Seat Belt Use and Crash-Related Injury Severity. Accident Analysis and Prevention 39(5), 1037–1049 (2007)
7. Joe, H.: Multivariate Models and Dependence Concepts. Chapman & Hall, London (1997)
8. Kasantikul, V., Ouellet, J.V., Smith, T., et al.: The role of alcohol in Thailand motorcycle crashes. Accident Analysis and Prevention 37, 357–366 (2005)

9. Krull, K., Khattak, A., Council, F.: Injury effects of rollovers and events sequence in single-vehicle crashes. Presented at the 80th Annual Meeting of the Transportation Research Board, Washington D. C. (2000)
10. Lee, L.-F.: Self-Selection. In: Baltagi, B.H. (ed.) A Companion to Theoretical Econometrics. Blackwell Publishing (2003)
11. Luechinger, S., Stutzer, A., Winkelmann, R.: Self-selection Models for Public and Private Sector Job Satisfaction. Research in Labor Economics 30, 233–251 (2010)
12. Nelsen, R.B.: An Introduction to Copulas, 2nd edn. Springer, New York (2006)
13. van Ophem, H.A.: Modeling selectivity in count-data models. Journal of Business & Economic Statistics 18(4), 503–511 (2000)
14. Panagiotelis, A., Czado, C., Joe, H.: Pair Copula Constructions for Multivariate Discrete Data. Journal of the American Statistical Association 107(499), 1063–1072 (2012)
15. Prieger, J.E.: A flexible parametric selection model for non-normal data with application to health care usage. Journal of Applied Econometrics 17(4), 367–392 (2002)
16. Roy, A.D.: Some thoughts on the distribution of earnings. Oxford Economic Papers, New Series 3(2), 135–146 (1951)
17. Santolino, M., Bolance, C., Alcaniz, M.: Factors affecting hospital admission and recovery stay duration of in-patient motor victims in Spain. Accident Analysis and Prevention 49, 512–519 (2012)
18. Sklar, A.: Fonctions de repartition a n dimensions et leurs marges. Publications de l'Institut de Statistique de l'Universite de Paris 8, 229–231 (1959)
19. Smith, M.D.: Modelling sample selection using archimedean copulas. Econometrics Journal 6(1), 99–123 (2003)
20. Spissu, E., Pinjari, A.R., Pendyala, R.M., Bhat, C.R.: A copula-based joint multinomial discrete-continuous model of vehicle type choice and miles of travel. Presented at 88th Annual Meeting of the Transportation Research Board, Washington, DC (2009)
21. Trivedi, P.K., Zimmer, D.M.: Copula Modeling: An Introduction for Practitioners. Foundations and Trends in Econometrics 1(1) (2007)
22. Winship, C., Mare, R.D.: Endogenous Switching Regression Models for the Causes and Effects of Discrete Variables. In: Long, S.J. (ed.) Common Problems in Quantitative Social Research. Sage Press (1988)
23. World Health Organization. Drinking and Driving: a road safety manual for decision-makers and practitioners. Global Road Safety Partnership, Geneva (2007)
24. Zimmer, D.M., Trivedi, P.K.: Using trivariate copulas to model sample selection and treatment effects: Application to family health care demand. Journal of Business & Economic Statistics 24, 63–76 (2006)

How Macroeconomic Factors and International Prices Affect Agriculture Prices Volatility?-Evidence from GARCH-X Model

Gong Xue and Songsak Sriboonchitta

Abstract. This study explains China's agricultural commodities volatility by using the short-term deviations along with the domestic macroeconomic factors as well as the international price factors. The GARCH-X model shows that the short-term deviations make significant and positive effect on volatility, and so, it can be taken as an important factors in estimating and forecasting the agricultural prices. However, it is disappointing that some of the macroeconomic factors are not significant in our model. This is because China is in a transition process, and many macroeconomic factors are not freely moved. Our study also analyzes China's policy and macroeconomic changes in last decades. To give a more thorough understanding about China's recent macroeconomic reform is also one of our objectives.

1 Introduction

Agriculture has, for a long time, played a dominant role even in China's modern history. Although the percentage of agriculture is lower comparing to industries recently, the importance of agriculture to China has never been insignificant. (Cheng, 2005) [1] Previously, China was a net exporter, but now it has even started to import many kinds of farm products to satisfy the different requests and demands for the various kinds of nutrition needs. The export of agricultural commodities increased from 10.23 billion dollars in 1996 to 50.49 billion dollars in 2011 (an average annual growth rate is around 11.6 %), while the imports increased from 5.67 billion dollars to 28.77 billion dollars (an average annual growth rate is about 13.2 %). (NBSC, 2012) [2]

China has been going the market economy way for quite some time now,(Lin et al., 1996; Lin, 2003) [3, 4] and the previous practice of direct agricultural price control is gradually disappearing. First, the agricultural market is already formed. The

Gong Xue · Songsak Sriboonchitta
Faculty of Economics, Chiang Mai University, Chiang Mai 50200 Thailand
e-mail: {gongxue.cmu,songsak}@gmail.com

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_32,

scope for governmental action is limited. And, second, with the increasing volume of agricultural trade, the pricing power of domestic authorities was largely reduced, especially in the case of imported agricultural products, such as corn and soybean, as they are closely related to the prices in the world market. (Wang and Xie, 2012; Yang and Li, 2008) [5, 6] Here, there emerges a few questions: How do the macroeconomic factors influence the agricultural price? Do their effects work only on the price level or on the volatility level also? And, second, with the channel made for import and export, a relationship has been built between the domestic food prices and the international food prices; under these circumstances, in what way does the international agricultural price or energy price impact China's agricultural prices? The first objective of our study is to give a brief introduction to China's macroeconomic reform progress, in support of our empirical econometric study. The second objective is to investigate the impact of the short-run deviations between the agricultural prices and the domestic macroeconomic factors and international prices on food price volatility. The forecasting and measurement of food price volatility is of great importance since the volatility causes uncertainty to producers, consumers and other stakeholders. In the section 3, we found a significant positive effect of short-run disequilibrium on the conditional mean and also variance. The results states that the short-term deviation effects can be used to better forecast the volatility of agricultural price, and also implies that the further the disequilibrium of the agricultural commodity price and other factors in the long term, the harder to predict the volatility in the short term.

The paper is organized as follows: In section 2, we present the literature reviews about the different macroeconomic factors that impact the agriculture commodities prices and the latest Chinese macro reforms will be presented. In section 3, the methodology about the GARCH-X is introduced. The empirical results are presented in section 4. This is followed by a few concluding remarks in section 5.

2 Literature Review

2.1 China Agricultural Commodity Prices, Macroeconomic Variables and International Price Index

(*a*) Exchange Rate

Many studies have investigated the role of exchange rate play in the prices of agricultural commodities. They believe that it impacts directly through the channel of international purchasing power (import agricultural commodities) and export producer cost (export agricultural commodities) and indirectly through the recent high oil price. (relationship between energy and agricultural commodities) (Gilbert, 1989) [7] The theory starts with Schuh (1974) [8] who pointed out the importance of the macroeconomic and financial factors, and especially the exchange rate on the real prices of the agricultural commodities. Rogoff, Rossi and Chen (2008) [9] exemplify the role of exchange rates in determining agricultural prices. They show that

"commodity currency" exchange rates have remarkably robust power in predicting global commodity prices, both in-sample and out-of-sample, and against a variety of alternative benchmarks.

China's exchange rate policy started reforming in the last two decades. On January 1st 1994, a managed floating exchange rate system was established. In the year 1999, the IMF also tagged China from a "fixed exchange rate system" to a "managed floating exchange rate". The climax of the exchange rate reform occurred in 2005, on July 21st, with the Chinese government announcing that China will adopt the new managed floating currency structure which is based on the market. On the next day, *The Washington Post* also wrote that:

> "China on Thursday took an important step forward in its move toward a market economy, announcing it would increase the value of its currency, the yuan, and abandon its decade-old fixed exchange rate to the U.S. dollar in favor of a link to a basket of world currencies." (*The Washington Post*, 2005) [10]

Wen (2012) [11] studied the channel of the Chinese Renminbi exchange rate impacting on the international import and export volume, and further to the price. Ni and Qin (2012) [12] estimate the transmission of Renminbin exchange rate to the agricultural commodity prices, and the results show that the transmission effect is limited. China is going to be a major player in the international agricultural market. Therefore, the study on the Chinese exchange rate on China's agricultural prices is urgent.

(*b*) Interest Rates

Agricultural commodities are almost homogenous, quite storable and also transportable. In an absolutely free market, agricultural products can be treated as competitive products. Bosworth and Lawrence (1982) [13] and Okun (1981) [14] argued that agricultural commodities have flexible prices since the information regarding the supply and the demand can reach the price quickly. An empirical study (Bordo, 1980) [15] also verified that the prices of raw goods respond more quickly to changes in the money supply than do prices of manufactured goods.

Frankel (1986) [16] explicitly explained the channel of the interest rate's impact on the commodities price. Frankel (1986) [16] argues that because the price of agricultural commodities is flexible, while the other goods' prices are comparatively rigid, a change in the interest rate can influence the commodity prices more than what people would expect. He also summarize three major channels: first, it decreases the farm's desire to carry inventories; second, it encourages people to invest into non-physical products, such as treasury bills; and third, it appreciates the domestic currency and reduces the price of internationally traded commodities.

As for the China's economy, interest rate policy is the least affected traditional macroeconomic policy.(Lin, 1996) [3] This is because China government believes that only the heavy industry-oriented development strategy can develop the Chinese economy. In this context, the interest rates should be set as low as possible to satisfy the industry development. In fifteen years of recent past, from 1996 to 2003, the interest rates have been adjusted and maintained to keep low, however, since 2004,

the Chinese government started expanding the floating interval of the interest rate. The reform of interest rate can be checked by the following: the deposit reserve ratio is the best index which is represented by the interest rate. The government first started to adjust the deposit reserve ratio from the year 1984, however, until 2003, in a period f almost 20 years passed, the deposit reserve ratio was adjusted only six times. In contrast, in the last ten years, adjustments were made over 40 times. Consequently, the interest rate started to make its effects show on the economies.

There is almost no study of the interest rate effect on the agricultural price in China. This is due to the interest rate in China cannot move freely, the usual econometric method is difficult to model. In our studies, we do not incorporate it directly into our empirical study. However the money supply (M1), which can work on both exchange rate channel and interest rate channel are included as a variable in our model. In a free economy, the higher money supply will lead to a lower interest rate and lower exchange rate. Therefore through different ways we just discussed, the money supply can finally impact on the agricultural commodity.

2.2 International Prices

(*a*) International Energy Price

With the development of biofuel, agricultural commodity prices are gradually getting connected to energy prices. Agricultural commodities are the major feedstock for producing biofuel; because of this, the prices of agricultural commodities are invariably getting linked to the domestic energy price, and, further, with the international energy price. (Tyner, 2008) [17] As the price of crude oil increased, so did the price of corn and other agricultural commodities. And when the price of crude oil started to decline in the summer of 2008, so did the prices of agricultural commodities. The basic mechanism is that higher crude oil price leads to higher gasoline price, which increases the demand for corn ethanol as a substitute for gasoline. An increase in the demand for corn ethanol causes an increase in the demand for corn, which, in turn, leads to an increase in the price of corn.

The China's biofuel industry has rapidly developed since 2001. With the aim of reducing greenhouse gas (GHG) emissions and the energy security considerations, China government implemented an amount of subsidies policies. Hence the bioethanol production increased and reached 1.35 million tons in 2007. Not mentioned to the private small plants, four large stated-own plants located in Heilongjiang, Jilin, Henan, and Anhui use approximately 1.5 million tons corn annually. (Qiu et al., 2010) [18] The largely usage of agricultural commodities in the biofuel even force China have to import crops from abroad. (FAO, 2013) [19] The direct and indirect channels of international energy, such as the trade of oil and large-scale biofuel production could impact on China agricultural commodity price.

(*b*) International Agricultural Commodity Prices

The transmission of international agricultural commodity prices to domestic markets has been found in many developing countries . (FAO, 2009) [20] China's transmission channel has been found to be as follows: Wang and Xie (2012) [6] use monthly data to estimate the transmission effect of international agricultural prices on China's domestic prices. The paper verified that the international prices have significant impact on the domestic prices in China, but that the effects vary greatly for different products. Wang and Zhao (2012) [21] adopt the four important agricultural products: cotton, wheat, corn, and soybean to estimate the transmission effects of international prices by monthly data from January 2002 to December 2010. They conclude that the international price has positive and significant effect on China's prices but China's price did not impact much on the international prices.

Many Chinese researches have shown the linkage of the domestic price of the agricultural products and international price. (Wang and Xie, 2012; Wang and Zhao, 2012) [6, 21] However, how does the international price impact on the agricultural commodity prices in China? Do the international agricultural commodities also impact on the volatility of China's agricultural commodities? If the effect existed, it is therefore timely and important to quantity the magnificence.

3 Methodology

3.1 Method of GARCH and GARCH-X Models

(*a*) GARCH Models

To model the volatility of agricultural commodities, autoregressive conditional heteroscedasticity (ARCH) of Engle (1982) [22] and generalized ARCH (GARCH) of Bollerslev (1986) [23] is a obvious way to measure volatility. This study adopted an extension of the GARCH model for the error correction models (ECM) of cointegrated series, called GARCH-X models. (Lee, 1994) [24] To understand the GARCH-X models, first we introduce the ECM in the context of cointegration.

The ECM was proposed by Sargan (1964) [25] to deal with the short-run dynamics. Since there are only five cointegrated variables as in our model, we define the error correction term by:

$$\xi = y_t - \beta_1 x_1 - \beta_2 x_2 - \beta_3 x_3 - \beta_4 x_4 \tag{1}$$

Here β_i (i = 1, 2, 3, 4) is a cointegration coefficient and ξ is the short-term deviation from the regression of y_t on x_1, x_2, x_3, x_4, therefore we define the ECM as:

$$\Delta y_t = \alpha \xi_{t-1} + \lambda_1 \Delta x_{1t} + \lambda_2 \Delta x_{2t} + \lambda_3 \Delta x_{3t} + \lambda_4 \Delta x_{4t} + u_t \tag{2}$$

Here u_t is i.i.d. The ECM implies that Δy_t can be explained by the Δx_{1t} and also lagged ξ_{t-1}, that is, suppose now the $\xi_{t-1} > 0$, the y_{t-1} is too high beyond the equilibrium value, therefore next period y_t should be lower, and vice versa. This

is due to the variables are relatively stable in cointergrating vectors. This idea was extended into the variance equation by Lee (1994) [24]. He argued that if the short-run deviation from a long-run cointegrated relationship, has powerful predictive for conditional mean of the cointegrated series, it could also work on the conditional variance. (Engle and Yoo, 1987) [26] The GARCH-X model was first applied in measuring food volatility in the study of Apergis and Rezitis (2010) [27]. They analyzed the monthly data of Greece's food price in the period from 1985 to 2007. The results show that the GARCH-X model performs well and there exists a positive and significant effect between the deviations from macroeconomic factors and food price volatility. The short-run error from the cointegrating long-run relationship is a useful variable in modeling conditional variance in food price. Our study applies the GARCH-X model on agricultural commodity price and according to the China's situation, we also included the international prices to better model commodity volatility.

Suppose there is a dependent variable as r_t which is the return on an asset. The mean value is m_t and the variance is h_t at time t, they are defined by past information. Therefore the r in the time t can be represented as:

$$r_t = m_t + \sqrt{h_t}\varepsilon \tag{3}$$

Note that mean and variance of ε is 0 and 1. Therefore, r_t is equal to the mean value plus volatility at time t. (Engle, 1982) [22] The GARCH model for the variance is:

$$h_{t+1} = \omega + \alpha(r_t - m_t)^2 + \beta h_t = \omega + \alpha\varepsilon_t^2 + \beta h_t \tag{4}$$

Since in the long-run $h_{t+1} = h_t = h$, the variance of ε_t^2 is 1, therefore the long-run average variance is $\sqrt{\omega/(1-\alpha-\beta)}$, and also this only works if $\alpha+\beta<1$ and $\alpha>0, \beta>0, \omega>0$. (Bollerslev, 1986) [23] The GARCH model implied that the volatility prediction is a weighted average of the long-run average volatility, the volatility predicted for this period, and the most recent squared residual in last period.

The GARCH-X model aimed to model the conditional variance by the error correction (EC) terms. Hence, the specification equation of GARCH model changes into:

$$h_{t+1} = \omega + \alpha(\varepsilon_t)^2 + \beta h_t + \gamma {ect_t}^2 \tag{5}$$

Moreover, like the GARCH model, in order to make the GARCH-X model be stationary, We also need: $\alpha+\beta<1$ and $\alpha>0, \beta>0, \omega>0$. Besides, the short-run deviations is represented by the squared and lagged error-correction term ${ect_t}^2$, which obtains from the equation (1). The parameter γ indicates the effects of the short-run deviations on the conditional variance of the agricultural commodity prices equation. If error correction term is responsible for uncertainty i.e., conditional variance, the agricultural commodity price are more volatile when the disequilibrium is larger.

In last section, we also use other three GARCH based models to do the robust test, they are Exponential GARCH (EGARCH), Glosten-Jagannathan-Runkle GARCH (GJR GARCH), and also GARCH with student t distribution (GARCH-t). GJR models are the typical asymmetric impact model, which are widely used in the financial literature. GJR model includes leverage terms for modeling asymmetric volatility clustering. However, to our knowledge, there is no studies applied the GJR in estimating agriculture price, although it is has nice property and can capture some of the agriculture commodities' characteristics. We also use the GARCH-t model. This is because in the data description section, we found that the data is skew. The conditional variance of GJR-GARCH is:

$$h_{t+1} = \omega + (\alpha_1 + \alpha_2 I_t)\varepsilon_t^2 + \beta h_t \tag{6}$$

where $I = 0$ when $r_t \geq \mu$ and $I = 1$ when $r_t < \mu$.

The GJR GARCH models different impacts of the positive and negative shocks on the variance equation.(Grier and Perry, 1998) [28]

4 Empirical Results

4.1 Data Description

We adopt the monthly data from August 2001 to October 2011, taking a total of 130 observations. The China agricultural commodity index (AGR.China) is collected from the National Bureau of Statistics of China (NBSC), the international food price (AGR. world), money supply (M1), and world energy price index (ENE), and real effective exchange rate index (REER) are from the IMF database, the updated version in November 2012. Instead of the interest rate, we include M1 to estimate in the empirical part since the reform of the interest policy is still slow.

The trend of the statistics is shown in the following Figure 1. Different factors seems moving together. Intuitively, they could be a cointegrating vectors. In the period of 2001 to 2010, China government relaxed various controls, from price reforms to macroeconomic policies. The data movements in the figure also show this. All the variables are transformed by logarithm form.

4.2 Unit Root Test

To exclude the problem of spurious regression, we test the non-stationarity in price level and the difference by four unit root tests. They are the Augmented Dickey-Fuller (ADF) test, Kwiatkowski et al (KPSS) test, Elliott, Rothenberg and Stock (ERS) test, and Zivot and Andrews (ZA) test. The ADF test is adopted in our study first. (Dickey and Fuller, 1981) [29] The results are as given in Table 1. The lag length is determined by Akaike information criterion (AIC) principle. All the variables in level fail to reject the unit root hypothesis, but the variables in the first difference all reject the unit root hypothesis at 1% significance level. However, the

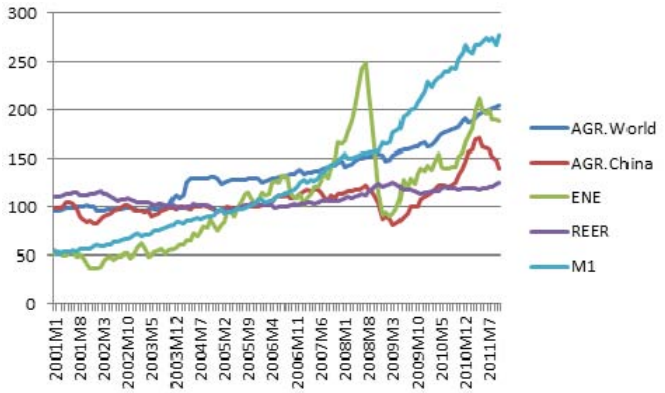


Fig. 1 The Trend of Chinese Agricultural Price and Other Factors

Note: The data span is from August 2001 to October 2011.

ADF test also has shortcomings; to improve the power of ADF test, ERS test was proposed in 1996, and this test applied ADF test to the detrending data without intercept. (Elliott et al., 1996) [30] The results are similar to the ADF test. However, the first difference of M1 cannot reject the null hypotheses at 5% significance level.

Since the unit root test results show inconsistencies, to further check the stationarity in our data, we carry out the KPSS test as well. This test overcomes the problem of the ADF test which becomes invalid in small samples. (Kwiatkowski et al., 1992) [31] The null hypothesis is that the series does not have a unit root which is opposite to the other tests. From Table 1, it can be seen that the results show that the first difference of M1 is a stationary series.

The last test we performed is the Zivot and Andrews (ZA) test. This test has a null hypothesis of the unit root process with drift which excludes exogenous structural change. (Zivot and Andrews, 2002) [32] Since in our data set, the variables may have structural changes. The results also show that although the variables in the level such as Agr.China, REER, and M1 are not stationary, all the variables in the first difference are stationary which is consistent with the ADF test and the KPSS test.

To sum up, after performing four unit root tests with trend and without trend, we can conclude that all of the variables are stationary and, therefore, reasonable to conduct the cointegration analysis, which is in the next section.

4.3 Cointegration Analysis

Now we ensure that all of the variables in the regression are I(1), and then next question is whether all the variables cointegrating. The intuition of cointegration is that the I(1) time series with a long-run equilibrium relationship cannot drift too far

Table 1 Unit Root Tests Results

ADF Test	Without Trend		With Trend	
	Levels	First Differences	Levels	First Differences
agr.China	1.624	-3.756(⋆⋆⋆)	-2.244	-4.123(⋆⋆⋆)
agr.world	-2.417	-5.146(⋆⋆⋆)	0.559	-5.163(⋆⋆⋆)
ene	0.901	-5.896(⋆⋆⋆)	-2.843	-5.962(⋆⋆⋆)
reer	0.678	-7.047(⋆⋆⋆)	-1.502	-7.324(⋆⋆⋆)
M1	7.603	-5.857(⋆⋆⋆)	-2.239	-10.140(⋆⋆⋆)
ERS Test	Without Trend		With Trend	
	Levels	First Differences	Levels	First Differences
agr.China	1.624	-3.756(⋆⋆⋆)	-2.244	-4.123(⋆⋆⋆)
agr.world	-2.122(⋆⋆)	-4.024(⋆⋆⋆)	-3.269(**)	-4.060(⋆⋆⋆)
ene	-0.278	-4.504(⋆⋆⋆)	-2.937(⋆⋆)	-4.541(⋆⋆⋆)
reer	-0.880	-3.958(⋆⋆⋆)	-1.301	-4.538(⋆⋆⋆)
M1	-1.301	-4.538(⋆⋆⋆)	-1.468	-1.693
KPSS Test	Without Trend		With Trend	
	Levels	First Differences	Levels	First Differences
agr.China	2.554(⋆⋆⋆)	0.124	0.122	0.058
agr.world	1.304(⋆⋆⋆)	0.080	0.152(⋆⋆)	0.080
ene	2.299(⋆⋆⋆)	0.043	0.262(⋆⋆⋆)	0.044
reer	1.235(⋆⋆⋆)	0.342	0.541(⋆⋆⋆)	0.061
M1	2.664(⋆⋆⋆)	0.133	0.378(⋆⋆⋆)	0.061
ZA Test	Without Trend		With Trend	
agr.China	-3.079		-7.654(⋆⋆⋆)	
agr.world	-5.808(⋆⋆⋆)		-5.030(⋆)	
ene	-6.107(⋆⋆⋆)		-6.127(⋆⋆⋆)	
reer	-3.351		-7.858(⋆⋆⋆)	
M1	-3.296		-8.097(⋆⋆⋆)	

Note: The asterisks in the brackets show the significance of the test statistics:
⋆ represents 10 %, ⋆⋆ represents 5 %, ⋆⋆⋆ represents 1 %

apart from the equilibrium because economic forces will act to restore the equilibrium relationship. In our case, the question changes into whether agricultural commodity price go together with other macroeconomic factors and international prices. Therefore, the usual statistical results will hold since we can exclude the spurious regression problems.(Johansen, and Juselius, 1990) [33]

Both the eigenvalue test statistic and the trace test statistic indicate that there exists one long-run relationship between relative food prices and macroeconomic variables, it makes sense to conduct an error correction analysis. The following equation also shows that the results of long-term relationship among the different macroeconomic variables and world price, world energy price in Table 3.

