Dynamic Modeling and Econometrics in Economics and Finance 24

Fredj Jawadi *Editor*

Uncertainty, Expectations and Asset Price Dynamics

Essays in Honor of Georges Prat

Springer

Dynamic Modeling and Econometrics in Economics and Finance

Volume 24

Editors
Stefan Mittnik
Ludwig Maximilian University of Munich
Germany

Willi Semmler
Bielefeld University
Germany
and
New School for Social Research
New York, USA

More information about this series at http://www.springer.com/series/5859

Fredj Jawadi
Editor

Uncertainty, Expectations and Asset Price Dynamics

Essays in Honor of Georges Prat

Springer

Editor
Fredj Jawadi
Lille University
Lille Cedex, France

ISSN 1566-0419 ISSN 2363-8370 (electronic)
Dynamic Modeling and Econometrics in Economics and Finance
ISBN 978-3-319-98713-2 ISBN 978-3-319-98714-9 (eBook)
https://doi.org/10.1007/978-3-319-98714-9

Library of Congress Control Number: 2018959836

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

An Interview with Georges Prat

Abstract The following interview was conducted by Dr. Fredj Jawadi (the University of Lille, France, Email: fredj.jawadi@univ-lille.fr) in April 2018 at the University of Paris Nanterre and covers Professor Georges Prat's career and his fields of research. The interview is divided into five parts. The first part concerns some of the salient features of Georges Prat's career, describing the background context and some general questions about the motivating factors of his research, and his experience with Maurice Allais in particular. The second part briefly covers his academic activities. The third part presents a succinct discussion of Georges' contributions to the fields of the monetary market and business cycles. The fourth part briefly looks at his current ongoing research on the labour market. The last part concerns his main contributions to the dynamics of financial asset prices. The interview of Georges' work on this last field of research emphasizes the importance of representations of uncertainty and expectations to understand financial asset prices dynamics. We also take the opportunity to examine some key notions of suitable econometric models for handling financial data. We hope that this interview will provide a better understanding not only of Georges Prat's research trajectory, but also of financial asset price dynamics.

Introduction: Overview of Georges Prat's Career and Work

Before moving on to the interview with Professor Georges Prat, I briefly introduce his career and work below.

After obtaining a Master in Economics and a Licence in Sociology from the University of Paris 10, Georges became an economics teacher at a high school in Poissy before working at the *Faculté Libre de Droit, d'Economie et de Gestion de Paris* as an assistant lecturer. In 1976, Georges joined the French Center for Scientific Research (Centre National de la Recherche Scientifique, *CNRS*) as an associate researcher to make a PhD in economics at the *Centre Clément Juglar d'Analyse Monétaire*, which was a department of the *Centre d'Analyse Economique* attached to both the *CNRS* and the *Ecole Nationale Supérieure des Mines de Paris*. He wrote his PhD, entitled "The Stock Market Price Dynamics and the Economic Conjuncture", under the supervision of Professor Maurice Allais (Nobel Prize in Economics in 1988), defending it in 1981 at the University of Paris 10. After being tenured as a Senior Researcher at the CNRS, Georges was promoted to Full Research Professor in 1988. Georges has been an Emeritus Research Professor since 2013.

Georges was Director of the *Institut d'Economie Appliquée et d'Econométrie* (IEAE) at the University of Paris 10 between 1986 and 1989. He also managed the *Centre d'Economie Monétaire Appliquée* (CEMA) that belonged to the IEAE over the period 1989–1993. Georges has supervised many master dissertations and PhD theses on a range of topics such as the construction of weighted monetary aggregates, money demand and supply, international portfolio choice, financial integration, expectation and risk in the Forex, level and term structure of interest rates, heterogeneity of behaviour in stock market, non-linear stock price adjustment towards fundamentals, etc.

Georges has also held several responsibilities, including member of the *National Committee of the CNRS*, member of the *Conseil National de l'Information Statistique*, President of the *Société de Statistique de Paris* and President of the *Association des Chercheurs Economistes* at the CNRS. He also serves as a member of the Scientific Committee of the international network on *Money, Banking and Finance* (GdRE) designed to enhance scientific collaborations between national and international researchers from numerous universities and institutions across Europe. Georges is also a member of the *Maurice Allais Foundation Scientific Committee* (Paris Tech) and a member of the jury for the Maurice Allais Prize in Economics. He also sits on the admissions jury at the business school HEC Paris (Section "Economics, Sociology and History").

Georges Prat's scientific contributions cover a range of topics and fields: money demand and supply and monetary control, psychological time and economic fluctuations and unemployment and wages in France. However, his principal research area aligns with the main topic of this volume: "Expectations, Uncertainty and Asset Price Dynamics".[1] His output related to this topic has won various scientific prizes and distinctions: the Gaëtan Pirou Prize in 1982 for his PhD awarded by the *Chancellerie des Universités de Paris*, the *Claude-Étienne Bourdin* Prize for the best paper published in the *Journal de la Société de Statistique de Paris* in the period 1984–1986 and the *Jacques Rueff* Prize in 1995 from the *National Industry Incentive Society* for his whole work on financial markets. Georges was also recently given the 2018 Albert Nelson Marquis Achievement award for "his hard work and dedication to his profession".

[1] A list of Georges Prat's publications related to this area is provided at the end of this interview.

Georges Prat continues to pursue his research at EconomiX-CNRS at Paris Nanterre University, where he co-supervises a research seminar entitled "Dynamics of Financial Asset Prices" with Remzi Uctum, in the *Asset Management* and *Banking, Money* and *Markets* Master programmes.

Context and Motivations

Fredj: **Q1.** When and why did you decide to conduct research on Economic Science?

Georges: At first, I was interested in studying sociology, but I soon realized that economic matters are a key to understanding social issues. That's why I began studying both economics and sociology. Then, when I began my research during my Master year, I decided to switch to economics.

Fredj: **Q2.** When and why did you decide to join the *French National Center for Scientific Research* (CNRS) and the *Paris Nanterre University*?

Georges: After completing my master's dissertation under the supervision of Professor Maurice Allais at Paris 10 University (today Paris Nanterre University), Professor Allais suggested to me that I apply for a researcher's post at the *CNRS.* There were a lot of applicants, but I got the job. That's how I joined the *Centre d'Analyse Economique*, which was attached to the *Ecole des Mines de Paris* and the *CNRS.* This two-year contract was renewed two times and allowed me to fund my PhD. I realize that I was very lucky to be in such great conditions to do my PhD.

Fredj: **Q3.** When did you first meet Professor Maurice Allais?

Georges: I met Professor Allais for the first time when he gave a seminar entitled "Economics as a Science" at the University of Paris 10 in 1971. Using different empirical charts, Professor Allais defended the idea whereby social science involves significant numerical regularities, as in physics. Compared to what I'd learned in my four undergraduate years in economics, I remember that it was a real shock. I was amazed that such phenomena might exist, and it was like a new door to my economic science studies where the word "science" was fully justified and required scientific research. It was a turning date, and I applied to join Professor Allais's research seminar. I got his agreement to attend his seminar, but the selection was rather hard after a long interview and I had to agree to work on a topic of his choice. Next, I prepared my master's dissertation on "Stock price and interest rates" under his supervision, which was the preliminary step before joining the PhD programme, also under his supervision. During the years I worked under his direction, Professor Allais helped me to develop a taste for research, and he taught me the thoroughness needed. I'm grateful for all those human values I learnt from him. It should also be noted that his wife Ms. Jacqueline Allais, who was a research engineer at the CNRS, played an important role, devotedly acting as a liaison between Professor Allais and his students as well as teaching complementary courses in mathematics and statistics.

Fredj: **Q4.** What's the main memory of your work under the supervision of Professor Maurice Allais?

Georges: It was the preparation of my PhD, which concerned various countries and focused on the relationship between stock prices and different economic indicators such as interest rates, money supply, inflation, production, consumer sentiment, etc., as well as profits and dividends, of course. My PhD added up to a total of three volumes and 1000 pages! I always presented my results and discussed them at the Allais' research seminar with other PhD students (there were six PhD students). The criticisms were without concession. Sometimes, Professor Allais would call me at 5 am to tell me: "*Prat, what you said on page X is absurd and you have to delete it*". I confess that I held on to and valiantly defended some of my ideas, as I didn't always follow his recommendations. This led to some frictions, but I have never felt a grudge, neither on his side nor mine.

Teaching, Research Supervision and Collaboration

Fredj: **Q5.** What were your favourite courses and classes?

Georges: I was particularly interested in two classes. First, a course called "*Economic Dynamics*" taught by Professor Gilbert Abhraham-Frois at the University of Paris 10, where I learnt a lot about economic growth models and more generally on the conditions of a dynamic equilibrium regime. Second, a course taught by Professor Allais at the *Ecole Nationale Supérieure des Mines de Paris* on "Equilibrium and Economic Efficiency". In this course, Professor Allais discussed the conditions of maximum economic efficiency, where the "star" was Vilfredo Pareto, but he also spoke a lot about his own work. Both of these courses were quite remarkable and gave me a great grounding that has been used as a mainstay throughout my career.

Fredj: **Q6.** You supervised several PhD theses and master dissertations in economics between 1988 and 2012. How did you select your students? Did you adopt Allais's method to supervise your students?

Georges: For master dissertations, students chose the teachers according to their desired research topics. Contrariwise, I couldn't take all the PhD applicants I received, so apart from particular cases, I only considered applications from students who'd already done a master's dissertation with me. This means that I already knew their capacity to carry out a scientific project, as well as their determination and ability to complete a PhD. For all theses that I supervised, there was a starting pact, namely, a theoretical approach that led to a formal model that can be verified using appropriate data and econometric methods. In this regard, I applied with my doctoral students the same methods that Professor Allais applied to his students!

Fredj: **Q7.** What's your role on the Scientific Committee of the *Maurice Allais Foundation*?

Georges: The *Maurice Allais Foundation* was created thanks to his daughter Christine Allais, with the aim of raising awareness of her father's contribution. Indeed, part of Allais work is not well known, as it was underlined in the book published as early as 1986 by Marcel Boiteux, Thierry de Montbrial and Bertrand Munier entitled "Markets, capital and uncertainty", which includes essays in honor of Maurice Allais. In particular, it is the case for the Allais' hereditary and relativistic theory of money supply and demand, as well his "fundamental equation of monetary dynamics", even though the book published by Eric Barthalon in 2014 and the PhD by Ramzi Klabi defended at Aix University two years ago aimed to show the originality and importance of this monetary theory. It is worth noting that Allais's main work was written in French, so the Foundation wants to translate some of his texts into English, especially his famous book "Economy and Interest". Otherwise, with the creation of the Maurice Allais Prize in Economics, the Foundation wants to encourage research carried out in line with Allais's work and respecting his approach that is based on the confrontation of theoretical hypotheses with the observed data, without conceptual or ideological preconceived ideas due to fashion effects. Like all the members of the Foundation's Scientific Committee, I'm committed to working within these guidelines.

Fredj: **Q8.** What type of work does the Maurice Allais Prize jury select?

Georges: The prize has been awarded every other year since 2013, and I have to say that the prizewinners' work was really excellent and in line with Allais's scientific approach that I recalled earlier. Generally speaking, the work submitted (either papers or books) deals with conditions of economic efficiency and equilibrium, risk behaviour, theory of cycles, monetary theory and policy, international economics and behavioural economics. Obviously, when evaluating the applications, the jury looks for a clear and intrinsic originality of the ideas as well as a rigorous demonstration, and this is true whatever the publication support of the submitted work. For example, the fact that an article is published in the *American Economic Review* will not necessarily impress the jury. The impact of the submitted work on the international scientific community is also taken into account. It's not a necessity to quote Allais among the authors of the bibliography, although it is of course welcome. There's no limit to the number of co-authors for a piece of work submitted, even if the jury prefers a limited number of co-authors. Finally, to apply for this prize, there are some nationality conditions that are set out in the charter of the prize.

Fredj: **Q9.** You've contributed to different areas (financial markets, money and business cycles and, recently, the labour market). How did you select your topics and your co-authors?

Georges: The topics that you mentioned were in the research programme of the *Centre Clément Juglar d'Analyse Monétaire* that I belonged to, which was an extension at the *University of Paris 10* of the *Centre d'Analyse Economique* located at the *Ecole Nationale Supérieure des Mines de Paris*. On the themes of this research centre, and for many years, I wrote a number of papers or books in French as a sole author. I should have written in English earlier; it's something I regret. Afterwards, theories and econometrics methods became more complex, so that to keep efficient I collaborate with other colleagues interested in similar research topics. I didn't really choose my co-authors so to speak; it happened naturally through meeting new colleagues, by discussing research questions with them that had not previously been fully explored in the literature. Obviously, the complementarity of the authors' scientific expertise, agreement on the general approach as well as mutual confidence are important conditions for a successful collaboration. However, due to the human qualities of the co-authors, each collaboration is specific, and this is good thing of course. Finally, what I did with my co-authors, I couldn't have done alone.

Contributions to Monetary Theory and Business Cycles

Fredj: **Q10.** Your first papers on money supply and demand suggested there was a need to review monetary control. Why?

Georges: I started working on these topics at the end of the 1980s. Contrary to the main stream of existing literature on monetary control, I showed that monetary multiplier variability is not enough on its own to decide to give up a policy of quantitative control of money supply based on interest rates. Otherwise, whatever the intervention instrument used, controlling money supply is a necessary but not sufficient condition to pursue an efficient monetary policy. This is because one also need knowledge not only of a stable money demand function (a condition always mentioned by authors), but also of a more accurate money demand function than the usual function. This issue had not previously been studied. Indeed, according to the Allais's "fundamental equation of monetary dynamics", we showed that a small error in measuring the demand for money might involve a substantial forecasting error in the growth rate of the global expenditure. This underscores the difficulty involved in conducting an efficient monetary regulation policy in the short term, excluding of course the effects of central bank announcements for which efficiency, associated with the degree of credibility, is difficult to assess.

Fredj: **Q11.** You suggested introducing uncertainty into money demand in 1988 to make the money demand function more accurate and then to reduce the effects of further measurement errors. Measurement errors are at the centre of the scientific project of the US Society for Economic Measurement launched by Bill Barnett in 2014. How important do you feel this topic/area is now?

Georges: As I just said, the stability and precision of the money demand function are important features that have to be taken into account when conducting monetary policy. For this reason, in line with pioneering authors who focused on the precautionary and speculative grounds, I introduced an indicator of economic uncertainty into the money demand function in order to improve the accuracy of this latter. The originality was to show that the spread between bond yields from different classes of default risk is a significant argument in addition to the two traditional variables that are income and short-term interest rate. Accordingly, I believe that I made a modest contribution to improve the money demand assessment. More generally, I think that any effort aimed at limiting measurement errors in macroeconomic variables is important, as there's often a substantial gap between an economic concept and the statistical indicator used to measure it. Measurement errors in variables can be greater than the residuals of models. For example, how can we define money in a demand for money model: M1? M2? M3? ...? The size of the error made in the choice of an aggregate perhaps exceeds that of the residuals of the demand for money model! With respect to this question, I feel the approach that consists of defining monetary aggregates by weighting assets with regard to their degree of substitutability with the M1 money type, as notably proposed by William Barnett and Maurice Allais, is a very relevant approach. Indeed, each asset has a liquidity characteristic that is more or less important. The representation of inflation expectations can also illustrate the importance of measurement issues. Indeed, even if we assume that inflation is correctly measured, any inflation expectations assessment made in a macroeconomic model will include a more or less important error with respect to the true expectations. These two examples highlight that evaluations of variables involved in models depend on assumptions that are often rather strong, so econometricians may make important measurement errors, which in an extreme case can even invalidate a valid theory!

Fredj: **Q12.** You used Allais's concept of "psychological time" to show further evidence of regularities and similarities in the economic behaviour of some macroeconomic and financial variables (interest rate, inflation, etc.). What's the main lesson from this concept? Why is there this psychological dimension?

Georges: I consider the "psychological time" concept as one of the gems of Allais's ideas, but unfortunately this jewel has been put aside. This concept seems very intuitive since everyone has experienced the feeling that time is longer when his/her life is disrupted than when everything is going well. Anyone will behave differently in these two situations. For example, a day in the 1920s during the German hyperinflation was, from the viewpoint of economic transactions, like a period of six months in a normal situation. What I find most interesting are not necessarily the hypotheses introduced by Allais to switch from the physical time to the psychological time, because even if we can understand their meaning, they of course remain debatable. What I think is most interesting is the very original idea that it is possible to evaluate any variable having a dimension with respect to time in a psychological time frame. The strength of this paradigm is that the switch in the time referential can help us identify some "hidden regularities" in agents' behaviours and could help us to highlight stable relationships between macroeconomic variables in the physical time. For instance, comparing a "normal" period to a hyperinflation one, Allais showed that there's a stable relationship between money supply and global expenditure. This so-called unitary character of his "hereditary and relativistic" theory is very appealing. However, even though his work has been published in English, Allais has not been able to convince economists, apart from a few exceptions. I think that the complexity of the concepts and maybe the arbitrary character of some hypotheses are the main reasons. Anyway, this paradigm is unique.

Fredj: **Q13.** When talking about the formation of inflation expectations in your papers published in 1985, 1988 and 1995, you claimed that expectations are neither rational nor naive. Why? Why should we believe more in a mixed model of traditional processes? How far does this link between the formation of inflation and uncertainty correspond to Friedman and Ball's theory of inflation and inflation uncertainty?

Georges: When I started working on this topic, most macroeconomic studies stipulated the hypothesis of rational expectations, following the thought of John Muth that was endorsed and renovated by Robert Lucas. I acknowledge that this hypothesis is very "aesthetic" and that it helps us to develop precise theories in the general equilibrium paradigm. Thus, when, in line with some related studies, I presented my first results showing that the inflation expectations provided by surveys (householders or experts in economics) are biased and therefore non-rational, I systematically received criticism suggesting that this type of approach is not valid (measurement error, non-representatively of surveys, peso effect, etc.). Today, I think that things have changed!

Indeed, non-rationality of expectations has been acknowledged in many contributions, which can be justified either through the presence of information costs (which makes it economically rational not to expect rationality) or by cognitive bias, as shown in the research thrust of the "behavioural economics" in line with Tversky and Kahneman's work. In this respect, it's interesting to note that, in a paper published in 1984, Muth himself showed that macroeconomic expectations of experts are not rational, but this contribution was not taken up by the literature... It's here important to recall that the concept of temporary equilibrium proposed by John Hicks and remarkably extended by Jean-Michel Grandmont allows us to consider any type of expectation in the definition of equilibrium. But the non-rationality of expectations raises the issue of how expectations are formed and the issue of how to take into account heterogeneity of beliefs in asset prices modeling. When considering the "consensus" representing the average of heterogeneous expectations, we need to define a process that is general enough to take into account all together naive behaviour (expected inflation = observed inflation), fundamentalist behaviour (mean-reverting towards a "normal" value), chartist behaviour (bandwagon), adaptive behaviour (correction of observed forecasting error) as well as the behaviour of some exceptional rational agents. These considerations suggest using a mix of these different behaviours, where the different components are weighted to better explain the dynamics of the consensus. According to this representation, the expected inflation rate appears to depend largely on current and lagged inflation, and the recent literature on the concept of "sticky expectations" is in line with this result. This implies that forecasting errors increase when inflation increases. From this perspective and to answer your last question, this property joins the Freidman-Ball's theory whereby high inflation implies high uncertainty about inflation.

Contributions to the Labour Market

Fredj: **Q14.** Since 2012, your recent ongoing work has deviated towards the labour market. In particular, you focused on the unemployment equilibrium. Why this deviation?

Georges: I got involved in this new research area for different reasons, all of them closely linked. First, at the end of the 1970s, Maurice Allais asked some of his students, including me, to work on the causes of the French unemployment, which led me to explore this topic for the first time. In line with Jacques Rueff, Allais distinguishes "chronic" unemployment, "conjonctural " unemployment and "frictional" unemployment, the first component being close to the modern concept of equilibrium unemployment. Second, there has been my contribution to a conference on Jacques Rueff organized by IPAG Business School in Paris in 2013, where I examined how this author explained unemployment in the UK during the 1920s using his famous concept of "permanent" unemployment. Third, when I met a specialist of the history of wages in France during an admission jury at HEC Paris (namely, Michel-Pierre Chélini who is Professor of contemporary history at the University of Arras), we discussed different topics regarding the labour market. He told me that he had built long macroeconomic series for France and that he was interested in investigating the data using econometric methods, in particular to analyse the causality relationships between wages and prices. I told him that in my opinion, it wasn't suitable to analyse wages and prices without taking unemployment into account, these three variables being linked both theoretically and empirically. Michel-Pierre agreed with me on this. The last reason was of course the persistence of French unemployment over the last thirty years or so: I wished to understand the causes of this enormous waste of resources. In fact, I wanted to explore the "real economy". It's important to point out here the great complementarity between us: Michel-Pierre knows a lot about historical facts and institutional changes since the 1950s (without this information, we would soon be talking nonsense), while my focus is on economic theory and econometrics modelling. Quite honestly, it was very hard for us to choose the theoretical framework at the beginning of this project as the related literature is very varied and highly elaborate, while our knowledge of this area was clearly limited. Although neither of us is an economist specialized in the labour market, this multidisciplinary experience was exciting and mutually rewarding.

Fredj: **Q15.** How does your recent analysis of the French labour market using the WS-PS model help us to better understand the dynamics of unemployment in France?

Georges: We showed that an econometric specification deduced from a simple WS-PS negotiation model could help us to understand the main features of the historical evolution regarding wages and unemployment in France at macroeconomic level. With regard to the related literature, the two main novelties of our approach are, first, the introduction of a degree of global rigidity in the labour market that is time-varying (represented by a stochastic state variable) and, second, the measure of reservation wage (represented by a function of the minimum legal wage). Our specification enabled us to distinguish the "chronic", "conjonctural" and "frictional" components of French unemployment (which is in line with Rueff and Allais's ideas that I mentioned earlier). Interestingly, no component appears negligible with regard to the other two. Our results suggest that the negotiation power of firms dominates that of unions. Moreover, in accordance with the theory, we confirm that change in employment plays a role in the convergence between the wage required by unions and the wage offered by firms, which allows to reach a wage contract (our model helps us to measure this deviation at each date). Overall, our model describes the main time patterns of French wages and unemployment. A detailed EconomiX's working paper in English is now available online.

Contributions to Expectations, Uncertainty and Asset Pricing

Fredj: **Q16.** Your recent work on financial markets stipulates that markets are inefficient, and expectations are irrational and based on limited information. You also confirm that stock prices deviate too much to be justified only by fundamentals. And you explain this inefficiency in terms of the effects of transaction, arbitrage and information costs. At the same time, in 2013, Eugene Fama, considered as the father of informational efficiency (even if Paul Samuelson focused on the efficiency question before Fama), shared the Nobel Prize in Economics with Robert Shiller and Lars Hansen. How do you feel about your conclusions on market efficiency with regard to the work of these authors: are you close to or far from their results?

Georges: It's an important question as the answer always conditions the general thrust of an author's research. Personally, I find it hard to believe that the financial markets are informationally efficient even if they're competitive, because the market can't eliminate irrational agents. That's why I agree with Robert Shiller as he argued there is a major influence of psychological factors in the way financial asset prices are determined. More generally, I agree with the ideas underlying the "Behavioural Finance" notably popularized by Richard Thaler, where decisions are conditioned both by the information usable and by investors' characteristics. The fact that the well-known Grossman–Stiglitz paradox was only solved when information costs are taken into consideration endorsed my opinion that the efficient market hypothesis does not correspond to the reality of financial markets. For me, stock prices are not rationally expected, and in my work, I always found that when the rational expectations assumption is introduced into an asset valuation model, it fails to explain the observed price. Further, concerning the information contained in risk premiums, we often have to add something like the "state of confidence" that Keynes popularized to the variance–covariance matrix. That's why the approaches involving heuristics for the representation of expectations and uncertainty seem relevant to explain market price dynamics. However, whatever the hypotheses used to represent expectations and uncertainty, it's worth noting that the deviation between the market price and the estimated theoretical value is never a white noise, leading to consider an adjustment process of the market price towards its theoretical value. In fact, such a process might be justified by the presence of arbitrage and transaction costs. Of course, approaches supposing the presence of a bubble offer an alternative, although it asks what information is contained in the bubble, if any.