Table 2 Cointegration Tests Results

	r	n-r	Eigenvalues	1% Confidence	Trace	1% Confidence
lags=8	r=0	r=1	41.38	37.52	111.21	87.31
	r≤ 1	r=2	28.12	31.46	69.83	62.99
	r≤ 2	r=3	19.48	25.54	41.71	42.44
	r≤ 3	r=4	13.50	18.96	22.23	25.32
	r≤ 4	r=5	8.73	12.25	8.73	12.25

Note: The null hypothesis H_0 of the trace test is that there are r cointegration vectors, and the alternative hypothesis is that there are n cointegration vectors; for the maximum eigenvalue test, the null hypothesis H_0 is r cointegration vectors against the alternative hypothesis is r+1 cointegration vectors; The number of lags was determined through the AIC principle.

Table 3 Long-term Relationship

	Estimate	Standard Error	t value	Probability
constant	2.44	0.35	6.89	0.00
agr.World	0.08	0.04	2.08	0.03
ene	0.05	0.02	2.18	0.03
m1	0.3	0.03	8.48	0.00

Note: $R^2 = 0.955$; D.W.=0.115(0.00); the model equation is agr.China = a + b_1·agr.World + b_2·ene + b_3·reer + b_4·m1

In the long-term relationship, the D.W statistics is 0.115, p value is 0.00 such that the null hypothesis is rejected, and the same results are found in the LM test, the statistic is 115.71 and p value is 0.00. There exist serial correlations in data set. Therefore even we get significant estimates. The estimates could be artificially higher. The results of the long-term relationship are not explainable, we turn to the error correction model.

4.4 Error Correction Model

After detecting cointegrating relationship between China's prices and different variables, we adopt an error correction vector autoregressive mechanism, which adds the residuals from the cointegrating vector. We only report variables which turn out to be significant in Table 4. We choose the lag length followed Enders (2008) [34], first assume the lag length is 12, and then use the AIC and BIC method to select appropriate lags.

The important findings are as following. First, the speed of agricultural price's adjustment to the disequilibrium although significant, the lag term of China's agricultural commodities is powerful to explain the price itself, the real exchange rate, money supply, international agriculture and energy price are also effective to

predict the price. The real exchange rate (reer) has negative impact, while money supply (M1) has positive impacts on the agricultural price.

As mentioned in the methodology section, to fully understand the relationship among the China agricultural price, international price and also macroeconomic variables, we measure different variables work on the volatility level via the error correction terms as equation (2).

Table 4 The Short-term Relationship

	Estimate	Standard Error	Probability
$ect(-1)^2$	0.05	0.02	0.03
constant	-1.27	0.57	0.03
$agr.China(-1)$	0.40	0.11	0.00
$agr.China(-2)$	-0.25	0.12	0.04
$agr.China(-4)$	0.49	0.13	0.00
$agr.world(-7)$	0.13	0.06	0.03
$agr.world(-10)$	-0.11	0.05	0.05
$ene(-10)$	0.08	0.03	0.01
$reer(-3)$	-0.26	0.14	0.06
$reer(-10)$	-0.30	0.13	0.03
$m1(-4)$	0.32	0.14	0.02
$m1(-6)$	0.42	0.13	0.00
$m1(-11)$	0.69	0.15	0.00
LM test	0.39(0.53)		
D.W. test	1.87(0.17)		

Note: We only report the significant variables;
the number in the bracket is the p-value.

4.5 GARCH-X Models

Recall the GARCH (1, 1)-X model in section 3. Here we do not consider other lag length due to AIC value in GARCH (1, 1) is smallest. In financial literature, GARCH (1, 1) is usually believed to be the most efficient one. (Tsay, 2005) [35]. The results are interesting as in the Table 5. The error correction term (ect), which we incorporate into the variance is positive and significant, indicating a direct relationship between volatility and short-run deviations. In terms of the log-Likelihood value, the GARCH (1, 1)-X model performs better than the standard GARCH(1,1) model. For saving the space we do not present the results of GARCH(1, 1) model. All the parameters in the variance equation are positive, and $(\alpha + \beta)$ is lower than one, implying that our model is workable.

We adopt three other GARCH models to confirm the above results; they are the EGARCH, the GJR-GARCH, and GARCH-t model. (Nelson, 1993; Glosten et al., 1993) [36, 37]From the Table 5, we can see that the results in the variance equation is almost consistent. However, reer and m1 variables are not significant in the mean

equation. Even some of the lag terms are significant, it is not consistent with last section's results. This show that the domestic macroeconomic factors may not work much on the agricultural commodity price, China is still in a transition economy. The power of macroeconomic factors is limited.

Table 5 Resutls of GARCH-X Model Comparing with Other Three Models

	GARCH-X	EGARCH	GJR-GARCH	GARCH-t
agr.China(-1)	0.335(⋆⋆⋆)	0.14(⋆⋆⋆)	0.480(⋆⋆⋆)	0.400
agr.China(-2)	0.358(⋆⋆⋆)	0.71(⋆⋆⋆)	0.740	0.688
agr.China(-4)	0.451(⋆⋆⋆)	0.41(⋆⋆⋆)	0.457(⋆⋆)	0.440
agr.World(-7)	0.681(⋆⋆⋆)	0.49(⋆⋆⋆)	0.156	0.166
agr.World(-10)	0.453(⋆⋆⋆)	0.90(⋆⋆⋆)	0.474	0.579(⋆⋆⋆)
energy(-10)	0.220(⋆⋆⋆)	0.99(⋆⋆⋆)	0.986	0.708
reer(-3)	0.225	0.165(⋆⋆⋆)	0.220(⋆⋆⋆)	0.428
reer(-10)	0.468	0.40(⋆⋆⋆)	0.479	0.685
m1(-4)	0.125	0.899 (⋆⋆⋆)	0.103	0.148
m1(-6)	0.059	0.21(⋆⋆⋆)	0.098	0.048
m1(-11)	0.648(⋆⋆⋆)	0.102 (⋆⋆⋆)	0.122	0.021
ω	0.000(⋆⋆⋆)	0.486(⋆⋆⋆)	0.000(⋆⋆⋆)	0.004(⋆⋆⋆)
α	1.102(⋆⋆⋆)	0.094(⋆⋆⋆)	0.600	0.14
β	0.591(⋆⋆⋆)	0.384(⋆⋆⋆)	0.329(⋆⋆⋆)	0.467(⋆⋆⋆)
ect(-1)	0.294(⋆⋆⋆)	-	-	-
γ	-	0.543(⋆⋆⋆)	-0.09	-
D.F	-	-	-	3.303(⋆⋆)
Loglikelihood	179.621	-112.11	173.52	152.50
AIC	-327.24	242.00	-329.04	-286.99
BIC	-314.881	300.33	-270.71	-228.66

Note: AIC is calculated by $AIC = 2k - 2ln(L)$, k is the number of parameters in the statistical model, and L is the maximized value of th likelihood function for the estimated model. BIC is calculated by $BIC = -2ln(L) + kln(n)$, n is the sample size.

5 Conclusions and Policy Implications

In this study, we use the GARCH-X method to analyze the relationship of China's agricultural price volatility and the short term deviations with a series of macroeconomic variables and international price indexes. The results show that there exists a cointegrating relationship in some of macroeconomics and international and domestic price variables. However, since China is still in the stages of reform and on the path to becoming an open market economy, some of the macroeconomics variables did not show a significant impact on the agricultural commodity prices, unlike in other studies which were based on open developed countries, the most influential factors is the international agricultural commodities prices,while the real exchange

rate, money supply and international energy price have comparatively less effects on the prices of agricultural commodities.

The empirical results of this study are quite interesting and important. The international food and energy price, real exchange rate, money supply influences the domestic agricultural price not only by the mean level but also by the volatility through the drift. As a result, we find out that the changes in international food price can affect the farmers, producers and consumers much more than we had actually thought they could.

The result is critical to the China's policy makers, too, as now they can include the short-term deviation of the macroeconomic variables and international price to explain the volatility of the Chinese price. It becomes clear that China's government has increasingly less control over the agricultural prices. According to our study, when there is an international food crisis, the domestic agricultural price will be following the international price, both in the mean and in the variance level. Therefore, the government would do well to take some measures and intervene in the allocation of the resources, thereby improving the welfare of the whole nation.

References

1. Cheng, G.Q.: Chinas Agriculture in WTO (2005), http://finance.people.com.cn/BIG5/8215/32688/32690/3256154.html
2. NBSC (2012), http://www.stats.gov.cn/
3. Lin, Y.F.J., Cai, F., Li, Z.: Lessons of China's Transition to a Market Economy. The. Cato J. 16, 201 (1996)
4. Lin, J.Y.: WTO accession and Chinese agriculture. In: International Rice Research Conference, Beijing, China, September 16-19 2002. International Rice Research Institute (2003)
5. Yang, D.T., Li, Y.: Agricultural Price Reforms in Rural Price Reforms in China: Experiece from the Past Three Decades. Agroalimentaria (27), 13–23 (2008)
6. Wang, X.S., Xie, S.X.: How do Prices of Foreign Agricultural Products Affects Prices of Chinese Agricultural Products? Economic Research Journal 3 (2012)
7. Gilbert, C.I.: The Impact of Exchange Rates and Developing Country Debt on Commodity Prices. The Economic Journal, 773–784 (1989)
8. Schuh, G.E.: The Exchange Rate and US Agriculture. American Journal of Agricultural Economics 56, 1–13 (1974)
9. Rogoff, K., Rossi, B., Chen, Y.C.: Can Exchange Rates Forecast Commodity Prices? In: 2008 Meeting Papers (No. 540). Society for Economic Dynamics (2008)
10. Washington Post, China Ends Fixed-Rate Currency (2005)
11. Wen, Z.W.: The study on the transmission effects of Renminbi exchange rate on Chinas import and export price. PhD thesis, Chongqing University, Chongqing, China (2012)
12. Ni, Y., Qin, Z.: The empirical analysis on the impacts of Renminbi exchange rate on Chinas agricultural commodity price. In: The 7th China Soft Science Conference, Beijing, China (2012)
13. Bosworth, B., Lawrence, R.Z.: Commodity Prices and the New Inflation. Brookings Institution (1982)
14. Okun, A.M.: Prices and Quantities: A Macroeconomic Analysis. Brookings Institution Press (1981)

15. Bordo, M.D.: The Effects of Monetary Change on Relative Commodity Prices and the Role of Long-term Contracts. The Journal of Political Economy, 1088–1109 (1980)
16. Frankel, J.A.: International capital mobility and crowding out in the US economy: imperfect integration of financial markets or of goods markets? NBER working paper (1986)
17. Tyner, W.E.: The US Ethanol and Biofuels Boom: Its Origins, Current Status, and Future Prospects. BioScience 58, 646–653 (2008)
18. Qiu, H., Huang, J., Yang, J., Rozelle, S., Zhang, Y., Zhang, Y., Zhang, Y.: Bioethanol development in China and the potential impacts on its agricultural economy. Applied Energy 87(1), 76–83 (2010)
19. FAO, http://www.fao.org/docrep/004/y3557e/y3557e08.htm
20. FAO. Crop Prospects and Food Situation 2009, Rome, vol. 1 (2009)
21. Wang, S.F., Zhao, X.D.: The impact of international agricutlrual commodity to Chinas agricultural commodity price. Macroeconomic Research 9 (2012)
22. Engle, R.F.: Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica 50, 987–1006 (1982)
23. Bollerslev, T.: Generalized Autoregressive Conditional Heteroscedasticity. Journal of Econometrics 31, 307–327 (1986)
24. Lee, S.W., Hansen, B.E.: Asymptotic theory for the GARCH (1, 1) quasi-maximum likelihood estimator. Econometric Theory 10, 29 (1994)
25. Sargan, J.D.: Wages and Prices in the United Kingdom: a Study in Econometric Methodology. Econometric Analysis for National Economic Planning 16, 25–54 (1964)
26. Engle, R.F., Yoo, B.S.: Forecasting and testing in co-integrated systems. Journal of Econometrics 35(1), 143–159 (1987)
27. Apergis, N., Rezitis, A.: Food Price Volatility and Macroeconomic Factors: Evidence from GARCH and GARCH-X Estimates. Journal of Agricultural and Applied Economics 43(1) (2011)
28. Grier, K.B., Perry, M.J.: On inflation and inflation uncertainty in the G7 countries. Journal of International Money and Finance 17(4), 671–689 (1998)
29. Dickey, D.A., Fuller, W.A.: Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association 74(366a), 427–431 (1979)
30. Elliott, G., Rothenberg, T.J., Stock, J.H.: Efficient tests for an autoregressive unit root. Econometrica: Journal of the Econometric Society, 813–836 (1996)
31. Kwiatkowski, D., Phillips, P.C., Schmidt, P., Shin, Y.: Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root? Journal of Econometrics 54(1), 159–178 (1992)
32. Zivot, E., Andrews, D.W.K.: Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. Journal of Business and Economic Statistics 20(1), 25–44 (2002)
33. Johansen, S., Juselius, K.: Maximum Likelihood Estimation and Inference on Cointegration with Applications to the Demand for Money. Oxford Bulletin of Economics and Statistics 52, 169–210 (1990)
34. Enders, W.: Applied econometric time series. John Wiley and Sons (2008)
35. Tsay, R.S.: Analysis of financial time series, vol. 543. Wiley. com (2005)
36. Nelson, D.B.: Conditional heteroskedasticity in asset returns: A new approach. Econometrica: Journal of the Econometric Society, 347–370 (1991)
37. Glosten, L.R., Jagannathan, R., Runkle, D.E.: On the relation between the expected value and the volatility of the nominal excess return on stocks. The Journal of Finance 48(5), 1779–1801 (1993)

Co-movement of Prices of Energy and Agricultural Commodities in Biofuel Era: A Period-GARCH Copula Approach

Gong Xue and Songsak Sriboonchitta

Abstract. This study examines volatility and co-movement structures of coal and agricultural commodities index returns in China's bioful era. After taking into account the periodicity of changes in coal and agriculture prices, we show that the Period-GARCH (P-GARCH), which captures the characteristics of two commodities is more adequate in contrast to the previously proposed models where the residuals were skewed and had kurtosis, here the resulting residuals are almost Gaussian. Finally, our proposed P-GARCH time-varying copula models indicate that the dependence between energy and agricultural commodities index returns is positive and increasingly stable.

1 Introduction

Traditionally, it has always been a low correlation that existed between the agricultural market and the energy market in China. However, the recent increases in the technology changes in the field coupled with biofuel production have altered the agriculture-energy relationship in a fundamental way. (Hertel and Beckman, 2011) [1]

The agriculture-energy relationship works two ways: First, the agricultural commodities price could increase the energy price via the bioethanol industry. As the third largest bioethanol producer in the world after the United States and Brazil, China mainly uses grains such as corn and wheat as the feedstock for bioethanol production. (Qiu et al., 2010) [2] The development of China's biofuel industry could be traced back to early 2000. With the aim of reducing greenhouse gas (GHG) emissions and due to energy security considerations, the Chinese government implemented an host of subsidy policies; hence the bioethanol production increased rapidly and reached 1.35 million tons in 2007. (Qiu et al., 2010) [2] Although,

Gong Xue · Songsak Sriboonchitta
Faculty of Economics, Chiang Mai University, Chiang Mai 50200 Thailand
e-mail: {gongxue.cmu,songsak}@gmail.com

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_33,

in 2007, the government prohibited the use of cereals such as corn and wheat in bioethanol plants, four large formerly stat-owned plants still allowed the use of these. These four plants use approximately 1.5 million tons of corn annually. Second, since the industrialization and the subsequent development, more and more young laborers went away to the western part of China to be workers, and the old laborers adopted more machines in their field. Also, the increasing adoption of chemical fertilizers in China calls for more energy usage. Thus, a rise in energy price is bound to increase farm fertilizer production cost, machine production, transportation cost, and, finally, the farm rise of produce price. (Hang and Tu, 2007) [3]

China is the major coal producer and exporter in the world. (Gregg et al., 2008) [4] Coal is still the main energy source for the Chinese people, although there are many kinds of emerging energy. In the past, the coal price and the prices of other raw materials were largely kept low because of the planned economy in China, and a dual price system was adopted, one price for the market and another price for the power plants. (Wright, 2009) [5] However, after thirty years of reform, the coal prices are now decided by the market. Since 2002, the government stopped publishing the "electricity and coal guidance price", which published every year until 2002. During the period from 2002 to 2005, the coal market price reform has been moving both backward and forward. The government announced the price control schedules temporarily; for example, in 2005, the government announced that the coal price should be based on last year's price, and that the growth rate cannot be larger than 8%. Until the end of 2005, the Chinese government had maintained that the government will no longer control the coal price. The coal price is finally going to the market. Without any controls and with a large demand from the industries in China, the coal prices had experienced unprecedented high levels of fluctuations. For example, the coal price rose steadily from $36.69 per ton in January 2001 to $74.23 per ton in October 2006. (Fridley and Eden, 2008) [6] At the same time, the agricultural commodities prices are also raising drastically, with the price index of agricultural commodities jumping from 96.6 to 139.3. The positive co-movement of the two prices in China is an interesting phenomenon. Since energy and agriculture are the two most important sectors for China, and increasing evidences show that various factors are causing the agricultural and energy markets to be more highly integrated. What is the dependence structure between the agriculture commodity price and coal price in China? Specifically, how do we model price and forecast accurately the volatility and dependence between the agriculture products price and the coal price? Understand these relationships can help the Chinese policy makers formulate better make food and energy policies.

The objective of this study is to investigate the volatility and dependence structures between China's energy price and agricultural commodity price using period-GARCH copula models. The contribution of this study to the relevant literature are the following: (1) this is the first study to attempt adopting the copula-based GARCH models to elastically describe the dependence structures of the coal price and the agricultural commodities price return in China, since China's coal price reform has been just reaching its final stages in last decades. (2) The P-GARCH modeling was first used to model the seasonal changes in China's commodities

price volatility, and since the commodities have a persistent demand and supply pattern, we believe that this method can capture the characteristics of the commodities price well. Moreover, taking into consideration the Chinese government's "back and forth" behavior, we also include a variable to represent the government adjustment on the food price and also the coal volatility function, respectively.

The paper is organized as follows: In Section 2, literature Reviews about relationship of energy and agricultural commodities price and also different methods will be presented. In Section 3, the methodology about the Period-GARCH and Copula will be introduced. The empirical results are presented in Section 4. It is followed by some concluding remarks in Section 5.

2 Literature Review

Recently, many methods have been adopted to estimate the relationship between the energy and agriculture commodities markets. Most of them use the vector error correction (VECM) model. Campiche et al. (2007) [7] examine the covariability between crude oil prices and several agricultural commodities prices from 2003 to 2007. The results show that corn and soybean prices are cointegrated with crude oil price during the 2006-2007 period, but not during 2003-2005 period. And when Harri, Nalley and Hudson (2009) [8] added a new variable, exchange rate, into the cointegrating vector, the results showed that a cointegrating relationship exists between the agricultural commodities prices and the crude oil prices in April 2006 and that the exchange rates also have an effect on the cointegrating vectors. There are several studies that focus on China. Zhang and Reeds (2008) [9] studied the cointegrating relationship between the oil price and the pork prices, and showed that the crude oil price is not the most influential factor for the continuing rise of Chinese feed grain and pork prices. Wang and Xie (2012) [10] explored whether prices of foreign agricultural products and other factors, including the international coal futures, affect Chinese domestic prices of agricultural products. They use the monthly data by using the cointegrating method. The results show that the international coal price has positive effect on China's agricultural commodity prices.

The study of price and return co-movements has both economic and statistical significance. Therefore, it is of interest to both academicians and practitioners. Previous research, such as those done by Du and McPhail (2012) [11], Baillie and Myers (1991) [12], examined the dynamic evolution of the prices of agricultural commodities and energy over the period of March 2005 to March 2011 via the multivariate GARCH model. However, the shortcoming of this approach is that it is a severe assumption based on the multivariate normal and t distributions within an elliptical world. (Wu et al., 2012) [13] However, the energy and agricultural commodities returns are skewed and heavy-tailed, and have excess kurtosis, and seasonal changes, in our empirical observations. Bester (1999) [14] analyzed the seasonal volatility by using the Period-GARCH (P-GARCH) model. Six futures price series such as those of oil, corn, and soybean were examined, and found to have a significant seasonal component in volatility. This P-GARCH model can eliminate the

excess kurtosis in the commodities. In this study, we also modify the P-GARCH model to exclude the skewness and heavy tail by replacing the normal error term with the student t distribution. (Baillie and Myers, 1991) [12]

Moreover, the actual dependence between the agricultural and energy commodities returns can be best captured by asymmetrical modeling. For example, the left downside risk of coal returns comovment with agricultural commodities is more obvious than the right side. The comovement could be not symmetric. To overcome these shortcomings, we use time varying Copula models to capture the dependence structures of coal and agricultural commodities returns. The usage of the copula makes the modeling more flexible than joint distributions of bivariate normal or Student-t distributions, and combined with the P-GARCH, it can better capture the characteristics of the commodities price. The model will hopefully explain the behavior of the dynamic dependence of these two commodities.

3 Methodology

3.1 Period-GARCH Modeling for Marginal

The prices of agricultural commodities, such as crops and even energy, for example, coal, all suffer seasonal changes. Although China is big country, mainland China is dominated by warm temperate monsoon climate zone. With this type of climate, the country has four significant seasons, especially from the point of view of agriculture: the grain is planted during the spring time, it grows in the summer, is harvested in the autumn, and is in hibernation in the winter. The information regarding the changes in grain supply reaches grain price. The same is the case with coal, in the northern part of China, where the lowest temperature can be -30°in the winter, while in the summer time, the highest temperature can touch the 40°in almost the whole of China. (Zhai and Pan, 2003) [15]In these extreme temperature conditions, the Chinese residents would invariably need the air conditioning as well as the room-heating facilities. In the traditional method, the Chinese would burn coals to keep themselves warm since in the northern China, there is a large number of coalmines. Nowadays, even though people use a variety of sources to get electricity, coal still remains the most important energy source for China. Therefore, it is reasonable to use the P-GARCH model to capture the price characteristics. Furthermore, we choose to add some variables in order to model the government control behaviors.

Autoregressive conditional heteroskedasticity (ARCH) models and generalized autoregressive conditional heteroskedasticity(GARCH) models model the price series by allowing variance to evolve through the time. (Engle, 1982; Bollerslev 1986) [16, 17] This model is widely used in the financial market since it considers the independent properties of the return series. The P-GARCH model, which was proposed by Bollerslev and Ghysels (1996) [18], accounts for the seasonal changes or the period cycle; at first it was developed for the high frequency financial data, the opening

and closing time will make the price has seasonal change property. However, later, this model was introduced to estimate and forecast the prices of the commodities, such as those of energy, agriculture, and so on. (Koopman and Carnero, 2007; Winniford, 2003) [19, 20]. The modified P-GARCH model is as follows:

The mean function is

$$y_t = E_{t-1}[y_t] + \varepsilon_t \tag{1}$$

where $\varepsilon_t = \sqrt{h_t} \cdot z_t$, z_t denote the independent individual variables (i.i.d.) with expectation zero, and variance one. The variance function is as follows:

$$h_t = \omega_{s(t)} + \sum_{i=1}^{q} \alpha_{is(t)} \varepsilon_{t-1}^2 + \sum_{i=1}^{p} \beta_{is(t)} h_{t-1} \tag{2}$$

where s(t) denotes the stage of the seasonal cycle at time t, implying that a different GARCH parameter will be estimated in different seasons. In our study, we define the $s(t) = k$ (k=1,2,3,...,12), k represents different calendar months, that is, k=1 denotes January, k=2 denotes February,..., and so on. Here, we only consider the simple version of the P-GARCH model as there are seasonal changes in $\omega_{s(t)}$ but α and β are constant. This is because first, in this model, the α measures the immediate impact of the news on volatility, while β measures the smooth, long-term change on volatility. (Bollerslev and Ghysels, 1996) [18] Second, empirically, it is difficult to estimate β. The estimations are always not converged. (Bester, 1999) [14] In our study, our data observations are small also.

Here, we insert the Producer Price Index (PPI) and the Consumer Price Index (CPI) as the indexes, the reason is, as we mentioned in the last section, because the government announced that it will not control the price but that it still feels its duty to stabilize the price; for this reason, we insert these two variables into the variance functions of coal and agricultural commodities respectively, which implies that the higher PPI, the higher the distorted pressure from the government. The same is the case for the agricultural commodities; however, because the food prices are usually reflected in the CPI, we used CPI as an indicator.

Model One:

$$h_t = \omega + \sum_{i=1}^{q} \alpha_{is(t)} \varepsilon_{t-1}^2 + \sum_{i=1}^{p} \beta_{is(t)} h_{t-1} + \gamma \cdot index \tag{3}$$

Model Two:

$$h_t = \omega_{s(t)} + \sum_{i=1}^{q} \alpha_{is(t)} \varepsilon_{t-1}^2 + \sum_{i=1}^{p} \beta_{is(t)} h_{t-1} + \gamma \cdot index \tag{4}$$

3.2 Copula Method

(*a*) Essence of Copula

Previously, when we wished to know the dependence of two series, we would have to assume that the marginal distribution is normal, and that the multivariate normal distribution is then the joint distribution. This is quite inconvenient because normal distribution is not widely fit for the price return data which has heavy tail. The invention of the copulas solved this problem. (Sklar, 1958) [21]

The central idea of the copula is to link different margins into a multivariate distribution functions. Such functions make the joint distribution more flexible, and therefore can explain the real distribution thoroughly. The highlight of the copula function allows to the defining of a multivariate distribution with the marginal distributions $F_1(x_1)$ and $F_2(x_2)$ but not the realizations x_1 and x_2.

According to Sklar's theorem, if H is a joint distribution between the marginals $F_1(x_1)$ and $F_2(x_2)$, which are defined on the realizations of x_1 and x_2, and the copula function C links directly between the marginals $F_1(x_1)$ and $F_2(x_2)$, for simple notations, we define $F_1(x_1) = u_1$ and $F_1(x_1) = u_1$. Therefore, it follows that

$$H(x_1, x_2) = C(u_1, u_2) \tag{5}$$

There are two copula functions in elliptical copula , they are Gaussian copula and t copula, which are exactly the bivariate normal and the student t distribution, the pdf, cdf forms and other characteristics can be found in Jondeau and Rockinger (2006) [22], and Lee and Long (2009) [23] in details. The difference between two elliptical copulas is the tail dependence. The normal copula has no tail independence, however the t copula can capture tail dependence but not asymmetric. Since, in the financial empirical analysis, the heavy tails always exist and since only tail can bring about loss, the study on the relationship between the tails is more interesting. To overcome two problems of elliptical copula: symmetric tail and static dependence, we introduce the Archimedean copulas and also in addition to the time-varying copula modeling.

We use several Archimedean copulas to model the dependence between the energy and agricultural returns, they are, Clayton copula, Survival Clayton copula, mixed Clayton copula ,Frank copula and Ali-Mikhail-Haq(AMH) copula, the details of these copulas can be checked in the references. (Jondeau and Rockinger, 2006; Lee and Long, 2009) [22, 23] The advantage of the Archimedean copula is that they can explain the asymmetric tails. Even if the co-movements of the tail exist, the left-side and the right-side movement could be asymmetric. Therefore the application of Archimedean copula could help to improve the accuracy of the model.

To relax the assumption of the static dependence parameters, we introduced the time-varying copulas also into our estimation. This was done because of the fact that the relationship between the coal price and the agricultural commodities price could change all the time. Patton (2006) [24] introduced the conditional copula function

to model time-varying conditional dependence. Assume the marginal conditional probability density functions are:

$$u_{1,t} = F_{1,t}(y_{1,t}|\psi_{t-1}) \tag{6}$$

and

$$u_{2,t} = F_{2,t}(y_{2,t}|\psi_{t-1}) \tag{7}$$

where $y_{1,t}$ and $y_{2,t}$ are the coal price return and the agricultural commodities index, and $F_{1,t}$ and $F_{2,t}$ are the period GARCH filters, and ψ_{t-1} is the information set in the last periods. If we plug equation (7) into the equation in Sklar's theory (6). Hence we obtain the following time-varying copula equal to the bivariate conditional CDF:

$$C_t(u_{1,t}, u_{2,t}|\psi_{t-1}) = F(y_{1,t}, y_{2,t}|\psi_{t-1}) \tag{8}$$

To estimate the copula function, we differentiate and log equation (8), and then maximize the log-likelihood functions:

$$l(\theta) = \sum_{t=1}^{t}\sum_{i=1}^{n} logc\{F_1(Y_{i1};\beta_1), F_2(Y_{i2};\beta_2;\alpha_t) + \sum_{i=1}^{n}(logf_1(Y_{i1};\beta_1), logf_2(Y_{i2};\beta_2)) \tag{9}$$

where F_1 and F_2 are the CDF of the P-GARCH marginal, f_1 and f_2 are the PDF of the P-GARCH marginal,and β_1 and β_2 are the vectors of the P-GARCH marginal parameters, and α_t is the time-varying copula parameter vector. In our study, four different copula functions are used in time varying context. Hopefully, they can provide the most accurate dependence structure between China's coal and agricultural commodities prices.

Here we adopt the two-stage estimation method which was proposed by Joe (1997) [25], and this estimation is also called inference functions for margins (IFM). The reason to use this method but not the maximum likelihood (ML) is that with the increase in the number of parameters increases in estimation, the optimization problem becomes difficult to implement: in our P-GARCH model the estimator is sixteen, much more than the usual GARCH(1,1) model. The process of this method is to estimate, in the first stage, the marginal parameters and then to estimate the copula function with the marginal parameters estimated in first stage. In the empirical study, the results of the IFM and ML are always found to be consistent. (Yan, 2007) [26]

To describe the time-varying dependence structure, we assume that the dependence parameters ρ_t or τ_t depend on dependence on the last period ρ_t or τ_t and a drift between the $u_{1,t-1}$ and $u_{2,t-1}$. Here, we adjust the previous method to better solve our problem. Since we believe that these two variables are cointegrating, the distance between the two$((u_{1,t-1} - u_{2,t-1})^2)$ could explain the correlation, that is, in the last period the drift is smaller, and then, in the next period, the dependence will be larger.

$$\rho_t = \Lambda(a + b\rho_(t-1) + c(u_{1,t-1} - u_{2,t-1})^2) \tag{10}$$

Following the literature (Patton, 2009 and Wu, 2012) [24,13], Λ is the logistic transformation, which assures the correlation and kendall's tau to be always in the range of [-1, 1].

4 Empirical Results

4.1 Data and Descriptive Statistics

In our study, the monthly agricultural commodity index were abstracted from the National Bureau of Statistics of China (NBSC), and the monthly coal prices were sourced from the China Energy Databook 7.0 which published by the Lawrence Berkeley National Laboratory, in cooperation with China in 2008, which is the standard reference in use by the international energy community. The data span is from January 2001 to December 2006. In the coal price data, some of the November and December are unavailable due to discontinued data in sources. Noted that the Pearson correlation of these two series is up to 0.92, which implying the strong correlation of two data series. The comovement of the two series can be observed in Figure 1. In addition, we included two variables, CPI index and PPI index, into the estimation. The data for both were recovered from the NBSC monthly database. To get a stationary series, we follow the literature and logarithm the price: $y_t = log(p_t/p_{t-1}) \times 100$. The summary of basic statistics is shown in Table 1.

Fig. 1 The Co-movement of the Coal and Agricultural Price Indexes

Note: This figure shows the co-movement of two series: the upper line is the agricultural index and the lower line is the coal price. The coal data is discontinued due to the data sources being incomplete.

It can be read from Table 1 that the standard deviation of coal is higher than the agricultural index, that is, the coal return is more volatile; however, the agricultural index return is more skewed toward the right tail. The dominant factor in the two series is the excess kurtosis statistics, the values of both the coal and agricultural com-

Table 1 Data Description

	Coal	Agricultural Commodity Index
Mean	0.997	0.501
S.D.	4.039	1.978
Skewness	0.487	2.245
Kurtosis	10.217	11.302
Max	19.35	10.74
Min	-16.49	-4.162
JB	313.348(0.00)	437.449(0.00)

Note: This table shows the descriptive statistics for monthly coal and agricultural commodities index returns from January 2001 to December 2006. SD is standard deviation value. JB is the JarqueBera statistic, a normality test. The p value is in bracket, 0.00 means the null hypothesis is rejected, and two series are not normal distributions.

modities indexes are significantly positive, implying that the distribution of returns has heavier tails than the normal distribution. The results of the JarqueBera statistics also confirm that the two distributions are not normal. Also, the Augmented Dickey-Fuller (ADF) test rejects the null hypothesis of a unit root in both series; therefore the two series are stationary.

4.2 Estimation Results

(*a*) GARCH Modeling with Government Control Variable

The models presented here are estimated via maximum likelihood in the R program. The results of the GARCH modeling (Model 1) are presented in Table 2. It is easy to read from the table that all the estimated coefficients are highly significant, and robust with respect to initial values. When we compared this with the fGarch

Table 2 Estimation Results of GARCH

	Coal	Agricultural Commodity Index
ω	2.33($\star\star\star$)	0.38($\star\star\star$)
α	0.40($\star\star\star$)	0.02($\star\star\star$)
β	0.77($\star\star\star$)	0.90($\star\star\star$)
γ	-.90($\star\star\star$)	-0.24($\star\star\star$)
Loglikelihood	-167.26	-101.93
Standard Residuals		
Skewness	-0.95	2.57
Kurtoiss	8.99	13.50

Note:$\star\star\star$ in bracket show the 1% significance.

package GARCH(1,1) model, we found that our estimates were consistent. The log-likelihood is bigger in both the series, which means the inclusion of our government control variable is reasonable. The sign of γ is negative, which implies that the price adjustment has an effect on the volatility.

(*b*) P-GARCH Model with Flexible α

To estimate the P-GARCH model, we follow the literature and sum up the conditional log likelihood in each season, and employ the MLE to get the estimates as the estimation of typical GARCH model as discussed in the previous section. (Bollerslev and Ghysels, 1996; Bester, 1999) [18,14] We adopt not only the normal distribution but also the student t distribution, since after GARCH filtering the agricultural commodity index still had high skewness. The results of the second model are presented in Table 3:

Table 3 Estimation of P-GARCH

	Coal		Agricultural Commodity Index	
	Normal	Student t	Normal	Student t
ω_1	1.27($\star$)	1.11	0.83($\star$)	0.80
ω_2	-0.37	0.13	-0.05	-0.02
ω_3	-0.57	-0.04	-0.05	-0.02
ω_4	0.14	0.17	0.01	0.00
ω_5	0.39	0.18	-0.04	0.02($\star$)
ω_6	-0.08	-0.04	0.04	-0.03($\star$)
ω_7	0.58	0.10	0.09	0.18
ω_8	0.12	0.09	0.07	0.05
ω_9	-0.13	-0.07	0.00	-0.02
ω_{10}	0.48	-0.10	0.03	0.01
ω_{11}	-0.13	-0.12	-0.02	-0.03
ω_{12}	-0.26	0.07($\star\star$)	-0.01($\star$)	-0.03
α	-0.20($\star\star\star$)	0.04($\star\star\star$)	0.20	0.20($\star\star\star$)
β	1.07($\star\star\star$)	0.82($\star\star\star$)	0.90($\star\star\star$)	0.85($\star\star\star$)
γ	-1.19($\star\star\star$)	-0.74($\star\star\star$)	-0.89($\star\star\star$)	-0.78($\star\star\star$)
df	-	-0.74($\star\star\star$)	-	-0.78($\star\star\star$)
Loglikelihood	-153.88	-159.59	-92.90	-100.11
Standard Residuals				
Skewness	0.11	-1.37	1.08	1.34
Kurtosis	0.97	5.97	2.72	4.12

Note: $\star\star\star$ in bracket show the 1% significance, $\star\star$ in bracket show the 5% significance, $\star$ in bracket show the 10% significance.

As mentioned in the last section, the seasonal change variable is the constant in the GARCH equation. Note that $\widehat{\omega_1} = \omega_1$; $\widehat{\omega_{s(t)}} = \omega_1 + \omega_{s(t)}$, in our study, s(t) represents the seasons in the period of a month. Therefore, t = 1, 2,..., 12; when t = 1,

s(1) corresponds to January, when t = 2, s(2) is February, etc. The best improvement of P-GARCH model when compared with the GARCH model is that the Kurtosis value is largely reduced in all the cases. However, the student t distribution is not superior to the normal distribution. These results are highly consistent with Bester (1999). [14]

(*c*) Static Copula

Some of the Archimedean copulas have complicated Kendall's tau transformation functions; hence, it is difficult to make the dependence, which is represented by Kendall's tau, change with the time. Therefore, we estimate the static copula just to roughly understand the dependence between the coal price and the agriculture index return. As a benchmark, we also include the Gaussian copula to represent the elliptical copula. The results, as shown in Table 4, are quite disappointing. According to the AIC and the BIC principles, the Gaussian dependence structure is superior to the others. These results indicate that the tail dependence may not form the core of the modeling in our analyses. The result is also consistent with Wu et al. (2012) [13], who studied the dependence between oil and exchange-rate returns during the period from 1990 to 2009. The similar outcome concludes that the tail dependence does not add any explanatory ability to the estimations in static context.

Table 4 Estimation Results of Static Copula

	Parameter	Loglikelihood	AIC	BIC
Gaussian	0.21	0.16	1.68	3.87
Frank	0.48	0.04	1.91	4.10
AMH	0.11	0.02	1.96	4.15

Note: The AIC is calculated by $AIC = 2k - 2ln(L)$; k is the number of parameters in the statistical model, and L is the maximized value of the likelihood function for the estimated model. BIC is calculated by $BIC = -2ln(L) + kln(n)$; n is the sample size, and k is the same as in the AIC.

(*d*) Time-Varying Copula

Recall the discussion in section 3. We follow the previous studies to assume that the dependence relies on the last period dependence and a certain relationship between the transformed uniform data. The specific function can be found in equation (10). When estimating, we give a restricted form for b, since b is correlation between next period and this period, we assume that the correlation should be between (-1, 1). And also for c, the difference between the two series has persistence effect, which is between (-1, 1).

Table 4 reports the parameter estimates for the different copula functions. There are totally four copula functions adopted, and they are Gaussian copula, Clayton

copula, survival Clayton copula, mixed Clayton copula in the time-varying context. We do not use the Gumbel copula due to its limited to describe a positive dependence structure. Hence, we tend to use the survival Clayton copula which is similar to the Gumbel copula. And we also include the time-varying Gaussian copula as a benchmark to compare. The results are encouraging: looking at the values of the log-likelihood, Guassian copula in time varying context is higher than the static copula, that means time-varying copula may outperform. The Clayton dependence structure is better than other Archimedean copulas and also the Gaussian copula. And therefore the survival Clayton copula may not fit well for our data since it has opposite tail dependence with Clayton copula. The coal and agricultural commodities price has lower tail dependence, when the coal price is moving down, the agricultural commodities also moves together; this lower tail dependence can be also found in many financial literature, which modeling by the Clayton copula. (Bartram et al., 2006; Wu, 2012) [13, 27]

We draw the time-varying dependence of Clayton in Figure 2, which is the best fit in our estimation. The Figure illustrates the dynamic changes of the dependence relationship clearly. The dependence is as high as our expected: the dependence becomes stable in recent year.

Table 5 Estimation Results of Time-Varying Copula

	a	b	c	w	Loglikelihood	AIC	BIC
Gaussian	0.24(⋆⋆⋆)	0.95	0.00	-	0.41	5.16	11.73
Clayton	0.99	0.94	1.00	-	3.19	-0.38	6.18
Survival Clayton	0.15	0.95	-1.00	-	1.54	2.91	9.48
Mixed Clayton	0.15	0.95	-1.00	0.00	1.54	2.91	9.48

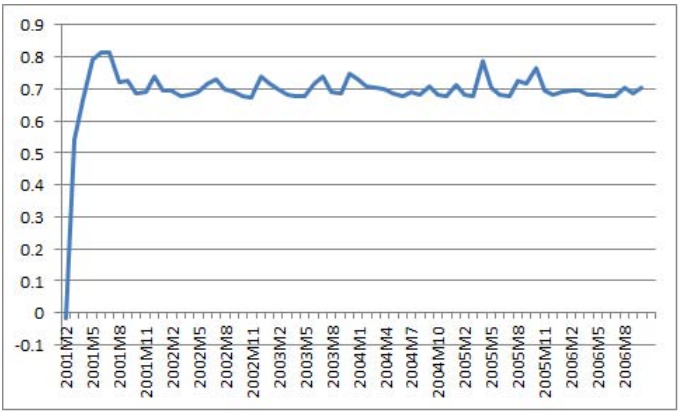


Fig. 2 The Dependence Estimates (Kendall's tau) between Coal and Agricultural Index from January, 2001 - October, 2006

5 Conclusion and Policy Implication

With China's advancement in coal reform and price control release, coal, as the most important energy source in China gradually went to the market. Since 2001, the government started to promote the emerging energy policy: prices of agricultural commodity, with agricultural commodities becoming the main feedstock, started to connect with energy price. Coal price, as well as prices of agricultural commodities, has exhibited significant co-movement. This relationship of co-movement has enabled the coal and agricultural products to serve as useful tools in forecasting for each other. Hence, this study is an attempt to estimate the volatility and co-movement structures of coal price and agricultural commodities return by using appropriate copula-based models.

However, it has been demonstrated that energy price and agricultural commodities index returns has some characteristic that very different from the financial series, say seasonal changes. With the demand and supply shocks regularly comes, the returns also has regular pattern. To capture these characteristics, we adopted the P-GARCH model to model the volatility. In statistical sense, these changes can eliminate the kurtosis and skewness in the return data. Unsurprisingly, when we filter the data by this improved GARCH model, the standard residuals show nice property than others. Moreover, we include a government control variable to the variance function, since the policy in our study period is always back and forth, the results show that the variable is significant.

The dependence structure between coal price and agricultural index returns may also exhibit an asymmetric or tail dependence structure. To overcome the shortcomings of multivariate GARCH model in elliptical world, we use more flexible Archimedean copula to model the dependence, the results in the context of time varying copula is promising. We find that the dependence structure between coal and agricultural commodities returns becomes increasingly positive.

Future work on this topic could extend the data span in order to make more accurate estimations. Since China's coal price reform was in its final stages, and the data set was difficult to obtain, our study could use only limited data. Trying a new and long data set may yield better results. Another point to bear in mind would be to improve the copula estimation. Further studies should be conducted to find better ways of incorporating the time-varying parameters.

References

1. Hertel, T.W., Beckman, J.: Commodity price volatility in the biofuel era: An examination of the linkage between energy and agricultural markets(No. w16824). National Bureau of Economic Research (2011)
2. Qiu, H., Huang, J., Yang, J., Rozelle, S., Zhang, Y., Zhang, Y., Zhang, Y.: Bioethanol development in China and the potential impacts on its agricultural economy. Applied Energy 87(1), 76–83 (2010)

3. Hang, L., Tu, M.: The impacts of energy prices on energy intensity: evidence from China. Energy Policy 35(5), 2978–2988 (2007)
4. Gregg, J.S., Andres, R.J., Marland, G.: China: Emissions pattern of the world leader in CO2 emissions from fossil fuel consumption and cement production. Geophysical Research Letters 35(8), L08806 (2008)
5. Wright, T.: Price reform in the Chinese coal industry. Asia Research Centre (2009)
6. Fridley, D., Eden, N.: China Energy Databook 7.0 (2008)
7. Campiche, J.L., Bryant, H.L., Richardson, J.W., Outlaw, J.L.: Examining the evolving correspondence between petroleum prices and agricultural commodity prices. In: AAEA Proc., Portland, OR (July 2007)
8. Harri, A., Nalley, L., Hudson, D.: The relationship between oil, exchange rates, and commodity prices. Journal of Agricultural and Applied Economics 41(2), 501–510 (2009)
9. Zhang, Q., Reed, M.R.: Examining the impact of the world crude oil price on China's agricultural commodity prices: the case of corn, soybean, and pork. In: 2008 Annual Meeting, Dallas, Texas (No. 6797), February 2-6. Southern Agricultural Economics Association (2008)
10. Wang, X.S., Xie, S.X.: How do Prices of Foreign Agricultural Products Affects Prices of Chinese Agricultural Products? Economic Research Journal 3 (2012)
11. Du, X., McPhail, L.L.: Inside the Black Box: the Price Linkage and Transmission between Energy and Agricultural Markets. Energy Journal-Cleveland 33(2), 171 (2012)
12. Baillie, R.T., Myers, R.J.: Bivariate GARCH estimation of the optimal commodity futures hedge. Journal of Applied Econometrics 6(2), 109–124 (1991)
13. Wu, C.C., Chung, H., Chang, Y.H.: The economic value of co-movement between oil price and exchange rate using copula-based GARCH models. Energy Economics 34(1), 270–282 (2012)
14. Bester, C.A.: Seasonal patterns in futures market volatility: a P-GARCH approach. Duke Journal of Economics 11, 65–102 (1999)
15. Engle, R.F.: Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica: Journal of the Econometric Society, 987–1007 (1982)
16. Bollerslev, T.: Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31(3), 307–327 (1986)
17. Bollerslev, T., Ghysels, E.: Periodic autoregressive conditional heteroskedasticity. Journal of Business and Economic Statistics 14, 139–151 (1996)
18. Zhai, P., Pan, X.: Change in Extreme Temperature and Precipitation over Northern China During the Second Half of the 20th Century. Acta Geographica Sinica 58 (2003)
19. Koopman, S.J., Ooms, M., Carnero, M.A.: Periodic seasonal reg-ARFIMAGARCH models for daily electricity spot prices. Journal of the American Statistical Association 102(477), 16–27 (2007)
20. Winniford, M.: Real estate investment trusts and seasonal volatility: a periodic GARCH model. working paper, Duke University (2003)
21. Sklar, M.: Fonctions de rpartition n dimensions et leurs marges. Publications de lInstitut Statistique delUniversite de Paris 8, 229–231 (1959)
22. Jondeau, E., Rockinger, M.: The copula-garch model of conditional dependencies: an international stock market application. Journal of International Money and Finance 25(5), 827–853 (2006)

23. Lee, T.H., Long, X.: Copula-based multivariate garch model with uncorrelated dependent errors. Journal of Econometrics 150(2), 207–218 (2009)
24. Patton, A.J.: Modelling asymmetric exchange rate dependence. International Economic Review 47(2), 527–556 (2006)
25. Joe, H.: Multivariate models and dependence concepts, vol. 73. Chapman and Hall/CRC (1997)
26. Yan, J.: Enjoy the joy of copulas: with a package copula. Journal of Statistical Software 21(4), 1–21 (2007)
27. Bartram, M., Taylor, S.J., Wang, Y.H.: The euro and European financial market integration. Journal of Banking and Finance 51(5), 1461–1481 (2005)

Wage Determination and Compensating Wage Differentials in the Informal Sector

A Quantile Regression with Multi-level Sample Selection

Pisit Leeahtam, Supanika Leurcharusmee, and Peerapat Jatukannyaprateep

Abstract. This study investigates Chiang Mai informal workers' wage determination in the equilibrium focusing on the compensating wage differential from taking occupational hazard risks with the presence of unemployment risk. Since there is a substantial heterogeneity among different groups within the informal sector, this study applies the quantile regression analysis with multi-level sample selection. The results show evidences for the compensating wage differentials in the lower and middle quantiles, but not the higher quantiles. The introduction of the unemployment risk variable into the wage equation proves the significance of the job mobility assumption. With unemployment risk, the workers not only are not compensated for their occupational hazards, but also face with an inefficient job matching outcomes. This emphasizes the significant spill-over benefit from the improvement of the job mobility condition.

1 Introduction

This paper is a part of The Informal Worker Analysis and Survey Modelling for Efficient Informal Worker Management Project with an objective to study the structure and nature of the informal sector in Thailand. The core project collected informal workers data in four provinces in four different regions of Thailand in 2012. This paper only focuses on the wage determination and compensating differentials in the informal sector in Chiang Mai Province.