Fredj: **Q17.** Regarding the expectation hypothesis, you feel you are in accordance with James Tobin and trust the data based on Opinion Surveys. Why this choice? And why should we believe that measurement errors would be less important when using such data?

Georges: I think we must distinguish two cases. First, if the aim is to determine how a category of agents (householders, companies, traders, experts, etc.) form their expectations, then data provided by surveys made on these agents seem a priori relevant. Here, the measurement errors regarding agents' expectations are captured by the residuals of the model, so that, if appropriate tests show that the residuals are "clean", then we can conclude that the proposed expectations process corresponds to agents' behaviours when they answer the surveys. It's interesting to note that analyses using individual survey data generally lead to rather similar conclusions as studies based on consensus, suggesting that the aggregation bias seems not very important despite the heterogeneity of beliefs. It's worth noting that we aim here to model "observed" expectations and that we are not trying to forecast at best a variable using a quantitative method. On the other hand, if the consensus regarding the expected price of a financial asset is introduced in a valuation model that aims to explain the asset's market price, we should of course ask whether this consensus might allow us to represent market expectations. That's the difficulty. The nature and number of respondents, especially their proximity to the market, as well as the confidential character of responses are important factors to discuss. Nonetheless, I think that, after all, we can come back to empirical criteria. For example, if the asset price model considered is not verified with the rational expectations hypothesis but is validated using expectations provided by surveys, why can't we admit, at least temporarily, that surveys provide to the econometrician some approximation of market expectations? Is not it better to be in a mist than in an opaque fog? One must add here that the existence of future markets helps investors to not confuse expectations with risk. Overall, even if I don't fully trust opinion surveys, I think that it would be a real pity not to work using such data on the pretext that respondents are not confused with the market.

Fredj: **Q18.** Your empirical work often applies non-linear time series models (switching models, state-space models, threshold models, models with structural breaks, ARCH model, etc.). Why do you apply these methodologies? How do you justify this non-linearity?

Georges: Non-linear econometrics has made very significant progress in the last thirty years thanks to contributions from exceptional scholars, and many applications have been made to analyse price dynamics in financial markets. Further, the propagation of econometrics has been facilitated thanks to progress in IT, either through hardware or software that integrates these new techniques. I'm very happy to have been able to benefit from this progress. Like many other colleagues, I try to apply the appropriate methods to best relate theory to reality. We shouldn't forget here that the linear model is in principle a limited case of the non-linear model, which implies that the application of non-linear models might lead us to conclude that linearity is relevant to describe reality. But in this case, it would be a result and not an a priori assumption, so that the non-linear model keeps all of its power. The economic relevance of non-linear models can be justified for many economic reasons : (1) to capture structural breaks in the behaviours due to institutional changes or crises, (2) to take agent heterogeneity into account, (3) to respect the variability of the variance–covariance matrix, (4) to take into account the fact that arbitrage and transaction costs imply the presence of thresholds that point to a change in behaviours (e.g. the importance of the deviation between the price and the fundamental value), (5) to detect bubbles through their explosive character, (6) to represent a latent and non-observable variable by a stochastic process (e.g. the degree of rigidity in the labour market), etc. We can note here that the Allais' paradigm according to which the objective is to seek permanent relationships over time and space only appears realistic if the model in itself allows for a change in the agents' behaviour, which Allais proposed with the non-linear variability of the ratio between psychological time and physical time.

Fredj: **Q19.** Your main work on expectations and uncertainty concern different markets (stock markets, foreign exchange market, debt markets and oil market). What are the shared characteristics of these markets with regard to uncertainty and expectations? Does price formation differ on these markets?

Georges: Generally speaking, one can say that the value of an asset is equal to the expected receipts (price plus revenue) corrected by a discount factor including an impatience rate and a risk premium. Consequently, leaving aside revenue that may be zero, constant or increasing at a given rate, if we know how the expected price and the risk premium are formed, we can understand how the market price is formed. Regarding price expectations in the four markets that you mentioned, my work showed that, while exploring the responses from expert's opinion surveys, first, expectations are not rational while we don't observe a learning process towards rationality, and second, there's an important common core that might help to explain the formation of expected prices. This common core is represented by processes mainly based on the lagged values of observed and expected prices, to which one can add the forward price if any. Mixing traditional processes (adaptative, extrapolative, regressive, forward component) that capture behavioural heterogeneity allows us to obtain rather good representations of expectations, even if each market has its own specificities (for example, contrary to other markets, due to the direct effect of Fed announcements, we find a small but significant rational component for short-term interest rate expectations). Considering this mix, we often find that the weights of components are time-varying, which suggests that the importance of the different forecasting groups is unstable or/and that agents mix the process at the individual level conditional to the state of the nature. Both the mix of processes and the variability of weights can be understood with regard to the theory of economically rational expectations of Feige and Pearce, who argued that the optimal quantity of information selected by a forecaster is such that the marginal benefit associated with a decrease in forecasting error equals the unitary information cost. Note that, according to this paradigm, the rational expectations hypothesis holds as a limit case when information costs are null. Regarding the risk premium modeling, our results show that pricing models are validated with experts' expectations but invalidated with the rational expectation hypothesis, whether for stock prices, exchange rates or interest rates. These results suggest that surveys provide useful information on the agents' beliefs about the upcoming asset price. This also suggests that modelling must distinguish between the "cognitive" rationality of expectations and the "operative" rationality of choices involved in market making (e.g. absence of arbitrage opportunity, intertemporal choices consistency). For example, the Euler equation applied to the S&P index was roughly validated using expectations from surveys, but strongly rejected under the rational expectations hypothesis. Such a result suggests that one can admit operative rationality but not cognitive rationality. Regarding individual equity prices, we have coupled the "Dividend Discount Model" (in which we consider a simple hypothesis for expected dividends growth) with the "Arbitrage Pricing Theory" that enables us to identify common factors of the long-term risk premiums (interest rate spreads, oil price, exchange rate, consumer sentiment, etc.): here again, our results are in accordance with an operative rationality although not with a cognitive rationality. However, in every case, whatever the hypothesis to represent expectations and whether we consider stock price indices or individual equity prices, we found that stock prices adjustments towards fundamental values are gradual. Overall, our results suggest that the allocative efficiency hypothesis seems rather acceptable while the informational efficiency hypothesis seems not in accordance with the reality of markets.

Fredj: **Q20.** Since the aftermath of the recent global financial crisis, the oil market has experienced considerable volatility. How does your model of oil price expectations in Prat and Uctum (2011) take this excess oil price volatility into account?

Georges: Our analysis in this paper concerns data up to the end of 2008 so we don't have enough hindsight to isolate the effects of the recent global financial crisis. However, we observed a high increase in oil price in 2007 and at the beginning of 2008, followed by a strong decrease at the end of 2008 that cancelled out the previous oil price increases: a volatility shock is clearly observed during the crisis. As for the expected change in oil price by experts (which is our endogenous variable), it appeared strongly negative in 2007 and at the beginning of 2008 but positive in late 2008, while the mixed expected process we proposed correctly reproduced these dynamics. According to our model, such are mainly explained by the presence of a negative mean-reverting component in 2007 which became positive in late 2008 as oil price dropped below its target value represented by the marginal cost of production in the mean-reverting component. Thus, we only can say that our model didn't lose its ability to explain experts' oil price expectations during the beginning of the crisis.

Fredj: **Q21.** Your research conducted on the determinants of stock market risk premium stipulates the presence of a risk premium term structure. Can this conclusion be extended to the oil market? How?

Georges: In principle, I guess yes. If we look at oil as a financial asset, we can consider that, as for stock prices, the expected oil return equals the free risk rate plus a risk premium that depends on return volatility, risk aversion and possibly some economic factors. In such a simple framework, unless the oil market is efficient (this would be the case if oil return was totally unpredictable), one can show that expected returns and variance depend both on time and investment horizon. This means that the premium is time-varying and depends on the horizon. For example, if the oil return is positively autocorrelated, it can be shown that variance and risk premiums increase with the term of the investment.

Fredj: **Q22.** What do you think are some of the most important and pressing issues in the field of financial markets? What advice would you give to a PhD student interested in finance? Where would you advise them to focus their research?

Georges: As well as determining the price of financial assets, we have to remember that financial markets are essential because they help to smooth consumption over time, which increases the satisfaction of economic agents over the course of their lives. These markets also mobilize savings and make them liquid, channelling them, in principle, towards the more useful purposes while reducing transaction costs. If I had to advise a PhD student, I would suggest that they keep in mind these essential functions, which fully justify the research work in this field. In particular, it is important not to lose sight of the link between the financial and the real spheres, because if we lose sight of this, the crisis will come back with all of its harmful effects. In my opinion, the global financial crisis was mainly due to a disconnect between these two spheres. About that, if automatic trading, liquidity and default in markets regulation are responsible, theoretical models that regard financial markets as only depending on themselves and forgetting fundamentals have also some responsibility. I would also tell a PhD student that Robert Lucas' superb model of general equilibrium based on the rational expectation hypothesis and the expected utility maximization of the representative agent satisfies the link between the financial and the real world, but with heroic assumptions about the reality of behaviours. I would say that this model is useful as it gives a rigorous view of the intertemporal choices, but it can't reproduce reality as shown by several studies that pointed to different "puzzles", suggesting strong deviations between the predictions from this model and the reality. This means that researchers should be clear-headed about the effective relevance of any model, because even if a model is very appealing, it may be based on bold hypotheses which are often accepted because they have internal consistency, which is a necessary but insufficient condition of validity. Researchers must strive to assess the degree of human realism of hypotheses on which any model is based and, as far as possible, must not hesitate to question them to try to bring the model closer to reality, of course without losing its consistency, and this is the main difficulty. In this regard, I think a good starting point is to accept the idea that it's not possible to forecast financial asset prices without committing important errors, which implies that we need to introduce biased expectations in models to understand the dynamics of observed prices. In such a context, the real challenge is to identify the types of relevant bias, and attempts to identify the effective expectations processes contribute to this goal. For this purpose, one must not hesitate to separate the operative rationality and the cognitive rationality which are all too often joint hypotheses. Indeed, operative rationality is not such a strong hypothesis as the cognitive rationality one. The first type of rationality involves assumptions such as the absence of arbitrage opportunities in the markets, the construction of an optimal portfolio as well as the coherence of intertemporal choices. These hypotheses can be accepted as they appear rather realistic with respect to the human abilities and the abilities of markets. Considering the cognitive rationality, it supposes that agents have an ability that the modeller does not have, so that the researcher has to distance himself from the rational expectation hypothesis and must look for representations of expectations that best validate the asset price model considered. Obviously, such an hybrid approach has an impact on the type of equilibrium to be considered as it leads the researcher to the temporary equilibrium paradigm rather than to the general equilibrium one. Anyway, the separation between these two types of rationality is a general sense that I give to my contributions. This perspective reinforces the relevance to analyse the various forms of heterogeneity characterizing the behaviour of individuals in the market, such as adopting a hedging behaviour or a risky behaviour, imitating others or following fundamentals, extrapolating the past dynamics or adopting a mean-reverting behaviour, etc. The characterization of heterogeneities is an infinite and passionate field of research. The analysis of contagion between markets and their degree of integration is also an important research area because these phenomena condition the importance of systemic risk. I haven't directly worked on these topics although I did supervise some PhD students tackling them. Finally, whether it's to identify asset fundamentals, characterize behaviours in the market or explore market interdependencies, I would certainly advise any PhD student to take the time required to draw adequacy between the data and the theoretical concepts and to well analyse the statistical properties of the data. This is necessary to determine the relevant specifications of the relationships and will help any researcher to justify the model and method used.

Fredj: **Q23**. What do you think of the Springer volume edited in your honour? Is it representative of your main research focus on Expectations, Uncertainty and Asset Prices?

Georges: This book is of course a great honour for me. First, because the co-editors thought about me when developing the project, but I was also pleasantly surprised to see how quickly this project got a number of contribution proposals from famous authors. Unfortunately, due to the book's limited size, we had to turn down several other proposals. The book's contents fully conform with my main line of research. The issues dealt with on "heterogeneity of beliefs", "uncertainty and volatility" and "fundamentals and bubbles" are closely related to my research concerns, while the topic of "transmission and integration" is in line with my exploration when supervising some of my PhD students. Moreover, all the papers in the book are interested in relevant issues and use quantitative methods that are appropriate to the objectives pursued by the authors, and this links up with what, for me, is a constant concern. Any one of the chapters could have been published in an international review, and I'm all the more touched that the authors chose this book to publish their work in. I sincerely thank the co-editors and all of the authors, and I'm really happy to see a good balance between French and foreign colleagues of all generations.

Publications of Georges Prat on "Expectations, Uncertainty and Asset Price Dynamics": Stock Market, Foreign Exchange Market, Debt Market, Oil Market

- "Term structure of interest rates: modelling the risk premium using a two horizons framework" [2018], *submitted paper* (*with R. Uctum*)
- "Do markets learn to rationally expect US interest rates? Evidence from survey data" [2018], *Applied Economics*, forthcoming (*with R. Uctum*)
- "Persistence of announcement effects on the intraday volatility of stock returns: evidence from individual data" [2017], *Review of Financial Economics,* 35, 43–56 (*with S. Lecarpentier, P. Renou, R. Uctum*)
- "Equity prices and fundamentals: a DDM-APT mixed approach" [2017], *Journal of Quantitative Finance and Accounting,* 49 (3), 661–695 (*with F. Jawadi*)
- "Expectation Formation in the Foreign Exchange Market: a Time-Varying Heterogeneity Approach Using Survey Data" [2015], *Applied Economics*, 47, 3673–3695 (*with R. Uctum*)
- "Modelling the horizon-dependent risk premium in the Forex: evidence from survey data" [2013], *Journal of International Financial Markets, Institutions & Money*, 23, 33–54 (with *R. Uctum*)
- "Equity risk premia and time horizon: what do US secular data say?" [2013], *Economic Modelling*, 34, 76
- « Fisher, Macaulay et Allais face au Paradoxe de Gibson » [2012], *Recherches Economiques de Louvain /Louvain Economic Review*, 78(2), 1–106 *(with J.-J. Durand*)
- "Arbitrage costs and nonlinear stock price adjustment in the G7 countries" [2012], *Applied Economics***,** vol. 44, n°12, 1561–1582 (*with F. Jawadi*)
- "Modelling oil price expectations: evidence from survey data" [2011], *Quarterly Review of Economics and Finance,* 51(3), 236–247 (*with R. Uctum*)
- « Anticipations, prime de risque et structure par terme des taux d'intérêt: une analyse des comportements d'experts » [2010], *Recherches Economiques de Louvain /Louvain Economic Review,* 76(2), 195–217 *(with R. Uctum)*
- "The dynamics of U.S. equity risk premia: lessons from a professional's view" [2010], *Bankers, Market and Investors*, n°104, 4–20 (*with Alain Abou*)
- "Switching between expectational processes in the foreign exchange market: a probabilistic approach using survey data" [2007], *Review of International Economics***,** *5(*4), 700–719 (*with R. Uctum*)
- « Les comportements boursiers sont-ils eulériens? » [2007], *Revue Economique*, 58(2), 427–453
- « Une Analyse de la dynamique des primes de risque des actions suivant l'horizon de placement » [2001], *Revue d'Economie Politique,* N°2, 291–329
- "Modelling stock price expectations: lessons from microdata" [2000], in *Price expectations in goods and financial markets,* F. Gardes & G. Prat Eds., Edward Elgar, 313–346 (*with A. Abou*)
- "The evidence of a mixed expectation generating process in the foreign exchange market" [2000], in *Price expectations in goods and financial markets,* F. Gardes & G. Prat Eds., Edward Elgar, 313–346 (*with R. Uctum*)

- "Trends of interest rates term structure in US secular data" [1999], *Advances in Investment Analysis and Portfolio Management, JAI Pres Inc n°*6, 109–132
- "Stock price and business cycles" [1997], *Encyclopedia of Business Cycles and Depressions,* David Glasner ed., Federal Trade Commission, Bureau of Economics, Washington, 659–662
- « Formation des anticipations boursières: consensus versus données individuelles » [1997], *Journal de la Société de Statistique de Paris,* Tome 138, n°2, 3–22 *(with A. Abou)*
- « A propos de la rationalité des anticipations boursières: quel niveau d'agrégation des opinions? » [1997], *Revue d'Économie Politique,* 107 (5), 647–669 *(with A. Abou)*
- « A propos des anticipations de change: nouveaux enseignements des données d'enquêtes » [1996], *Economie et Prévision*, *n*°4, 107–108 (*with F. Gardes*)
- « Le modèle d'évaluation des actions confronté aux anticipations des agents informés » [1996], *Revue Économique*, n°l, 85–110
- « La formation des anticipations et l'hypothèse d'un agent représentatif: quelques enseignements issus de simulations stochastiques » [1995], *Revue d'Économie Politique*, 105 (2), 197–222
- « La formation des anticipations boursières, États-Unis, 1956 à 1989 » [1994], *Économie et Prévision*, n°l, 101–125
- « Anticipations, prime de terme et maturité du titre long: que nous enseignent les données séculaires sur la structure des taux d'intérêt? États-Unis, 1873 à 1975 » [1992], *Revue Économique*, n°6,1037–1069
- *« Anticipations et évaluation des actions »* [1992], in « Monnaie, taux d'intérêt et anticipations », H. Kempf et W. Marois eds, Économica
- "La Bolsa y el ciclo economico (La bourse et le cycle économique)" [1990], *Information Comercial Espanola, Revista de Economia*, numéro spécial "Mercados de Capitales: analisis e innovaciones", n°688, 75–91
- "Dynamique des marchés financiers: notes de lectures sur F.R. Macaulay (1938)" [1990], *Journal de la Société de Statistique de Paris*, n°3-4, 57–68
- « Prévisions de la tendance du cours moyen des actions dans le cadre d'une formulation, États-Unis, 1981-1984 » [1987], *Revue d'Économie Politique*, n°5, 575–591
- « Reformulation du cours moyen des actions, États-Unis, 1956-1986 » [1987], *Journal de la Société de Statistique de Paris,* n°4, 201–222
- « Simulations de la tendance du cours moyen des actions dans le cadre d'une formulation, États-Unis, 1981-1984 » [1986], *Journal de la Société de Statistique de Paris*, n°2, 86–110.
- *« Essai pour une formulation générale du cours des actions: Un modèle d'évaluation pour l'ensemble du marché »* [1984], Économica, France, p. 286 (Preface by André Babeau)

- *« La bourse et la conjoncture économique - Essai sur les facteurs déterminants du cours des Actions »* [1982], Economica, France, 315 p. (Preface by Maurice Allais)
- « Le taux d'intérêt monétaire et le cours moyen des actions: interdépendance ou causalité? » [1981], *Journal de la Société de Statistique de Paris*, n°4, 215–231

Lille University, France Fredj Jawadi

Introduction

This book is co-edited in honour of Georges Prat, Emeritus Research Professor at the French National Center of Scientific Research (*CNRS*) and at the University of Paris Nanterre. Georges Prat began conducting his research in the 1970s under the supervision of Professor Maurice Allais, winner of the 1988 Nobel Prize in Economics. Georges' work has covered four main areas: (1) money supply and demand, inflation and monetary control; (2) economic fluctuations and psychological time; (3) long-run dynamics of unemployment and wages in France; and (4) expectations, uncertainty and dynamics of asset prices. The book focuses on this fourth area of research, undoubtedly Georges' main field of research. This book includes an overview of his career and work, an interview I conducted with him and a list of his publications on "Expectations, Uncertainty and Dynamics of Asset Prices".

The book also includes recent research conducted by eminent international researchers from institutions in Asia, Europe and the United States. There are eight chapters in all, organized into four parts. The first part, called "**Uncertainty and Volatility**", has two chapters. The first chapter is entitled "*Uncertainty or Stationarity in Macroeconomic Aggregates*?" and is co-authored by William A. Barnett (University of Kansas, USA) and Qing Han (University of Kansas, USA). The authors investigate non-stationarity in macroeconomic and financial time series for China and examine the size in unit root tests by Fourier approximation, which converts the estimation of location and style of breaks into the problem of appropriate frequency selection. They find that only financial series have good reason to be regarded as unit root processes, while most of the other series are better regarded as trend stationary with smooth transitions. The inference based on these results affirms that China's real business cycles are indeed fluctuations around different deterministic trends. The results also underscore the importance of handling data carefully to reduce econometric uncertainty and misspecification when modelling the dynamics and properties of financial and macroeconomic data. The second chapter entitled "*Oil Market Volatility: Is Macroeconomic Uncertainty Systematically Transmitted to Oil Prices"* is co-authored by Marc Joëts (Banque de France and EconomiX-CNRS, France), Valérie Mignon (EconomiX-CNRS, University of Paris Nanterre and CEPII, France) and Tovonony Razafindrabe (EconomiX-CNRS, CREM-CNRS and University of Rennes 1, France). This contribution also focuses on the analysis of uncertainty in macroeconomic data and deals with its transmission into the oil market. The authors study the impact of macroeconomic uncertainty on oil market volatility. To this end, they use a robust measure of uncertainty based on monthly macroeconomic and financial indicators and estimate a structural threshold vector autoregressive (TVAR) model. Interestingly, the authors show that a significant component of oil price volatility is due to macroeconomic uncertainty. Further, the authors point to a moderate increase in oil volatility in the aftermath of the global financial crisis.

The second part of this volume is called "**Heterogeneity of Beliefs and Information**" and also includes two chapters. The third chapter entitled "*Heterogeneous Beliefs and Asset Price Dynamics: An Overview*" is co-authored by Saskia ter Elleny (Norges Bank, Norway) and Willem F.C. Verschoor (VU University, Netherlands). The main focus of this chapter is on asset price dynamics in a framework of heterogeneous beliefs with regard to the traditional rational agent model. In particular, supported by the agent-based literature, the authors estimate a dynamic heterogeneous agents model and demonstrate its capacity to describe, explain and often forecast asset price dynamics. The superiority of an agent-based model is validated for different assets: equities, foreign exchange, credit, housing, derivatives and commodities. Interestingly, the model points to further evidence of market inefficiency (which is in line with Georges Prat's work) and appears valuable for reproducing different stylized facts and properties of financial data. "*High Frequency Trading in the Equity Markets during U.S. Treasury POMO*" is the title of the fourth chapter, co-authored by Cheng Gao (Rutgers University, USA) and Bruce Mizrach (Rutgers University, USA). The authors analyse high-frequency trading (HFT) activity in equities during U.S. Treasury permanent open market operations (POMO) by the Federal Reserve. In particular, they develop a model to study HFT quote and trade behaviour when private information is released and validate it empirically. Interestingly, the authors show that HFT firms decrease their inside quote contribution by up to 8% during POMO auctions. Further, the market impact also increases during Treasury POMO. These findings show that access to HFT and HF information can improve performance in stressful market conditions, confirming the usefulness of such information for HF traders.

The third part, called "**Transmission and Market Integration**", also includes two chapters. The fifth chapter, entitled **"***Crude Oil and Biofuel Agricultural Commodity Prices",* is co-authored by Semei Coronado (Universidad de Guadalajara, Mexico), Omar Rojas (Universidad Panamericana, Mexico), Rafael Romero-Meza (Universidad Autónoma de Chile, Chile), Apostolos Serletis (University of Calgary, Canada) and Leslie Verteramo Chiu (Cornell University, USA). The authors investigate the hypothesis of price transmission from oil prices to agricultural commodity prices. To this end, they study the relationship between oil price and the prices of three agricultural commodities that are used for biofuel production (corn, soybeans and sugar) using linear and non-linear causality and interdependence tests. Accordingly, while their findings do not find evidence of linear causality linkage between oil and commodity markets, they point to further evidence of strong bidirectional non-linear causality relationships and non-linear integration between oil and commodity markets, especially for the period from 2006 to 2016. This conclusion suggests that non-linear dynamics between the series studied have changed in recent years. This is in line with Georges Prat's work which also found evidence of non-linearity in financial price adjustment dynamics. Julien Acalin (Bank of France, France), Bruno Cabrillac (Bank of France, France), Gilles Dufrénot (Aix-Marseille School of Economics and CEPII, France), Luc Jacolin (Bank of France, France) and Samuel Diop (Bank of France, France) are the co-authors of the sixth chapter, entitled "*Financial Integration and Business Cycle Synchronization in Sub-Saharan Africa"*, which focuses on the hypothesis of market integration. In particular, the authors deal with the relationship between financial integration and business cycles in sub-Saharan African countries. Considering asymmetric dynamics during expansions and recessions, they de-synchronize fluctuations that capture the costs and benefits of financial integration. Their main findings show a significant effect of financial integration, but which varies across groups of countries. Indeed, while this relationship appears positive for some countries, financial integration increases the de-phasing of business cycles for WAEMU and SADC.