The definition of informal worker varies across organizations. Therefore, it is important to first state the definition used for this study. In the context of Thai labor law and the government-provided work benefit, the core project defines the term informal workers as workers aged 15 or over who do not receive occupational welfare benefits from the government or do not have license for professional practice

Pisit Leeahtam · Supanika Leurcharusmee · Peerapat Jatukannyaprateep
Department of Economics, Chiang Mai University
e-mail: {pisitleeahtam,airmito,p.jatukannyaprateep}@gmail.com

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_34,

or enterprise owners with no business registration. In particular, the informal sector includes all workers not included in the formal sector. The formal sector includes 1) members of government organized occupational benefit fund; 2) workers with license for professional practice and; 3) enterprise owners with business registration. Specifically, the informal sector in Thailand is predominated with self-employed workers, freelance workers and unregistered employees such as farmers, street vendors, construction workers and housemaids.

In 2012, Thailands National Statistical Office[21] reported that informal workers were accounted for approximately 62.2 percent of the entire labor force. Despite the large number of informal workers, the labor laws and policies are not yet well suited for the nature of workers in this sector. Furthermore, comparing to the minimum wage rate at the time of data collection of 251 baht per day or 31.37 baht per hour, a little above 50 percent of informal workers from the survey data in Chiang Mai lived on wage lower than the minimum wage. This result is not entirely surprising because the minimum wage law cannot be applied to the majority of the informal workers. For example, self-employed and freelance workers do not have permanent employers and, thus, do not receive formal wages that can be bind by the minimum wage law. Moreover, the law is violated in several other groups of informal workers. Those employees with no employment contract are prone to face under-standard working condition and low wages.

In addition to the low wage problem, informal workers also faced several occupational hazards and many of them reported injuries from work. The data show that 53.46 percent of the informal workers in Chiang Mai believe that they are facing occupational hazards that affect their health and 68.37 percent reported that they had been injured from their job in the past 12 months.

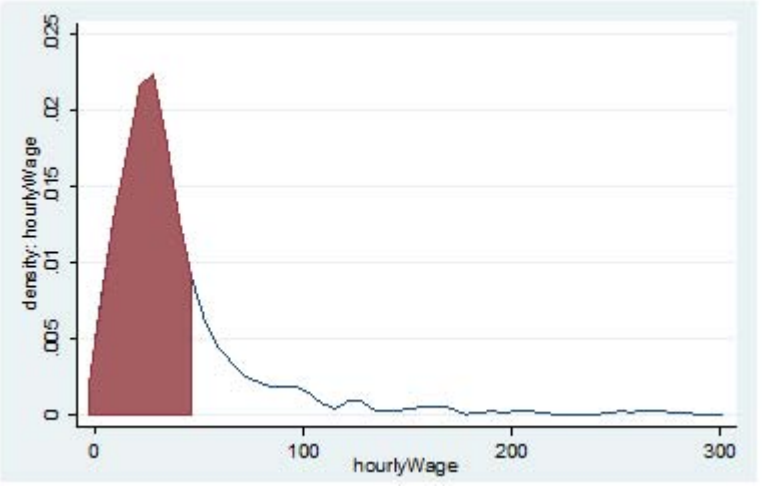


Fig. 1 The informal worker wage distribution in Chiang Mai and the mean wage of 51.72 baht per hour

In this research, we examine whether and when informal workers in Chiang Mai receive compensation for their risk from occupational hazards. That is, the key objective is to indicate in which cases that the compensating wage differential effect is significant. In particular, we construct an equilibrium wage determination model with occupational hazards and unemployment risk for informal workers in different quantiles of the wage distribution. From the model, we estimate the compensating wage differentials and the effect of unemployment risk on compensating wage differentials in all quantiles. The results then indicate in which wage ranges that we can observe the compensating wage differential effect. Moreover, the results can also indicate the effect of the labor market condition on the compensating wage differential effect. That is, when the labor market does not exhibit perfect job mobility or there exists unemployment risk, the compensating wage differential effect becomes insignificant.

The concept of compensating wage differentials introduced by Rosen (1974[15], 1986[16]) suggests that multiple equilibrium wage levels and the wage differentials should reflect workers willingness to accept the compensation for negative non-wage job characteristics such as the presence of occupational hazards. That is, the concept predicts a higher equilibrium wage for a riskier job and a lower equilibrium wage for a safer job. The magnitude and significance of the compensating wage differentials is valuable because it provides policy makers with information on potential effects of safety policies. Moreover, the presence or absence of the compensating wage differential effect is also an indicator of labor market efficiency.

The standard compensating wage differential relies on three main assumptions which are: (1) Workers maximize their utility (2) There is no asymmetric information among workers and firms, and (3) There is perfect job mobility. The first assumption implies that workers decisions also depend on other job characteristics, not wage alone. The second assumption states that workers know about the desirable and undesirable job characteristics before choosing the jobs, and the third assumption implies that there is no unemployment risk. Violating the above assumptions causes a bias in the compensating wage differential estimation. In our paper, however, we relax assumption (3) to capture a more realistic nature of the actual labor market. In this case, workers face unemployment risk causing imperfect mobility i.e. there is a cost in changing jobs. To allow the imperfection of the labor market, we include the unemployment risk variable to the wage equation. As the unemployment risk variable in most literatures is included in the wage equation to study the compensating wage differential due to the risk itself, it is important to clarify that the variable is added here to control for the market condition. The unemployment risk variable used in this study measures whether the worker believes that she can find a new job within three months if she loses her current job. Therefore, the variable does not directly imply the workers risk to lose her current job. It is interesting to note that the coefficients of the unemployment risk are not statistically significant in any models. However, the coefficients of the interaction terms between the hazard and the unemployment risk variables are significantly negative, especially in the lower quantiles. Consistent with the theory, the compensating wage differentials shrink in

the presence of unemployment risk. When there is a friction in the labor market, workers are less likely to switch jobs to gain a better wage-risk combinations.

As suggested by Viscusi and Aldy (2003)[22], the choice of the risk measures significantly affects the compensating wage differential estimates. The standard approach to measure occupational risks is to use industry-specific or occupational-specific risk measures. Only few papers use workers subjective perception toward risks and no paper use firms risk perception. Viscusi and Aldy (2003)[22] also mentions that, from the theory of compensating wage differential, the ideal measure of risks should reflect the perception of both workers and firms. In contrast to most papers, this study uses individual-level data with variables on each individuals perception on her occupational hazard and her unemployment risk which is an advantage for the estimation of compensating wage differentials.

Since there is a substantial heterogeneity among different groups within the informal sector, this study applies the quantile regression analysis introduced by Koenker and Bassett (1978)[9] to capture different structures of wage determinations for informal workers in different quantiles of the conditional wage distribution. This is crucial as the estimation at the conditional mean wage may not represent the entire wage distribution. The results reflect the contribution of the quantile analysis as the estimates for the compensating wage differential significantly varies across quantile. The ordinary least squared (OLS) estimates are consistent with the quantile regression results in the middle and higher quantiles. Nonetheless, the OLS method does not represent the estimates in the lowest quantiles very well.

In addition to the better illustration of the informal labor market, the quantile regression is robust to outliers, heteroskedasticity and misspecification of the model. However, the estimates still suffer from sample selection. In this study, the wage equation for informal workers faces two-stage sample selection problemthe selection into the labor force and the selection into the informal sector. To handle the selection bias, the estimation is executed in two stages. The first stage uses the probit model with sample selection to acquire the inverse Mills ratio (IMR). The second stage follows Buchinsky (1998)[4] s quantile regression method. The results show that the coefficient of the IMR is significant indicating the necessity for the sample selection treatment.

2 Literature Reviews

A common approach to empirically study the compensating wage differentials is the hedonic wage model. The model examines the equilibrium wage-risk combinations without explaining the underlying labor demand and supply structure behind the equilibrium outcomes . Empirical evidences for the compensating wage differentials in the case of occupational hazards differ in many dimensions. The main differences specified by Viscusi and Aldy (2003)[22] include the measure of risks, the measure of wage, the estimation and the technique to control for unobserved heterogeneities and other biases. Viscusi and Aldy (2003)[22] concludes that different risk measures result in different estimates for compensating differentials. In

particular, the models that use occupational specific risk measure are more likely to face measurement error problem. Different wage measures have a smaller effect on the estimates. For example, Moore and Viscusi (1988)[12] and Shanmugam (1996 [17], 1997[18]) find that the Box-Cox wage specification is significantly different from the semi-logarithmic specification. However, both specifications give approximately the same results for the compensating differential estimates. Techniques to control for unobserved heterogeneities are also an important aspect that significantly affects the results. While an inability to control for other job characteristics biases the estimate, the inclusion of too many job characteristics causes multicolinearity problem.

Viscusi and Aldy (2003)[22] reviews more than 60 studies of mortality risk premiums. The study finds that, while roughly 40 studies show the evidence of the risk premiums, the other 20 studies did not find the effect significant. Shanmugam (2001)[19] estimates the return to risk under self-selection bias in India. The results show a significant evidence for the compensating wage differentials. They also show substantial difference in the return to risk before and after correcting for the bias from self-selecting into a risky job. Leeth and Ruser (2003)[11] finds a strong evidence of the compensating wage differentials for fatal risk in white and Hispanic males and that for nonfatal injury risk in all groups with the largest effect for white females. All the results suggest that the compensating wage differentials are highly heterogeneous.

There are a limited number of empirical studies on the compensating wage differentials in an imperfect labor market. Guo and Hammitt (2009)[6] estimates the value of mortality risk in China using the compensating-wage-differential method. The study finds a significantly positive correlation between wages and occupational fatality risk and the correlation reduces with the unemployment rate. Bender and Mridha (2011)[1] finds that a probability of job loss significantly reduces the compensating wage differentials.

As for the estimation method, this study applies the quantile regression analysis introduced by Koenker and Bassett (1978)[9] to capture different structures of wage determinations for informal workers in different quantiles of the conditional wage distribution. In addition, the quantile regression is considered a robust estimation in several aspects. First, similar to median regressions, the quantile regression is robust to outliers. As long as the sign of the error term does not change, a change in the value of the dependent variable of that observation does not change the result. Second, with the models semi-parametric nature, it does not rely on the normality assumption of the error terms. Finally, with the pairs-bootstrap method for the covariance estimation suggested by Buchinsky (1995)[3], the model is robust to heteroskedasticity and misspecification of the quantile regression function.

The estimation of the model also suffers from the sample selection bias because the wage variable is observed only for a non-random subsample of the population. In this case, an hourly wage is observed only if the individual is working in the informal sector. That is, the model faces two levels of sample selection. The first level is the selection into workforce and the second level is the selection into the informal sector. For the solution to the selection problem, this study follows

Buchinsky (1998)[4] with a modification to support the two-stage sample selection issue. The estimation in Buchinsky (1998)[4] applies Ichimura (1993)[8]s semiparametric least-squares method for the selection equation and Newey (1991)[14]s series expansion of the inverse Mills ratio to estimate the conditional quantile regression. For this study with the two-stage sample selection, this study uses the probit model with sample selection to acquire the inverse Mills ratio in the first stage instead of the semiparametric least-squares method. The second stage to estimate the wage equation follows Buchinsky (1998)[4] entirely.

3 Data

Data used in this study is from the core project, The Informal Worker Analysis and Survey Modelling for Efficient Informal Worker Management Project. The sampling frame used in the project was from Chiang Mais Household Listing Survey by the National Statistical Office and the sampling procedures were conducted using the method parallel to the Household Socio-Economic Survey by the National Statistical Office. In particular, the samples were selected in two stages. In the first stage, sub-regions are selected using the stratified sampling method. In the second stage, households within each sampling sub-regions are selected using the systematic sampling method.

The questionnaire for the survey was composed of two main parts. The first part asked fundamental household information and the informal sector screening questions. All household members were interviewed for this part. If the respondent was an informal worker, he/she was asked the second part of the questionnaire. If the respondent was not an informal worker, then he/she did not have to answer the second part. The second part of the questionnaire asks in depth questions regarding workers working and living conditions. Therefore, demographic variables are observed for all individuals and the wage and working variables are only observed for informal workers.

For the variables characterizing occupational hazard, this study uses the workers self-report on occupational hazard where workers were asked whether they faced at least one hazard at work and whether the hazard had a negative effect on their health. Unemployment risk is measured by the question asking the workers whether they would be able to find a new job within 3 months if they lose their current job. Therefore, this variable measures the unemployment risk from the perspective of the workers themselves.

4 Model and Methodology

Assume that the wage equation in the informal sector is linearly dependent on a set of labor market characteristics. A traditional compensating wage differential regression in the form of hedonic wage equation is:

$$Wage_i = x_i'\beta + u_i \tag{1}$$

where the subscript i indexes individuals. $Wage_i$ is an hourly wage in its natural log form. x_i is a vector of labor market characteristics including the intercept and an occupational hazard dummy variable which takes value 1 if the individual is exposed occupational hazard and 0 otherwise. u_i is the disturbance term. The compensating wage differential from accepting occupational hazard is measured by the coefficient of the occupational hazard variable.

In our model, an hourly wage is observed only if the individual is working in the informal sector. However, we can observe characteristics of individuals who are working in the formal sector and those who are not working. Therefore, the model becomes:

$$Wage_i|(LF_i = 1, IS_i = 1) = (x_i'\beta + u_i)|(LF_i = 1, IS_i = 1) \tag{2}$$

where LF_i and IS_i are binary dummy variables.

$$LF_i = \begin{cases} 1, \text{ if individual i works} \\ 0, \quad \text{otherwise} \end{cases} \tag{3}$$

$$IS_i = \begin{cases} 1, \text{ if individual i works in the informal sector} \\ 0, \text{ if individual i works in the formal sector} \end{cases} \tag{4}$$

By taking expectations over both sides of equation (2):

$$E(Wage_i|LF_i = 1, IS_i = 1) = E(x_i'\beta + u_i|x_i, LF_i = 1, IS_i = 1) \tag{5}$$

$$= x_i'\beta + E(u_i|x_i, LF_i = 1, IS_i = 1) \tag{6}$$

Then, if u_i is not independent of LF_i and IS_i i.e. $E(u_i x_i, LF_i, IS_i)0$, the disturbance term is biased. In our case, the bias comes from the sample selection which has two sources; one is the unobserved wage of those who are not working ($LF_i = 0$) and the other is of those who are working in the formal sector ($LF_i = 1, IS_i = 0$). In particular, the model faces two levels of sample selection problem as shown in Figure 2. The first level is the selection into workforce and the second level is the selection into the informal sector.

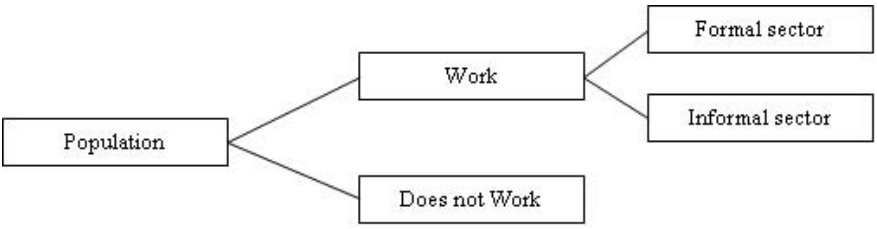


Fig. 2 Multi-level sample selection

To deal with the bias in the disturbance term arisen from the sample selection, we combine Heckman two-stage method (Heckman, 1979)[7] with Heckman probit estimation method (probit with sample selection).

In the first stage of Heckman two-stage, instead of just using a typical probit model, we start with a probit model with sample selection. In this stage, we try to determine the probability that an individual will choose to work in an informal sector given the individuals labor market characteristics,

$$Pr(IS_i = 1|z_i) \tag{7}$$

where z_i is a column vector of the labor market characteristics which affect the sector choice of an individual including the intercept. However, we do not observe the choice of those who are not working, i.e. the individuals whose $LF_i = 1$. Hence, the estimated coefficients may be biased from the sample selection process. To eliminate the bias, we follow Heckman probit estimation method by assuming that LF_i and IS_i follow the following rules:

$$LF_i = \begin{cases} 1, \text{ if } LF_i^* > 0 \\ 0, \text{ if } LF_i^* \leq 0 \end{cases} \tag{8}$$

$$IS_i = \begin{cases} 1, \text{ if } IS_i^* > 0 \\ 0, \text{ if } IS_i^* \leq 0 \end{cases} \tag{9}$$

where LF_i^* and IS_i^* are assumed to be linearly dependent to some relevant labor market characteristics:

$$LF_i^* = w_i'\gamma + \varepsilon_i \tag{10}$$

$$IS_i^* = z_i'\theta + \eta_i \tag{11}$$

with

$$E(\varepsilon_i) = 0, E(\eta_i) = 0, E(\varepsilon_i\eta_i) = \rho \tag{12}$$

where w_i a column vector of labor market characteristics that influence the decision to enter the labor force, i.e. decision to work or not to work, including the intercept. By assuming that η_i is normally distributed, we have that:

$$Pr(IS_i = 1 w_i, LF_i = 1) = \Phi((z_i')/(\sigma_\eta)|\varepsilon_i > -w_i'\gamma) \tag{13}$$

$$= \Phi_2((z_i'\theta)/(\sigma_\eta), (w_i'\gamma)/\sigma_\varepsilon, \rho) \tag{14}$$

where $\Phi_2(,,\rho)$ is the joint cumulative distribution function of 2 standard- normally distributed random variables with the correlation between the two variables equals to .

Let Φ_2 be a joint cumulative distribution function of 2 standard-normally distributed random variables and let $E(u_i\eta_i\varepsilon_i) = \omega$.

$$E(u_i|x_i, LF_i = 1, IS_i = 1) = E(u_i|x_i, \varepsilon_i > -w_i'\gamma, \eta_i > -z_i'\theta) \tag{15}$$

$$= \omega\sigma_u \frac{\phi_2(\frac{z_i'\theta}{\sigma_\eta}|\varepsilon_i > -w_i'\gamma)}{\Phi_2(\frac{z_i'\theta}{\sigma_\eta}|\varepsilon_i > -w_i'\gamma)} \tag{16}$$

$$= \omega\sigma_u \frac{\phi_2(\frac{z_i'\theta}{\sigma_\eta}, \frac{w_i'\gamma}{\sigma_\varepsilon}, \rho)}{\Phi_2(\frac{z_i'\theta}{\sigma_\eta}, \frac{w_i'\gamma}{\sigma_\varepsilon}, \rho)} \tag{17}$$

$$= \beta_\lambda \lambda(\frac{z_i'\theta}{\sigma_\eta}, \frac{w_i'\gamma}{\sigma_\varepsilon}, \rho) \tag{18}$$

where $\gamma()$ is the inverse Mills ratio function.

Hence, equation (3) becomes[1]:

$$E(Wage_i LF_i = 1, IS_i = 1) = x_i'\beta + \beta_\gamma\gamma((z_i'\theta)/(\sigma_\eta), (w_i'\gamma)/\sigma_\varepsilon, \rho) + E(v_i|x_i) \tag{19}$$

where v_i is the demeaned disturbance term i.e. $E(v_i x_i) = 0$.
Re-writing equation (19) into the conventional quantile regression equation, the final equation becomes:

$$Q_\tau(Wage_i x_i, LF_i^* = 1, IS_i^* = 1) = x_i'\beta_\tau + \beta_(\lambda_\tau)\gamma((z_i'\theta)/(\sigma_\eta), (w_i'\gamma)/\sigma_\varepsilon, \rho) + Q_\tau(v_i x_i). \tag{20}$$

5 Results and Discussions

5.1 The Sample Selection Regression

Table 1 shows the results from the probit model with sample selection which is the first stage of the Heckman two-stage model. Consistent with the literatures, the result in regression (1) shows that age has a positive but diminishing effect on decision to work while the coefficients of education variables are not statistically significant implying that education does not affect decision to work. The result also indicates that, on average, a married male is more likely to work than a single male; on the other hand, marriage does not affect the decision to work of female. Furthermore, having kids also affects male and female differently. Male with kids tends to work while female with kids tends not to. In addition, disabled individuals are less likely to work.

Given that they are in the labor force, regression (2) shows that the effect of age on decision to work in the informal sector is U-shape, that is, young and old individuals, as oppose to the middle age group, are more likely to be in the informal

[1] Buchinsky (1998) [4] suggests that the series expansion of the inverse Mills ratio should be included to estimate the conditional quantile regression. However, in our estimation, the higher-order terms are not statistically significance and therefore omitted from the model.

Table 1 Regression (1) describes the decision to work using the probit model. Regression (2) describes the decision to work in the informal sector conditioning on being in the labor force using the probit model with sample selection.

	(1)	(2)
VARIABLES	Informal	Labor Force
female	0.057	-0.072
	(0.071)	(0.062)
age	-0.093**	0.258***
	(0.036)	(0.007)
age2	0.001***	-0.003***
	(0.000)	(0.000)
edu1	-0.436***	0.068
	(0.104)	(0.067)
edu2	-0.783***	-0.076
	(0.062)	(0.054)
edu3	-0.999***	-0.012
	(0.067)	(0.056)
pregnant		-0.332
		(0.306)
handicap		-1.181***
		(0.151)
kids	0.026	0.062**
	(0.030)	(0.028)
married		0.416***
		(0.070)
marriedFemale		-0.420***
		(0.087)
kidsFemale	0.072*	-0.099***
	(0.043)	(0.034)
Constant	3.301***	-4.463***
	(0.823)	(0.162)
Observations	6,186	6,186

Standard errors in parentheses
*** p<0.01, **p<0.05, *p<0.1

sector. The education variables are negative and highly significant indicating that individuals with higher education are less likely to work in the informal sector. Lastly, female with kids are more likely to work in the informal sector.

5.2 The Wage Equation

The wage regressions using OLS are shown in Table 2 while the wage regressions using quantile regression are shown in Table 3 and 4. As can be seen from Table 2, although most of the estimated coefficients have the predicted sign and are statistically significant at high level, the estimated coefficients using OLS seems to be highly biased from sample selection.

After adjusting for the sample selection bias by including the inverse Mills ratio term as a control variable, the estimated coefficient of the common wage regression variables, in both regressions with and without unemployment risk, are consistent

Table 2 Wage regressions using Ordinary Least Squares (OLS)

	(1)	(2)	(3)	(4)
	olsH	olsHJ	olsHSS	olsHJSS
VARIABLES	lnHW	lnHW	lnHW	lnHW
female	-0.080*	-0.082*	-0.172***	-0.162***
	(0.046)	(0.048)	(0.059)	(0.062)
age	0.032***	0.027**	0.079***	0.068***
	(0.011)	(0.011)	(0.022)	(0.022)
age2	-0.000***	-0.000**	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
edu1	0.061	0.046	0.245**	0.206**
	(0.069)	(0.074)	(0.099)	(0.105)
edu2	0.145**	0.147**	0.728***	0.654***
	(0.057)	(0.060)	(0.225)	(0.234)
edu3	0.445***	0.473***	1.513***	1.409***
	(0.113)	(0.115)	(0.436)	(0.452)
owner	0.143***	0.160***	0.139***	0.157***
	(0.048)	(0.050)	(0.047)	(0.050)
agriculture	-0.038	-0.032	-0.047	-0.042
	(0.057)	(0.061)	(0.057)	(0.061)
unempRisk3mo		0.046		0.043
		(0.087)		(0.087)
hazard	0.117***	0.165***	0.115**	0.161***
	(0.045)	(0.055)	(0.045)	(0.055)
hazardUnempRisk3mo		-0.194*		-0.187*
		(0.112)		(0.112)
IMR			-1.547***	-1.348**
			(0.587)	(0.611)
Constant	2.489***	2.587***	1.788***	1.981***
	(0.246)	(0.257)	(0.372)	(0.387)
Observations	1,723	1,557	1,723	1,557
R-squared	0.036	0.039	0.040	0.042

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3 Wage regression which does not control for unemployment risk using quantile regression

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	bsqreg10H	bsqreg20H	bsqreg30H	bsqreg40H	bsqreg50H	bsqreg60H	bsqreg70H	bsqreg80H	bsqreg90H
VARIABLES	lnHW	lnHW	lnHW	lnHW	lnHW	lnHW	lnHW	lnHW	lnHW
female	-0.119	-0.101	-0.139**	-0.110***	-0.153***	-0.148***	-0.170***	-0.197***	-0.312***
	(0.153)	(0.094)	(0.064)	(0.028)	(0.045)	(0.057)	(0.055)	(0.064)	(0.106)
age	0.097**	0.070*	0.053*	0.047***	0.059***	0.061***	0.055***	0.055**	0.109***
	(0.048)	(0.040)	(0.029)	(0.016)	(0.013)	(0.016)	(0.018)	(0.022)	(0.040)
age2	-0.001**	-0.001*	-0.001*	-0.001***	-0.001***	-0.001***	-0.001***	-0.001**	-0.001***
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
edu1	0.167	0.258	0.191	0.115	0.110*	0.147*	0.137*	0.116	0.405**
	(0.268)	(0.163)	(0.141)	(0.103)	(0.067)	(0.078)	(0.078)	(0.154)	(0.161)
edu2	0.864**	0.509	0.485*	0.431**	0.515***	0.582***	0.576**	0.539*	1.157***
	(0.406)	(0.382)	(0.263)	(0.178)	(0.140)	(0.162)	(0.253)	(0.326)	(0.404)
edu3	1.806**	1.219*	0.988*	0.821**	1.006***	1.150***	1.073**	0.896**	2.320**
	(0.805)	(0.692)	(0.564)	(0.329)	(0.272)	(0.354)	(0.457)	(0.424)	(0.944)
owner	-0.136	-0.084	-0.010	0.069	0.125**	0.194***	0.314***	0.338***	0.402***
	(0.161)	(0.063)	(0.058)	(0.043)	(0.050)	(0.052)	(0.051)	(0.063)	(0.099)
agriculture	-0.129	-0.167**	-0.153**	-0.196***	-0.192***	-0.099*	-0.009	0.076	0.239**
	(0.145)	(0.065)	(0.064)	(0.032)	(0.063)	(0.057)	(0.070)	(0.056)	(0.108)
hazard	0.191	0.114*	0.072	0.080**	0.073**	0.094***	0.060	0.083**	0.118
	(0.136)	(0.060)	(0.044)	(0.039)	(0.029)	(0.029)	(0.054)	(0.041)	(0.088)
IMR	-1.773	-1.210	-1.028	-0.732	-1.020***	-1.200***	-1.088*	-0.996	-2.768***
	(1.155)	(0.958)	(0.745)	(0.458)	(0.354)	(0.447)	(0.658)	(0.802)	(1.059)
Constant	0.568	1.609**	2.134***	2.295***	2.257***	2.332***	2.574***	2.780***	2.226***
	(0.805)	(0.698)	(0.568)	(0.302)	(0.224)	(0.280)	(0.287)	(0.360)	(0.680)
Observations	1,723	1,723	1,723	1,723	1,723	1,723	1,723	1,723	1,723

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 4 Wage regression which controls for unemployment risk using quantile regression

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	bsqreg10HJ	bsqreg20HJ	bsqreg30HJ	bsqreg40HJ	bsqreg50HJ	bsqreg60HJ	bsqreg70HJ	bsqreg80HJ	bsqreg90HJ
VARIABLES	lnHW	lnHW	lnHW	lnHW	lnHW	lnHW	lnHW	lnHW	lnHW
female	-0.062	-0.122	-0.113	-0.108	-0.142***	-0.128**	-0.165***	-0.198***	-0.324**
	(0.140)	(0.102)	(0.079)	(0.068)	(0.040)	(0.054)	(0.057)	(0.064)	(0.136)
age	0.079**	0.066**	0.044	0.044**	0.063***	0.055***	0.050***	0.046**	0.080
	(0.036)	(0.029)	(0.027)	(0.020)	(0.019)	(0.019)	(0.017)	(0.018)	(0.058)
age2	-0.001**	-0.001**	-0.001*	-0.001**	-0.001***	-0.001***	-0.001***	-0.001**	-0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
edu1	0.026	0.227**	0.105	0.150	0.177**	0.143	0.140*	0.080	0.273
	(0.318)	(0.097)	(0.109)	(0.105)	(0.088)	(0.096)	(0.078)	(0.119)	(0.224)
edu2	0.808*	0.487*	0.365	0.480*	0.609**	0.494**	0.535***	0.531**	0.953*
	(0.432)	(0.254)	(0.270)	(0.247)	(0.241)	(0.208)	(0.189)	(0.256)	(0.500)
edu3	1.502**	1.320***	0.830	0.984**	1.300***	1.152***	1.027***	0.906*	1.883*
	(0.656)	(0.418)	(0.554)	(0.492)	(0.404)	(0.353)	(0.323)	(0.478)	(1.076)
owner	-0.123	-0.071	0.011	0.092**	0.121***	0.222***	0.328***	0.352***	0.456***
	(0.145)	(0.073)	(0.056)	(0.037)	(0.045)	(0.055)	(0.061)	(0.074)	(0.108)
agriculture	-0.126	-0.108	-0.174***	-0.170***	-0.188***	-0.115*	-0.023	0.098*	0.300***
	(0.182)	(0.084)	(0.065)	(0.055)	(0.055)	(0.064)	(0.071)	(0.055)	(0.106)
unempRisk3mo	-0.184	-0.174	-0.108*	-0.133***	-0.123***	-0.099	-0.018	0.026	0.138
	(0.130)	(0.117)	(0.062)	(0.047)	(0.044)	(0.067)	(0.043)	(0.053)	(0.114)
hazard	0.201*	0.127**	0.085	0.056	0.046	0.085**	0.045	0.076*	0.022
	(0.120)	(0.062)	(0.056)	(0.036)	(0.039)	(0.041)	(0.042)	(0.041)	(0.111)
IMR	-1.357	-1.134*	-0.719	-0.905	-1.314**	-1.005*	-0.977**	-0.939	-2.230
	(0.881)	(0.603)	(0.745)	(0.682)	(0.563)	(0.516)	(0.440)	(0.617)	(1.416)
Constant	0.825	1.677***	2.340***	2.449***	2.232***	2.433***	2.634***	2.971***	2.834***
	(0.798)	(0.444)	(0.401)	(0.313)	(0.360)	(0.349)	(0.360)	(0.347)	(0.921)
Observations	1,557	1,557	1,557	1,557	1,557	1,557	1,557	1,557	1,557

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

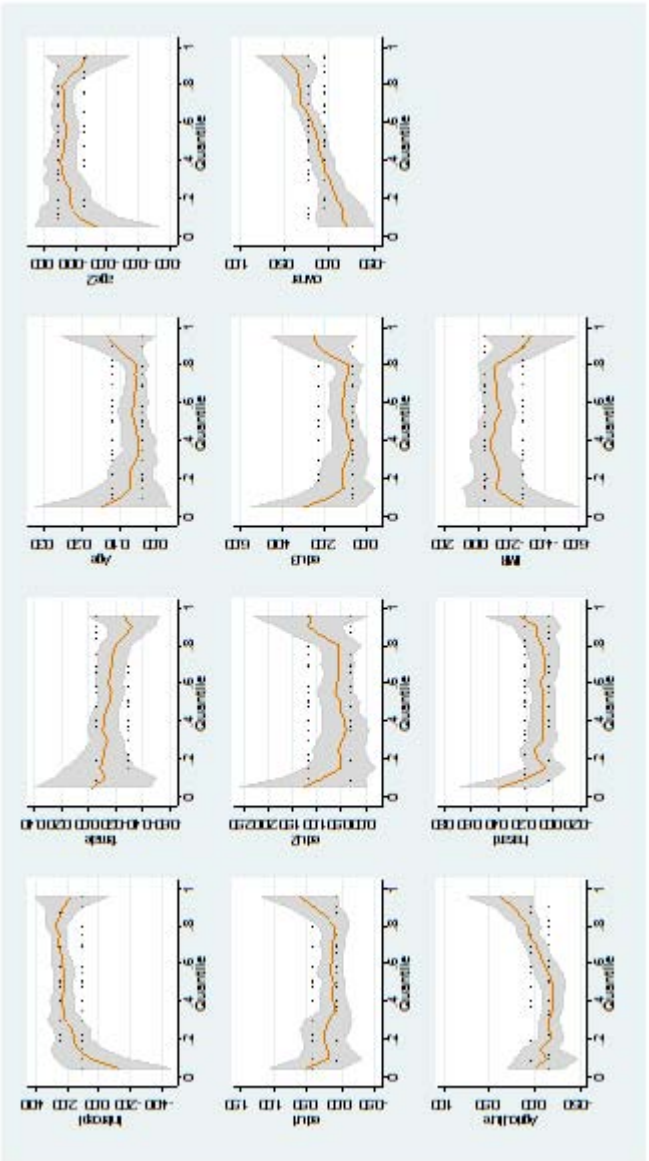


Fig. 3 The graphical illustration of the estimated coefficients in Table 3

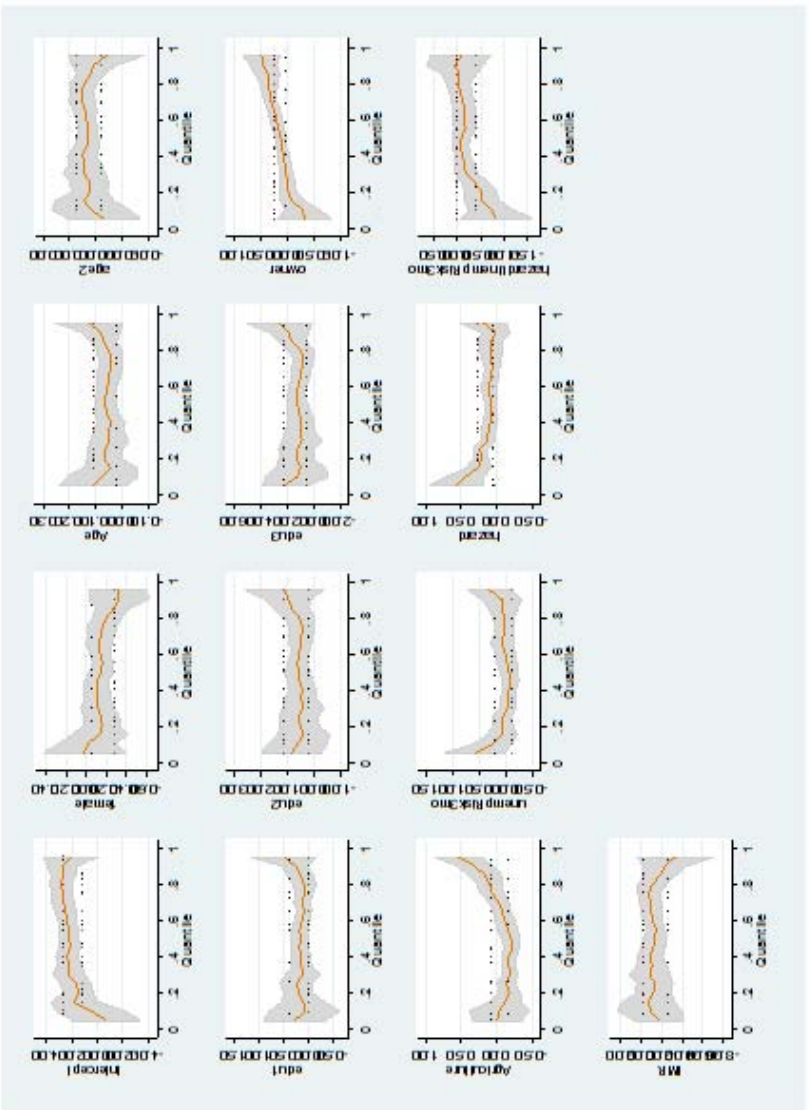


Fig. 4 The graphical illustration of the estimated coefficients in Table 4

with previous literatures. That is, age has a positive effect on wage with diminishing return and individuals with higher education tend to receive higher wage. The result also shows that there is, to some extent, gender discrimination in the labor market, that is, female workers on average has wage significantly lower than male workers.

Regarding to the variables related to the compensating wage differentials, the coefficients of hazard, unemployment risk, and their interaction term indicate that workers who face occupational hazard receive higher wage i.e. compensating wage. However, this benefit is reduced to null if those facing occupational hazard has unemployment risk.

Although OLS seems to be a reasonable method to estimate the condition wage, it fails to capture the difference across quantile. As shown in Table 3 and 4, although the sign of the estimated coefficients are consistent with those from OLS, their values vary across quantile.

The two wage regressions illustrated in Table 3 and 4 show some common results. While there is no evidence for gender discrimination in the lower quantiles, female earn significantly lower wages than male in the higher quantiles. Age has a positive and significant effect on wage in all quantiles. However, the positive effect is diminishing as the coefficients on the variable age squared are negative. Education has positive effect and the effect becomes stronger in the higher quantiles. The variable owner has no effect in the lower quantiles. Nevertheless, owners earn higher income comparing to employee or other types of independent workers in the higher quantiles. This illustrates a higher wage gap between owners and employees. Workers in agricultural sector earn lower wages in the lower quantiles and higher wages in the higher quantiles. The coefficients of inverse Mills ratio (IMR) are negative and more significant in the higher quantiles indicating that the sample selection bias is more severe for the estimates in the higher quantiles.

However, the two regressions give different results on the estimates for compensating wage differentials. In both regressions, the compensating wage differential effect is significant and positive in the lower and middle quantiles but not significant in the two highest quantiles. Nevertheless, the estimates in Table 3 are smaller and have a U-shape effect when the estimates in Table 4 are larger in the lower quantile and decline with the quantiles. When the unemployment risk variable is included into the model as well as its interaction term with occupational hazard, the results indicate that individuals whose wages are in the lower quantile are penalised heavily from having unemployment risk.

The main results in Table 4 are consistent with Guo and Hammitt (2009)[6] and Bender and Mridha (2011)[1] where there exists the evidence for the compensating wage differentials and the effect is weaker in the presence of the unemployment risk. The compensating wage differential effect is strong in the lower quantiles

In addition, this study illustrates the significance of controlling for the unemployment risk. Comparing the results in Table 3 and 4, the compensating wage differentials are highly underestimated in the quantiles that the unemployment risk effect is strong.

References

1. Bender, K., Mridha, H.: The Effect of Local Area Unemployment on Compensating Wage Differentials for Injury Risk. Southern Economic Association 78, 287–307 (2011)
2. Buchinsky, M.: Quantile Regression, Box-Cox Transformation Model and Changes in the Returns to Schooling and Experience. Journal of Econometrics 65, 109–154 (1995a)
3. Buchinsky, M.: Estimating the Asymptotic Covariance Matrix for Quantile Regression Models: a Monte Carlo Study. Journal of Econometrics 68, 303–338 (1995b)
4. Buchinsky, M.: The Dynamic of Changes in the Female Wage Distribution in the USA: A Quantile Regression Approach. Journal of Applied Econometrics 13, 1–30 (1998)
5. Buchinsky, M.: The Dynamics of Changes in the Female Wage Distribution in the USA: A Quantile Regression Approach. Journal of Applied Econometrics 13, 1–30 (2003)
6. Guo, X., Hammitt, J.K.: Compensating Wage Differentials with Unemployment: Evidence from China. Environmental and Resource Economics 42, 187–209 (2009)
7. Heckman, J.: Sample Selection Bias as a Specification Error. Econometrica 47, 153–161 (1979)
8. Ichimura, H.: Semiparametric Least Squares (SLS) and Weighted SLS Estimation of Single-Index Models. Journal of Econometrics 58, 71–120 (1993)
9. Koenker, R., Bassett, G.: Regression Quantiles. Econometrica 46, 33–50 (1978)
10. Lavetti, K.: The Estimation of compensating Differentials and Preferences for Occupational Fatality Risk. Working paper (2012)
11. Leeth, J.D., Ruser, J.: Compensating Wage Differentials for Fatal and Nonfatal Injury Risk by Gender and Race. Journal of Risk and Uncertainty 27, 257–277 (2003)
12. Moore, M.J., Viscusi, W.K.: Doubling the Estimated Value of Life: Results Using New Occupational Fatality Data. Journal of Policy Analysis and Management 7, 476–490 (1988a)
13. Moore, M.J., Viscusi, W.K.: The Quantity-Adjusted Value of Life. Economic Inquiry 26, 369–388 (1988b)
14. Newey, W.: Two Step Series Estimation of Sample Selection Model. MIT Unpublished Manuscript (1991)
15. Rosen, S.: Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. Journal of Political Economy 82, 34–55 (1974)
16. Rosen, S.: The Theory of Equalizing Differences. In: Handbook of Labor Economics, pp. 641–692 (1986)
17. Shanmugam, K.R.: The Value of Life: Estimates from Indian Labor Market. Indian Economic Journal 44, 105–114 (1996)
18. Shanmugam, K.R.: Value of Life and injury: Estimating Using Flexible Functional Form. Indian Journal of Applied Economics 6, 125–136 (1997)
19. Shanmugam, K.R.: Self Selection Bias in the Estimates of Compensating Differentials for Job Risks in India. Journal of Risk and Uncertianty 22, 263–275 (2001)
20. Smith, A.: The Wealth of Nations, W. Strahan and T. Cadell, London (1776)
21. Thailand, National Statistical Office of, Informal Sector Survey, National Statistical Office of Thailand (2012)
22. Viscusi, W.K., Aldy, J.A.: The Value of a Statistical Life: a Critical Review of Market Estimates Throughout the World. Journal of Risk and Uncertainty 27, 5–76 (2003)

Optimal Combination of Energy Sources for Electricity Generation in Thailand with Lessons from Japan Using Maximum Entropy

Tatcha Sudtasan and Komsan Suriya

Abstract. This study uses maximum entropy method to find an optimal combination of energy sources for electricity generation in Thailand. It sets three targets including unit cost, risk and pollution. In the optimization process, it forms three constraints according to these three targets. It solves the system following the guideline of Golan, Judge and Miller (1996). It analyses six scenarios of the targets. For the major results, it finds that hydropower, nuclear, wind and solar energy are major sources of electricity generation. The country cannot avoid adopting nuclear energy for its electricity generation in order to meet all the three targets that are optimal for its electricity generation and economic development.

1 Introduction

Thailand has never generated electricity by nuclear power. It is a controversial issue for the Thai society whether Thailand should adopt nuclear power plant. The plant will drive the cost of electricity low and avoid the shortage of electricity. However, its risk of explosion cannot be ignored.

Anti-nuclear people in Thailand fight for a nuclear-free country. They argue that Thailand can survive the energy crisis without nuclear power plant. It was indeed that this argument might be not true when the country and its people had to face a credible threat in April 2013 after Myanmar decided to pause the transportation of natural gas to Thailand under a maintenance reason. Thailand lost one-fourth of its electrical capacity during the days. Many factories were advised to stop their manufacturing.

Clean energy such as solar and wind power are new to Thailand. Even though Thailand is located in a tropical zone, the sunshine is unstable because of rapidly moving cloud all over the sky. Therefore, the capacity of solar energy to generate

Tatcha Sudtasan · Komsan Suriya
Department of Economics, Chiang Mai University
e-mail: {miracle.zaza,suriyakomsan}@gmail.com

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_35,

electricity is still low. This drives the cost of solar power high. Wind is also costly when the wind speed in Thailand is not so fast compared to that in Europe.

Coal and oil are less costly but release more pollution. Even though, advanced technology can capture CO2 and reduce its emission to the atmosphere, the emission level is still higher than other kinds of energy. Moreover, the threat of running out of coal in next 10 years makes the cost even higher.

When Thailand has to face at least three objectives in electricity generation to trade-off; cost, risk and pollution, the country has to decide the optimal combination of energy sources for its electricity generation. A critical issue is at nuclear power. A major research question is whether Thailand should adopt its first nuclear power plant. Is there any other source of power that can substitute nuclear power plant and keep Thailand a nuclear-free zone? This study aims to find out the answers for these questions.

2 Literature Reviews

After Golan, Judge and Miller (1996)[3] introduce the concept of maximum entropy, we search for literatures on using the technique for the portfolio optimization. Literatures using maximum entropy in portfolio optimization include several works such as Park (2007)[10], Rodriguez (2007)[13], Bera and Park (2008)[1], Jiang, He and Li (2008)[6], Qin, Li and Ji (2009)[11], Roeddner, Gartner and Rudolph (2009)[14], and Gartner (2012)[2]. These studies focus on mathematical issues of using maximum entropy in portfolio optimization and applications on risk management in financial sector. However, there is still no study that applies maximum entropy on the optimization of energy sources for electricity generation.

For studies on energy portfolio optimization, there are several works such as Liu (2007)[7] who use quadratic programming to find the solution, Rebennack, Kallrath and Pardalos (2010)[12] which use linear programming to solve the optimization problem, Hochreiter, Pflug and Wozabal (2005)[4] which apply stochastic programming to find the optimality. These studies do not use maximum entropy in the optimization.

The difference between this study and other literatures on energy optimization is that this study uses maximum entropy to find optimal combination of energy sources for electricity generation while other studies use mathematical programming and focus on the combination between buying electricity from spot market and bilateral contract between the buyer and electricity producers. The reason why this study uses maximum entropy is at the multi-objective optimization. It tries to set three objectives which are cost, risk and pollution. However, it has seven parameters to optimize from seven sources of energy which are solar, wind, hydro, oil, gas, coal and nuclear power. In this case, mathematical programming does not work when the number of equations is less than the number of parameters. The only way to solve this problem is to use maximum entropy.

Another reason why this study does not minimize cost under risk and pollution constraint is that the cost is not the only objective for the optimization. It tries to fit the three objectives at the same time. Even though the cost minimization under constraints is valid, the method is hard to process under limited information. However, this problem can be handled by maximum entropy. Therefore, the usage of maximum entropy is to serve the multi-objective optimization with limited information.

3 Methodology

This study uses maximum entropy to calculate the optimal combination of the energy sources for electricity generation for Thailand. It aims at three targets that the country needs to trade-off. They are cost, risk and pollution. Seven sources of energy are taken to the analysis. They are solar, wind, hydro, oil, gas, coal and nuclear power.

It forms three information equations as follows:

Information 1:

$$\sum_{k=1}^{7} UnitCost_k Energy_k = TargetCost \tag{1}$$

where

UnitCost is the cost of producing 1 MWh of electricity,

Energy is portion of each source of energy that generates electricity which is unknown,

TargetCost is the target of unit cost of producing 1 MWh of electricity set by policy makers.

Information 2:

$$\sum_{k=1}^{7} RiskIndex_k Energy_k = TargetRisk \tag{2}$$

where

RiskIndex is the human dead toll of producing 1 MWh of electricity,

Energy is portion of each source of energy that generates electricity which is unknown,

TargetRisk is the target of human dead toll of producing 1 MWh of electricity set by policy makers.

Information 3:

$$\sum_{k=1}^{7} CO_2Emission_k Energy_k = TargetPollution \tag{3}$$

where

$CO_2Emission$ is the emission of carbon dioxide to the atmosphere from producing 1 MWh of electricity,

$Energy$ is portion of each source of energy that generates electricity which is unknown,

$TargetPollution$ is the target of the emission of carbon dioxide to the atmosphere from producing 1 MWh of electricity set by policy makers.

Maximum entropy equation

To solve for seven unknowns of portion of each source of energy that generates electricity when we have only three information equations, it is impossible to do with other techniques but maximum entropy. The maximum entropy will construct a Lagragian function that try to maximizes the entropy function by using all the information equations as constraints. The Lagrangian function can be written as follows:

$$\begin{aligned} L = &- \sum_{k=1)}^{7} Energy_k lnEnergy_k + \lambda_1 (TargetCost - \sum_{k=1}^{7} UnitCost_k Energy_k) \\ &+ \lambda_2 (TargetRisk - \sum_{k=1}^{7} RiskIndex_k Energy_k) \\ &+ \lambda_3 (TargetPollution - \sum_{k=1}^{7} CO_2Emission_k Energy_k) \end{aligned} \tag{4}$$

where

L is Lagrangian function,

$Energy$ is portion of each source of energy that generates electricity which is unknown,

ln is natural logarithm,

$TargetCost$ is the target of unit cost of producing 1 MWh of electricity set by policy makers,

$TargetRisk$	is the target of human dead toll of producing 1 MWh of electricity set by policy makers,
$TargetPollution$	is the target of the emission of carbon dioxide to the atmosphere from producing 1 MWh of electricity set by policy makers,
$UnitCost$	is the cost of producing 1 MWh of electricity,
$RiskIndex$	is the human dead toll of producing 1 MWh of electricity,
$CO_2Emission$	is the emission of carbon dioxide to the atmosphere from producing 1 MWh of electricity,
λ	is Lagrange multiplier.