The fourth and last part of this volume is called "**Fundamentals and Bubbles**" and is also organized into two chapters. Entitled "*Informational Efficiency and Endogenous Rational Bubbles*", the seventh chapter is authored by George Watters (Illinois State University, USA). The author focuses on the assumption of rational bubbles that form and collapse endogenously and points to the inadequate predictability of tests of return to deal with these bubbles. Rather, he suggests a weighted replicator dynamic model that describes the switching of agents between a forecast based on fundamentals to a state with a rational bubble forecast. In line with the work of Georges Prat on market efficiency and asset prices, the author proposes a concise analysis of market efficiency/inefficiency and rational bubbles. Interestingly, this model explains multiple stylized facts about asset markets such as excess variance and GARCH effects and provides a clear discussion of the consequences in terms of rational bubbles and informational efficiency. The eighth chapter is entitled "*Stock Market Bubble Migration: From Shanghai to Hong Kong*". It is co-authored by Eric Girardin (Aix-Marseille School of Economics, CNRS & EHESS, France), Roselyne Joyeux (Macquarie University, Australia) and Shuping Shi (Macquarie University, Australia) and also focuses on bubbles. In particular, the authors examine the diffusion of bubbles from the Shanghai to the Hong Kong stock markets. They offer an exciting research piece that deals with bubble migration between these two markets over the period 2005–2017. Accordingly, they apply recursive explosive-root test to distinguish and date speculative episodes in both markets. Further, they check the migration assumption between these two markets, which confirms a significant, but declining, bubble migration from Shanghai to Hong Kong.

Lille University, France Fredj Jawadi

Contents

Part I Uncertainty and Volatility

Uncertainty and Stationarity in Financial and Macroeconomic Time Series—Evidence from Fourier Approximated Structural Changes 3
William A. Barnett and Qing Han

Oil Market Volatility: Is Macroeconomic Uncertainty Systematically Transmitted to Oil Prices? 31
Marc Joëts, Valérie Mignon, and Tovonony Razafindrabe

Part II Heterogeneity of Beliefs and Information

Heterogeneous Beliefs and Asset Price Dynamics: A Survey of Recent Evidence 53
Saskia ter Ellen and Willem F. C. Verschoor

High Frequency Trading in the Equity Markets During US Treasury POMO 81
Cheng Gao and Bruce Mizrach

Part III Transmission and Market Integration

Crude Oil and Biofuel Agricultural Commodity Prices 107
Semei Coronado, Omar Rojas, Rafael Romero-Meza, Apostolos Serletis, and Leslie Verteramo Chiu

Financial Integration and Business Cycle Synchronization in Sub-Saharan Africa 125
Julien Acalin, Bruno Cabrillac, Gilles Dufrénot, Luc Jacolin, and Samuel Diop

Part IV Fundamentals and Bubbles

Informational Efficiency and Endogenous Rational Bubbles 149
George A. Waters

Stock Market Bubble Migration: From Shanghai to Hong Kong 173
Eric Girardin, Roselyne Joyeux, and Shuping Shi

Part I
Uncertainty and Volatility

Uncertainty and Stationarity in Financial and Macroeconomic Time Series—Evidence from Fourier Approximated Structural Changes

William A. Barnett and Qing Han

Abstract The idea that many macroeconomic variables are unit root processes serves voluminously as a preliminary result in empirical works, but it is just a result of misspecification or weak identification with respect to the structural breaks. This contribution raises the size or power in tests of a null of a stationary process/unit root by Fourier approximation which converts the estimation of location and style of breaks into the problem of appropriate frequency selection. An examination of China's 15 representative macroeconomic series indicates that only the financial series have good reason to be regarded as unit root processes; most of others are better regarded as trend stationary with smooth transitions. The inference based on these results affirms that China's real business cycles are indeed fluctuations around different deterministic trends, and it is not the noise component rather the historical events corresponding to the breaks that have persistent effects. The results also support that large government-initiated shocks aimed at improving fundamentals are indeed capable of positive effects on the balanced growth path.

1 Introduction

Ever since the initiation of Nelson and Plosser (1982), followed by a burgeoning complementary elaborations such as Wasserfallen (1986), Phillips and Perron (1988), Cochrane (1988), and Kwiatkowski et al. (1992), the comprehension that most much-handled macroeconomic variables are unit root processes are deeply rooted. Still, the proclaimed dominance of a stochastic trend which usually serves as a preliminary result for further analysis in time series prevails voluminously in

W. A. Barnett
University of Kansas, Lawrence, KS, USA

Center for Financial Stability, New York, NY, USA

Q. Han (✉)
The University of Kansas, Lawrence, KS, USA

F. Jawadi (ed.), *Uncertainty, Expectations and Asset Price Dynamics*,
Dynamic Modeling and Econometrics in Economics and Finance 24,
https://doi.org/10.1007/978-3-319-98714-9_1

empirical work, even though Perron (1989, 1990) loquaciously argue that most of the $I(1)$ series are better to be modeled as $I(0)$ series when we take into account of structural changes. The lack of consensus regarding Perron's (1989, 1990) conclusion is not haphazard as the exact break date(s) and even the number of breaks are largely unknown and are believed to be more convincing if estimated endogenously, yet nearly all the successive tests based on endogenous estimation of the structural break property suffer from low power or size distortion in terms of insufficient break acquisition. The unit root or stationary tests incorporating structural changes are remotely satisfactory.

The aim of this chapter is twofold: firstly, it provides evidence as per the idea of approximating structural changes using Fourier series adopted by Becker et al. (2006) and Enders and Lee (2012) to argue that, besides financial series, most macroeconomic time series are in essence stationary just with structural breaks. It is the breaks or various smooth transformations instead of the error terms that have permanent effect. The method adopted in this work transcends estimating the number and shape of breaks by shifting these into an optimal frequency selection problem with respect to the Fourier approximations, a merit worth particular attention. Secondly, it contributes to the extant literature an internationally comparable investigation to Nelson and Plosser (1982)'s work by focusing on Chinese data; the work and results are comparable in terms of data selection and methodological comparison. The Nelson-Plosser data set has been used and reused by gobs of successive researchers within this realm, among others are Zivot and Andrews (1992) and Lumsdaine and Papell (1997). This contribution deliberately makes its data set comprehensive and representative so as to facilitate in the same way to those who are engrossed in Chinese empirics. Chinese economy which features various structural breaks is an ideal one for such an empirical demonstration because, being a typical state of government interference on business matters, a clear differentiation between structural break unit root and structural break stationary process helps vindicate the role that government should play in perspectives of long-term growth. Unit root in the real GDP series means the effect from government intervention can be easily canceled or diverted by a random shock in the error terms, a drastically different story if the series were stationary with governmental initiated breaks. Since people rarely have a priori information about the true data-generating processes, it is natural, and right people form a combination of different processes and make it into a mixed model. Even in this situation, as Prat and Uctum (2011) demonstrate, the mixed model could also depend on macroeconomic fundamentals whose effects are subject to structural changes, so break-dependent process is of equal relevance for mixed models.

The remainder of this chapter is structured as follows: Sect. 2 reviews the literature with comments. Section 3 elaborates methodologically how the tests are constructed and why the tool set is desirable for the purpose. Section 4 applies the methods on 15 mostly used Chinese macroeconomic time series manifesting both threshold breaks and smooth transitions, and while complementing what really happened during times of twists and turns, it helps understand it is the particular historical event rather than the stochastic error process that leaves a non-degenerating scent. The final section concludes the paper.

2 Literature and Comments

For lucidity we'd like to categorize structural changes into four varieties according to the shape of breaks and our comprehension with respect to their presence: exogenous threshold breaks, endogenous threshold breaks, exogenous smooth transitions, and endogenous smooth transitions. Threshold refers to sudden changes in mean or slope, whereas the series' systematic properties do not change substantively before or after the break point(s). Smooth transition touches on gradual changes which would take sort of visually sizable periods to have their effects fully released. While whether the change is abrupt or gradual may be a matter of data frequency available or the grid length of our foci, whether the breaks are captured exogenously or endogenously is the abyssal reason of all the dissatisfaction and refinements. As we shall see from the application on Chinese international trade data in Sect. 4 of this chapter, there are cases even researchers are inclined to overlook or not sure whether a certain part of the data is a transitional stage or not.

Exogenous structural break unit root tests specify the number of breaks a priori, thus incurring spurious rejections when data fail to cooperate. For exogenous smooth transition tests, one might refer to Luukkonen et al. (1988), Leybourne et al. (1998), Saikkonen and Lütkephol (2002), Lanne et al. (2002), and Kapetanios et al. (2003). Simulations made by Hecq and Urbain (1993) indicate that there will be both size distortions and loss of power in exogenously structural break in mean test if the pre-specified break date does not conform to the real one, though Montañés (1997) shows the distortion disappears in large samples, Montañés and Olloqui (1999) further point out the problem of low test power cannot be eliminated even asymptotically if the break date misspecification happens in trend.

Endogenous break tests estimate break date(s) first, then detrend accordingly, and construct unit root or stationarity test statistics using the residuals or detrended series. This general strategy applies to both the threshold changes and smooth transitions, and the criteria used to single out the date(s) are no other than minimal sequential t statistics, minimal summed squared residuals, and maximized F statistics, things like that. Christiano (1992), Banerjee et al. (1992), Perron (1997), and Zivot and Andrews (1992) develop threshold break tests that accommodate single break, and Lumsdaine and Papell (1997) and Ohara (1999), among others, extend the tests to allow more breaks. A universal problem in this line is that break is only allowed under the alternative hypothesis but not the null, insufficient exploitation of the structural break information leads to the coexistence of size distortion and low test power. Kim and Perron (2009) shun the problem of asymmetry and develop a test whose asymptotic distribution is the same as Perron's (1989, 1990) exogenous cases, but their mere accommodation of one break impedes its popularity. Harvey and Mills (2004) develop an endogenous smooth transition test but the number and mode of transition have to be pre-specified.

We learn from a retrospection of the literature that unspecified break properties do not lead to improvement of power compared with standard ADF unit root tests; erroneously specified number of breaks or mode of changes has basically no difference from completely overlooking structural changes in the context of a test. This be madness, yet there is method in it. Different from using dummy variables to capture threshold changes, from using logistic or exponential density functions to delineate smooth transitions, Fourier approximation uses the iteration of sinusoidal components to seize structural changes. No matter how many changes are there and no matter how unbelievable the breaks look like, all the break effects can be fully accounted for under a frequency that is both sufficiently large and doesn't bring in side effects. In this way the estimation of the number and shape of breaks boils down to a selection of optimal frequency. Employing this idea, Bierens (1997), Enders and Lee (2004, 2012), and Rodrigues and Taylor (2012) propose unit root tests, and Becker et al. (2006) come up with a stationarity test using Fourier series. Since stationarity tests take stationary processes as the null hypothesis and thus subject to type two errors in case of non-rejection, unit root tests serve to attest the results under such circumstances. This contribution makes a formal sharpening of the integration order of 15 most frequently used China's macroeconomic series by approximating the structural changes using Fourier series under the methodological discipline of Becker et al. (2006) and further vindicates the results by resorting to Fourier unit root tests of Enders and Lee (2012) for the sake of robustness.

Existing discussions about the unit-root property of China's GDP, such as Li (2000) and Smyth and Inder (2003), suffer from low test power in the methodologies used, an essential reason why the results are highly mixed. While these discussions focus on aggregate as well as provincial level of GDP, this chapter contributes a cross country comparable investigation parallel to Nelson and Plosser (1982)'s work in terms of variable coverage and methodological addressing.

3 Principles of Fourier Approximation and the Test

Contrary to the idea of estimating the latent number of breaks and specific type of changes, for the maneuver of endogenous breaks, Fourier approximation is not to estimate these specific features of structural changes but to transfer the problem into choosing the appropriate frequency when fitting the latent breaks via iterations of sinusoidal functions. The trade-off of the optimal frequency that appropriately fits the data without overplaying its hand is the core of Fourier approximation and its corresponding test.

3.1 Principle and Method of Fourier Approximation

Consider a regression with Fourier series, and assume there's a trend in the dependent variable:

$$f(t) = c_0 + \beta t + \sum_{\omega=1}^{k} a_\omega \sin\left(\frac{2\pi\omega t}{T}\right) + \sum_{\omega=1}^{k} b_\omega \cos\left(\frac{2\pi\omega t}{T}\right) + \varepsilon_t; \quad k < \frac{T}{2}.$$

Theoretically, no matter how many or what type of breaks *f(t)* incorporates, it can be approximated to any degree of accuracy and reduce each ε infinitely approaches to zero provided the Fourier series are sufficiently long. In this equation, ω is some specific frequency, k is the number of frequencies, and T represents the sample size. The degree of approximation accuracy increases as k becomes larger. Structural changes are captured by sinusoidal terms. Zero amplitudes $a_\omega = b_\omega = 0$ ($\omega = 1, \ldots, k$) indicate there is no nonlinearity in the function; au contraire, if there is nonlinearity, it must correspond to one particular frequency. Too many frequencies deplete degrees of freedom quickly and lead to the over-fitting problem pointed out by Enders and Lee (2004). Sinusoidal components which are used here to capture the structural changes lead to the fact that it is best to apply Fourier approximation to gradual process, and the smoother the process is, the less necessarily higher frequency is needed. It can be proved that structural changes shift the spectral density toward frequency zero; thus the optimal frequency for a break locates most probably at the low end of the spectrum. Becker et al. (2004) further show that high frequencies are prone to bring about stochastic variability of parameters, and the common sense up to now is that ω can best be chosen from the integer interval of [1, 5]; actually single frequency $\omega = 1$ (or $\omega = 2$) is sufficient for the delineation of a majority of breaks. Consider the following data-generating process:

$$\begin{aligned} y_t &= X_t'\beta + Z_t'\gamma + u_t + \varepsilon_t, \\ u_t &= u_{t-1} + v_t, \quad v_t \sim WN\left(0, \sigma_v^2\right). \end{aligned} \tag{1a}$$

where ε_t are stationary disturbances allowing for heterogeneity, u_t are random walks, v_t are white noise processes, and ${\sigma_v}^2$ is the variance. Furthermore, $X_t = [1]$ is used for level series of y_t and $X_t = [1, t]'$ for processes with trend. $Z_t = [\sin(2\pi\omega t/T), \cos(2\pi\omega t/T)]'$ capture the breaks in deterministic trend (or other forms of nonlinearity), and $\gamma = [\gamma_1, \gamma_2]'$ measures the amplitude. The null of ${\sigma_v}^2 = 0$ corresponds to y_t which is a $I(0)$ process with structural changes. If Z_t is absent, Eq. (1a) degenerates into standard KPSS stationarity test.

One favorable property of Fourier approximation in capturing the breaks is that for a given size and duration, the location of a break does not affect the fitness of the data. To be concrete, consider the following two DGPs: suppose the sample size $T = 100$, $y_t = 1.5$ when $33 \le t \le 66$, and $y_t = 2$ otherwise, as charted in Fig. 1a. Another DGP has the identical break size and duration except that the break happens at the low end of the sample for $12 \le t \le 45$ as in Fig. 1b. Dashed lines are Fourier

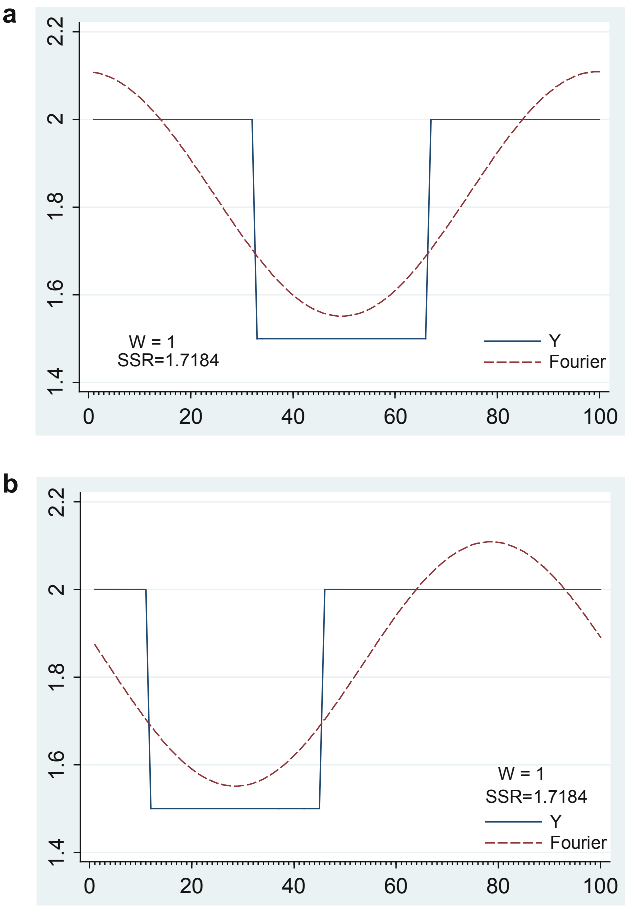


Fig. 1 (**a**) U-shaped break, $33 \le t \le 66$; (**b**) U-shaped break, $12 \le t \le 45$

approximations with $\omega = 1$, and the selection criterion for ω is the integer frequency from the interval [1, 5] that minimizes the SSR. Then the Fourier regressions for each DGP are:

$$y_t = 1.830 - 0.009 \sin\left(\frac{2\pi t}{T}\right) + 0.279 \cos\left(\frac{2\pi t}{T}\right) + u_t, \quad \text{SSR} = 1.7184, \quad R^2 = 0.6937 \tag{1b}$$

$$y_t = 1.830 - 0.273 \sin\left(\frac{2\pi t}{T}\right) + 0.061 \cos\left(\frac{2\pi t}{T}\right) + u_t, \quad \text{SSR} = 1.7184, \quad R^2 = 0.6937 \tag{1c}$$

Evidently the location of the break only affects the trigonometric estimator γ but does not change the value of the sum of squared residuals in each of the model. The importance of such property is twofold: Firstly, it guarantees breaks near the end of the series do not fester into test power.[1] Secondly, it can be proved (see Enders and Lee 2004; Becker et al. 2006) that test maintains invariance with regard to the value of β and γ yet depends merely upon the frequency ω, which avoids spurious inferences lured by some paradoxical breaks which themselves are highly subject to negligence.

The construction of structural change stationarity test statistics from Eq. (1a) is based on the standard KPSS statistics; denote $\hat{e}_t$ as the OLS residuals from the following regression (1d) or (1e):

$$y_t = c_0 + \gamma_1 \sin\left(\frac{2\pi\omega t}{T}\right) + \gamma_2 \cos\left(\frac{2\pi\omega t}{T}\right) + \varepsilon_t \tag{1d}$$

$$y_t = c_0 + \beta t + \gamma_1 \sin\left(\frac{2\pi\omega t}{T}\right) + \gamma_2 \cos\left(\frac{2\pi\omega t}{T}\right) + \varepsilon_t \tag{1e}$$

And the test statistics is:

$$\tau_i\left(\omega\right) = \frac{1}{T^2}\frac{\sum_{t=1}^{T} {S_t}^2\left(\omega\right)}{\sigma^2}, \quad i = \mu, \tau. \tag{2a}$$

Let $\tau_\mu\left(\omega\right)$ denote the test statistics for the level Eq. (1d) and $\tau_\tau\left(\omega\right)$ for the trend Eq. (1e). $S_t\left(\omega\right)$ itself is the sum of cumulative residuals from period 1 to t; different from the standard KPSS is that S_t now relies upon frequency ω; σ^2 is the long run population variance, and its sample counterpart can be figured out via nonparametric method as the standard KPSS does. Specifically, truncate the sample and choose the weight w_j, $\tilde{\gamma}_j$ is the jth autocovariance of the sample residuals, and l is the truncation lag. Then the sample counterpart of $S_t\left(\omega\right)$ and σ^2 can be calculated as follows, where $K(.)$ is the kernel function and h is the bandwidth:

$$\widehat{S}_t\left(\omega\right) = \sum\nolimits_{j=1}^{t} \widehat{e}_j \tag{2b}$$

$$\widehat{\sigma}^2 = \tilde{\gamma}_0 + 2\frac{\sum_{j=1}^{l} K\left(\left(e_j - e_0\right)/h\right)\tilde{\gamma}_j}{\sum_{j=1}^{l} K\left(\left(e_j - e_0\right)/h\right)}. \tag{2c}$$

$\tau_i\left(\omega\right)$ do not follow the usual probabilistic distribution, but due to the test invariance with regard to the parameters β and γ, the DGP for the simulation of critical values can be set with $\beta = \gamma = 0$,[2] and Becker et al. (2006, p. 389) give the Monte Carlo simulated critical values on the frequencies [1, 5].

[1]Some tests, for instance, Bai and Perron (1998), have little power when a break happens near the end of a series.

[2]Simulation results are actually the same for the other β and γ.

3.2 Frequency Selection

Frequency is the crux of the test, and there are logically three strategies of selecting the frequency component(s) in the test: single frequency, cumulative frequencies, and endogenous frequency. The selection criterion is still based on the rationale that the candidate yields the highest power and sound size. Single frequency means a pre-specified $\omega = 1$ (or $\omega = 2$) is sufficient to replicate the essentials of many break. It can be verified via simulation; however, once the real latent frequency is higher than 2, undervalued frequency often leads to severe oversized problem. So there is risk in utilizing $\omega = 1$ (or $\omega = 2$) without exception since researchers usually do not have definite information about it.

Gallant (1984) and Bierens (1997) recommend cumulative frequencies, and the reason is straightforward: if $\omega = 1$ captures an unknown functional form well, then a compound of $\omega = 1$ and $\omega = 2$ can do better. Under the circumstances of cumulative frequencies, test statistics still depend on the frequencies used because of the orthogonality of the trigonometric components on each frequency. Only the critical values and the distributions move toward the origin a bit as the frequency dimension increases.

No matter for the single frequency or cumulative frequencies, the frequency component(s) need(s) to be specified a priori, whereas endogenous frequency, although confines to an estimation of merely one frequency, estimates the unknown frequency and thus is a data-driven method. The estimation is usually conducted by choosing the very frequency that minimizes SSR within $\omega \in [1, 5]$. This is because comparing with other principles, such as t or F tests of related coefficient(s), the convergence speed of the test statistics has not been fully investigated, whereas the consistency and convergence speed of those based on the minimization of SSR can be guaranteed from the discussion of Hatanaka and Yamada (1999) as well as Perron and Zhu (2005). Becker et al.'s (2006) critical value simulated for single frequency can still be used in case of estimated frequency since the estimation of frequency is consistent.

The test for the optimal cumulative frequencies is not feasible according to this line since adding one more frequency undoubtedly decreases SSR. There are also oversized problems if the compound frequencies used deviate from the real ones. The SSR frequency, however, also subjects to a mild size distortion in small γ and T circumstances, which is due to the lack of precision for the estimation of frequency. It needs to be pointed out the cumulative frequencies still maintain reasonable test size if there's no break at all, but SSR frequency suffers size distortion. As for the test power, prior single frequency rates the highest, and then SSR frequency and cumulative frequencies score the minimum. This is because the endogenous method contains a procedure of searching the optimal frequency, and this may lead to a non-rejecting-the-null bias, with its power still higher than cumulative frequencies though. The comparisons between cumulative frequencies and SSR are thus straightforward: the former is latent for the risk of specification bias but is capable of sound test size even if there's no break; the latter sustains conservative test size but maintains higher power. Within this trade-off, there is a caveat for using cumulative frequencies, as it can be subject to oversized problem as well as loss of power.[3] This being said, cumulative frequencies are not always futile, and there are two cases where the application of cumulative frequencies is favorable in my eyes: no break at all and the break curvature is large in degrees. In the former case, cumulative frequencies remain reasonable size, while single frequency doesn't; in the latter case, oversized problem is rampant when single frequency is used but doesn't fit the break sufficiently. When researchers have no prior information about the appropriate frequency, SSR strategy is a reasonable starting point as less parameters are estimated.

3.3 Test the Break Components

One thing that cannot be neglected is the null $H_0{:}{\sigma_v}^2 = 0$ and the alternative hypothesis $H_1{:}{\sigma_v}^2 > 0$ of the test in Eq. (1a) have not specified there must be structural changes in the DGP. If the DGP doesn't contain any breaks, standard KPSS guarantees better test power. So it is necessary to test whether there are break components (nonlinearities in general) in the series, which equals to whether Eq. (1a) contains some specific frequency. When a single frequency is used, the null hypothesis is $H_0{:}\gamma_1 = \gamma_2 = 0$, and H_1 corresponds to some form of structural changes under the frequency used. Such a test can be performed according to the usual F statistics:

$$F_i(\omega) = \frac{[\mathrm{SSR}_R - \mathrm{SSR}_U(\omega)]/2}{\mathrm{SSR}_U(\omega)/(T-k)}, \quad i = \mu, \tau \tag{3}$$

where k is the number of regressors and SSR_R is the restricted sum of squared residuals with its unrestricted counterpart denoting SSR_U which is dependent upon the frequency, so the F statistics also depends on the frequency. Equation (3) applies only when ω is given, if ω is unknown—thus appears in the test as an unidentified nuisance parameter, the regular critical values of F test cannot be used even though ω can be figured out via minimization of SSR. In this case the F statistics is specified as follows:

$$\begin{gathered} F_i(\widehat{\omega}) = \max_{\omega} F(\omega), \quad i = \mu, \tau. \\ \widehat{\omega} = \arg\inf_{\omega} \mathrm{SSR}_i(\omega) \end{gathered} \tag{4}$$

Namely, the frequency is obtained by minimization of SSR of Eq. (1d) or (1e). Becker et al. (2006, p. 389) also give the critical values of the F statistics under such circumstances and further indicate that the test power is very low for the

[3] I'm indebted to the anonymous reviewer for pointing this out.

non-stationary series, that is, the test is defected for its inclination for the absence of nonlinearity under the unit root circumstances, whereas there is indeed nonlinearity. So this contribution only makes use of this F test that justifies existence of breaks when the null of a stationarity is not rejected.

3.4 Corroboration from a Fourier Unit Root Test

The major maneuver we employ takes stationarity as its null hypothesis, concrete as the evidence from a stationarity test could be, it is better to attest the result further by a unit root test which takes the null of a unit root against a stationary process.[4] For this purpose this work utilizes the Fourier unit root test proposed by Enders and Lee (2012) who introduce Fourier series to approximate structural breaks on an ADF basis.[5] With the above articulation, it is handy and beneficial to take a glance at this test principle which is based on the LM regularity.

Take the single frequency as a case in point, if the data-generating process is represented by Eq. (1e); regressing it using the first-order differences yields:

$$\Delta y_t = \alpha_0 + \alpha_1 \Delta \sin\left(\frac{2\pi\omega t}{T}\right) + \alpha_2 \Delta \cos\left(\frac{2\pi\omega t}{T}\right) + u_t.$$

Denoting by $\widehat{\alpha}_0, \widehat{\alpha}_1$, and $\widehat{\alpha}_2$ the estimated coefficients, a detrended series has been constructed by Enders and Lee using these coefficients:

$$\xi_t = y_t - y_1 + \widehat{\alpha}_0 (1-t) + \widehat{\alpha}_1 \left[\sin\left(\frac{2\pi\omega}{T}\right) - \sin\left(\frac{2\pi\omega t}{T}\right)\right] + \widehat{\alpha}_2 \left[\cos\left(\frac{2\pi\omega}{T}\right) - \cos\left(\frac{2\pi\omega t}{T}\right)\right],$$

where y_1 is the first observation of y_t. The test is based on regressing first differences of y on the detrended series as well as first differences of the trigonometric components:

$$\Delta y_t = \theta_0 + \phi\xi_{t-1} + \theta_1 \Delta \sin\left(\frac{2\pi\omega t}{T}\right) + \theta_2 \Delta \cos\left(\frac{2\pi\omega t}{T}\right) + \sum_{i=1}^{k} \phi_i \Delta\xi_{t-i} + \eta_t.$$

[4] I'm grateful to the referee for these refinements.

[5] Rodrigues and Taylor (2012) also come up with a unit root test using a Fourier series to approximate smooth breaks on an ADF basis. The difference is their test statistics is established according to the DF-GLS method, while Ender and Lee construct their test statistics according to the LM principle. Though DF-GLS is associated with higher test power for nonstructural settings, the penalty of sticking to this idea in Fourier approximated breaks is for the test statistics to suffer from asymptotically rank deficiency.

Non-stationarity corresponds to $H_0 : \phi = 0$, and the LM test statistic is naturally the t-statistic for this null hypothesis.[6] Lagged values of $\Delta\xi_{t-i}$ are used to correct for serial correlation.

Like Becker et al. (2006), Enders and Lee (2012) also recommend lower frequencies that don't exceed 5 in approximating the breaks out of the same reason and a data-driven method of minimizing SSR in selecting the optimal frequency. Furthermore, they use the same max F statistic to verify the existence of nonlinearities. So the optimal frequency singled out and the max F test are valid for both the Fourier stationarity and the unit root test, with the latter that serves as a robustness check for the results obtained from the major stationarity approach.

4 Fourier Approximation Tests and Analysis of China's Macroeconomic Time Series

This part moves to an exposition and discussion of China's macroeconomic time series; the following rests on the general assumption that China's entire macroeconomic system is a stochastic process that follows a certain distribution. This contribution singles out 15 commonly used variables which cover a wide range of output, employment, price, exchange rates, money, security, and trade. All the variables are in logarithms; detailed information and data sources are summarized in Table 1. Because this data set is reusable for future methodologies, put aside the idea this chapter is trying to argue for, this data set itself could contribute to a canonical set from which, like Nelson and Plosser (1982)'s, cross country comparison might come.

Since non-stationary series are cumulated disturbances based on one period ahead value of each date, an obvious property of a non-stationary series is that their autocorrelation coefficients decay to zero pretty slowly. Detrending the series either according to Eqs. (1d) or (1e) reveals that only the NEER, REER, and stock prices have a speed of decay analogous to that of a random walk. Other series decay to zero much quickly,[7] which forms an implicit manifestation of a property other than a unit root.

Nine of the fifteen concerned series take on obvious trend in the entire sample; still four of the series present trends within some specific subsamples, and these series are tested according to Eq. (1e) that allows for trend. Since economic theory does not suppose deterministic trend in stock prices, Shanghai composite index and Shenzhen component index are performed under Eq. (1d).

The first six columns of Table 2 summarize the main analytical results. Figure 2a–d plots the outputs of China with Fourier approximations in dashed lines.

[6] Please refer to Enders and Lee (2012, pp. 580, 582) for critical values under single frequency and cumulated frequencies.

[7] Such an informal demonstration is available upon request.

Table 1 Variables and explanatory notes

Series	Periods and remarks	T	Data source
Nominal GDP	Yearly: 1952–2008	57	GTA
Real GDP	Yearly: 1952–2008; base year = 1978	57	GTA (nominal)
Real per capita GDP	Yearly: 1952–2008; base year = 1978	57	GTA (nominal)
Real industrial production	Yearly: 1952–2008; base year = 1978	57	GTA (nominal)
Employment	Yearly: 1952–2008	57	GTA
Money supply (M2)	Monthly: 1996.1–2009.12	168	RESSET
CPI index	Monthly: 1995.2–2009.12; 1995.2 = 100	179	GTA (chain data)
Nominal exchange rate of RMB against USD	Quarterly: 1994.I–2009.IV	64	RESSET
Nominal effective exchange rate (NEER)	Monthly: 1980.1–2009.12	360	IFS
Real effective exchange rate (REER)	Monthly: 1980.1–2009.12; 2000 = 100	360	IFS
Shanghai composite index	Monthly: 1992.1–2009.12; closing rate	216	GTA
Shenzhen component index	Monthly: 1995.1–2009.12; closing rate	180	GTA
Export and import	Monthly: 1990.1–2009.12; seasonally adjusted	240	RESSET
Export	Monthly: 1990.1–2009.12; seasonally adjusted	240	RESSET
Import	Monthly: 1990.1–2009.12; seasonally adjusted	240	RESSET

Note: Obtained from the GTA database are only the nominal values of GDP, per capita GDP, and industrial production; the real values of the triple variables are denominated according to the 1978 based price index which come from the *China Statistical Yearbook* of relevant years; the CPI index obtained from the GTA are chain data, and they are transformed into fixed base index with February 1995 as their basic month; the real effective exchange rate of RMB is based on the year 2000; Shanghai composite index and Shenzhen component index are closing rate stock prices with different initial date; export and import from the RESSET database are nonseasonally adjusted data, and I remove the seasonal factors using Tramo/Seats method

Similarly, it is interesting to notice the optimal frequencies that fit these various output indexes best are all $\omega = 1$. While most of the parts look relatively smooth in real GDP, real per capita GDP, and real industrial production, the period of 1960–1962 in contrast corresponds to evident downward transitions which are not so evident in the non-inflation adjusted nominal GDP though. Every Chinese can instantly recognize them as the result of the Great Leap Forward and the following Great Famine which has been attributed to the 3 years of natural disasters, a term preferred by official announcements and mandated in textbooks even today. The Great Leap Forward was an ideologically pathological national movement

Table 2 Tests for stationarity with corroborations and comparisons

						Corroboration	Tests suffer from spurious rejection			
Series	T	Type	ω	$\tau_i\ (\widehat{\omega})$	*max F*	Fourier UR	KPSS	DF-GLS	LSTAR	Zivot-Andrews
Nominal GDP	57	τ	1	0.0429	393.69	−4.80[a]	0.53[a]	−0.45	−3.83	−3.87
Real GDP	57	τ	1	0.0467	63.56	−4.93[a]	0.45[a]	−0.36	−4.91	−4.10
Real per capita GDP	57	τ	1	0.0514	104.71	−4.57[a]	0.49[a]	−0.40	−4.52	−3.89
Real industrial production	57	τ	1	0.0353	7.84	−4.73[a]	0.14	−1.52	−4.76	−6.36[a]
Employment	57	τ	1	0.0660	19.89	−5.35[a]	0.23[a]	−1.31	−4.07	−2.86
Money supply (M2)	168	τ	1	0.0678	67.58	−4.77[a]	0.42[a]	−2.19	−3.41	−4.27
CPI index	179	τ	1–3	0.0371	448.70	−7.05[a]	0.43[a]	−2.73	−2.94	−4.70[b]
Nominal exchange rate of RMB against USD	64	τ	1–2	0.0275	273.61	−2.38	0.32[a]	−3.20[b]	−3.39	−4.89[b]
Nominal effective exchange rate (NEER)	360	τ	1	0.0874[a]		−1.88	1.51[a]	−0.72	−2.57	−3.38
Real effective exchange rate (REER)	360	τ	1	0.1163[a]		−2.01	1.43[a]	−1.14	−2.71	−3.79
Shanghai composite index	216	μ	2	1.9163[a]		−3.75	3.65[a]	−1.10	−2.65	−4.18
Shenzhen component index	180	μ	2	1.3557[a]		−1.00	2.31[a]	−0.06	−1.53	−3.80
Export and import	240	τ	1–2	0.0319	187.14	−7.06[a]	0.88[a]	−1.78	−2.99	−2.54
Export	240	τ	1–2	0.0277	167.57	−7.09[a]	0.83[a]	−1.98	−1.75	−2.29
Import	240	τ	1–2	0.0327	151.97	−7.81[a]	0.83[a]	−2.28	−3.78	−3.27

Note: All the single frequencies are endogenously determined optimal ones from the minimizations of SSR, whereas the cumulative frequencies are exogenously appointed. One percent critical values of max F statistics for the former six variables under endogenously determined frequency are all 6.873; for the avoidance of spurious inferences that may be induced by low test power, F tests are omitted since NEER, REER, and stock prices reject the null significantly. DF-GLS reports the test statistics of the optimal lag selected by Ng-Perron sequential t statistics. For LSTAR estimation, the initial transition parameter is set to 10. In Zivot and Andrews' (1992) test, 5% trimming of the data is used

[a]Rejecting the null at a significance of or higher than 1%

[b]Rejecting the null at 5% significance level

whose aim was to jog into Communism, when an enthusiasm had been roused as every political campaign would be capable of and local officials were twisted with an incentive of overstating the crop productions under dictatorship. In a haste of showing devotion or allegiance, lower echelons had a fondness of raising the ante of the output volume sequentially in a single-direction hierarchical system. Since tax was proportional to output, the more was produced or was assumed to be produced, exactly speaking, the more was collected by the country. And when the overstatement was rolled down to a preposterous exaggeration, peasants were expected to hand over not only all the yields but nearly everything including the necessary amount of food for maintaining sustenance and the seeds for sowing. That's how demagoguery results in the Great Famine during 1959–1962; it is estimated 37.55 million Chinese people starve to death,[8] an estimation that is in close accordance with the population reduction inferred from demographic records. So the ups in output correspond to intense devotion of the Great Leap Forward; the downs in output correspond to the starvation and the hindrances to reproduction. That's enough for the background.

Test results show non-rejections at the significance of 5% unanimously, which means although the Great Famine exerts significant shocks toward the latent data-generating process of China's GDP, there are still evidences which indicate that the growth of Chinese economy follows a tractable trend. Such a trend may not be properly handled by simple linear or quadratic trends due to the various and infinite randomness of the entire economic system; all the significance this work intends to address methodologically is that delineating the trend nonlinearly by the thought of Fourier approximation may yield a better exposition. Based on these observations, it is more pervasive to regard China's GDP as trend stationary process with structural changes of smooth transitions, even the shocks as large as 3 years of nationwide starvation could not lead the economy to deviate from its Fourier trend for a long time, and business cycles in this sense are just fluctuations around this smooth trend with some drastic yet gradual ups and downs. Finally, all the max F tests for nonlinearity in outputs reject the null of $\gamma_1 = \gamma_2 = 0$ significantly, which justifies the obvious feature of gradual changes. Omitting the breaks or detrending the series by simple trend may easily yield the inference of a unit root about China's GDP.

There is a little bit subtlety in employment and money supply, although the two variables reject the null at 5% significance level, they both fail to reject when the significance level tightens up to 1%. Under this condition, graphs once again help to recognize the problem or build confidence in the results. Figure 2e shows an upsurge and down of the employment during the year 1957–1960; the Great Leap Forward enhances employment initially, but the following Great Famine drags it down. A large "N-shape" fluctuation emerges during 1989–1992; the growth rate whereafter slows down, which is presumably a result of the reform of state-owned enterprises (SOEs). Even though the open-up reform has started as early as 1978, the SOEs have taken dominance in nearly all industries by the end of 1980s. Then a relative

[8]Source (In Chinese): http://jiuliyougancheng.blogchina.com/1373459.html

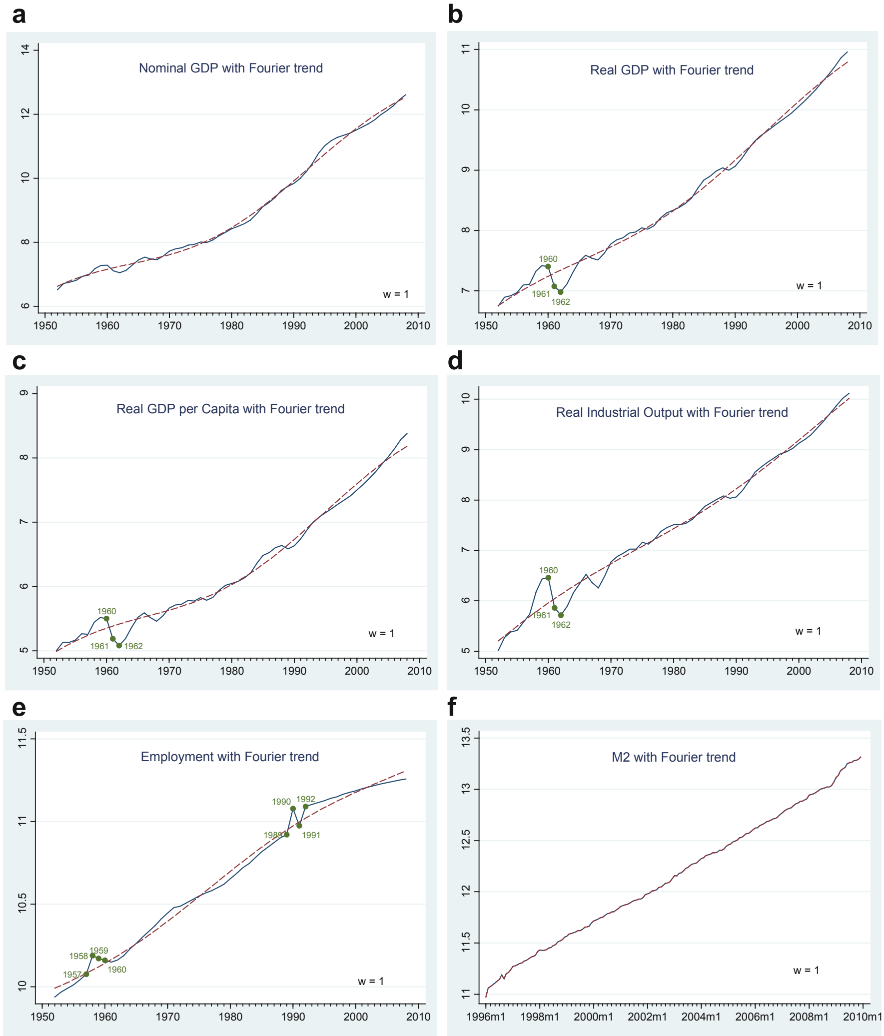


Fig. 2 (**a**) Nominal GDP; (**b**) real GDP; (**c**) real per capita GDP; (**d**) real industrial production; (**e**) employment; (**f**) money supply (M2)

fierce reform of SOEs was initiated in many industries to give rise to privatization which brought about a large amount of unemployment in the beginning; since SOEs were redundant and inefficient, it wasn't uncommon at that time in cities that a former SOEs' employee had become a laid-off worker at his/her 30s or 40s. The effects of the two volatile deviations cannot be canceled, which I suppose may account for the rejection at 5% significance level. But the rest of the processes other than the two breaks are comparatively stationary, including the slow down after 1992. Thus it is feasible to consider China's employment as trend stationary with structural changes. Figure 2e offers just another case; mere eye glances can hardly distinguish between Fourier fit and money supply. This is due to the small variance of China's monetary supply, and there are not any recognizable fluctuations since 1996; besides, it is ponderable from the graph that China's monetary supply of each period may considerably take into account of the amount issued one period ahead or even earlier and thus takes on the effect of intensive persistency. Small variance as well as intensive persistency exerts an effect that a small deviation is actually a large shock, which give rise to the inclination for a unit root inference. The amazing straightforward trend, however, still locks the inference of a trend stationary at 1% significance level, and as witnessed, there are reasons in such an inference.

The case of CPI is a bit more complicated yet illustrative about the focal point. China's CPI reaches its peak as high as 24.1% in 1994[9] and still suffers an annual average of 17.1% and 8.3% separately in 1995 and 1996 though alleviated gradually. Because the monthly data this contribution utilized are available only since February 1995 on which is also the data based, there is an obvious price increasing trend during 1995–1996, as plotted in Fig. 3a–c. Since then China steps into mild deflation during 1997–2002, and the fixed base CPI somewhat goes down. The entire shift from inflation to deflation stands prominently as a smooth transition process with distinctive growth rates on each period and neat stationarities on each trend. If CPI is fitted with the single optimal frequency $\omega = 1$, as plotted in Fig. 3a, distinctive change of growth rates results in insufficient approximation for single frequency, and the test rejects the null at 1% significance and is in favor of a unit root inference. But the approximation improves when cumulative frequencies $\omega = 1$ and $\omega = 2$ are employed, as plotted in Fig. 3b; the test now rejects the null at 5% significance level but fails to reject at 1%. One more step further, if a compound frequencies of $\omega = 1$, $\omega = 2$, and $\omega = 3$ are fitted in the equation, the approximation improves yet again, and the test result listed in Table 2 shows the null of a stationarity is not rejected even at 10% significance level. The fact revealed in the investigation of CPI is that insufficient fit makes possible rooms for size distortion, which is prone to be favorable for a unit root inference. A reasonable inference is conditional on sufficient fit (to the exclusion of over fit, of course), whereas there is still lack of a universal rule to differentiate the circumscription between sufficient and insufficient fit. And the distinction between the two, under many circumstances, depends on the meticulous command of the researchers to a substantial extent. For this reason, it makes sense to deem the monthly fixed base CPI as trend stationary process with smooth transitions.

The same problem appears in the nominal exchange rate of Renminbi against US dollar. Because commodity prices were controlled under planned economy

[9]Chinese economy is an investment-driven economy up until today. Investment, irrespective of private or public, rather than consumption, plays a determinant role in economic growth. In 1992, Mr. Deng Xiaoping, the former chairman of the CPC, made one of his most cited speeches in South China whose content was to encourage the existence of private economy for the sake of efficiency. The speech instigated an upsurge of investments, and the inflation during 1993–1996 was brought about by this round excessive investing.

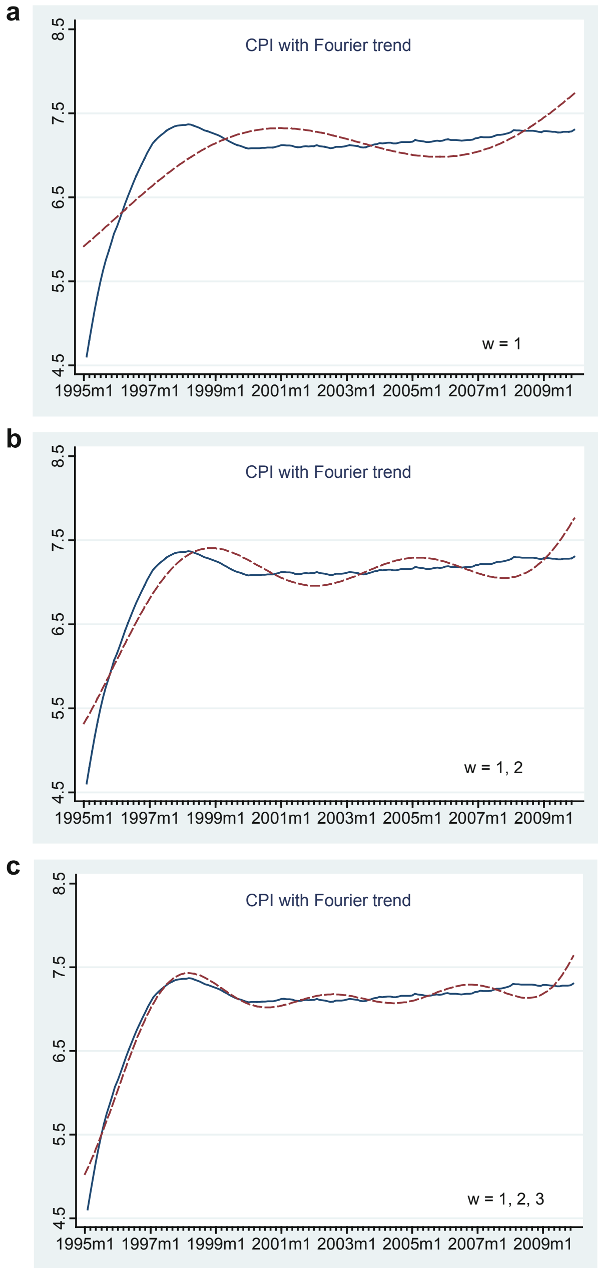


Fig. 3 (**a**) CPI, $\omega = 1$; (**b**) CPI, $\omega = 1, 2$; (**c**) CPI, $\omega = 1, 2, 3$

for a long term, the same commodity might have a drastically different world price compared to its domestic market, which caused international trade to be in a red. At that time the State Council of China presumed this to be the fact that the RMB exchange rate could not cater both the non-trade aspect and the international trade aspect simultaneously, so a double regime of RMB exchange rate was designed and put to application until 1994. That is, a new exchange rate of RMB which, 1 USD = 2.80 RMB initially, is designated for trade settlement is established; meanwhile the old official RMB exchange rate which is approximately 1 USD = 1.50 RMB served the so called non-trade settlement. Later as the trade regime gravely depreciated to 1 USD in exchange of nearly eight RMB Yuan, together with the prevalence of exchange rate arbitrage in the black market, those strength finally forced the merger of official rate and the trade regime rate in 1994. As graphed, the RMB exchange rate firmly pegged USD at the level of 8.27 from the fourth quarter of 1996 to the second quarter of 2005. And this exchange rate is undoubtedly stationary if the sample is confined within this period. However if the period falls to be the subset of the entire sample, such an embarrassing tranquility only helps to boost pseudo-inference. The test rejects the null at 5% but fails to reject at 1% if the single optimal frequency $\omega = 1$ is used to fit the data (Fig. 4a). If the frequencies are relaxed to a compound of $\omega = 1$ and $\omega = 2$ altogether (Fig. 4b), the test does not reject the null even at 10% besides fitness improvement. Although the break points on the corresponding quarters of 1995, 2005, and 2008 are by no means smooth, there is still evidence to treat RMB exchange rate against USD as trend stationary with structural changes.