The technique to estimate parameters Energy can be presents step by step as follows:
Step 1: Use the formula of the concentrate maximum entropy of Golan, Judge and Miller (1996)[3] as follows to find .

$$l(\lambda) = \sum_{t=1}^{3} \lambda_t * Target_t + ln(\Omega(\lambda)) \tag{5}$$

where
$\Omega(\lambda) = \sum_{k=1}^{7} exp(-\lambda_1 \cdot UnitCost_k - \lambda_2 \cdot RiskIndex_k - \lambda_3 \cdot CO_2Emission_k)$

Step 2: Find $\frac{\partial l}{\partial \lambda_1}$, $\frac{\partial l}{\partial \lambda_2}$ and $\frac{\partial l}{\partial \lambda_3}$ and set them to zero.
The derivatives are as follows:

$$\begin{aligned}\frac{\partial l}{\partial \lambda_1} = 0 = TargetCost + \frac{1}{\Omega(\lambda)} \cdot \\ (-e^{(-\lambda_1 * UnitCost_1 - \lambda_2 * RiskIndex_1 - \lambda_3 * CO_2Emission_1)} \cdot UnitCost_1 \\ - e^{(-\lambda_1 * UnitCost_2 - \lambda_2 * RiskIndex_2 - \lambda_3 * CO_2Emission_2)} \cdot UnitCost_2 \\ - e^{(-\lambda_1 * UnitCost_3 - \lambda_2 * RiskIndex_3 - \lambda_3 * CO_2Emission_3)} \cdot UnitCost_3 \\ - e^{(-\lambda_1 * UnitCost_4 - \lambda_2 * RiskIndex_4 - \lambda_3 * CO_2Emission_4)} \cdot UnitCost_4 \\ - e^{(-\lambda_1 * UnitCost_5 - \lambda_2 * RiskIndex_5 - \lambda_3 * CO_2Emission_5)} \cdot UnitCost_5 \\ - e^{(-\lambda_1 * UnitCost_6 - \lambda_2 * RiskIndex_6 - \lambda_3 * CO_2Emission_6)} \cdot UnitCost_6 \\ - e^{(-\lambda_1 * UnitCost_7 - \lambda_2 * RiskIndex_7 - \lambda_3 * CO_2Emission_7)} \cdot UnitCost_7)\end{aligned} \tag{6}$$

$$\frac{\partial l}{\partial \lambda_2} = 0 = TargetRisk + \frac{1}{\Omega(\lambda)} \cdot \begin{aligned} &(-e^{(-\lambda_1 * UnitCost_1 - \lambda_2 * RiskIndex_1 - \lambda_3 * CO_2Emission_1)} \cdot RiskIndex_1 \\ &- e^{(-\lambda_1 * UnitCost_2 - \lambda_2 * RiskIndex_2 - \lambda_3 * CO_2Emission_2)} \cdot RiskIndex_2 \\ &- e^{(-\lambda_1 * UnitCost_3 - \lambda_2 * RiskIndex_3 - \lambda_3 * CO_2Emission_3)} \cdot RiskIndex_3 \\ &- e^{(-\lambda_1 * UnitCost_4 - \lambda_2 * RiskIndex_4 - \lambda_3 * CO_2Emission_4)} \cdot RiskIndex_4 \\ &- e^{(-\lambda_1 * UnitCost_5 - \lambda_2 * RiskIndex_5 - \lambda_3 * CO_2Emission_5)} \cdot RiskIndex_5 \\ &- e^{(-\lambda_1 * UnitCost_6 - \lambda_2 * RiskIndex_6 - \lambda_3 * CO_2Emission_6)} \cdot RiskIndex_6 \\ &- e^{(-\lambda_1 * UnitCost_7 - \lambda_2 * RiskIndex_7 - \lambda_3 * CO_2Emission_7)} \cdot RiskIndex_7) \end{aligned} \quad (7)$$

$$\frac{\partial l}{\partial \lambda_3} = 0 = TargetPollution + \frac{1}{\Omega(\lambda)} \cdot \begin{aligned} &(-e^{(-\lambda_1 * UnitCost_1 - \lambda_2 * RiskIndex_1 - \lambda_3 * CO_2Emission_1)} \cdot CO_2Emission_1 \\ &- e^{(-\lambda_1 * UnitCost_2 - \lambda_2 * RiskIndex_2 - \lambda_3 * CO_2Emission_2)} \cdot CO_2Emission_2 \\ &- e^{(-\lambda_1 * UnitCost_3 - \lambda_2 * RiskIndex_3 - \lambda_3 * CO_2Emission_3)} \cdot CO_2Emission_3 \\ &- e^{(-\lambda_1 * UnitCost_4 - \lambda_2 * RiskIndex_4 - \lambda_3 * CO_2Emission_4)} \cdot CO_2Emission_4 \\ &- e^{(-\lambda_1 * UnitCost_5 - \lambda_2 * RiskIndex_5 - \lambda_3 * CO_2Emission_5)} \cdot CO_2Emission_5 \\ &- e^{(-\lambda_1 * UnitCost_6 - \lambda_2 * RiskIndex_6 - \lambda_3 * CO_2Emission_6)} \cdot CO_2Emission_6 \\ &- e^{(-\lambda_1 * UnitCost_7 - \lambda_2 * RiskIndex_7 - \lambda_3 * CO_2Emission_7)} \cdot CO_2Emission_7) \end{aligned} \quad (8)$$

Step 3: Use Newton method to solve for λ_1, λ_2 and λ_3. This is done by using fsolve function in Matlab.

Step 4: Ensure that the derivatives in step 2 are all zero. This is to check the convergence of the optimization.

Step 5: Calculate the Lagragian, $l(\lambda)$.

Step 6: Use the formula of Golan, Judge and Miller (1996)[3] to find the parameter *Energy* by plugging λ_1, λ_2 and λ_3 into the formula.

$$Energy_k = \frac{exp(-\lambda_1 \cdot UnitCost_k - \lambda_2 \cdot RiskIndex_k - \lambda_3 \cdot CO_2Emission_k)}{\sum_{k=1}^{7} exp(-\lambda_1 \cdot UnitCost_k - \lambda_2 \cdot RiskIndex_k - \lambda_3 \cdot CO_2Emission_k)} \quad (9)$$

or it can be written as:

λ_1, λ_2 and λ_3 into the formula.

$$Energy_k = \frac{exp(-\lambda_1 \cdot UnitCost_k - \lambda_2 \cdot RiskIndex_k - \lambda_3 \cdot CO_2Emission_k)}{\Omega(\lambda)} \quad (10)$$

Step 7: Calculate the Entropy, $-\sum_{k=1}^{7} Energy_k lnEnergy_k$

Scenarios to be analyzed includes 6 scenarios as shown in table 1.

Table 1 Scenarios to be analyzed in the study

Scenario	Target cost (JPY per MWh)	Target risk (Human dead toll per MWh per year)	Target pollution (CO_2 emission per MWh)
1	14	64	341
2	7	10	120
3	10	5	100
4	10	20	30
5	10	10	30
6	10	10	20

Note: JPY = Japanese Yen (Money currency of Japan).
MWh = Megawatts (Unit of electricity).

The first scenario illustrates the average value of cost, risk index and pollution from the seven sources of energy. The second scenario drops all the targets down enormously. The third scenario trades-off between cost and risk plus pollution compared to the second scenario. The fourth scenario allows the risk rise but drives the pollution even lower than the third case while keeps the cost unchanged. The fifth scenario attempts to find the optimal combination by negotiates the lower risk compared to the fourth scenario. The last scenario finds whether it is possible to reduce more pollution than the fifth scenario.

4 Data

Data of unit cost are derived from Sudtasan and Suriya (2012)[16]. The unit cost is in Japanese Yen. This is because Thailand has never used nuclear power. Therefore, we would like to compare relative cost of nuclear power and other energy sources. We use the data and lessons from Japan in this study because of the complete set of data and Japan is in Asia like Thailand.

Data of risk index are collected from Inhaber (1982)[5]. The paper evaluates the total risk per unit energy output (one megawatt-year) by the total deaths caused by each energy system. Data of pollution are according to Sovacool (2008)[15]. The paper measures the CO2 emission for electricity generators.

All the data are shown in table 2 as follows:

Table 2 Data of unit cost, risk index and pollution of each energy source

Energy sources	Unit cost (JPY per MWh)	Risk index (Human dead toll per MWh per year)	Pollution (CO_2 emission per MWh)
Solar (Photovoltaic)	46	55	32
Wind	12	70	10
Hydropower	10.75	5.5	10
Natural gas	6.45	0.4	443
Coal without scrubbing	5.75	170	1,050
Nuclear	5.50	1.55	66

Sources: Unit cost from Sudtasan and Suriya (2012), risk index from Inhaber (1982) and pollution from Sovacool (2008).

5 Results

Table 3 illustrates the results from the optimization process using maximum entropy. It shows that in the first case where the targets are set at the average value of all seven energy sources, the portions of energy sources to generate electricity are quite the same. Each of them shares around 14% -15% of total electricity generation.

In the second scenario where the target cost is lowest among scenarios, nuclear power plays a major role in electricity generation, accounting around 61% of total power. However, when we trade-off the cost with the risk plus pollution in the third scenario, hydro power replaces the major role with the portion around 68% leaving nuclear powers portion dropped to around 11%. In this scenario, natural gas appears to be the second largest combination in total electricity generation with the portion of 19%.

In the fourth scenario, when we keeps the cost constant at JPY10 per MWh as in the third scenario and trade-off between risk and pollution, the result show that hydropower is still the most important energy source in the combination, around 42%. The second largest source is nuclear, around 32%. Wind comes to contribute to the countrys electricity generation when its portion exceeds 23%.

Keeping the cost and pollution constant as in the fourth scenario, the negotiation on risk in the fifth scenario to be just half of the previous one is possible. The system can still be solved. This can be noticed by the zero derivatives that all show that the system is converged. Now, the portion of hydropower increases to be around 57% when nuclear remains around 34%. Wind power drops to around 7% but solar power increases a little bit to be around 3%. Oil, gas and coal are likely to disappear from the usage for electricity generation.

In the last scenario when we negotiate for even less pollution compared to the fifth scenario, the system is yet possible to be solved. The significance of

Table 3 Results from the optimization process using maximum entropy

	Scenario1	Scenario2	Scenario3	Scenario4	Scenario5	Scenario6
Target						
TargetCost	14	7	10	10	10	10
TargetRisk	64	10	5	20	10	10
TargetPollution	341	120	100	30	30	20
Lamda						
Lamda1	0.0018	0.2909	-0.8108	0.0691	0.0332	0.0937
Lamda2	0.0001	0.0045	0.6491	0.0081	0.0317	0.0329
Lamda3	0.0000	0.0032	0.0025	0.0119	0.0146	0.0366
Derivatives						
Derivative1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Derivative2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Derivative3	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Lagrangian	1.9456	1.1416	0.8947	1.1620	0.9802	0.7436
Entropy	1.9456	1.1416	0.8947	1.1620	0.9802	0.7436
Combination (%)						
Solar	13.53	0.00	1.84	1.91	2.65	0.24
Wind	14.37	8.07	0.00	23.02	7.01	7.88
Hydro	14.50	15.48	67.99	42.42	56.58	74.12
Oil	14.12	0.31	0.00	0.00	0.00	0.00
Gas	14.58	14.11	19.28	0.34	0.14	0.00
Coal	14.26	1.20	0.00	0.00	0.00	0.00
Nuclear	14.64	60.82	10.88	32.31	33.63	17.76
Sum (%)	100.00	100.00	100.00	100.00	100.00	100.00

Source: Calculation using Matlab.

hydropower increases dramatically from around 57% to 74%. Nuclear reduces its role from around 34% to 17%. Wind power remains quite constant at around 8%. Unfortunately, solar power is likely to disappear from the scene.

6 Discussions

Before Fukushima incident of the blasting nuclear power plant, the combination of energy sources in Japan in 2009 for electricity generation combines 25% of coal, 10.7% of nuclear, 39.5% of natural gas, 14.4% of oil, 9% of hydropower and 1.4% of other sources (Mitsubishi Corporation, 2012[9]). A question arises why this fact is very different from the results of this paper. Hydropower that should be the largest source of power shares only 9% when it should be 74%. Nuclear shares 10.7% in

reality while it should be reduced to only 17%. Moreover, coal is still in active for electricity generation in Japan when it should be eliminated away from the system.

The small portion of hydroelectric is usual in many developed countries (Mitsubishi Corporation, 2012[9]). U.S.A. generates only 6.6% of its electricity from hydropower in 2009. In the same year, Germany uses hydroelectric just less than 5%. Even in the natural-resource rich country as India, the portion of hydropower is just around 12%. Therefore, hydropower is a good source of energy that satisfies almost every target. Unfortunately, a country may not be able to construct dams as many as it wants. The construction of a new dam is controversial among local people as well as NGOs.

For the lessons learnt from Japan to Thailand, it can be seen in 2011 that Thailand depends its 71% of total electricity generation on natural gas, 21% from coal, 5% from hydroelectric, 2% from renewable energy and 1% from oil (Ministry of Energy of Thailand, 2011[8]). From the results, it can be expected that the portion of coal will be reduced enormously in the next decades. Hydropower is still a hope to generate cheaper and cleaner electricity but the construction of new dams will face severe protests. The most important issue raised by the results of this study is that Thailand needs a nuclear power plant to substitute the gradually reduced electricity generated by coal.

7 Conclusions

It is clear but cautious from the analysis that nuclear power seems to be a must for Thailand when no scenario shows that the portion of nuclear power should be reduced to zero. However, the caution is that the study includes only short-term costs and risks for nuclear power.

Dirtier sources of power such as coal, gas and oil will be faded away from electricity generation of the country. Clean energy will replace the role. Hydropower will be the most significant energy source in keeping the cost, risk and pollution low. Wind power will be more significant than solar power. It will become the third largest source of energy.

In conclusion, Thailand will move toward clean energy with only 4 sources of energy that will be included into the combination of the electricity generation. They are hydro, nuclear, wind and solar power. The country might not be able to avoid nuclear power. Otherwise the country cannot meet the targets that favor all dimensions of the economic development.

Further studies on this issue should include other long-term costs and risks such as storing the nuclear waste and possible radiation leaks caused by accidents or natural disasters at nuclear power plants like what happened in Fukushima incident. After including these costs and risks, the results may reduce the likelihood to use nuclear power. However, the trend that Thailand will move toward clean energy will still persist.

References

1. Bera, A.K., Park, S.Y.: Optimal Portfolio Diversification Using Maximum Entropy Principle. Econometric Reviews 27, 484–512 (2008)
2. Gartner, I.R.: Differentiated Risk Models in Portfolio Optimization: A Comparative Analysis of the Degree of Diversification and Performance in the Sao Paulo Stock Exchange (Bovespa). Pesquisa Operacional 32(2), 271–292 (2012)
3. Golan, A., Judge, G.G., Miller, D.: Maximum Entropy Econometrics: Robust Estimation with Limited Data. John Wiley & Sons (1996)
4. Hochreiter, R., Pflug, G.C., Wozabal, D.: Multi-stage stochastic electricity portfolio optimization in liberalized energy markets. Working Paper on Optimal Energy Portfolios, Department of Statistics and Decision Support System, University of Vienna (2005)
5. Inhaber, H.: Is Solar Power More Dangerous Than Nuclear? IAEA Bulletin 21(1), 11–17 (1982)
6. Jiang, Y., He, S., Li, X.: A Maximum Entropy Model for Large Scale Portfolio Optimization. In: Proceedings of the International Conference on Risk Management and Engineering Management 2008, pp. 610–615 (2008)
7. Liu, M.: Portfolio optimization in electricity markets. Electric Power Systems Research 77, 1000–1009
8. Ministry of Energy of Thailand. Thailand electricity generation by source in 2011. Ministry of Energy, Bangkok (2011)
9. Mitsubishi Corporation. Annual Report, Power Business (2012), `http://www.mitsubishicorp.com`
10. Park, S.Y.: Optimal Portfolio Diversification Using Maximum Entropy Principle. Chapter 3 in Sung Yong Park, Essays on Maximum Entropy Principles with Applications to Econometrics and Finance. ProQuest (2007)
11. Qin, Z., Li, X., Ji, X.: Portfolio selection based on cross-entropy. Journal of Computational and Applied Mathematics 228, 139–149 (2009)
12. Rebennack, S., Kallrath, J., Pardalos, P.M.: Energy Portfolio Optimization for Electric Utilities: Case Study for Germany. In: Energy, Natural Resources and Environmental Economics Energy Systems, pp. 221–246 (2010)
13. Rodriguez, J.: A New Portfolio Optimization Based on Entropy. Master thesis, Section of Mathematics, Faculty of Sciences, University of Geneva (2007)
14. Roeddner, W., Gartner, I.R., Rudolph, S.: Entropy-Driven Portfolio Selection: A Downside and Upside Risk Framework. Discussion Paper Number 437. Faculty of Economic Sciences, University of Hagen (2009)
15. Sovacool, B.K.: Valuing the greenhouse gas emissions from nuclear power: A critical survey. Energy Policy 36, 2940–2953 (2009)
16. Sudtasan, T., Suriya, K.: Nuclear power plant after Fukushima incident: Lessons from Japan to Thailand for choosing power plant options. The Empirical Econometrics and Quantitative Economics Letters 1(3), 1–8 (2012)

Valuation of Interest Rate Derivatives under CSA Discounting

Amy R. Daniels, Coenraad C.A. Labuschagne, and Theresa M. Offwood-le Roux

Abstract. Standard pricing theory assumes that traders can borrow and lend at a unique risk-free rate, ignoring the intricacies of the collateralization market. Since 2007, the market has adopted an advanced methodology for valuing interest rate derivatives, based on the standard Credit Support Annex (CSA), which is a document used to define the terms under which collateral is posed between counterparties. This change however, has not yet been implemented in South African markets due to the difficulty created by the lack of a liquid overnight indexed swap (OIS) market in South Africa. In this paper, we propose two proxies, which could be used to approximate an OIS market. We compare the implied forward rates as well as the pricing of a vanilla swap under these OIS methods to the classical case.

1 Introduction

Preceding the financial crisis that started in the second half of 2007, the common rate used for both discounting and forecasting in derivative valuation was the 3-month rate such as LIBOR[1]. The focus of interest rate derivative valuation was on

Amy R. Daniels
Deloitte, Deloitte Place, The Woodlands,
20 Woodlands Drive, Woodmead, 2052, South Africa
e-mail: amdaniels@deloitte.co.za

Coenraad C.A. Labuschagne
Department of Finance and Investment Management,
University of Johannesburg, South Africa
e-mail: coenraad.labuschagne@gmail.com

Theresa M. Offwood-le Roux
Standard Bank, 3 Simmonds Street, Johannesburg, South Africa
e-mail: theresa.offwood-leroux@standardbank.co.za

[1] LIBOR is the average interbank rate at which a selection of banks on the London Money Market are prepared to lend to one another.

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_36,

the term structure of interest rates, and aspects such as credit risk, liquidity risk, collateral agreements, and funding costs were ignored [10]. However, as the basis between the 3-month lending rate and the overnight rate was small and relatively stable, the impact between using a 3-month funding curve to an overnight funding curve to discount was inconsequential [3].

In 2008, credit risk was brought to the fore and at the height of the financial crisis the difference between 3-month LIBOR and the federal funds rate increased from about 8 basis points to around 366 basis points (See Figure 1). The problem is that even today, the difference between these rates is significant. Analogously, such patterns were also found between forward rate agreements (FRA) rates and the forward rates implied by two consecutive deposits, as well as among swaps[2] rates with different adjustable (also known as floating) leg maturities (also known as tenors).

Thus, the market has been forced to develop a new framework and to re-evaluate the no-arbitrage models used for derivative pricing and risk analysis. The traditional no-arbitrage framework developed to price derivatives, originating from Black and Scholes (1973) and Merton (1973) [2], has become out-dated, see Piterbarg (2010) [8] and Piterbarg (2012) [9]. The concept of constructing a single risk-free yield curve, which reflects both the present value of funding future cash flows as well as the level of forward rates, has been rejected.

An overnight indexed swap (OIS) is a financial instrument which swaps a rolled overnight rate for an interbank rate such as 3-month LIBOR.

The vast majority of derivative traders now agree that collateralized trades should be discounted at the appropriate OIS rate, while non-collateralized trades should be discounted at the interbank rate. The OIS rate eliminates the bank credit and liquidity risk that a 3-month rate includes.

Although OIS discounting has universal approval, in South Africa this has proved difficult to implement. It is still the case that both collateralized and uncollateralized trades are discounted using the same curve, which is derived from the Johannesburg Interbank Agreed Rate (JIBAR). Even the interest rate swap clearing service, Swap-Clear, run by London-based clearing house, LCH.Clearnet uses the JIBAR curve to value rand-currency swaps, as a result of the absence of an applicable substitute.

The main difficulty lies in the lack of a liquid, local currency OIS curve. Although market leaders recognize the need for an OIS market in South Africa, there is no consensus on how the initiation of such a market should be approached. Industry participants have called attention to the fact that the financial crisis did not hit South Africa as hard as it did the rest of the world. For example, the spread between

[2] A swap is an agreement between two parties to exchange sequences of cash flows for a set period of time. Usually, at the time the contract is initiated, at least one of these series of cash flows is determined by a random or uncertain variable, such as an interest rate, foreign exchange rate, equity price or commodity price. Conceptually, one may view a swap as either a portfolio of forward contracts, or as a long position in one bond coupled with a short position in another bond.

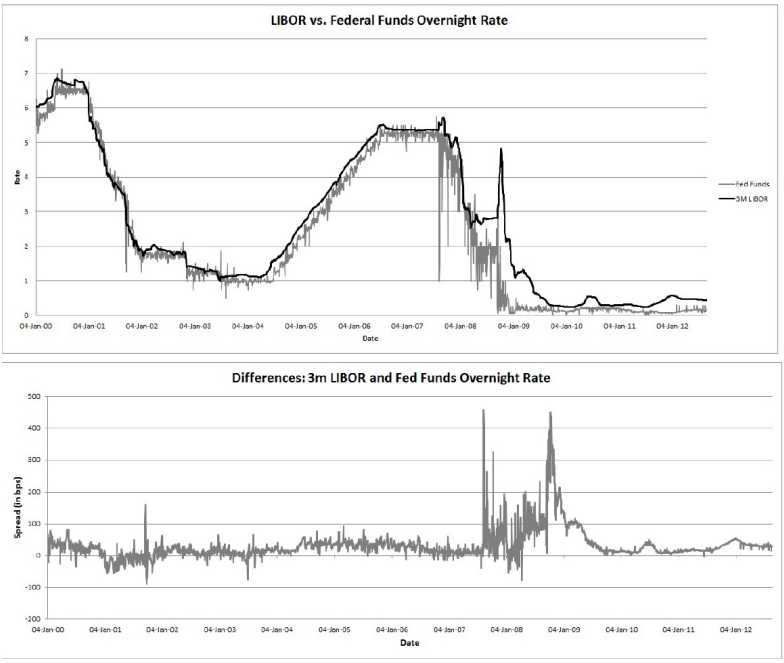


Fig. 1 3-month LIBOR and the Federal Funds Rate Plots and Spreads for the Period January 2000-January 2012

3-month JIBAR and the Safex overnight rate rose from 65 basis points at the start of 2008 to a high of 141.5 basis points on May 29 that year (see Figure 2).

So although the need for an OIS curve is not that urgent, and while the construction of an OIS market is not a high priority, is there a way to price derivatives, without waiting for the conception of a formal, liquid OIS market? One possible solution would be to use the South African Futures Exchange (Safex) rate as our overnight rate. This rate is the average rate that Safex receives on its deposits with the banks, weighted by the size of the investments placed at each bank. However, Safex deposits represent a very small portion of the overnight funding in South Africa. The Safex overnight rate may therefore not be a good reflection, or representative of the weighted average call rates paid on Rand deposits by all banks.

Another option would be the use of the South African benchmark overnight rate (Sabor). Published by the South African Reserve Bank (SARB), it is the volume weighted average of interbank funding at a rate other than the current repurchase rate, and the twenty highest rates paid by banks on their overnight and call deposits, plus a five percent weight for funding through foreign exchange swaps [12]. A drawback to using Sabor is that it is known merely as a point of reference, and cannot be traded.