The nominal effective exchange rate (NEER) and real effective exchange rate (REER) of RMB primarily undergo a long way of persistent devaluation before 1994 and appreciate in fluctuations ever since (Fig. 5a, b). Exchange rate reforms in 1994 once again dovetail with the stylized facts of structural changes. The optimal frequency for NEER and REER are both $\omega = 1$, and the tests reject the null both at 1% significance level. So NEER and REER of RMB are inclined to be inferred as unit root processes with structural changes. Shanghai composite index and Shenzhen component index correspond separately to Fig. 6a, b, with the former initiated on January 1992 and the latter on January 1995. The optimal frequencies for the stock prices represented by the two indexes are both $\omega = 2$, and the null hypotheses are rejected significantly. It is necessary to make clear that economic theory does not provide much evidence to back up the viewpoint that stock prices have long-term growth trend; thus the tests are performed according to Eq. (1d) that precludes the trend. The results conform to the commonly acknowledged perception that financial prices are generally processes of random walk. This result is consistent with the finding of Jawadi and Prat (2017) whose dividend discount model combined with arbitrage pricing theory yields major fluctuations of stock prices.

China's foreign trade statistical data provide excellent material for the comprehension of smooth transition stationarity test. The bygone subprime crisis leaves an impressive scent at the terminal end of each of the sample in Fig. 7a–f; also it can be found from a comparison between Fig. 7c and e that the crisis strikes the import heavier than the export. Since Fourier approximation maintains fitness invariance as

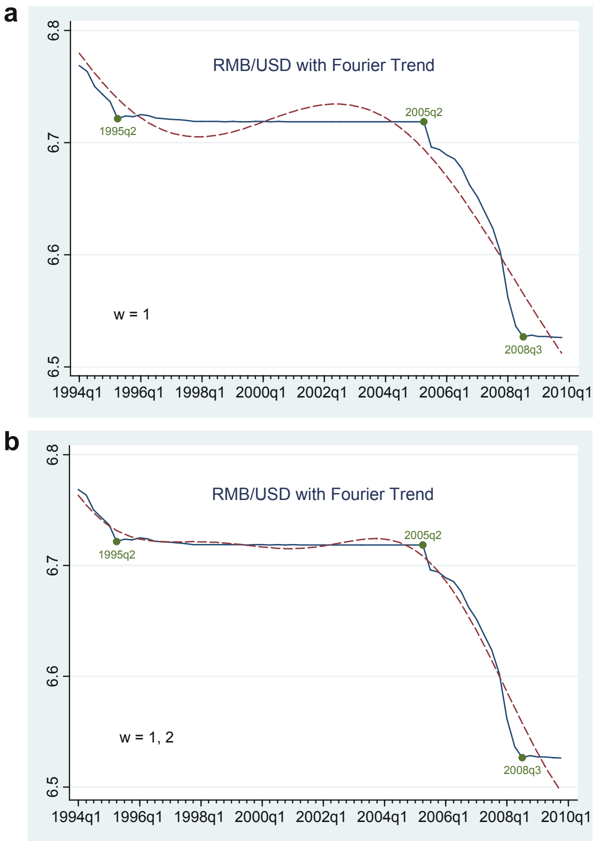


Fig. 4 (**a**) RMB/USD, $\omega = 1$; (**b**) RMB/USD, $\omega = 1, 2$

for the break position within the detrended series, a break located at the end of the sample will not lead to severe size distortion. If the data are fitted by optimal single frequency $\omega = 1$, then the total trade, the export, and the import all reject the null at 1% without exception and are in favor of unit root inferences. But when cumulative frequencies $\omega = 1$ and $\omega = 2$ are utilized, the tests unanimously fail to reject the null at 10% without exception. Reasons do not lie in the subprime shocks at the terminal ends, whether the marked parts in the middle of the sample are regarded as smooth transition processes as lined out in Fig. 7b, d, and f is the one that really counts. Comparing each graph with their counterpart under $\omega = 1$ situations, the intermediate sessions do not stand prominently as smooth transition processes under single frequency; thus the test itself regards them as realizations of purely stochastic trend, which is the reason that leads to unit root inferences. But under cumulative frequencies, along the traces following the approximated dashed lines are clearly sluggish then revived growth rates, though nothing changes in the original data plots but only the Fourier approximations. For the total trade, the middle part singled out

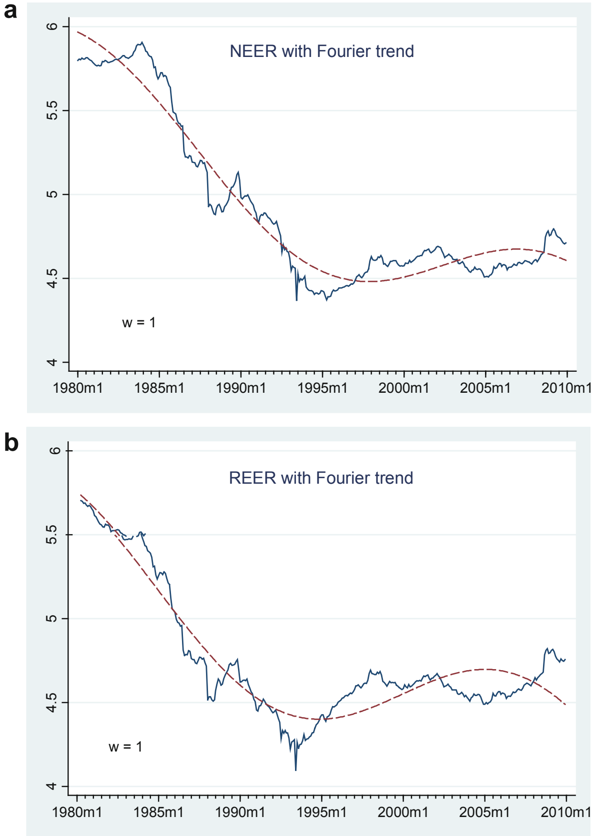


Fig. 5 (**a**) NEER; (**b**) REER

corresponds to the period from July 1997 to June 2000. For the export, it is July 1997 to February 2001. For the import, it is February 1996 to December 2000. The very historical event these periods coincide with is the Southeast Asian financial crisis! So it justifies considering these sessions as smooth transitions rather than the results of purely stochastic trend. And within the scope of gradual changes, the more persuading inference for China's trade data should be trend stationary with smooth transitions.

5 Robustness Check: Corroborations and Comparisons

Since non-rejection of the null undertakes a probability of making mistakes which is hard to tell, for the sake of robustness, we corroborate our results with a Fourier unit root test of Enders and Lee (2012) which has been briefly exposited

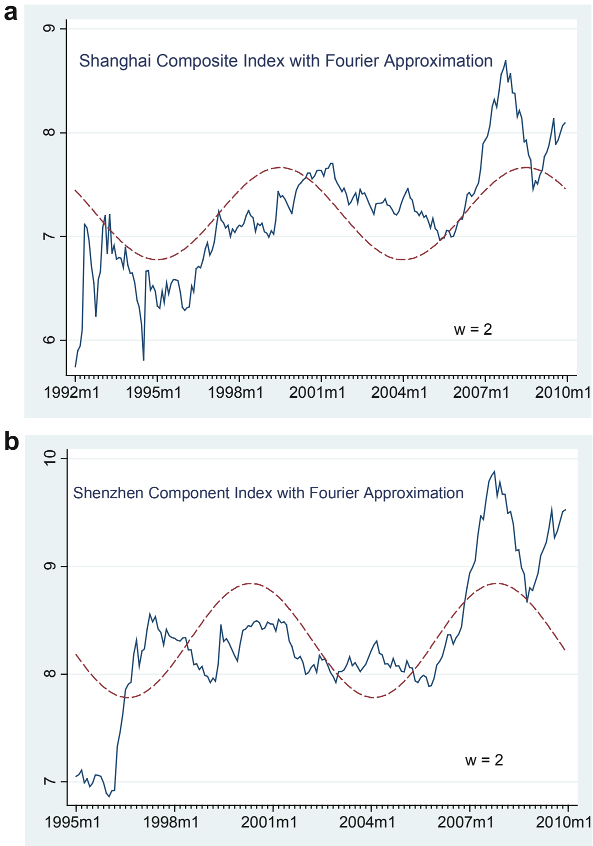


Fig. 6 (**a**) Shanghai composite index; (**b**) Shenzhen component index

in Sect. 3.4. The attesting results are shown in Table 2 under the column labeled with "Fourier UR." The fact that the optimal frequency, as well as the max F test statistic, applies to both Fourier stationarity and unit root tests greatly facilitate the cross reference. Besides all the exchange rates and stock prices, other variables unanimously reject the null at the 1% significance level in favor of the conclusion that all these variables are stationary. Except for the bilateral RMB/USD exchange rate, all our former conclusions under the stationarity test can be cross attested by this Fourier unit root test under the same trend type and frequency specification.

In comparison, to demonstrate the erroneous inference because of the spurious rejection, Table 2 also reports the test results obtained from the standard KPSS, the DF-GLS, the logistic smooth transition autoregressive (LSTAR) unit root test suggested by Saikkonen and Lütkephol (2002), and finally the Zivot and Andrews' (1992) test that uses dummy variables but allows for the endogenously determined

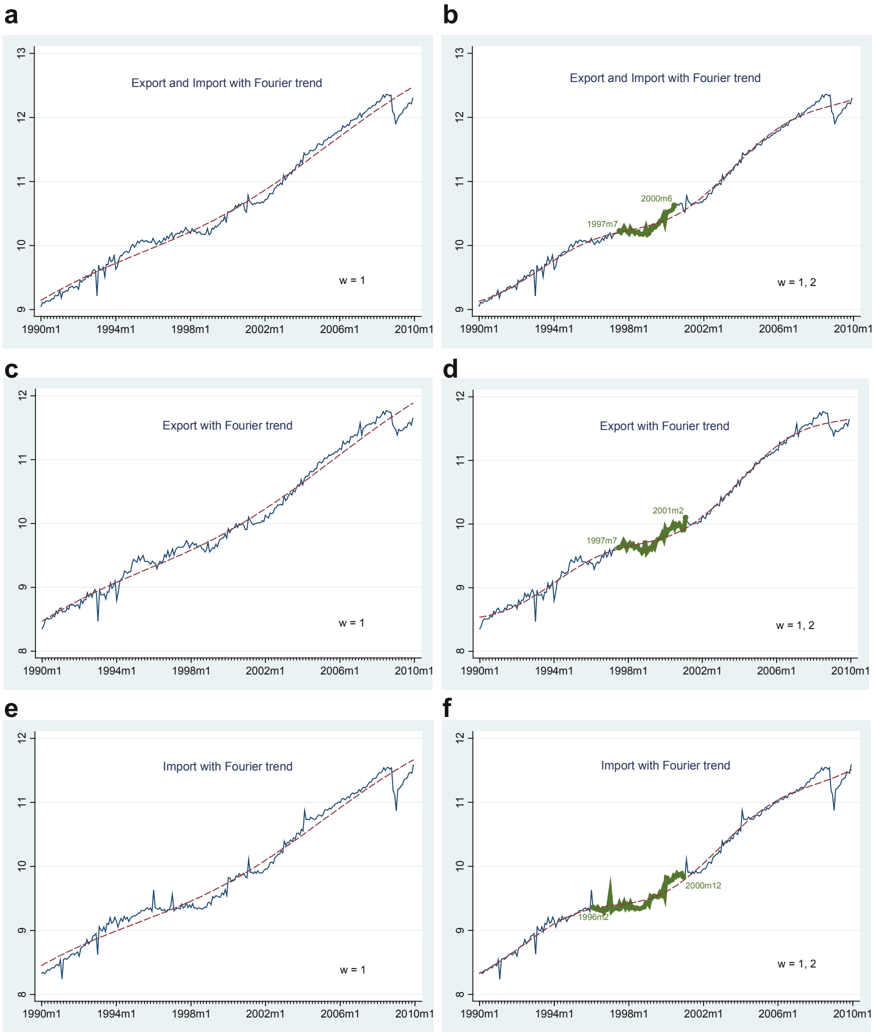


Fig. 7 (**a**) Export and import, $\omega = 1$; (**b**) export and import, $\omega = 1, 2$; (**c**) export, $\omega = 1$; (**d**) export, $\omega = 1, 2$; (**e**) import, $\omega = 1$; (**f**) import, $\omega = 1, 2$

break date. Situations can be quite different when the breaks are not catered or the nonlinearities are not controlled for by trigonometric components. The standard KPSS just significantly rejects the null of a stationary process to all except for the real industrial production; DF-GLS, the one that is considered to be of the highest power under no break circumstances, fails to reject all except for the bilateral

RMB/USD exchange rate.[10] Without due consideration of the structural breaks, nearly every variable we got is prone to be concluded as a unit root process. Yet even considering approximation using a smooth transitional function suffers from low test power, because LSTAR test rejects none of the null hypotheses of all the variables.[11] Finally, we can compare the results with the traditional method using dummies to account for breaks, as indicated in the Zivot and Andrews' test,[12] only the real industrial production, CPI, and the bilateral exchange rate could reject the null of a unit root, even though the real GDP and GDP per capita resemble the real industrial production largely, as can be seen from the figures.

6 Concluding Remarks and Discussion

This contribution aims to give an investigation on the stochastic trend that possibly lurks within China's macroeconomic system from the perspective of structural changes. Since the initiation of Perron (1989), research in this realm are flourishing yet still short of satisfaction; it is thus necessary to base this investigation on the detailed comments and comparisons among the former literature to reduce the possibility of pseudo-inferences to a maximum extent. And detailed treatment of the development in this realm leads to the observation that Fourier approximation intrigues the investigation in two aspects: Firstly, it shifts the determination of number of breaks and modes of changes into the appropriate frequency selection problem and thus enables a shortcut over the incessant dispute on the specific properties of the structural changes. Secondly, a Fourier detrended series maintains the fitness invariance as for the break locations, so breaks near the end of a series do not encroach on the inferences. For these reasons, this chapter explores Fourier principles and tests suggested by Becker et al. (2006) to investigate the structural change stationarity of the 15 variables chosen from China's macroeconomic system. The results are further corroborated by the Fourier unit root test of Enders and Lee (2012) and are compared with other parallel methodologies. We believe that only the stock prices, the nominal effective exchange rates, and the real effective exchange rates are dominated significantly by stochastic trend; it is more appropriate to model these variables according to structural change unit root processes; other variables, including the real GDP and foreign trade, however, are dominated by deterministic trend, and modeling them according to trend stationary processes with smooth transitions yields more accurate forecasts and perceptions.

[10]Here in the table, the DF-GLS test only reports among others the test statistic of the optimal lag selected by Ng-Perron sequential t statistics.

[11]The initial transition parameter is set to 10 for LSTAR estimation.

[12]5% data trimming is used.

We do not implement tests to cherish the intention of labeling a particular series with an $I(0)$ or $I(1)$ tag; we implement tests to gain improved epistemic grasp on how we should comprehend the latent data-generating process of the true economy. Take the real GDP, for instance; if it was inferred to be a unit root process irrespective of the smooth transitions, this implies what happens during 1960–1962 are 3 years of succeedingly large outliers which are due to the fat tail of the time-invariant probabilistic distribution that the latent DGP follows. Whereas if the real GDP was inferred to be a trend stationary process with smooth transitions as this work believes, then 1960–1962 should be comprehended as the shocks of the Great Leap Forward as well as the Great Famine changed the probabilistic distribution that the DGP follows during this period; nonetheless, either the distribution before or after these shocks remains stationary. It is the shock events that the smooth transition processes correspond to rather than the entire noise processes are of permanent effect. The same implications of perception apply to the other series, such as the Southeast Asian financial crisis for trade and the shift from inflation to deflation for CPI. The results of this work, in terms of policy implications, back up the effect of government-initiated structural reforms, because they cannot be easily offset by other noisy disturbances. Such interventions should either be of powerful intensity or be able to maintain transitional persistency institutionally.

As the problem all the hypothesis tests confront, of course, rejecting the null in a unit root test does not mean an instantaneous acceptance of a particular alternative, but for the purpose that tests are designed to raise power against a specific class of alternatives, forecast precision may gain improvement by comparing among the close alternatives and modeling according to a particular one. Analogously, failing to reject the null in a stationarity test does not mean adhering to this specific null hypothesis even seriously, but for the purpose that tests are developed to maintain reasonable test size, it helps to confine the forecast errors within reasonable ranges when modeling according to this null. However neither the structural change unit root tests nor structural change stationarity tests settle the problem of observational equivalence well in finite samples. Metaphorically, a trend stationary process with one break but the same slopes and zero variance in the errors is observationally equivalent to a unit root process with drift, but errors are zero with a high probability and nonzero occasionally. Namely, nonzero but finite variances correspond to fat-tail distributed errors of a unit root, and it is hard to differentiate the two kinds of models once the nonzero variance sprang out, or put it equivalently, such a differentiation is only feasible on the infinite horizons. For more general cases, any trend stationary processes are almost observationally equivalent to unit root processes with strong mean reversion.[13] The fact of non-rejection after

[13]Consider the stationary process $y_t = \varepsilon_t$ and the unit root process $(1 - L)y_t = (1 + \theta L)\varepsilon_t$ where $|\theta| < 1$; the observable implications are virtually the same as θ goes to -1. Again, consider the unit root process $y_t = y_{t-1} + \varepsilon_t$ and the stationary process $y_t = \rho y_{t-1} + \varepsilon_t$ under $|\rho| < 1$; it is also difficult to distinguish between the two when ρ is close to 1 at finite time horizons.

approximating the smooth transitions with Fourier series as this contribution does has an implication in this term that the disturbances must present strong mean reversion if the data are actually dominated by stochastic trend. Although for any given finite sample series, a representative can be found from either class of processes that are capable of delineating the observed features of the data, there are still good reasons that justify the distinction between the two classes of models. The first involves a trade-off between efficiency and consistency. If a restriction (a structural change trend stationarity, in this work's context) is true, imposing it in the estimation yields more efficient estimates and more accurate predictions. If the restriction is false, the estimates are unreliable, and the inconsistency cannot be wiped out even asymptotically. Besides the familiar trade-off between efficiency and consistency, two classes of models also correspond with different data perceptions. Whether a particular class of models should be used depends not only upon the data themselves but also upon the rationality of the perception. If a unit root is employed, there must be justifications to guarantee the realizations of fat-tailed disturbances. But for the major representative variables of real macroeconomy, described as hereinbefore, such justifications are seldom sufficient. Thus it is more reasonable to employ trend stationarity with structural changes (smooth transitions), which is all that this chapter means to address.

References

Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica, 66*(1), 47–78.

Banerjee, A., Lumsdaine, R. L., & Stock, J. H. (1992). Recursive and sequential tests of the unit-root and trend break hypothesis: Theory and international evidence. *Journal of Business and Economic Statistics, 10*(3), 271–287.

Becker, R., Enders, W., & Hurn, S. (2004). A general tests for time dependence in parameters. *Journal of Applied Econometrics, 19*(7), 899–906.

Becker, R., Enders, W., & Lee, J. (2006). A stationarity test in the presence of an unknown number of smooth breaks. *Journal of Time Series Analysis, 27*(3), 381–409.

Bierens, H. (1997). Testing the unit root with drift hypothesis against nonlinear trend stationarity, with an application to the US price level and interest rate. *Journal of Econometrics, 81*(1), 29–64.

Christiano, L. J. (1992). Searching for a break in GNP. *Journal of Business and Economic Statistics, 10*(3), 237–250.

Cochrane, J. H. (1988). How big is the random walk in GNP. *Journal of Political Economy, 96*(5), 893–920.

Enders, W., & Lee, J. (2004). *Testing for a unit-root with a nonlinear Fourier function*. Econometric Society 2004 far Eastern meetings, no. 457. http://www.nd.edu/~meg/MEG2004/Lee-Junsoo.pdf

Enders, W., & Lee, J. (2012). A unit root test using a Fourier series to approximate smooth breaks. *Oxford Bulletin of Economics and Statistics, 74*(4), 574–599.

Gallant, A. R. (1984). The Fourier flexible form. *American Journal of Agricultural Economics, 66*(2), 204–208.

Harvey, D., & Mills, T. (2004). Tests for stationarity in series with endogenously determined structural change. *Oxford Bulletin of Economics and Statistics, 66*(5), 863–894.

Hatanaka, M., & Yamada, K. (1999). A unit root test in the presence of structural changes in I(1) and I(0) models. In B. R. F. Engle & H. White (Eds.), *Cointegration, causality, and forecasting: A Festschrift in honour of Clive W. J. Granger*. Oxford: Oxford University Press.
Hecq, A., & Urbain, J. P. (1993). Misspecification tests, unit roots and level shifts. *Economics Letters, 43*(2), 129–135.
Jawadi, F., & Prat, G. (2017). Equity prices and fundamentals: A DDM-APT mixed approach. *Journal of Quantitative Finance and Accounting, 49*(3), 661–695.
Kapetanios, G., Shin, Y., & Snell, A. (2003). Testing for a unit root in the nonlinear STAR framework. *Journal of Econometrics, 112*(2), 359–379.
Kim, D., & Perron, P. (2009). Unit root tests allowing for a break in the trend function at an unknown time under both the null and alternative hypotheses. *Journal of Econometrics, 148*(1), 1–13.
Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics, 54*(1–3), 159–178.
Lanne, M., Lütkephol, H., & Saikkonen, P. (2002). Comparison of unit root tests for time series with level shifts. *Journal of Time Series Analysis, 23*(6), 667–685.
Leybourne, S. J., Newbold, P., & Vougas, D. (1998). Unit roots and smooth transitions. *Journal of Time Series Analysis, 19*(1), 83–97.
Li, X. M. (2000). The great leap forward, economic reforms, and the unit root hypothesis: Testing for breaking trend functions in China's GDP data. *Journal of Comparative Economics, 28*(4), 814–827.
Lumsdaine, R. L., & Papell, D. H. (1997). Multiple trend breaks and the unit root hypothesis. *Review of Economics and Statistics, 79*(2), 212–218.
Luukkonen, R., Saikkonen, P., & Terasvirta, T. (1988). Testing linearity against smooth transition autoregressive models. *Biometrika, 75*(3), 491–499.
Montañés, A. (1997). Level shifts, unit roots and misspecification of the breaking date. *Economics Letters, 54*(1), 7–13.
Montañés, A., & Olloqui, I. (1999). Misspecification of the breaking date in segmented trend variables: Effect on the unit root tests. *Economics Letters, 65*(3), 301–307.
Nelson, C., & Plosser, C. (1982). Trends and random walks in macroeconomic time series. *Journal of Monetary Economics, 10*, 139–162.
Ohara, H. I. (1999). A unit root test with multiple trend breaks: A theory and an application to US and Japanese macroeconomic time-series. *Japanese Economic Review, 50*(3), 266–290.
Prat, G., & Uctum, R. (2011). Modelling oil price expectations: Evidence from survey data. *Quarterly Review of Economics and Finance, 51*(3), 236–247.
Phillips, P. C. B., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika, 75*(2), 335–346.
Perron, P. (1989). The great cash, the oil price shock, and the unit root hypothesis. *Econometrica, 57*(6), 1361–1401.
Perron, P. (1990). Testing for a unit root in a time series with a changing mean. *Journal of Business and Economic Statistics, 8*(2), 153–162.
Perron, P. (1997). Further evidence on breaking trend functions in macroeconomics variables. *Journal of Econometrics, 80*(2), 355–385.
Perron, P., & Zhu, X. (2005). Structural breaks with deterministic and stochastic trends. *Journal of Econometrics, 129*(1–2), 65–119.
Rodrigues, P. M. M., & Taylor, A. M. R. (2012). The flexible Fourier form and local generalized least squares de-trended unit root tests. *Oxford Bulletin of Economics and Statistics, 74*(5), 736–759.
Saikkonen, P., & Lütkephol, H. (2002). Testing for a unit root in a time series with a level shift at unknown time. *Econometric Theory, 18*(2), 313–348.
Smyth, R., & Inder, B. (2003). Is Chinese provincial real GDP per capita nonstationary? Evidence from multiple trend break unit root tests. *China Economic Review, 15*(1), 1–24.

Wasserfallen, W. (1986). Non-stationarities in macro-econometric time series-further evidence and implications. *Canadian Journal of Economics, 19*(3), 498–510.

Zivot, E., & Andrews, D. W. K. (1992). Further evidence on the great crash, the oil price shock, and the unit-root hypothesis. *Journal of Business and Economic Statistics, 10*(3), 251–270.

Oil Market Volatility: Is Macroeconomic Uncertainty Systematically Transmitted to Oil Prices?

Marc Joëts, Valérie Mignon, and Tovonony Razafindrabe

Abstract The aim of this contribution is to analyze the impact of macroeconomic uncertainty on the oil market. We rely on a robust measure of macroeconomic uncertainty based on a wide range of monthly macroeconomic and financial indicators, which is linked to predictability rather than to volatility. We estimate a structural threshold vector autoregressive (TVAR) model to account for the varying effect of macroeconomic uncertainty on oil price returns depending on the degree of uncertainty, from which we derive a robust proxy of oil market uncertainty. Our findings show that a significant component of oil price uncertainty can be explained by macroeconomic uncertainty. In addition, we find that the recent 2007–2009 recession has generated an unprecedented episode of high uncertainty in the oil market that is not necessarily accompanied by a subsequent volatility in the price of oil. This result highlights the relevance of our uncertainty measure in linking uncertainty to predictability rather than to volatility.

JEL Classification Q02, E32, C32

M. Joëts
Banque de France, International Macroeconomic Division and EconomiX-CNRS, University of Paris Nanterre, Nanterre, France

V. Mignon (✉)
EconomiX-CNRS, University of Paris Nanterre and CEPII, Nanterre, France

T. Razafindrabe
EconomiX-CNRS and CREM-CNRS, University of Rennes 1, Rennes, France

F. Jawadi (ed.), *Uncertainty, Expectations and Asset Price Dynamics*, Dynamic Modeling and Econometrics in Economics and Finance 24,
https://doi.org/10.1007/978-3-319-98714-9_2

1 Introduction

Uncertainty has been widely documented in the economics literature, particularly its transmission mechanism to activity, which has been extensively discussed both theoretically and empirically. For instance, theories of investment under uncertainty explain why under irreversibility condition or fixed costs, uncertainty over future returns reduces current investment, hiring, and consumption through an "option value to wait."[1] At a micro level, increased uncertainty may diminish the willingness of firms to commit resources to irreversible investment and the readiness of consumers to spend or allocate their earning and wages. Such uncertainty at a micro level may also be transmitted to the macro level, as shown by Bernanke (1983), arguing that uncertainty about the return to investment at a micro level may create cyclical fluctuations in aggregate investment at a macro level. Regarding monetary theory, Prat (1988) investigates the effect of macroeconomic and financial uncertainty on the demand for money. He uses an indicator measuring the degree of economic uncertainty perceived by the agents, as revealed by the behavior of these agents on financial markets. Relying on data over the 1949–1982 period, he emphasizes the importance of uncertainty on the US demand for money.

In the wake of this literature, our aim in this chapter is to investigate the impact of uncertainty on the oil market. Regarding the previous literature, two classes of papers have highlighted the significance of uncertainty in oil prices in explaining economic fluctuations as well as its role in exacerbating asymmetry in the oil price–economic activity nexus.[2] The first strand of studies is based on the volatility of the real price of oil at medium- and long-term horizons, i.e., at horizons that are relevant to purchase and investment decisions. In this vein, the theoretical models of Bernanke (1983) and Pindyck (1991) suggest that oil price uncertainty was the main cause of the 1980 and 1982 recessions.[3] The second strand of the literature is empirical and based on short-term uncertainty. Elder and Serletis (2009, 2010) were among the first to provide a fully specified model for the impact of oil price uncertainty on real GDP. Using a structural vector autoregressive (SVAR) model for post-1980 data, they find empirical evidence that uncertainty about oil price

[1] See Henry (1974), Bernanke (1983), Brennan and Schwartz (1985), Majd and Pindyck (1987), Brennan (1990), Gibson and Schwartz (1990), Bloom et al. (2007), Bloom (2009), Edelstein and Kilian (2009), and Bredin et al. (2011) among others.

[2] Considering the uncertainty channel, the asymmetric responses of real output to oil price shocks may come from the fact that uncertainty tends to amplify the effect of unexpected oil price increases and offset the impact of unexpected oil price decreases (see Kilian, 2014). It is worth mentioning that Georges Prat is an internationally recognized specialist in the analysis of expectations. He has written several contributions on this topic, particularly on rational expectations (see Prat (1994, 1995) and Gardes and Prat (2000) among others). Among his numerous papers, his 2011's article (see Prat and Uctum, 2011) deals with the modeling of expectations on the oil market.

[3] It should be noticed that Bernanke (1983) and Pindyck (1991) consider the oil price as exogenous with respect to the US economy, which is not consistent with the recent literature about the endogenous component of oil prices (see Kilian, 2008a).

evolution tends to depress output, investment, and consumption in the USA and the G-7 countries in an asymmetric way. They also show that the net effect of an unexpected drop in the real price of oil is to cause recessions, a result which is not in line with the economic theory.[4] Relying on a quarterly vector autoregressive (VAR) model with stochastic volatility in the mean, Jo (2014) finds far much smaller uncertainty effect on world industrial production than Elder and Serletis (2010). The author points out that Elder and Serletis (2010)'s GARCH-in mean VAR is misspecified because the same model is driving both the conditional mean and the conditional variance. While providing interesting results, these models suffer from an important drawback regarding the supposed role of oil price shocks in explaining recessions. Indeed, as underlined by Kilian and Lewis (2011) and Kilian (2014), the empirical literature does not consider that recessionary episodes are driven by sequences of oil price shocks of different magnitude and sign. In addition, as argued by Kilian (2014), energy is not necessarily a key component of the cash flow of investment projects, making the effect of oil price uncertainty on output not plausible. Theoretically, the existent measures of price uncertainty also raise some important issues since they are constructed at short-run horizons (month-ahead) rather than at horizons relevant to purchase and investment decisions (years) (see Kilian and Vigfusson, 2011).

As can be seen, the previous literature has faced some limitations in explaining uncertainty in oil prices and its impact on economic activity. The reverse effect, namely, the influence of macroeconomic uncertainty on oil price fluctuations, has not been widely addressed through this question is of particular interest given the modern view that the price of oil is characterized by an important endogenous component. Some exceptions can, however, be mentioned. Regarding first the theoretical papers, (1) Pindyck (1980) discusses the theoretical implications of uncertainty associated with oil demand and reserves on the oil price behavior; (2) Litzenberger and Rabinowitz (1995) analyze backwardation behavior in oil futures contracts; and (3) Alquist and Kilian (2010) allow for endogenous convenience yield and endogenous inventories, and stress that it is uncertainty about the shortfall of supply relative to demand that matters. Turning to the empirical papers, one may refer to (1) Kilian (2009) and Kilian and Murphy (2014) who design as a precautionary demand shock a shock that reflects shifts in uncertainty and treat macroeconomic uncertainty as unobserved; and (2) Van Robays (2013) who investigates whether observed macroeconomic uncertainty changes the responsiveness of the oil price to shocks in oil demand and supply.

Considering the impact of macroeconomic uncertainty on the oil market, our work contributes to this literature and extends it in several ways. First of all and as previously mentioned, though most of the literature has focused on the incidence of oil price uncertainty on economic activity, we address the reverse effect by examining the influence of macroeconomic uncertainty on oil prices. Then, turning

[4]Other empirical papers exist in the literature, such as Favero et al. (1994), Lee et al. (1995), and Ferderer (1996), but they either treat oil as exogenous or use data from the pre-1973 period.

to methodological issues, our contribution is threefold. First, we retain a nonlinear, threshold vector autoregressive (TVAR) specification for modeling oil price returns to account for potentially different effects of macroeconomic uncertainty on the oil market depending on the environment. Second, because macroeconomic uncertainty is unobservable, assessing its effect on the oil market obviously requires us to find an adequate proxy. To this end, we rely on Jurado et al. (2015) and consider a robust approach to measuring macroeconomic uncertainty. The retained proxy uses a wide range of monthly macroeconomic and financial indicators and is based on the underlying idea of a link between uncertainty and predictability. In this sense, we go further than the previous literature[5]—particularly compared with Van Robays (2013)'s paper, which is the closest to ours—which generally relies on dispersion measures such as conditional volatility (e.g., conditional variance of world industrial production growth or of the US GDP growth estimated from a GARCH(1,1) process) or the popular VXO index proposed by Bloom (2009). An important drawback in using GARCH-type models to proxy uncertainty is that they are inherently backward-looking, whereas investors' expectations tend to be forward-looking. Third, we provide a complete analysis by investigating how macroeconomic uncertainty can generate uncertainty in oil prices. To this end, we construct a robust proxy of oil market uncertainty based on macroeconomic uncertainty[6] and provide a historical decomposition that allows us to determine the contribution of macroeconomic uncertainty to oil price uncertainty.

The rest of the chapter is organized as follows. Section 2 presents the methodology and data. Section 3 displays the results regarding the link between macroeconomic uncertainty and uncertainty on the oil market. Finally, Sect. 4 concludes the contribution.

2 Methodology and Data

2.1 Macroeconomic Uncertainty

2.1.1 Measuring Macroeconomic Uncertainty

Measuring uncertainty and examining its impact on market dynamics is a challenging question for economists because no objective measure exists. Although in a general sense uncertainty is defined as the conditional volatility of an unforecastable disturbance,[7] the empirical literature to date has usually relied on proxies. The most

[5] See references in Sect. 2.

[6] This approach is therefore theoretically robust to the endogenous component of commodity prices, in line with the recent literature (see references in Sect. 2).

[7] See Bloom (2009), Bloom et al. (2010, 2012), Gilchrist et al. (2010), Arellano et al. (2011), Bachmann and Bayer (2011), Baker et al. (2011), Basu and Bundick (2011), Knotek and Khan

common measures used are the implied or realized volatility of stock market returns, the cross-sectional dispersion of firm profits, stock returns, or productivity, and the cross-sectional dispersion of survey-based forecasts.[8] However, their adequacy to correctly proxy uncertainty is questionable, and such measures are even misspecified with regard to the theoretical notion of uncertainty, as highlighted by Jurado et al. (2015).

Indeed, stock market volatility, cross-sectional dispersion in stock returns and firm profits can vary over time due to several factors—such as risk aversion, the leverage effect, and heterogeneity between firms—even if there is no significant change in uncertainty. In other words, fluctuations that are actually predictable can be erroneously attributed to uncertainty, putting forward the importance of distinguishing between uncertainty in a series and its conditional volatility. Specifically, properly measuring uncertainty requires to remove the forecastable component of the considered series before computing the conditional volatility. In this sense, uncertainty in a series is not equivalent to the conditional volatility of the raw series.

Another important characteristic of Jurado et al. (2015)'s approach is that macroeconomic uncertainty is defined as the common variation in uncertainty across many series rather than uncertainty related to any single series. This is in line with the uncertainty-based business cycle theories which implicitly assume a common variation in uncertainty across a large number of series.

Accordingly, to provide a consistent measure of macroeconomic uncertainty, we follow the definition of Jurado et al. (2015) by linking uncertainty to predictability. Specifically, the h-period-ahead uncertainty in the variable $y_{jt} \in Y_t = \left(y_{1t}, \ldots, y_{N_y t}\right)'$ is defined as the conditional volatility $U_{jt}^{y}(h)$ of the purely unforecastable component of the future value of the series:

$$U_{jt}^{y}(h) \equiv \sqrt{E\left[\left(y_{jt+h} - E\left[y_{jt+h} \mid J_t\right]\right)^2 \mid J_t\right]}, \tag{1}$$

where $j = 1, \ldots, N_y$, $E(. \mid J_t)$ is the conditional expectation of the considered variable, and J_t denotes the information set available at time t. Uncertainty related to the variable y_{jt+h} is therefore defined as the expectation of the squared error forecast. Aggregating over j individual uncertainty measures $U_{jt}^{y}(h)$ equally weighted by w_j leads to the following expression of aggregate or macroeconomic uncertainty:

$$U_t^{y}(h) \equiv \operatorname{plim}_{N_y \to \infty} \sum_{j=1}^{N_y} w_j U_{jt}^{y}(h) \equiv E_w\left[U_{jt}^{y}(h)\right]. \tag{2}$$

(2011), Fernández-Villaverde et al. (2011), Schaal (2012), Leduc and Liu (2012), Nakamura et al. (2012), Bachmann et al. (2013), and Orlik and Veldkamp (2013) among others.

[8] See Jurado et al. (2015). The papers cited in Sect. 1 related to the study of the relationship between uncertainty and oil prices have used such indicators.

As discussed by Jurado et al. (2015), the estimation of Eqs. (1) and (2) requires three fundamental steps. The first step is to replace the conditional expectation $E\left[y_{jt+h} \mid J_t\right]$ in Eq. (1) by a forecast in order to compute forecast errors. It is a crucial step since the forecastable component should be then removed from the conditional volatility computation.[9] To do so, an as rich as possible predictive model based on factors from a large set of N predictors $\{X_{it}\}$, $i = 1, \ldots, N$, is considered, taking the following approximated form:

$$X_{it} = \Lambda_i^{F\prime} F_t + e_{it}^X, \tag{3}$$

where F_t is an $r_f \times 1$ vector of latent common factors, Λ_i^F is the vector of latent factor loadings, and e_{it}^X is a vector of idiosyncratic errors which allows for some cross-sectional correlations. To account for time-varying omitted-information bias, Jurado et al. (2015) further include estimated factors, as well as nonlinear functions of these factors in the forecasting model through a diffusion forecast index. The second step consists of: (1) defining the h-step-ahead forecast error by $V_{jt+h}^y = y_{jt+h} - E\left[y_{jt+h} \mid J_t\right]$, and (2) estimating the related conditional volatility, namely $E\left[\left(V_{t+h}^y\right)^2 \mid J_t\right]$. To account for time-varying volatility in the errors of the predictor variables, $E\left[\left(V_{t+h}^y\right)^2 \mid J_t\right]$ is recursively multistep-ahead computed for $h > 1$. In the third, final step, macroeconomic uncertainty $U_t^y(h)$ is constructed from the individual uncertainty measures $U_{jt}^y(h)$ through an equally weighted average.

Using large datasets on economic activity, Jurado et al. (2015) provide two types of uncertainty measures that are as free as possible from both the restrictions of theoretical models and/or dependencies on a handful of economic indicators. The first one is the "common macroeconomic uncertainty" based on the information contained in hundreds of primarily macroeconomic and financial monthly indicators, and the second one is the "common microeconomic uncertainty" based on 155 quarterly firm-level observations on profit growth normalized by sales.[10] Empirically, these measures have the advantage of providing far fewer important uncertainty episodes than do popular proxies. As an example, though Bloom (2009) identifies 17 uncertainty periods based on stock market volatility, Jurado et al. (2015) find evidence of only three episodes of uncertainty over the 1959–2011 period: the month surrounding the 1973–1974 and 1981–1982 recessions and the recent 2007–2009 great recession. As stressed above, this illustrates that popular uncertainty proxies based on volatility measures usually erroneously attribute to uncertainty fluctuations that are actually forecastable. In addition, with the proposed measures defined for different values of h, they allow us to investigate uncertainty transmission to the oil market for distinct maturities.

[9]Recall that removing the forecastable component of y_{jt} is crucial to avoid erroneously categorizing predictable variations as uncertain.

[10]Dealing with monthly data and focusing on macroeconomic uncertainty, we consider in this work the "common macroeconomic uncertainty" measure.

2.1.2 Endogenous and Exogenous Components of Uncertainty

One important issue when investigating the impact of macroeconomic uncertainty on oil prices is to understand the intrinsic nature of uncertainty with respect to prices. In other words, it is important to disentangle the endogenous and exogenous components of macroeconomic uncertainty (i.e., whether macroeconomic uncertainty is demand-driven or supply-driven with respect to oil prices). Since 1974, the price of oil—as the price of other commodities—has become endogenous with respect to global macroeconomic conditions (see Alquist et al., 2013).[11] Since then, the empirical literature has provided overwhelming evidence that commodity prices have been driven by global demand shocks.[12] As pointed out by Barsky and Kilian (2002), the 1973–1974 episode of dramatic surge in the price of oil and industrial commodities is the most striking example where the price increase was explained for 25% by exogenous events and for 75% by shifts in the demand side. With the predominant role of flow demand on prices, another important channel of transmission is the role of expectations in the physical market.[13] The underlying idea is that anyone who expects the price to increase in the future will be prompted to store oil now for future use leading to a shock from the demand of oil inventories. Kilian and Murphy (2014) demonstrate that shifts in expectations through oil inventories have played an important role during the oil price surge in 1979 and 1990, and the price collapse in 1986.

The aggregate specification of our proxy has the particularity to be "global," accounting for a lot of information regarding uncertainty in the supply and demand channels. While it is quite difficult in this framework to identify the proportion of unanticipated demand or supply, some reasonable assumptions about the effect of demand and supply shocks on prices may give us some insight about the mechanisms behind the relationship between macroeconomic uncertainty and oil prices. In our analysis, we follow the dominant view about the endogenous nature of oil prices with respect to macroeconomic conditions, considering the aggregate demand channel as a primary source of price fluctuations (see Mabro, 1998, Barsky and Kilian, 2002, 2004, Kilian, 2008a, and Hamilton, 2009). In line with the previous literature, we therefore assume that exogenous events coming from the supply channel—such as cartel decisions, oil embargoes, or the effects of political uncertainty from the Middle East—are secondary, being mainly an indirect consequence of the macroeconomic environment. By construction, our approach accounts for both channels, the demand channel being a direct effect of macroeconomic aggregate and the supply channel an indirect effect of macroeconomic conditions on exogenous events (see Barsky and Kilian, 2002, Alquist and Kilian, 2010, Kilian

[11] Before this date, there was no global market for crude oil and the price of oil in the USA was regulated by the government.

[12] One exception for the case of oil is the 1990s, where the flow supply shocks have played an important role (see Kilian and Murphy, 2014).

[13] See Kilian (2014) for a review.

and Vega, 2011, and Kilian and Murphy, 2014). In other words, our macroeconomic uncertainty proxy primarily reflects uncertainty about the demand side.

2.2 Measuring Oil Market Uncertainty

To investigate how macroeconomic uncertainty can affect oil market uncertainty, we need to: (1) define an uncertainty measure for the oil market, and (2) assess the transmission mechanism of macroeconomic uncertainty to oil market uncertainty.

Let us first consider the determination of the oil market uncertainty proxy. We rely on Eq. (1) and proceed in two steps. In a first step, we account for the result that macroeconomic uncertainty nonlinearly affects the oil price behavior depending on the level of uncertainty (see Joëts et al., 2017). Indeed, we consider that uncertainty may be a nonlinear propagator of shocks across markets, a property which is captured by a structural threshold vector autoregressive model.[14] In addition to providing an intuitive way to capture the nonlinear effects of uncertainty on markets, the TVAR model has the advantage of endogenously identifying different uncertainty states. Indeed, according to this specification, observations can be divided, for example, into two states delimited by a threshold reached by uncertainty, with estimated coefficients that vary depending on the considered state (low- and high-uncertainty states). In other words, the TVAR specification allows uncertainty states to switch as a result of shocks to the oil market.

From the estimation of the TVAR model,[15] we generate the h-period-ahead forecast of the oil price return series, accounting for the information about macroeconomic uncertainty. Let $E\left[y_{t+h}/J_t, u_t^u\right]$ be the obtained forecast, where y is the oil price return series, J_t the information set available at time t, and u_t^u the macroeconomic uncertainty shock at time t. As seen, our forecast value accounts for information about macroeconomic uncertainty. Given this forecast, we define in a second step the h-period-ahead forecast error as the difference between y_{t+h} and $E\left[y_{t+h}/J_t, u_t^u\right]$, the forecast that accounts for information about macroeconomic uncertainty. The underlying idea is that a way to understand the transmission mechanism of macroeconomic uncertainty to the oil market is to assess how the forecast of our considered variable changes if we add information about macroeconomic uncertainty. The oil market uncertainty measure is then given by the volatility of this forecast error.

To account for the volatility-clustering phenomenon, which is a typical feature of commodity markets, we rely on time-varying volatility specifications and consider the moving average stochastic volatility model developed by Chan and Jeliazkov

[14]See Balke (2000) for a detailed presentation of TVAR processes as well as Tong (2010) and Hansen (2011) for general developments on threshold models. Van Robays (2013) uses the TVAR methodology to examine uncertainty on the oil market.

[15]See the detailed results in Joëts et al. (2017).

(2009) and Chan and Hsiao (2013) given by:

$$x_t = \lambda + v_t, \tag{4}$$

where x_t denotes the forecast error, i.e., the difference between the forecast of y that does not account for information about macroeconomic uncertainty and the forecast that accounts for such information. The error term v_t is assumed to be serially dependent, following an MA(q) process of the form:

$$v_t = \varepsilon_t + \psi_1 \varepsilon_{t-1} + \cdots + \psi_q \varepsilon_{t-q}, \tag{5}$$

$$h_t = \lambda_h + \phi_h (h_{t-1} - \lambda_h) + \zeta_t, \tag{6}$$

where $\varepsilon_t \sim N\left(0, e^{h_t}\right)$ and $\zeta_t \sim N\left(0, \sigma_h^2\right)$ are independent of each other, $\varepsilon_0 = \varepsilon_{-1} = \ldots = \varepsilon_{-q+1} = 0$, and the roots of the polynomial associated with the MA coefficients $\psi = \left(\psi_1, \ldots, \psi_q\right)'$ are assumed to be outside the unit circle. h_t is the log-volatility evolving as a stationary AR(1) process. Following Chan and Hsiao (2013), under the moving average extension, the conditional variance of the series x_t is given by:

$$V\left(x_t \mid \lambda, \psi, h\right) = e^{h_t} + \psi_1^2 e^{h_t - 1} + \cdots + \psi_q^2 e^{h_t - q}. \tag{7}$$

This specification allows us to capture two nonlinear channels of macroeconomic uncertainty: (1) the one coming from the moving average of the $q + 1$ most recent variances $e^{h_t} + \cdots + e^{h_t - q}$, and (2) the other from the AR(1) log-volatility stationary process given by Eq. (6).