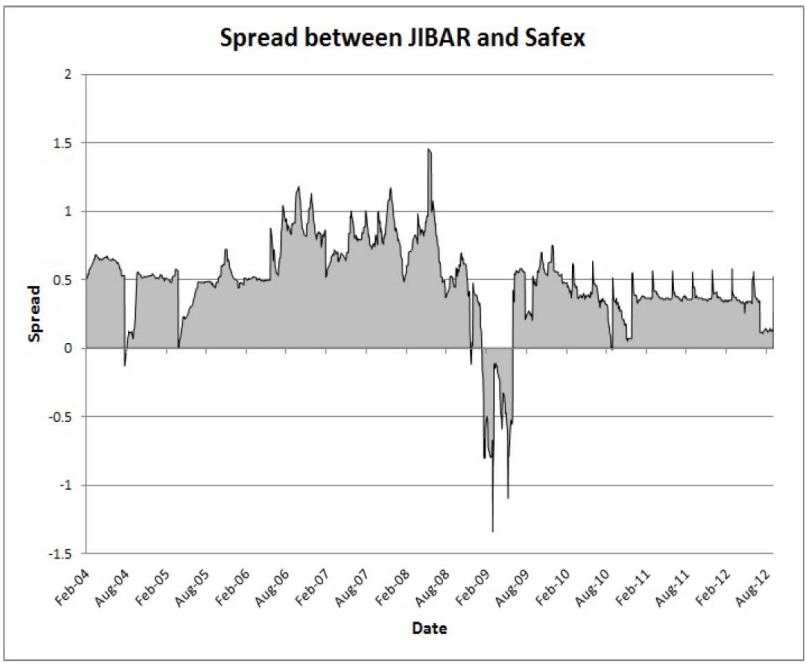


Fig. 2 3-month JIBAR-Safex Overnight Rate Basis for the Period February 2004-August 2012

In this paper, we explore two methods to approximate an OIS market. We compare the implied forward rates to the classical situation as well as the differences in valuation of a vanilla interest rate swap.

2 Classical vs OIS Swap Pricing

In this section we make use of a numerical example to show how the use of an OIS curve differs from using JIBAR in the discounting process. The distinction can be shown through the implied forward curves that we derive from each method.

We show the difference between the mark-to-market (MTM) of a hypothetical swap under both procedures. Our swap trade date is the 21st of September 2012, and our data runs quarterly, up until the 21st of September 2027. In other words, we have a fifteen year tenor.

We make the following assumptions:

- All transactions occur in a single currency economy.
- There are no taxes or transaction costs.
- Pricing and valuation occur on a settlement date.

- Our day count convention is $\frac{ACT}{365}$, and we adopt the Modified Following Rule[3] as our business day convention, in accordance with the South African market protocol.

2.1 Classic Approach to Price and Value Interest Rate Swaps

In the past, the 3-month swap curve was used to forecast and discount cashflows. The swap curve consists of observed market interest rates, derived from market instruments that represent the most liquid and dominant instruments for their respective time horizons, bootstrapped and combined using an interpolation scheme. For more details on how curve bootstrapping works see Alexander [1] or Hull [5, 6]. Under the classical or traditional framework used to price interest rate swaps, the movement between the JIBAR forward[4], spot[5] and swap curves[6], is done comfortably. We can begin with the observed forward curve, and use the technique of bootstrapping to obtain the implied spot rates and the swap fixed rates. Conversely, we could be supplied with the observed swap curve, that is, the fixed rates on par interest rate swaps, and then bootstrap the implied spot curve as well as the implied forward curve. Note, that this ease of movement is due to the assumption that the JIBAR-based implied spot rates can also be used to discount future cash flows. This assumption becomes imperative in what follows - the calculation of the mark-to-market of an existing swap.

Now consider a swap with a fixed rate of 7% swapped quarterly with 3 month JIBAR and a notional of 50 million. We want to value this swap on 21st September 2012 which matures on 21st December 2017, i.e. it only has 63 months remaining. Its MTM value is based on a comparison to the 5.8467% fixed rate on the 63-month at-market swap. The annuity is the difference between the contractual and the current market fixed rates, multiplied by the notional principal and day count factor.

$$(7\% - 5.8467\%) \times R\,50\,000\,000 \times 0.249315068 = R\,143\,767.72.$$

Notice that this is the explicit aspect of the valuation. The obscure aspect of this arises in discounting the annuity. Traditionally, under the classical model, we would utilize the sequence of implied spot rates. Then the value of the swap is R2 611 382.01.

Now observe that there are fundamental assumptions made in the computation of this mark-to-market value:

[3] Payment days that fall on holidays, Saturdays or Sundays roll forward to the next business day. If that day falls in the next calendar month, the payment rolls back to the day that precedes the payment date.

[4] A forward rate curve represents the no-arbitrage rate today that will be earned in the future.

[5] Spot is the price that is quoted for immediate settlement on a commodity, a security or a currency. Spot settlement is normally one or two business days from trade date.

[6] A swap rate curve shows the fixed-rate leg against the floating leg of 3-month JIBAR.

- Either this is an uncollateralized swap or we do not take into account the collateral in the valuation methodology, and
- the fixed rate payer is a "JIBAR-flat" borrower. This essentially, yet interestingly indicates that the fixed rate payer, or floating rate receiver, effectively possesses the same credit standing as the banks that are used to develop the JIBAR index.

2.2 CSA Approach to Pricing and Valuing Interest Rate Swaps

As South Africa does not have an OIS market, there are two proxies we would like to suggest to be able to approximate the value of a derivative under OIS discounting:

- We can assume that the OIS curve is just a parallel shift from the above-mentioned bootstrapped spot rate curve. Then we can use this curve to calculate discount factors, forward rates and OIS fixed swap rates. The reason for taking a spread below the ZAR spot rate curve is intuitive. We simply expect the overnight indexed swap rate to be less than the 3-month JIBAR curve that this swap is derived from, since we have decreased our exposure to credit risk through the inclusion of collateral, and this reduction in credit risk implies a reduction of yield.
- Instead of shifting the already bootstrapped curve, we can shift the normal par swap rates downwards in a parallel fashion to obtain the OIS par fixed rates. Then we can bootstrap an OIS curve from these swap rates.

To decide the amount by which we do this parallel shift, we looked at the historical average spread between the proxy (Safex or Sabor) rates and the 3-month JIBAR curve. This value comes to on average a 50 basis point difference.

Using both of these methods, we derived the implied forward curves and compared them to the classical forward curve (see Figure 3). The solid black line represents the implied forward curve from the classical approach, while the dotted line is from shifted curve - OIS approach. The implied forward rates from the shifted fixed rates OIS method were indistinguishable from the classical case, thus we only plotted one of them. It is clear that the differences are small. Figure 4 shows these differences in basis points.

Comparing the implied forward curves under the classical model and under OIS discounting, we can make the following observations: Following markets abroad and their analysis [11], we should have the result that the implied forward curve under the classical model, is on average higher than the implied forward under the OIS discounting method. This expectation does not hold true in our case, indicating that the OIS curves constructed here might not be entirely appropriate for practical use in industry.

Lastly, we plot the discount factors calculated from the classical method as well as both OIS proxy methods (see Figure 5).

The OIS discount factors are higher than the classical ones, as expected.

Next, we compare the mark-to-market value under OIS discounting and the mark-to-market value under the classical model of the above mentioned swap. We get the following values

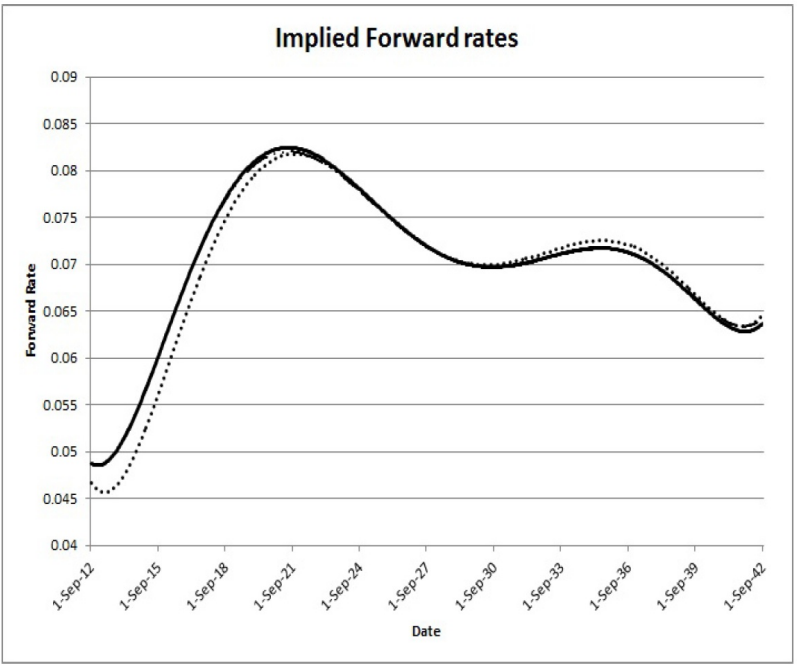


Fig. 3 Implied forward rates in the constructed OIS and the classical case

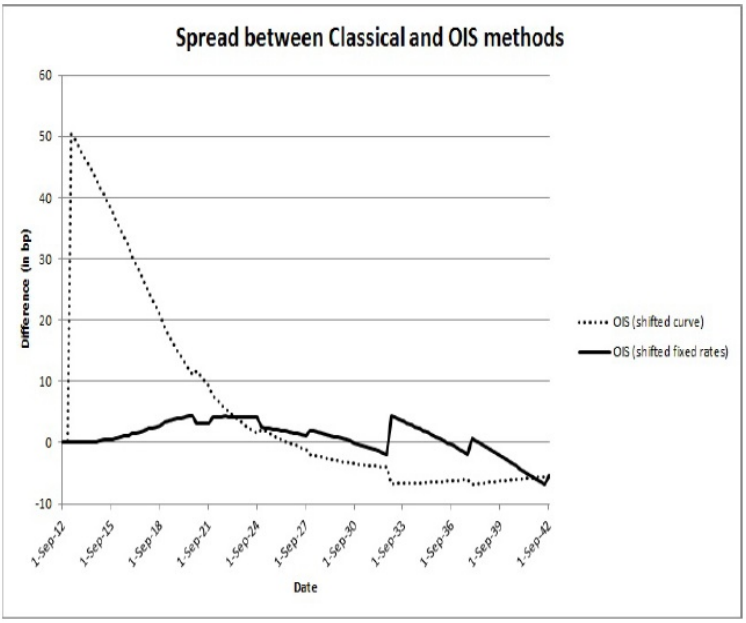


Fig. 4 Spread in basis points between the OIS forward curves and the classical forward curve

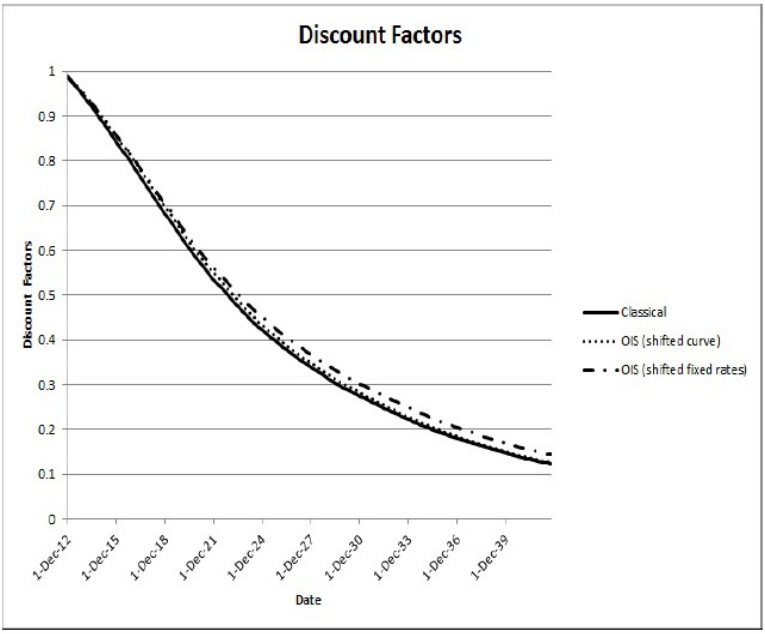


Fig. 5 Comparison of the discount factors calculated using the OIS curves and the classical curve

- R2 611 382.01 under the classical model,
- R2 640 984.89 under the shifted curve OIS method,
- R2 646 838.13 under the shifted fixed rate OIS method.

As expected, the MTM of the swap is higher under OIS discounting. This is due to the fact that collateral posting decreases the credit risk. It is interesting to see that the MTM under the shifted fixed rate OIS method is higher than under the shifted curve OIS method. But this could be specific to the swap we are pricing and further analysis is needed to determine a pattern in the swap MTM values.

3 Conclusion

The results of our proxy methods follow the behaviour we would expect from an OIS curve. Thus, these methods might be worth considering for the purpose of getting an approximation of the differences that OIS discounting may incur. However, the best solution in our opinion, would be to create a liquid OIS market. South Africa should maybe look to countries like Poland, which only recently (2004) started their own OIS market. They faced similar problems to South Africa with ‘some market participants not believing at all in the possibility of the market being created for the Polish currency. It suffices to mention that the first transactions were preceded by ca. 4 years of meetings, debates and agreements to realize how bad the situation

was' (as stated on the Polish Financial Markets Association website). Since the Polish OIS market was launched, the annual turnover of the Polish market rose from USD 330 million to 64.8 billion in 2010, according to central bank surveys [4]. Hungary is another country which created their own OIS market from scratch. South Africa could learn from these processes and hopefully one day there will be a local OIS market. A comparison between the OIS market in Thailand and South Africa could be helpful, but information about the OIS market in Thailand is difficult to obtain. In the meantime, proxies like the ones discussed in this paper will need to be investigated to understand the approximations that will have to be made in derivatives pricing.

Acknowledgements. Coenraad Labuschagne would like to thank the NRF for financial support.

References

1. Alexander, C.: Quantitative Methods in Finance, Market Risk Analysis, vol. 1. John Wiley and Sons, Ltd. (2008)
2. Bianchetti, M., Carlicchi, M.: Interest Rates After The Credit Crunch: Multiple-Curve Vanilla Derivatives and SABR. Quantitative Finance Papers (2012)
3. Cameron, M.: Behind the Curve. Risk Magazine (2011)
4. Erhart, S., Kollarik, A.: The launch of HUFONIA and the related international experience of overnight indexed swap (OIS) markets, http://www.mnb.hu/Root/Dokumentumtar/ENMNB/Kiadvanyok/mnben_mnbszemle/mnben_mnb_bulletin_april_2011/erhart-kollarik_ENG.pdf
5. Hull, J.: Options, Futures and Other Derivatives, 7th edn. Prentice Hall, New Jersey (2009)
6. Hull, J.: Fundamentals of Corporate Finance, 9th edn. McGraw Hill (2010)
7. Johannes, M., Sundaresan, S.: The Impact of Collateralization on Swap Rates. The Journal of Finance 62(1), 383–410 (2007)
8. Piterbarg, V.: Funding Beyond Discounting: collateral agreements and derivatives pricing. Risk Magazine, 97–102 (2010)
9. Piterbarg, V.: Cooking With Collateral. Risk Magazine, 58–63 (2012)
10. Rohan, D., Decrem, P.: Interest-Rate Models: OIS & CSA Discounting. Derivatives Week: Learning Curves (2011)
11. Smith, D.J.: A Teaching Note on Pricing and Valuing Interest Rate Swaps Using LIBOR and OIS Discounting. SSRN Working Paper (2012)
12. West, G.: South African Financial Markets. Financial Modelling Agency (September 2009)

Systemic Knowledge Synthesis for Product Recommendation

Yoshiteru Nakamori

Abstract. This paper considers the problem of systemic knowledge synthesis for product recommendation based on the theory of knowledge construction systems. This theory suggests actors to collect knowledge from scientific, social, and creative dimensions and to synthesize them systemically. It is believed that the pursuit of systematic, or mathematical approach in the scientific dimension is the role of a researcher. This paper mainly introduces mathematical information aggregation techniques for product recommendation, but these techniques usually give only partial answers. Finally, the paper returns to the theory of knowledge synthesis to suggest how to provide a better answer to the problem.

1 Introduction

Human beings have been troubled with complex decision-making problems since the ancient times. Most academic disciplines, including econometrics, have been developed to lighten the burden of decision making, introducing systematic problem-solving techniques. However, in such a situation of economic game, a decisive decision making is often made by the intuition using the experience-based knowledge. A decision is always related to the future events, while models constructed in academic disciplines usually depend on the past data, neglecting many complex matters. Any optimal solution based on any elaborated mathematical model has a possibility that it is born dead, as Checkland [1] stated. Thus, the mathematical model-based decision-making techniques do not always provide perfect answers to those who are suffering from complex issues. This is the reason why this paper claims the importance of knowledge synthesis rather than just information aggregation.

Yoshiteru Nakamori
Japan Advanced Institute of Science and Technology,
1-1 Asahidai, Nomi, Ishikawa 923-1292, Japan
e-mail: nakamori@jaist.ac.jp

V.-N. Huynh et al. (eds.), *Modeling Dependence in Econometrics*,
Advances in Intelligent Systems and Computing 251,
DOI: 10.1007/978-3-319-03395-2_37,

This paper first introduces a theory of knowledge construction systems [2], which consists of three fundamental parts: *a knowledge construction system* (the *i*-System) [3] [4] for collecting distributed knowledge, *a structure-agency-action paradigm* [5] for utilizing actors' abilities, and *evolutionary-constructive objectivism* [6] for justifying synthesized knowledge. The paper then tries to use this theory for product recommendation. This paper mainly introduces mathematical information aggregation techniques for product recommendation, understanding that these techniques give only partial answers. The paper finally returns to the theory of knowledge synthesis to suggest how to provide a better answer to the problem. An example treated in this paper is related to a traditional ceramic craft in Japan, with a hope of revitalizing this culturally important industry.

2 Theory of Knowledge Synthesis

The theory of knowledge construction systems [2] chooses three important dimensions from high-dimensional *Creative Space* [7], and requires actors to work well in each dimension in collecting and organizing distributed, tacit knowledge. These are *Intelligence* (scientific dimension), *Involvement* (social dimension), and *Imagination* (creative dimension). When the theory is interpreted from the viewpoint of sociology, the *Creative Space* is considered as *Social Structure* which constrains and enables human action, and which consists of *a scientific-actual front*, *a social-relational front* and *a cognitive-mental front* corresponding respectively to the three dimensions. The theory introduces two more dimensions: *Intervention* and *Integration*, which correspond to *social action* and *knowledge* from the sociological point of view.

The theory aims at integrating systematic approach and systemic (holistic) thinking; the former is mainly used in the dimensions *Intelligence*, *Involvement*, and *Imagination*, and the latter is required in the dimensions *Intervention* and *Integration*. Leading systems thinkers today often emphasize *holistic thinking* [8] [9], or *meta-synthesis* [10]. They recommend and require systems thinking for a holistic understanding of the emergent characteristics of a complex system, and for creating new systemic knowledge about a difficult problem confronted. Our theory aims at synthesizing objective knowledge and subjective knowledge, which inevitably requires intuitive, holistic integration.

These five ontological elements were originally interpreted as nodes, as illustrated in Fig. 1. Because the *i*-System is intended as a synthesis of systemic approaches, *Integration* is, in a sense, its final dimension. In Fig. 1 all arrows converge to *Integration* interpreted as a node; links without arrows denote the possibility of impact in both directions. The beginning node is *Intervention*, where problems or issues perceived by the individual or the group motivate their further inquiry and start the entire creative process. The node *Intelligence* corresponds to various types of knowledge, the node *Involvement* represents social aspects, and the creative aspects are represented mostly in the node *Imagination*.

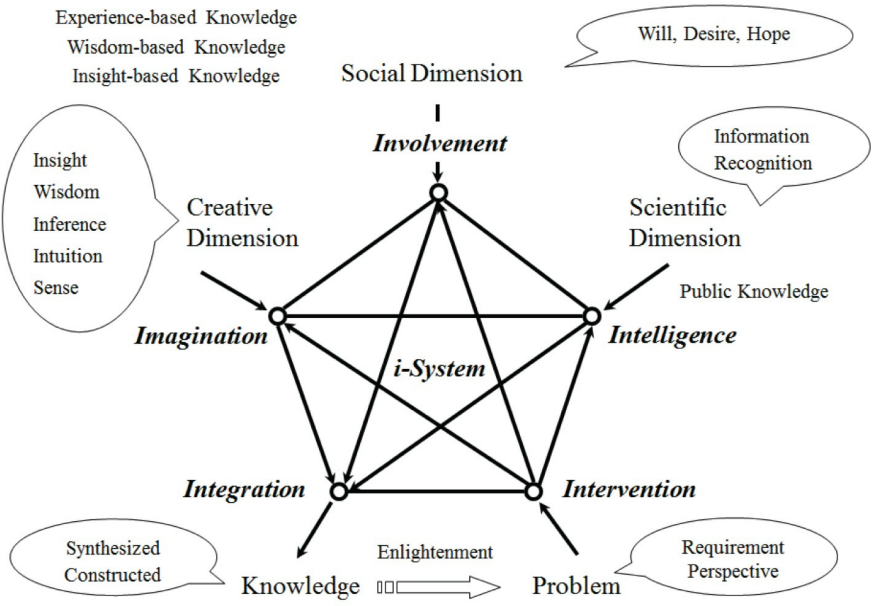


Fig. 1 The *i*-System from a systems scientific viewpoint

3 Product Recommendation

The introduced knowledge construction model can be applied to any social problems in principle. Here, let us consider the following situation. Suppose that a young woman visits a traditional craft dealer and says her hope, "I want a ceramic cup that is cute and smooth, quite modern, largish, but not too much expensive. It should be the most suitable gift for my grandmother." How could you help her?

If you have been involved in the sale of ceramics for a long time, say more than 20 years, you will immediately recommend several ceramic cups that fit her desired attributes. This is clearly an example of *systemic knowledge synthesis*. Figure 2 shows an approach to systemic knowledge synthesis. In an actual situation, the shop owner, synthesizing knowledge from the scientific dimension, the social dimension, and the creative dimension, would recommend some cups to this customer. But in this paper we will study how to create analytical knowledge mainly in the scientific dimension.

This paper is intended to create complete knowledge of *Intelligence* and partial knowledge of *Integration* as in Fig. 2. This will be done by:

- Creating knowledge of the degree of alignment between cups in the store and respective desired attributes such as cute, smooth, largish, etc. (Intelligence);
- Creating knowledge of the overall ranking of cups in the store by aggregating respective degrees of alignment (Integration).

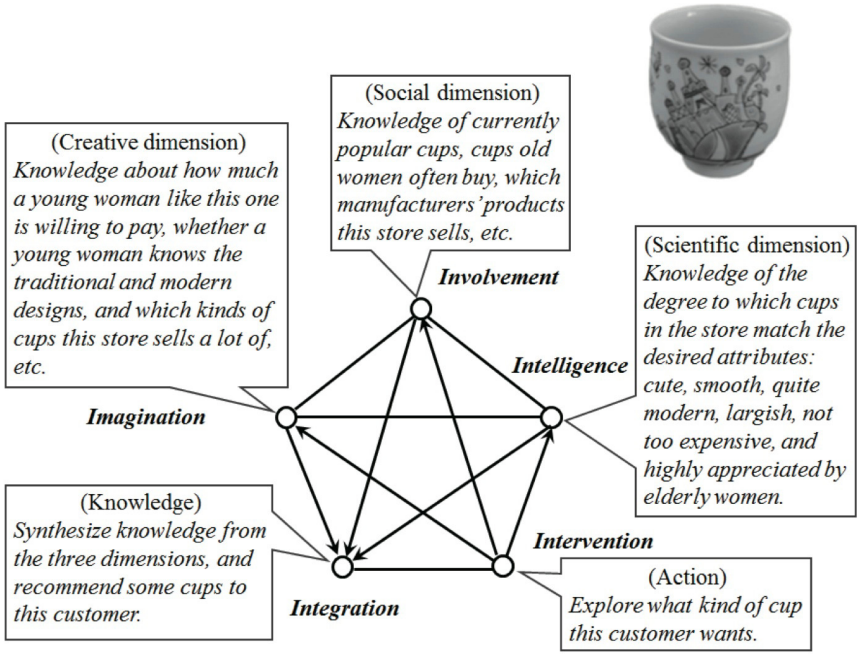


Fig. 2 An approach to systemic knowledge synthesis

Knowledge synthesis at *Integration* is actually done by taking into account knowledge from *Involvement* and *Imagination*. Here, let us consider partial tasks at *Intervention*, *Intelligence*, and *Integration* as shown in Fig. 3.