Given the challenge of estimating this kind of nonlinear model due to high-dimensional and nonstandard data—with the conditional density of the states being non-Gaussian, a Bayesian estimation using Markov chain Monte Carlo methods is hardly tractable. We follow Chan and Hsiao (2013) and estimate the conditional variance of forecast errors by band-matrix algorithms instead of using conventional methods based on the Kalman filter.[16]

[16]See Chan and Jeliazkov (2009) and Chan and Hsiao (2013) for more details. The Matlab code used to estimate the moving average stochastic volatility model is freely available from the website of Joshua Chan. We obtain 20,000 draws from the posterior distribution using the Gibbs sampler after a burn-in period of 1000.

2.3 Data

The oil price series is the monthly WTI crude oil spot price taken from NYMEX, spanning the period from October 1978 to December 2011. The series is transformed into first-logarithmic differences (i.e., price returns). Turning to data related to macroeconomic uncertainty measures for distinct maturities, they are freely available on Ludvigson's homepage.[17] Recall that we rely on the macroeconomic uncertainty measure which is based on several macroeconomic and financial monthly indicators. Specifically, 132 macroeconomic time series are considered, including real output and income, employment and hours, real retail, manufacturing and trade sales, consumer spending, housing starts, inventories and inventory sales ratios, orders and unfilled orders, compensation and labor costs, capacity utilization measures, price indexes, bond and stock market indexes, and foreign exchange measures. Turning to the financial indicators, 147 time series are retained, including dividend–price and earning–price ratios, growth rates of aggregate dividends and prices, default and term spreads, yields on corporate bonds of different ratings grades, yields on Treasuries and yield spreads, and a broad cross-section of industry equity returns. Both sets of data are used to estimate the forecasting factors, but macroeconomic uncertainty is proxied using the 132 macroeconomic time series only.

3 Results

3.1 Transmission of Macroeconomic Uncertainty to Oil Market Uncertainty

Figure 1 depicts the evolution of uncertainty in the oil market for 1 month (blue line),[18] together with the evolution of corresponding prices (black line) and volatility (green line). The horizontal bar corresponds to 1.65 standard deviation above the mean of oil uncertainty series. The gray bands correspond to episodes of important macroeconomic uncertainty: the months surrounding the 1981–1982 recession and the 2007–2009 great recession (see Joëts et al., 2017). When uncertainty in the oil market exceeds the horizontal bar, this refers to episodes of heightened uncertainty for the oil price return series. When oil price uncertainty coincides with the

[17] http://www.econ.nyu.edu/user/ludvigsons/. Since the submission of the present paper, an updated version of the database has been made available in February 2018 and can be downloaded at: https://www.sydneyludvigson.com/data-and-appendixes/.

[18] We focus on short-run uncertainty ($h = 1$) because the effects have been largely documented both theoretically and empirically in the literature (see Bloom, 2014, for a review). For the sake of completeness, we have also estimated uncertainty at longer horizons, namely, 3 and 12 months. The corresponding figures are available upon request to the authors.

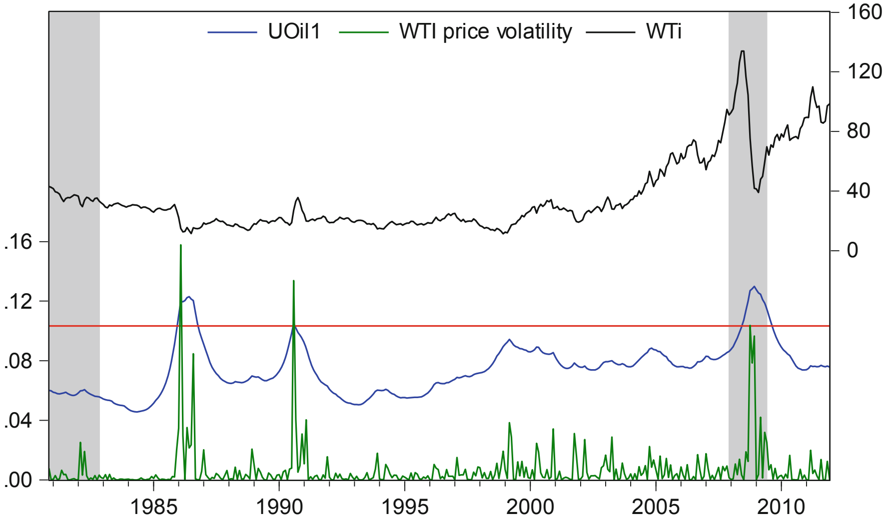


Fig. 1 Uncertainty in the oil market. Note: This figure depicts uncertainty proxy for the oil market at 1 month (left axis, blue line). The horizontal red bar corresponds to 1.65 standard deviation above the mean of the series (left axis). Vertical gray bands represent macroeconomic uncertainty periods. Volatility (green line) is proxied by the daily squared returns of oil prices (left axis). Black line refers to oil price series (right axis)

vertical gray bands, it indicates a potential transfer from macroeconomic to oil market uncertainty (with both uncertainty episodes occurring in the same period). Otherwise, uncertainty is attributable to the own characteristics of the oil market.

As shown in Fig. 1, the sensitivity of oil price uncertainty to macroeconomic uncertainty differs depending on the retained period, highlighting that oil shocks do not all follow the same pattern. For example, the period that just follows the invasion of Kuwait in 1990, the Afghan war in 2001, and the Iraq War in 2002–2003 are episodes characterized by sharp spikes in oil prices. The Iran–Iraq war in 1980 and the 1999 OPEC meeting are, in contrast, associated with small price movements. As stressed by Barsky and Kilian (2004), a simplistic view should be that major war episodes cause price uncertainty to increase through a rise in precautionary demand for oil. However, among all episodes of important fluctuations in oil prices, only two seem to be accompanied by uncertainty: (1) the 2007–2009 recession, and (2) the 1984–1986 period.

During the 2007–2009 recession, oil price uncertainty is indeed very sensitive to macroeconomic uncertainty, a result that is not surprising given the well-known relationship that exists between economic activity and the oil market.[19] This episode of high oil price uncertainty is accompanied by the biggest oil price spike in the

[19] See Barsky and Kilian (2002, 2004), Kilian (2008a,b, 2009), Kilian and Murphy (2012, 2014), and Kilian and Hicks (2013) to name a few.

postwar experience and results from various macroeconomic factors. A common explanation lies in the global economic growth starting in 2003, as illustrated by the increase in real gross world product combined with the stagnant oil production from Saudi Arabia from 2005 to 2007.[20] Whatever the origin of price surges, this period of macroeconomic uncertainty is reflected in oil market uncertainty by an unprecedented oil price increase.

The 1984–1986 period is also characterized by heightened oil price uncertainty, but it does not coincide with macroeconomic uncertainty. This episode seems to be related to the conjunction of two events: (1) the production shutdown in Saudi Arabia between 1981 and 1985, which caused a strong price decrease[21]; and to a lesser extent (2) the OPEC collapse in 1986. On the whole, our results are in line with the literature that has recently stressed the limited impact of exogenous events on oil price fluctuations.

Overall, the recent 2007–09 recession generated an unprecedented episode of uncertainty in the oil price. A key result is that, as clearly shown by Fig. 1, uncertainty episodes are not necessarily accompanied by high volatility in the price of oil. This major finding illustrates the interest of our retained measure of uncertainty, underlining that uncertainty is more related to predictability than to volatility.

3.2 Historical Decomposition Analysis

To assess the contribution of macroeconomic uncertainty to oil price uncertainty, we perform a historical decomposition analysis of oil market uncertainty with respect to macroeconomic uncertainty. Based on the estimation of a VAR model,[22] Fig. 2 reports the historical decomposition associated with oil price uncertainty. Our previous results are confirmed. Indeed, we find a strong proportion of macroeconomic uncertainty in oil price uncertainty during the recent 2007–2009 financial crisis (around 35% of oil price uncertainty is explained by macroeconomic uncertainty during this period). Recalling that our proxy of macroeconomic uncertainty is demand-driven, these conclusions are in line with the literature. During the 1986–1987 episode, oil price uncertainty is not explained by macroeconomic uncertainty and the proportion of macroeconomic uncertainty appears to be negative. This

[20] According to the US Energy Information Administration, the total Saudi Arabia crude oil production significantly decreases from 9550.136 thousand barrels per day in 2005 to 8721.5068 thousand barrels per day in 2007.

[21] At the beginning of the 1980s, the strategy of Saudi Arabia to shut down production (compensating higher oil production elsewhere in the world) was initiated to prevent an oil price decline, without success. Saudi Arabia finally decided to ramp production back up in 1986, causing an oil shock from $27/barrel in 1985 to $12/barrel in 1986 (see Kilian and Murphy, 2014).

[22] The lag order of the VAR specification is 3, as selected by usual information criteria.

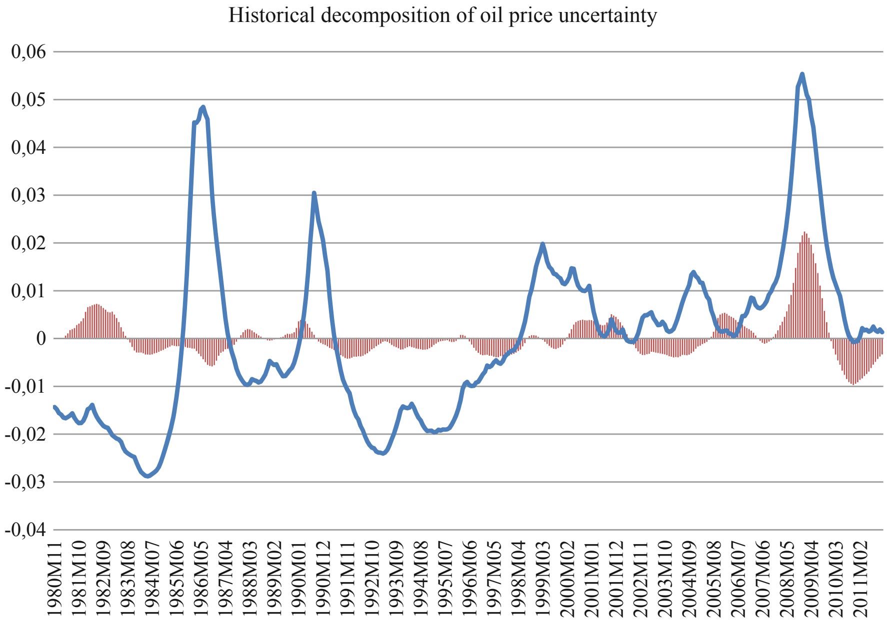


Fig. 2 Historical decomposition of oil price uncertainty with respect to macroeconomic uncertainty. Note: Oil price uncertainty corresponds to the blue line, the proportion of macroeconomic uncertainty is in red

suggests that other shocks not related to economic activity have been at play during this period.

Overall, these results show that a significant component of oil price uncertainty can be explained by macroeconomic uncertainty. A key finding is that the recent oil price movements in 2005–2008 associated with a rise in oil price uncertainty appear to be mainly the consequence of macroeconomic uncertainty, confirming the endogeneity of the oil price with respect to economic activity (i.e., the demand-driven characteristic).[23]

3.3 Distinguishing Between Different Types of Shocks

As stressed above, macroeconomic uncertainty contributes to a large extent to price uncertainty. This result is of primary importance since it shows that oil price uncertainty during the 2005–2008 period can be partly explained by macroeconomic uncertainty. Besides, as stressed by the literature on oil prices, four types of shocks

[23]See Kilian (2008a, 2009), and Kilian and Murphy (2012, 2014) among others.

can be distinguished[24]: (1) shocks to the flow supply of oil, (2) shocks to the flow demand for crude oil reflecting the state of the global business cycle, (3) shocks to the speculative demand for oil stocks above the ground, and (4) other idiosyncratic oil demand shocks. These different shocks may also be reflected in oil price uncertainty movements. Specifically, we aim here at investigating which type of shock contributes the most to oil price uncertainty.

As it is common in the literature, we proxy the flow supply in two ways: by the data on Saudi Arabia crude oil production, and by the global crude oil production—both series being from the Energy Information Agency (EIA).[25] Our measure of fluctuations in global real activity is the dry cargo shipping rate index developed by Kilian (2009). Finally, turning to the speculative component of the oil demand, we rely on data for the US crude oil inventories provided by the EIA.[26] Figures 3, 4, and 5 allow us to assess the quantitative importance of each type of shock (supply, demand, and speculation) on oil price uncertainty. Results show that the contribution of each shock to price uncertainty varies depending on the period, and that the nature of the shock matters in explaining oil price uncertainty. For instance, while Kilian (2009) identified the invasion of Kuwait in 1990 and the Iraq War in 2002–2003 as episodes of surges in speculative demand for oil responsible for sharp price increases, we find that these events are not associated with important oil price uncertainty. More importantly, the contribution of the speculative shocks appears to be very limited or even negative compared to the contribution of flow supply (around 17%) and flow demand (4%) shocks in 1990.

Several events in the oil market history have appeared during the period 1986–1987. The two most important are the decision of Saudi Arabia to shut down the crude oil production to prop up the price of oil, and the OPEC collapse. While the latter is known to have limited impact on the crude oil price (see Barsky and Kilian, 2004), the former created a major positive shock to the flow supply droving down the price of oil. As we have seen, this period has led to an important movement in oil price uncertainty. Figures 3, 4, and 5 show that this event is mainly supply-driven: around 18% of oil price uncertainty is explained by the shut down of Saudi Arabia crude oil production against less than 4% by the flow demand.

The most interesting episode over the last decades is obviously the unprecedented price surge after 2003 and, in particular, in 2007–2008, which led to the most heightened period of oil price uncertainty. According to a popular view, this price increase was the consequence of speculative behaviors on the market (i.e., growing financialization of oil futures markets) and could not be explained by changes in economic fundamentals (see Fattouh et al., 2013, for a discussion). The standard interpretation is that oil traders in spot markets buy crude oil now and store it in

[24]See Kilian (2009), Baumeister and Peersman (2013), and Kilian and Murphy (2012).

[25]We only report the results with Saudi Arabia crude oil production because they are more significant. Results from the global crude oil production are available upon request to the authors.

[26]Similar to Kilian and Murphy (2014), we scaled the data on crude oil inventories by the ratio of OECD petroleum stocks over the US petroleum stocks for each time period.

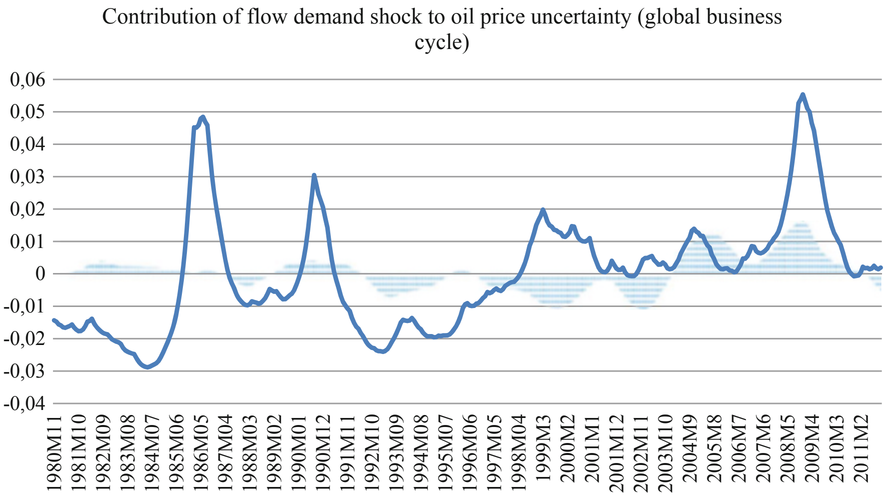


Fig. 3 Historical decomposition of oil price uncertainty with respect to supply and demand shocks. Case of demand shock. Note: This figure depicts the contribution (in blue) of the flow demand (global business cycle) shock to oil price uncertainty. The blue line corresponds to oil price uncertainty at 1 month

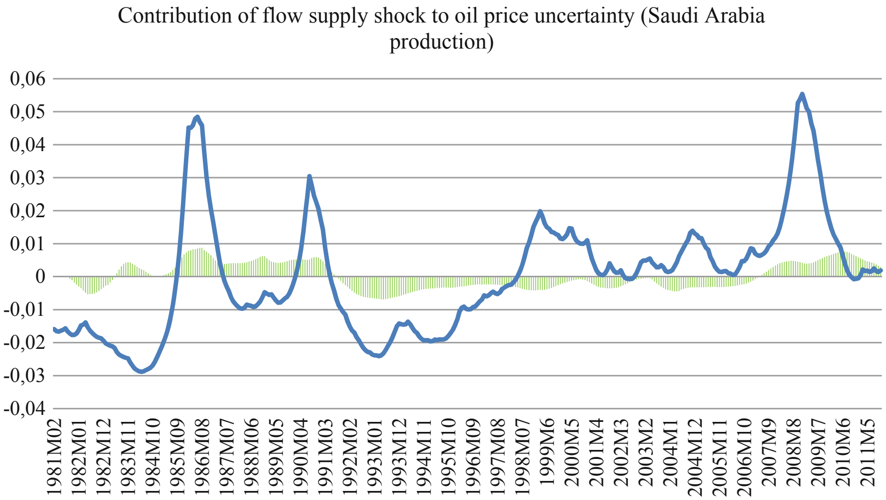


Fig. 4 Historical decomposition of oil price uncertainty with respect to supply and demand shocks. Case of supply shock. Note: This figure depicts the contribution (in green) of the flow supply shock (Saudi Arabia crude oil production) to oil price uncertainty. The blue line corresponds to oil price uncertainty at 1 month

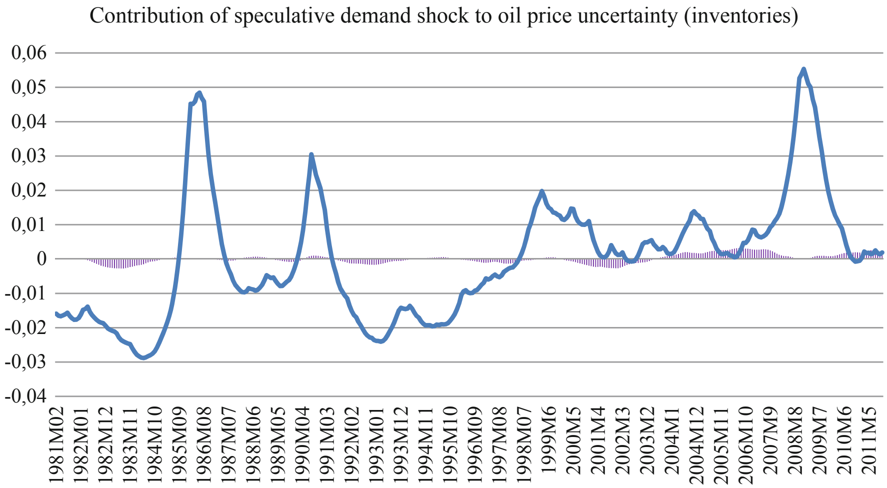


Fig. 5 Historical decomposition of oil price uncertainty with respect to supply and demand shocks. Case of speculative demand shock. Note: This figure depicts the contribution (in purple) of the speculative oil demand shock (inventories) to oil price uncertainty. The blue line corresponds to oil price uncertainty at 1 month

anticipation of higher future oil prices. On the contrary, the recent literature supports the conclusion that the surge in the oil price during this period was mainly caused by shifts in the flow demand driven by the global business cycle (see Kilian, 2009 and Kilian and Hicks, 2013). Our findings corroborate this view that oil price uncertainty has been primarily driven by global macroeconomic conditions. Indeed, as shown in Figs. 3, 4, and 5, the contribution of speculative demand to price uncertainty is very small (around 5%) compared to the proportion of the flow demand from the global business cycle in 2008 (around 40%). An alternative view regarding speculation is that OPEC held back its production by using oil below ground in anticipation of higher oil prices. As discussed by Kilian and Murphy (2014), this behavior would be classified as a negative oil supply shock. Our results provide no evidence that such negative oil supply shocks have significantly contributed to oil price uncertainty, contrary to demand shocks.

Finally, looking at Fig. 6, which reports the simultaneous contribution of each shock (flow demand, flow supply, speculative demand, and macroeconomic uncertainty) to oil price uncertainty, we find that the 2007–2008 period of heightened oil price uncertainty is mainly the consequence of shocks coming from the global business cycle and macroeconomic uncertainty.

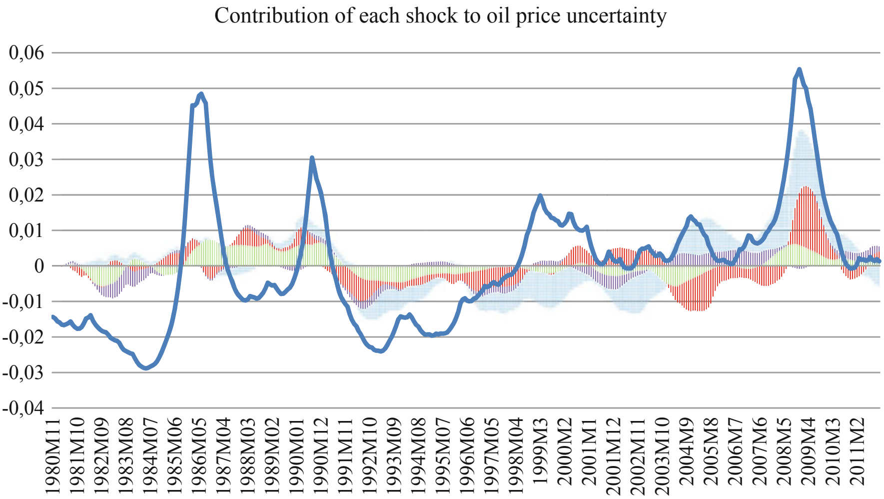


Fig. 6 Historical decomposition of oil price uncertainty with respect to supply and demand, macroeconomic uncertainty shocks. Note: This figure depicts the contribution of each shock to oil price uncertainty. The blue line corresponds to oil price uncertainty at 1 month. The contribution of the flow demand (global business cycle) is in blue, the contribution of the macroeconomic uncertainty at 1 month is in red, the contribution of the flow supply shock (Saudi Arabia crude oil production) is in green, and the contribution of speculative oil demand shock (inventories) is in purple

4 Conclusion

The aim of this contribution is to analyze the impact of macroeconomic uncertainty on the oil market. To this end, we rely on a robust measure of macroeconomic uncertainty based on a wide range of monthly macroeconomic and financial indicators. We also account for nonlinear effects of macroeconomic uncertainty on oil price returns depending on the degree of uncertainty, through the estimation of a threshold VAR model from which we derive a robust measure of oil market uncertainty. We show that the recent 2007–2009 recession generated an unprecedented episode of high uncertainty in the price of oil. In addition, our analysis puts forward that macroeconomic uncertainty episodes are not necessarily accompanied by high volatility in the oil price. This major finding illustrates the interest of our measure of uncertainty, underlining that uncertainty is more related to predictability than to volatility. The relevance of the predictability-based approach could be explained by some specific properties of the oil market. In particular, this market is known to be characterized by a low elastic demand together with a strong inertial supply, making any unexpected adjustment difficult and costly. This importance of factors that are specific to the oil market is in line with the conclusions of the recent World Economic Outlook published by IMF (IMF, 2015) underlining that greater than expected oil supply and some weakness in the demand for oil linked to improvements in energy efficiency have played a key role in explaining the recent oil price collapse.

Acknowledgements This contribution largely relies on Joëts, Mignon, and Razafindrabe (2017). We are grateful to Nathan Balke for providing us with his code for the TVAR estimations. We would like to thank Nick Bloom, Soojin Jo, and Lutz Kilian for their constructive comments and suggestions that helped us improve an earlier version of the work. The usual disclaimers apply.

References

Alquist, R., & Kilian, L. (2010). What do we learn form the price of crude oil futures? *Journal of Applied Econometrics, 25*, 539–573.