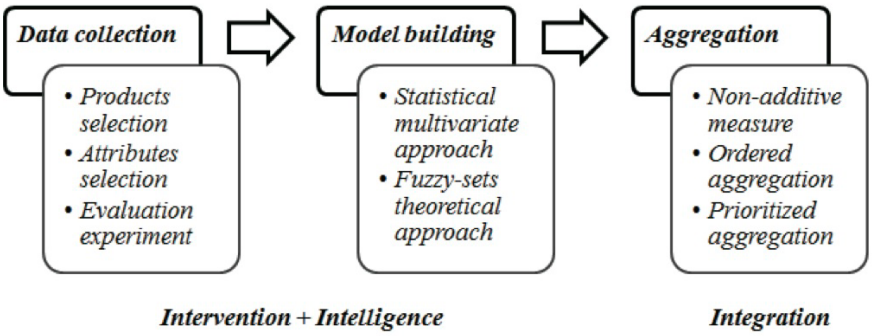


Fig. 3 The task flow for solving the problem

Let us first consider how to collect data and to make a model using the data.

4 Data Collection and Modeling

To build a model of the alignment between desired attributes and products, we need to collect a good data set. The procedure for data collection and data screening are summarized as follows:

1. *Preparation*: The first task is to prepare products for evaluation by the bipolar measures used in the *semantic differential method* [11], which is often used in subjective evaluation experiments. Here, the most difficult task is to choose words, mainly adjectives, to be used in the measure.
2. *Experiment*: The evaluation experiment should be designed carefully. Depending on the products to be evaluated, we have to gather appropriate evaluators, and teach them the purpose of the experiment, how to do the scoring, etc.
3. *Screening*: The data screening is sometimes necessary because of errors or biased scoring. Moreover, those who are familiar with the products and those who do not know them would make different scores in some measures. Therefore, we need to select data as well as appropriate bipolar measures before going into the modeling phase.

The next problem is modeling to calculate the degree of fit between requests and products. There are mainly three approaches to data processing:

1. *Statistical approach*: Among many statistical approaches, factor analysis is mostly used to obtain information about the gaps between objects, between words, and between objects and words. But here, we introduce correspondence analysis, which measures the gaps between objects and words directly. A fuzzy version of correspondence analysis is given in Nakamori and Ryoke [12].
2. *Probabilistic approach*: To treat the degree of the requirement, such as "a little cute" or "quite traditional" without making any models, we can use the frequencies obtained from the data directly. We can convert the frequencies to probabilities; for instance, the degree of "a little cute" of this cup is given by a certain probability.
3. *Fuzzy-set theoretical approach*: The above approach is acceptable if we have a plenty of data. In usual cases, we develop models that interpolate data distributions. The Gaussian-type probabilistic models are often used. But here, taking into account the possibility of data, we use the triangular fuzzy (possibility) model.

Let o_m denote a sample (ceramic cup) to be evaluated, and w_n denote a measure that is given by a pair of bipolar words:

$$w_n : \langle w_n^-, w_n^+ \rangle, \quad w_n^- : \text{left word}, \quad w_n^+ : \text{right word}.$$

The evaluated value z_{mnk} of the evaluator e_k, regarding the object o_m, from the standpoint of w_n is given by a $(2L+1)$-level value:

$$z_{mnk} \in \{-L, \cdots, 0, \cdots, L\}, \quad L : \text{ a positive integer}.$$

An example of bipolar words is $\langle smooth, rough \rangle$, and a scale of seven grades is given as follows:

{very smooth, smooth, a little smooth, neutral, a little rough, rough, very rough}

4.1 Correspondence Analysis

To consider the possibility of missing values, we denote the set of evaluators who evaluate the object o_m with the measure of evaluation w_n by E_{mn}. Assuming that every E_{mn} is not empty, we put[1]

$$z_{mn} = \frac{1}{|E_{mn}|} \sum_{k \in E_{mn}} z_{mnk}.$$

We define the average data matrix:

$$Z = \begin{pmatrix} z_{11} & z_{12} & \cdots & z_{1N} \\ z_{21} & z_{22} & \cdots & z_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ z_{M1} & z_{M2} & \cdots & z_{MN} \end{pmatrix}. \tag{1}$$

We normalize the average data $\{z_{mn}\}$ in (1) as follows. If we focus on the words to the right, let

$$z'_{mn} = (L+1) + z_{mn} \in [1, 2L+1]. \tag{2}$$

On the contrary, if we focus on the words to the left, let

$$z'_{mn} = (L+1) - z_{mn} \in [1, 2L+1]. \tag{3}$$

Then, letting

$$p_{mn} = \frac{z'_{mn}}{z_T}, \quad z_T = \sum_{m=1}^{M} \sum_{n=1}^{N} z'_{mn}, \tag{4}$$

we define a correlation matrix:

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1N} \\ p_{21} & p_{22} & \cdots & p_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ p_{M1} & p_{M2} & \cdots & p_{MN} \end{pmatrix}. \tag{5}$$

Using this matrix we shall consider the handling of data using correspondence analysis [13].

[1] $|E_{mn}|$ indicates the number of elements in the set E_{mn}.

Correspondence analysis is reduced to an eigenvalue problem. It is known that the eigenvector corresponding to the maximum eigenvalue is meaningless. Let $\tilde{x}_i$ and $\tilde{y}_i$, which are derived from the eigenvectors corresponding to the second and third largest eigenvalues, be

$$\tilde{x}_i = (\tilde{x}_{i1}, \tilde{x}_{i2}, \cdots, \tilde{x}_{iM})^t, \quad \tilde{y}_i = (\tilde{y}_{i1}, \tilde{y}_{i2}, \cdots, \tilde{y}_{iN})^t, \quad i = 2,3. \tag{6}$$

From this we plot

$$(\tilde{x}_{2m}, \tilde{x}_{3m}), \quad m = 1,2,\cdots,M, \tag{7}$$

$$(\tilde{y}_{2n}, \tilde{y}_{3n}), \quad n = 1,2,\cdots,N, \tag{8}$$

on a plane, and try to understand the relationships between objects and words.

The gap d_{mn} between the object o_m and the measure w_n can be calculated by

$$d_{mn}^2 = (\tilde{x}_{2m} - \tilde{y}_{2n})^2 + (\tilde{x}_{3m} - \tilde{y}_{3n})^2. \tag{9}$$

Then we can define the alignment s_{mn} between the object o_m and the word w_n by

$$s_{mn} = \exp\{-d_{mn}\}. \tag{10}$$

Assume that the alignment or match is calculated by a statistical method using the words to the right $w_n^+, n = 1,2,\cdots,N$. Let $S[w^+]$ be the alignment matrix given by

$$S[w^+] = \begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1N} \\ s_{21} & s_{22} & \cdots & s_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ s_{M1} & s_{M2} & \cdots & s_{MN} \end{pmatrix}. \tag{11}$$

The alignment of the object o_m with the words to the left can be defined by $1 - s_{mn}$. From this we define another alignment matrix $S[w^-]$:

$$S[w^-] = \begin{pmatrix} 1-s_{11} & 1-s_{12} & \cdots & 1-s_{1N} \\ 1-s_{21} & 1-s_{22} & \cdots & 1-s_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ 1-s_{M1} & 1-s_{M2} & \cdots & 1-s_{MN} \end{pmatrix}. \tag{12}$$

4.2 Fuzzy-Set Theoretical Data Processing

The method presented above cannot treat a word with a quantifier, for instance, "a little cute." For this purpose, we consider a modeling method for alignment between objects and words using the fuzzy-sets theory [14].

Let us define the frequency matrices:

$$Y_n = \begin{pmatrix} y_{1n(-L)} & y_{1n(-L+1)} & \cdots & y_{1nL} \\ y_{2n(-L)} & y_{2n(-L+1)} & \cdots & y_{2nL} \\ \vdots & \vdots & \ddots & \vdots \\ y_{Mn(-L)} & y_{Mn(-L+1)} & \cdots & y_{MnL} \end{pmatrix}, \quad n = 1, 2, \cdots, N. \tag{13}$$

The elements of Y_n are given by

$$y_{mnl} = |\{e_k \in E_{mn}; z_{mnk} = l\}|, \quad l \in \{-L, \cdots, 0, \cdots, L\}. \tag{14}$$

That is, y_{mnl} is the number of evaluators who gave the level l to the object o_m from the standpoint w_n.

From (13) we calculate

$$\bar{y}_{mn} = \frac{\sum_{l=-L}^{L} (y_{mnl} \times l)}{\sum_{l=-L}^{L} y_{mnl}}, \tag{15}$$

$$\sigma_{mn}^2 = \frac{\sum_{l=-L}^{L} \left\{ y_{mnl} \times (l - \bar{y}_{mn})^2 \right\}}{\sum_{l=-L}^{L} y_{mnl}}. \tag{16}$$

We need to define a membership function that represents the degree of alignment of the object o_m and the measure w_n : $\langle w_n^-, w_n^+ \rangle$ as follows:

$$\mu_{mn}(y) = \begin{cases} \frac{1}{c\sigma_{mn}} \{y - (\bar{y}_{mn} - c\sigma_{mn})\}, & y \le \bar{y}_{mn}; \\ -\frac{1}{c\sigma_{mn}} \{y - (\bar{y}_{mn} + c\sigma_{mn})\}, & y \ge \bar{y}_{mn}. \end{cases} \tag{17}$$

Here, c (> 1) is a tuning parameter.

Now we define membership functions $\{\mu_l(y) \mid l \in \{-L, \cdots, 0, \cdots, L\}\}$, each of which represents the l-level fuzzy set.

- For $l = -L$,

$$\mu_l(y) = \begin{cases} -y + 1 + l, & l \le y \le l + 1; \\ 0, & \text{otherwise.} \end{cases} \tag{18}$$

- For $-L < l < L$,

$$\mu_l(y) = \begin{cases} y+1-l, & l-1 \le y \le l; \\ -y+1+l, & l \le y \le l+1; \\ 0, & \text{otherwise.} \end{cases} \tag{19}$$

- For $l = L$,

$$\mu_l(y) = \begin{cases} y+1-l, & l-1 \le y \le l; \\ 0, & \text{otherwise.} \end{cases} \tag{20}$$

Now we can define the alignment or match between the object o_m and the measure w_n with degrees $l \in \{-L, \cdots, 0, \cdots, L\}$:

$$s_{mnl} = \max_y \min \{\mu_{mn}(y), \mu_l(y)\}.$$

Then, we have the alignment matrices:

$$S_n = \begin{pmatrix} s_{1n(-L)} & s_{1n(-L+1)} & \cdots & s_{1nL} \\ s_{2n(-L)} & s_{2n(-L+1)} & \cdots & s_{2nL} \\ \vdots & \vdots & \ddots & \vdots \\ s_{Mn(-L)} & s_{Mn(-L+1)} & \cdots & s_{MnL} \end{pmatrix}, \quad n = 1, 2, \cdots, N. \tag{21}$$

5 Information Aggregation

For information aggregation we can use the Choquet integral with non-additive measure [15] to cope with such a case where a customer expresses, "I want a ceramic cup, which is cute and modern, but also cheap. Cheap is most important. But it is best if it is also cute and modern." Or, we can use the ordered weighted averaging aggregation [16] to treat the case where a customer expresses, "I want a ceramic cup, which is cute and modern, but also cheap. The cup should meet as many of my desired attributes as possible." In theses techniques, we use the alignment matrices given in (11) and (12).

This paper briefly introduce the prioritized max-min aggregation [17] to deal with the case that a customer expresses, "I want a ceramic cup, which is cheap, a little cute, and quite modern. But, modern is the most important, and cute is of secondary importance." Here we use the alignment matrices given in (21).

Suppose that a customer's requirements are given by a set of words:

$$W = \left\{w'_1, w'_2, \cdots, w'_J\right\} \subset \left\{w_n^- \text{ or } w_n^+; n = 1, 2, \cdots, N\right\},$$

and suppose that we have the alignment matrices given in (21). Each $w'_j \in W$ has a level l and a priority p:

$$l \in \{-L, \cdots, 0, \cdots, L\}, \quad p \in \{1, 2, \cdots, P = \text{highest}\}.$$

We write the degree of fit between the object o_m and this requirement by $s_{mj(l)(p)}$. According to the priority, we use the following transformation:

$$g_p(x) = \begin{cases} \dfrac{2P-p}{p}x, & 0 \le x \le \dfrac{p}{2P}; \\ \dfrac{p}{2P-p}(x-1)+1, & \dfrac{p}{2P} \le x \le 1. \end{cases}$$

Using this we have

$$E_{mj(l)} = g_p\left(s_{mj(l)(p)}\right).$$

The comprehensive evaluation of the product o_m is then given by

$$CE(m) = \min_{j(l)} \left\{E_{mj(l)}\right\}.$$

Finally, the most recommended product o_{m^*} is given by

$$m^* = \arg\max_{m} \left\{CE(m)\right\}.$$

Let us consider an example to compare two products, where the attributes have levels and priorities as shown in Table 1.

Table 1 An example of evaluation of two objects

	Cheap (w_1') ($l=-2$) ($p=1$)	A little cute (w_2') ($l=1$) ($p=2$)	Quite modern (w_3') ($l=3$) ($p=3$)
o_1	$s_{11(-2)(1)} = 0.5$	$s_{12(1)(2)} = 0.9$	$s_{13(3)(3)} = 0.7$
o_2	$s_{21(-2)(1)} = 0.5$	$s_{22(1)(2)} = 0.6$	$s_{23(3)(3)} = 1.0$

Note that $p=3$ means *most important* in this case. The transformation functions are:

$$g_1(x) = \begin{cases} 5x, & 0 \le x \le \frac{1}{6}; \\ \frac{1}{5}x + \frac{4}{5}, & \frac{1}{6} \le x \le 1. \end{cases}$$

$$g_2(x) = \begin{cases} 2x, & 0 \le x \le \frac{1}{3}; \\ \frac{1}{2}x + \frac{1}{2}, & \frac{1}{3} \le x \le 1. \end{cases}$$

$$g_3(x) = x,\ 0 \le x \le 1.$$

From the above we have

$$CE(1) = E_{13(3)} = 0.7, \quad CE(2) = E_{22(1)} = 0.8,$$

$$m^* = \arg\max_{m}\{CE(m); m = 1,2\} = 2.$$

Thus, it is appropriate to recommend product o_2 in this example.

Figure 4 shows the transformations. Because the priority of *cheap* is the lowest, its evaluation values are transformed to larger values.[2] As a result, *cheap* is no longer involved in the decision.

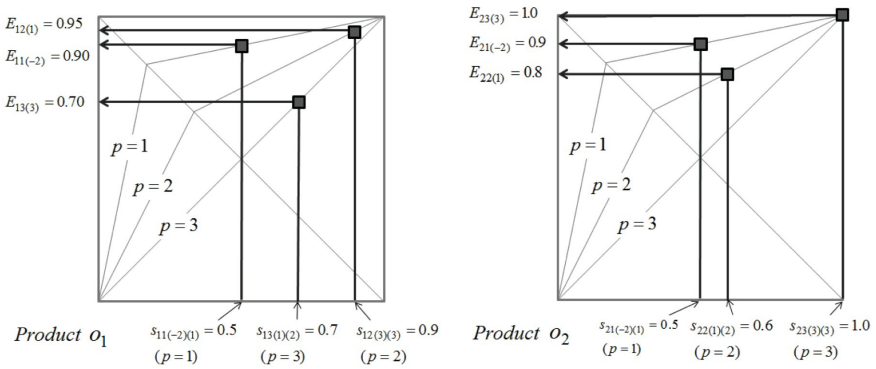


Fig. 4 Transformations according to priorities of desires

6 Knowledge Integration

This section introduces a government project[3] to develop a recommendation system to provide products to customers by recognizing their desires. It aims to support sales expansion and new product development in traditional crafts in Ishikawa Prefecture, Japan. In order to do this, the project is developing a technique for selecting and providing information according to an individual person's desires. This will be done by creating a search engine and an information aggregation system. See Fig. 5.

This recommendation system has already been installed on several websites of arts and crafts shops. If a user inputs his/her desired attributes, then the system will recommend several products, but the system prepares the bipolar scales in advance, which are different for respective stores. To deal with a large number of products, the present system uses the direct modeling approach by the shop owner, and it selectively uses the ordered weighted averaging operators and the prioritized max-min operators to aggregate information.

The method described above gives a partial answer to the subject of product recommendation by knowledge integration. Actually, this corresponds to *Intelligence* in Fig. 6, which shows an example of specified knowledge integration.

[2] This idea was suggested by Marek Makowski, International Institute for Applied Systems Analysis, Austria.

[3] This study was supported by SCOPE 102305001 of the Ministry of Internal Affairs and Communications, Japan.

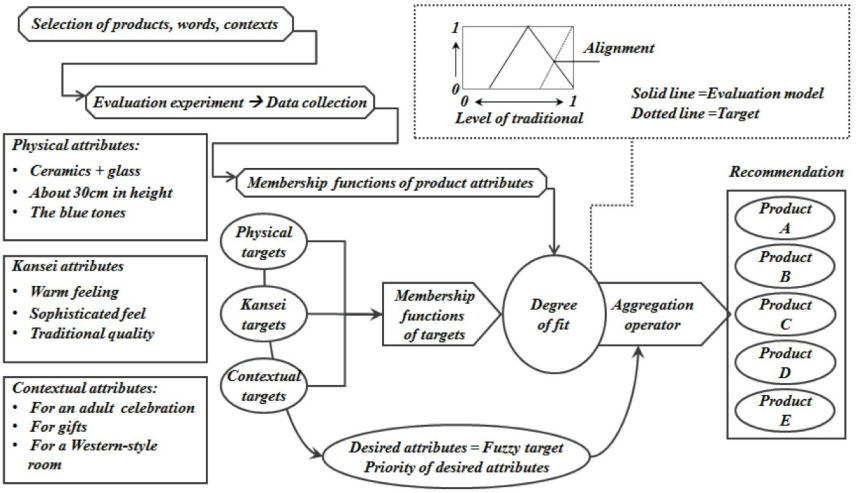


Fig. 5 Implemented functions of the recommendation system

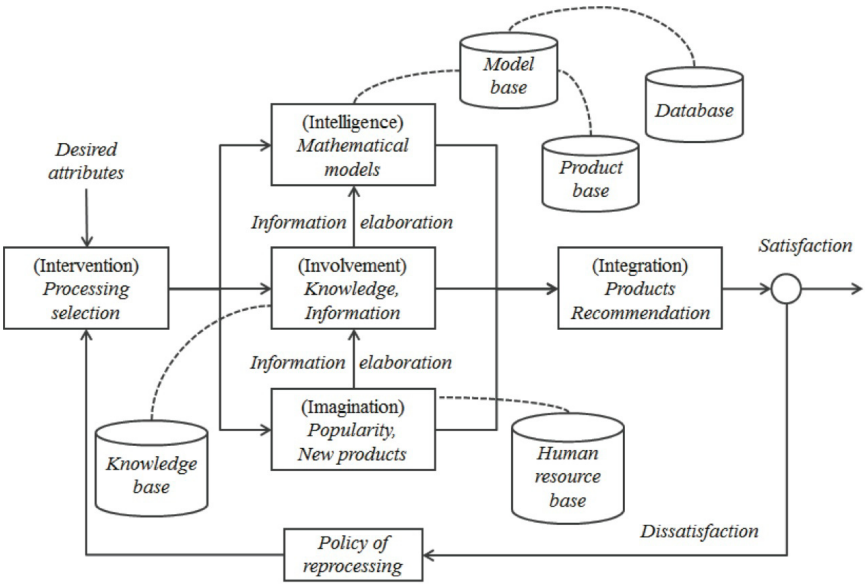


Fig. 6 A designed framework of knowledge integration

- *Intervention*: The user inputs his/her desired attributes, and if necessary selects a method of information aggregation.
- *Intelligence*: The computer system recommends several products that might fit with the desired attributes, based on the preinstalled aggregation methods and product information.

- *Involvement*: We can use the idea of widely deployed recommendation systems; examples given in Wikipedia [18] are:
  - "When viewing a product on Amazon.com, the store will recommend additional items based on a matrix of what other shoppers bought along with the currently selected item."
  - "Pandora Radio takes an initial input of a song or musician and plays music with similar characteristics (based on a series of keywords attributed to the inputted artist or piece of music). The genre stations created by Pandora can be refined through user feedback (emphasizing or de-emphasizing certain characteristics)."
  - "Netflix offers predictions of movies that a user might like to watch based on the user's previous ratings and watching habits (as compared to the behavior of other users), also taking into account the characteristics of the film (such as the genre)."
- *Imagination*: The information used in *Intelligence* and *Involvement* are based on past data. However, producers have ideas about current and future fashion trends, and also know their products. We can use such knowledge in *Imagination*.
- *Integration*: If we could install all the necessary information in a computer, we could recommend some products by using a certain integration rule. But usually the integrator, the shop owner, might have information from *Imagination* as a result of direct communication with the producers. So, we could invent an integration rule by taking into account information that might be regarded as tacit knowledge.

7 Conclusion

This paper considered issues of knowledge synthesis for product recommendation based on the theory of knowledge construction systems, which suggests actors to collect knowledge from scientific, social, and creative dimensions and to synthesize them systemically. This paper mainly introduced mathematical information aggregation techniques for product recommendation, but these techniques gives only partial answers. The paper finally returned to the theory of knowledge synthesis to suggest how to provide a better answer to the problem.

Developing knowledge synthesis methodologies, methods, and tools is most important for knowledge science, using a variety of knowledge and media. When we face a complex problem, based on our experience-based knowledge, we make a plan to collect knowledge from the three dimensions and then synthesize the collected knowledge to obtain knowledge for problem solving. Here, it is important to acquire a *systemic* view through *trained intuition*, and using methods of justifying new knowledge without simply relying on the scientific method in the narrow sense.

References

1. Checkland, P.B.: Systems Thinking, Systems Practice. John Wiley & Sons, New York (1981)
2. Nakamori, Y., Wierzbicki, A.P., Zhu, Z.C.: A theory of knowledge construction systems. Systems Research and Behavioral Science 28, 15–39 (2011)
3. Nakamori, Y.: Knowledge management system toward sustainable society. In: Proceedings of the 1st International Symposium on Knowledge and System Sciences, Ishikawa, Japan, September 25-27, pp. 57–64 (2000)
4. Nakamori, Y.: Systems methodology and mathematical models for knowledge management. Journal of Systems Science and Systems Engineering 12(1), 49–72 (2003)
5. Nakamori, Y., Zhu, Z.C.: Exploring a sociologist understanding for the *i*-System. International Journal of Knowledge and Systems Sciences 1(1), 1–8 (2004)
6. Wierzbicki, A.P., Nakamori, Y.: The importance of multimedia principle and emergence principle for the knowledge civilization age. Journal of Systems Science and Systems Engineering 17(3), 297–318 (2008)
7. Wierzbicki, A.P., Nakamori, Y.: Creative Space: Models of Creative Processes for the Knowledge Civilization Age. Springer, Berlin (2006)
8. Jackson, M.C.: Systems Thinking: Creative Holism for Managers. John Wiley & Sons, Chichester (2003)
9. Mulej, M.: Systems theory: A world view and/or a methodology aimed at requisite holism/realism of human's thinking, decisions and action. Systems Research and Behavioral Science 24(3), 347–357 (2007)
10. Gu, J.F., Tang, X.J.: Meta-synthesis approach to complex system modeling. European Journal of Operational Research 166(3), 597–614 (2005)
11. Osgood, C.E., Suci, G.J., Tannenbaum, P.H.: The Measurement of Meaning. University of Illinois Press, Urbana (1957)
12. Nakamori, Y., Ryoke, M.: Treating fuzziness in subjective evaluation data. Information Sciences 176, 3610–3644 (2006)
13. Hirschfeld, H.O.: A connection between correlation and contingency. Mathematical Proceedings of the Cambridge Philosophical Society 31, 520–524 (1935)
14. Zadeh, L.A.: Fuzzy sets. Information and Control 8(3), 338–353 (1965)
15. Choquet, G.: Theory of capacities. Annales de L'institut Fourier 5, 131–295 (1954)
16. Yager, R.R.: On ordered weighted averaging aggregation operators in multi-criteria decision making. IEEE Transactions on Systems, Man, and Cybernetics 18(1), 183–190 (1988)
17. Nakamori, Y.: Kansei information transfer technology. In: Proceedings of International Symposium on Integrated Uncertainty in Knowledge Modeling and Decision Making, Hangzhou, China, October 28-30, pp. 209–218 (2011)
18. Recommender system (June 22, 2013), `http://en.wikipedia.org/wiki/Recommender_system`

Author Index