Alquist, R., Kilian, L., & Vigfusson, R. J. (2013). Forecasting the price of oil. In G. Elliott & A. Timmermann (Eds.), *Handbook of economic forecasting* (Vol. 2, pp. 427–507). New York: Elsevier.

Arellano, C., Bai, Y., & Kehoe, P. *Financial markets and fluctuations in uncertainty*. (2011). Federal Reserve Bank of Minneapolis Research Department Staff Report.

Bachmann, R., & Bayer, C. (2011). Uncertainty business cycles − really? NBER working paper 16862, National Bureau of Economic Research.

Bachmann, R., Elstner, S., & Sims, E. R. (2013). Uncertainty and economic activity: Evidence from business survey data. *American Economic Journal: Macroeconomics, 5*, 217–249.

Baker, S. R., Bloom, N., & Davis, S. J. (2011). *Measuring economic policy uncertainty*. Unpublished paper, Stanford University.

Balke, N. (2000). Credit and economic activity: Credit regimes and nonlinear propagation of shocks. *The Review of Economics and Statistics, 82*, 344–349.

Barsky, R. B., & Kilian, L. (2002). Do we really know that oil caused the great stagflation? A monetary alternative. In B. Bernanke & K. Rogoff (Eds.), *NBER macroeconomics annual 2001* (pp. 137–183). Cambridge: MIT Press.

Barsky, R. B., & Kilian, L. (2004). Oil and the macroeconomy since the 1970s. *Journal of Economic Perspectives, 18*(4), 115–134.

Basu, S., & Bundick, B. (2011). *Uncertainty shocks in a model of effective demand*. Unpublished paper, Boston College.

Baumeister, C., & Peersman, G. (2013). The role of time-varying price elasticities in accounting for volatility changes in the crude oil market. *Journal of Applied Econometrics, 28*(7), 1087–1109.

Bernanke, B. S. (1983). Irreversibility, uncertainty, and cyclical investment. *The Quarterly Journal of Economics, 98*, 85–106.

Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica, 77*, 623–685.

Bloom, N. (2014). Fluctuations in uncertainty. *Journal of Economic Perspectives, 28*, 153–176.

Bloom, N., Bond, S., & Van Reenen, J. (2007). Uncertainty and investment dynamics. *Review of Economic Studies, 74*, 391–415.

Bloom, N., Floetotto, M., & Jaimovich, N. (2010). Really uncertain business cycles. Mimeo, Stanford University.

Bloom, N., Floetotto, M., Jaimovich, N., Saporta-Eksten, I., & Terry, S. J. (2012). Really uncertain business cycles. NBER working paper 18245, National Bureau of Economic Research.

Bredin, D., Elder, J., & Fountas, S. (2011). Oil volatility and the option value of waiting: An analysis of the G-7. *The Journal of Futures Markets, 31*, 679–702.

Brennan, M. J. (1990). Presidential address: Latent assets. *Journal of Finance, 45*, 709–730.

Brennan, M. J., & Schwartz, E. S. (1985). Evaluating natural resource investments. *The Journal of Business, 58*, 135–157.

Chan, J. C. C., & Hsiao, C. Y. L. (2013). Estimation of stochastic volatility models with heavy tails and serial dependence. CAMA working paper 2013–74, Centre for Applied Macroeconomic Analysis, Crawford School of Public Policy, The Australian National University.

Chan, J. C. C., & Jeliazkov, I. (2009). Efficient simulation and integrated likelihood estimation in state space models. *International Journal of Mathematical Modelling and Numerical Optimisation, 1*, 101–120.

Edelstein, P., & Kilian, L. (2009). How sensitive are consumer expenditures to retail energy prices? *Journal of Monetary Economics, 56*(6), 766–779.

Elder, J., & Serletis, A. (2009). Oil price uncertainty in Canada. *Energy Economics, 31*, 852–856.

Elder, J., & Serletis, A. (2010). Oil price uncertainty. *Journal of Money, Credit and Banking, 42*, 1137–1159.

Fattouh, B., Kilian, L., & Mahadeva, L. (2013). The role of speculation in oil markets: What have we learned so far? *The Energy Journal, 34*(3), 7–33.

Favero, C. A., Pesaran, M. H., & Sharma, S. (1994). A duration model of irreversible oil investment: Theory and empirical evidence. *Journal of Applied Econometrics, 9*, 95–112.

Ferderer, P. J. (1996). Oil price volatility and the macroeconomy: A solution to the asymmetry puzzle. *Journal of Macroeconomics, 18*, 1–16.

Fernández-Villaverde, J., Rubio-Ramírez, J. F., Guerrón-Quintana, P., & Uribe, M. (2011). Risk matters: The real effects of volatility shocks. *American Economic Review, 6*, 2530–2561.

Gardes, F., & Prat, G. (Eds.). (2000). *Price expectations in goods and financial markets, new developments in theory and empirical research*. Cheltenham: Edward Elgar.

Gibson, R., & Schwartz, E. S. (1990). Stochastic convenience yield and the pricing of oil contingent claims. *Journal of Finance, 45*, 959–976.

Gilchrist, S., Sim, J. W., & Zakrajšek, E. (2010). *Uncertainty, financial frictions, and investment dynamics*. Unpublished manuscript, Boston University.

Hamilton, J. D. (2009). Causes and consequences of the oil shock of 2007–08. *Brookings Papers on Economic Activity, 40*, 215–283.

Hansen, B. E. (2011). Threshold autoregression in economics. *Statistics and Its Interface, 4*, 123–127.

Henry, C. (1974). Investment decisions under uncertainty: The 'irreversibility effect'. *The American Economic Review, 64*, 1006–1012.

IMF. (2015). World Economic Outlook.

Jo, S. (2014). The effect of oil price uncertainty on global real economic activity. *Journal of Money, Credit, and Banking, 46*(6), 1113–1135.

Joëts, M., Mignon, V., & Razafindrabe, T. (2017). Does the volatility of commodity prices reflect macroeconomic uncertainty? *Energy Economics, 68*, 313–326.

Jurado, K., Ludvigson, S., & Ng, S. (2015). Measuring uncertainty. *American Economic Review, 105*(3), 1177–1215.

Kilian, L. (2008a). The economic effects of energy price shocks. *Journal of Economic Literature, 46*(4), 871–909.

Kilian, L. (2008b). Exogenous oil supply shocks: How big are they and how much do they matter for the U.S. economy? *Review of Economics and Statistics, 90*, 216–240.

Kilian, L. (2009). Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *American Economic Review, 99*, 1053–1069.

Kilian, L. (2014). Oil price shocks: Causes and consequences. *Annual Review of Resource Economics, 6*, 133–154.

Kilian, L., & Hicks, B. (2013). Did unexpectedly strong economic growth cause the oil price shock of 2003–2008? *Journal of Forecasting, 32*(5), 385–394.

Kilian, L., & Lewis, L. T. (2011). Does the fed respond to oil price shocks? *Economic Journal, 121*, 1047–1072.

Kilian, L., & Murphy, D. P. (2012). Why agnostic sign restrictions are not enough: Understanding the dynamics of oil market VAR models. *Journal of European Economic Association, 10*(5), 1166–1188.

Kilian, L., & Murphy, D. P. (2014). The role of inventories and speculative trading in the global market for crude oil. *Journal of Applied Econometrics, 29*, 454–478.
Kilian, L., & Vega, C. (2011). Do energy prices respond to U.S. macroeconomic news? A test of the hypothesis of predetermined energy prices. *Review of economics and statistics, 93*(2), 660–671.
Kilian, L., & Vigfusson, R. J. (2011). Nonlinearities in the oil price-output relationship. *Macroeconomic Dynamics, 15*(3), 337–363.
Knotek, E. S., & Khan, S. (2011). How do households respond to uncertainty shocks? *Federal Reserve Bank of Kansas City Economic Review, 96*, 5–34.
Leduc, S., & Liu, Z. (2012). Uncertainty shocks are aggregate demand shocks. Federal Reserve Bank of San Francisco, Working paper 2012–10.
Lee, K., Ni, S., & Ratti, R. A. (1995). Oil shocks and the macroeconomy: The role of price variability. *The Energy Journal, 16*, 39–56.
Litzenberger, R. H., & Rabinowitz, N. (1995). Backwardation in oil futures markets: Theory and empirical evidence. *The Journal of Finance, 50*, 1517–1545.
Mabro, R. (1998). *The oil price crisis of 1998*. Oxford: Oxford Institute for Energy Studies.
Majd, S., & Pindyck, R. S. (1987). The learning curve and optimal production under uncertainty. MIT Working paper 1948–87, Cambridge, MA: MIT Press.
Nakamura, E., Sergeyev, D., & Steinsson, J. (2012). Growth-rate and uncertainty shocks in consumption: Cross-country evidence. Working paper, Columbia University.
Orlik, A., & Veldkamp, L. (2013). *Understanding uncertainty shocks*. Unpublished manuscript, New York University Stern School of Business.
Pindyck, R. S. (1980). Uncertainty and exhaustible resource markets. *The Journal of Political Economy, 88*, 1203–1225.
Pindyck, R. S. (1991). Irreversibility, uncertainty and investment. *Journal of Economic Literature, 29*, 1110–1148.
Prat, G. (1988). Note à propos de l'influence de l'incertitude sur la demande de monnaie. *Revue Economique, 39*(2), 451–460.
Prat, G. (1994). La formation des anticipations boursières, Etats-Unis, 1956 à 1989. *Economie et Prévision, 1*, 101–125.
Prat, G. (1995). La formation des anticipations et l'hypothèse d'un agent représentatif : quelques enseignements issus de simulations stochastiques. *Revue d'Economie Politique, 105*(2), 197–222.
Prat, G., & Uctum, R. (2011). Modelling oil price expectations: Evidence from survey data. *Quarterly Review of Economics and Finance, 51*(3), 236–247.
Schaal, E. (2012). *Uncertainty, productivity, and unemployment in the great recession*. Unpublished paper, Princeton University, Princeton, NJ.
Tong, H. *Threshold models in time series analysis-30 years on*. (2010). Research Report 471, University of Hong Kong.
Van Robays, I. (2013). Macroeconomic uncertainty and the impacts of oil shocks. ECB working paper 1479, European Central Bank.

Part II
Heterogeneity of Beliefs and Information

Heterogeneous Beliefs and Asset Price Dynamics: A Survey of Recent Evidence

Saskia ter Ellen and Willem F. C. Verschoor

Abstract This contribution reviews the empirical literature on heterogeneous beliefs and asset price dynamics that challenges the traditional rational agent framework. Emphasis is given to the validation and estimation of (dynamic) heterogeneous agent models that have their roots in the agent-based literature. Heterogeneous agent models perform well in describing, explaining and often forecasting asset markets dynamics, such as equities, foreign exchange, credit, housing, derivatives and commodities. Our survey suggests that heterogeneous agent models have the ability to produce important stylised facts observed in financial time series and to replicate important episodes of financial turmoil.

1 Introduction

In recent decades, we have seen an increase in the number of studies that attempt to explain asset price dynamics in financial markets. Expectations are crucial in this respect, and theories of the expectations formation process have been at the forefront of economic research in the financial economic literature. Muth's (1961) 'rational expectations hypothesis' (REH) has attracted the greatest attention and states that market participants have equal access to information and form their expectations

This chapter should not be reported as representing the views of Norges Bank. The views expressed are those of the authors and do not necessarily reflect those of Norges Bank. We would like to thank Elisabeth Werenskiold and Lein M. Hansen for providing research assistance, and seminar participants at Norges Bank for valuable comments and discussion.

S. ter Ellen (✉)
Norges Bank, Oslo, Norway
e-mail: Saskia.ter-Ellen@Norges-Bank.no

W. F. C. Verschoor
Tinbergen Institute and VU University, Amsterdam, Netherlands

F. Jawadi (ed.), *Uncertainty, Expectations and Asset Price Dynamics*,
Dynamic Modeling and Econometrics in Economics and Finance 24,
https://doi.org/10.1007/978-3-319-98714-9_3

about future events in a uniform, rational manner based on the 'true' probability of the state of the economy. Whereas classical economic models often assume these expectations to be rational and therefore conveniently summarised by a representative, perfectly rational agent, there is an interesting and promising new literature in the direction of bounded rationality, and the accompanying heterogeneity of agents' expectations. The notion of rational expectations is losing more and more ground and new insights on how economic agents form their expectations is therefore warranted. As it turns out, economic models that incorporate a behavioural, agent-based approach are better able to explain financial market asset price dynamics than are models based on a representative rational agent.

In this work, we will provide an overview of the empirical literature that acknowledges and incorporates the heterogeneous agents approach that challenges the traditional rational agent framework. More specifically, our focus is on the validation and estimation of (dynamic) heterogeneous agent models (HAM) that have their roots in the agent-based literature. This branch of behavioural finance assumes that agents are at least boundedly rational (Simon, 1957), and that they use certain rules of thumb in order to form expectations about future asset prices. This setup goes back to Zeeman (1974), and was further advanced by, among others, Frankel and Froot (1987), Chiarella (1992), Brock and Hommes (1997, 1998), Lux (1998) and De Grauwe and Grimaldi (2006). Although different names are being used in the literature for different forecasting strategies, they roughly come down to two or three types of agents. One typical type of agent uses past (price) information in order to predict future returns. The strategy this agent uses is referred to as (trend) extrapolation, technical analysis, bandwagon (for positive trend extrapolation), contrarian (for trend reversion) or chartism. The second type of agent bases his expectations on the deviation of the asset price from its fundamental value. This agent is said to be mean reverting, regressive or fundamentalist. Third and fourth types differ among studies and markets, such as carry traders (Pojarliev and Levich, 2008; Spronk et al., 2013).

Although several studies survey the theoretical work on this type of models (Hommes, 2006; LeBaron, 2000; Chiarella et al., 2009, among others), there is a gap in the literature when it comes to surveying empirical work. Our purpose is to present a comprehensive review of the empirical findings and recent developments of estimation designs put forth over the past two decades. Heterogeneous agent models perform very well in describing, explaining, and often forecasting (financial) market's dynamics: they have been used to explain asset price dynamics in equities, foreign exchange, bonds, housing, derivatives, commodities and even macroeconomic variables.[1] In order to make the results comparable, Ter Ellen et al. (2017) estimate a generic heterogeneous agent model on a variety of asset classes and find support for heterogeneity of market participants for all asset classes but equities.

[1] They have also proven to be very well able to explain and replicate certain stylised facts of financial markets (Lux, 2009), such as volatility clustering, fat tails, and bull and bear markets.

Moreover, they find that heterogeneity is more pronounced for macroeconomic variables and that these are more prone to behavioural bubbles than financial assets.

The remainder of this chapter is as follows. Section 2 provides a brief description of how the field developed from rational agent models to models with boundedly rational, heterogeneous agents. Section 3 presents the first theoretical contributions that have been made and some of the empirical support from experiments and survey studies. In Sect. 4, the focus of attention is turned to the challenges in empirically measuring heterogeneous agent models for a variety type of asset classes and estimation methods. Section 5 concludes the survey.

2 From Rational Expectations to Bounded Rationality

2.1 Efficient Markets

The rationality of agents' expectations has been at the forefront of economic research in the financial economics literature. As such, expectations are the driving force in the (financial) marketplace. Modelling these expectations as rational has the convenient attribute that in such a case '[expectations] are essentially the same as the predictions of the relevant economic theory' (Muth, 1961). Fama (1965) argued that financial markets are efficient because of rational behaviour and expectations of economic agents, and that market efficiency (EMH) requires that actual prices (or rates of return) follow a 'fair game' process relative to expected equilibrium prices (or rates of return). The assumption of rational agents implies that agents incorporate all available information in their decision-making process and that they are able to do this in an efficient way because they have full knowledge about the economic models underlying financial markets. This means that all agents should have the same expectations and that all prices of (financial) products should reflect their fundamental values. It is acknowledged that some agents might not be rational and that therefore mispricing may occur. However, overreaction of some agents will be offset by underreaction of other agents. Moreover, according to Friedman (1953), possible mispricing caused by the so-called noise traders will soon vanish through the actions of rational agents. He argues that in such a way, speculators keep foreign exchange markets stable and efficient in case of a flexible exchange rate system. The concept of arbitrage, as described by Friedman, is one of the main fundaments of the EMH. It entails that rational agents will observe mispricing and take actions upon it. Therefore, noise traders do not have a significant effect on prices, and it is impossible to consistently beat the market and earn riskless returns. If arbitrage opportunities exist, rational agents would pick upon these and trade upon them. In other words, 'there's no such thing as a free lunch'.

Although the efficient market hypothesis has been the conventional way of thinking about asset pricing on financial markets at least since the seventies, it has also been a target of criticism since its publication. An important reason for the criticism is that the theory has some internal contradictions. If agents are rational and thus have the same expectations, there would be no trade in financial securities at all. With transaction costs taken into account and prices being perfect reflections of all (available) information no agent would either want to sell or buy its assets, since no extra returns can be made with that transaction. Milgrom and Stokey (1982) show that even when some agents have private information, this 'no-trade theorem' applies. The fact that trade does take place, and in large and growing amounts, is one of the observations that weaken the EMH.

2.2 Limits of the EMH

The debate regarding the validity of the efficient market hypothesis is a long and standing one. With the arrival of several anomalies that are puzzling from the perspective of purely rational models, such as the forward premium puzzle, the equity premium puzzle or the excess trade volume, the notion of the rational expectations hypothesis is losing more and more ground. The finding of excessive trading (Milgrom and Stokey, 1982) poses a challenge to the hypothesis that investors are rational. Other observed market anomalies that are difficult to explain in the conventional setup are, for example, momentum effect (Jegadeesh, 1990, on the short term recent losers tend to underperform the market, recent winners tend to outperform the market), post-earnings announcement drift (Ball and Brown, 1968, prices do not adjust to information immediately but adjust slowly, causing a positive drift after positive news and a negative drift after disappointing news), long-term reversal (De Bondt and Thaler, 1985, extreme past losers tend to outperform the market, past winners tend to underperform the market), size effect (Black et al., 1972, small-firm stocks outperform stocks of large companies), excess volatility (Shiller, 1981) and foreign exchange rate puzzles (e.g. reversed evidence on purchasing power parity and interest parity).

Another explanation for the persistence of mispricing that can be found in the literature is that there are serious limits to arbitrage. Grossman and Stiglitz (1980) argue that if arbitrage is costly (which it inherently is), it cannot be the case that a competitive economy is always in equilibrium, as that would mean that arbitrageurs would not be able to make returns. Among others, De Long et al. (1990) introduce noise trader risk to explain why arbitrage opportunities cannot always be fully exploited. They argue that the existence of noise traders (i.e. traders whose trading decisions are based on non-fundamental information: noise) in the market brings along a significant amount of uncertainty that affects the riskiness of arbitrage. After all, if the effect of noise traders was strong enough to create the mispricing, these traders could as well increase the gap even further. Therefore, noise traders can heavily destabilise the market. According to the EMH, mispricing cannot persist because it creates the possibility of a riskless return that would immediately be exploited. However, if the profit opportunity is not riskless because of the unpredictable behaviour of noise traders, the mispricing can persist. This limit to arbitrage is usually labelled 'noise trader risk', but there can be other risks that limit arbitrage opportunities.

Still, limits to arbitrage are no explanation of exchange rate puzzles, the inefficiency of markets and the inherent mispricing. After all, it does not explain how mispricing can occur in the first place. Results from psychology and sociology have given some insight in the non-rational beliefs of investors which may help to understand the observed anomalies in financial markets.

2.3 Survey Evidence and Bounded Rationality

Although these contributions from the field of psychology are an important insight in the actual behaviour of people and clearly show that agents do not behave in a rational way, they have generated quite some skepticism. After all, most economists already knew from the start that not all investors behave fully rationally, but they consider this as a necessary assumption to include investor behaviour in sophisticated economic models. They argued that behavioural economics and behavioural finance were impractical bifurcations of economics, since it was impossible to model the complex behaviour of human beings. On top of that, the results from psychology were mainly generated by laboratory experiments which did not always replicate the real world in a very accurate way. These difficulties were reinforced by the problem that we could only observe price reactions to human behaviour instead of observing actual expectations of future asset prices.

The latter problem was partly overcome in the eighties, when companies like Money Market Services International (MMSI) and Consensus Economics started to gather investors' expectations of future asset prices by means of surveys. The use of survey data allows researchers to directly observe investors' expectations about future prices and exchange rates,[2] therefore making it easier for them to test investor rationality and information efficiency and to detect possible expectation formation mechanisms that are used by institutional investors. Early work by Blake et al. (1972), Dominguez (1986) and Frankel and Froot (1987) utilises such survey-based expectations, and many studies have used some form of survey measures of expectations in explaining foreign exchange rate puzzles after that. For example, MacDonald (1990a), MacDonald and Marsh (1996), Cavaglia et al. (1993) and Ito (1990) have used foreign exchange rate survey data in examining the rationality of exchange rate expectations and have concluded that respondents give biased forecast that do not efficiently capture publicly available information such as past interest rate movements.

The EMH incorporates the joint hypothesis that expectations are formed rationally and that market participants are risk neutral with respect to investing in

[2]Jongen et al. (2008) provide an excellent overview of the literature on expectations in foreign exchange markets.

domestic or foreign assets (Jongen et al., 2008). Several possible explanations for the failure of the forward rate as an unbiased estimate for future spot rates have been put forward in the financial economics literature (see Engel (1996), MacDonald (1990b) and Jongen et al. (2008), for instance). The main competing views are that the unbiasedness stems from irrational behaviour of exchange rate forecasters (Bilson, 1981; Cumby and Obstfeld, 1984, for instance), versus the existence of a time-varying risk premium (Fama, 1984; Hsieh, 2017; Wolff, 1987). However, the inherently necessary use of joint tests of rationality and for the existence of a risk premium made it impossible to distinguish between these causes of the forward premium bias. Survey-based expectations are a useful tool in this respect, as they allow us to decompose the forward premium into an 'irrational expectations' component and a 'time-varying risk premium' component. The literature suggests that both irrational expectations and time-varying risk premiums account for the forward discount anomaly (Froot, 1989; Froot and Thaler, 1990; Cavaglia et al., 1994, for instance). With the arrival of irrational expectations, the focus is shifting in the direction of expectation formation mechanisms. Three alternative models of expectation formation are mainly considered in the literature—the extrapolative, the regressive and the adaptive—against the null hypothesis that expectations are static. Whereas many of the studies focus on expectations following one of these specifications at a time, Prat and Uctum (2007) show that survey respondents use a combination of these rules.

When analysing the process of expectations formation, it appears that the longer the forecast horizon, the more exchange rate expectations reverse recent price trends. At horizons exceeding one month, expectations appear to stabilise and regress towards their equilibrium values. However, at horizons up to approximately one month agents extrapolate the most recent trend and diverge from their hypothesised long-run equilibrium values (Frankel and Froot, 1987, 1990a; Cavaglia et al., 1993; Ito, 1990). Prat and Uctum (2015) find that although the share of fundamentalists indeed increases with forecasting horizon, chartists always dominate.

2.4 *Boundedly Rational Heterogeneous Agents Models*

Although survey studies provided evidence to reject the assumptions of rational expectation formation and information efficiency, the problem of modelling behaviour persisted. As a response, some authors started incorporating certain aspects of the investors' behaviour in their models. In their contribution, Barberis et al. (1998) propose a parsimonious model of how investors form beliefs that is consistent with the available statistical and psychological evidence. In their 'model of investor sentiment', they include conservatism and representativeness to explain under- and overreaction of stock prices. Almost parallel to that, boundedly rational heterogeneous agents models (BRHA models, or HAM) were developed. This heterogeneous agents theory, originally founded by Zeeman (1974), Beja and Goldman (1980) and Frankel and Froot (1987) and further developed by, among others, Brock and Hommes (1997, 1998), Day and Huang (1990), Chiarella (1992) and De Grauwe et al. (1993) rejects the idea that investors behave rationally.

With some exceptions, these investigations have in common that the distinction they make is one between a fundamental approach in forming expectations and an extrapolative approach, which is usually referred to as 'technical analysis' or 'chartist behaviour'. Furthermore, some of the models assume that agents switch between the two strategies, depending on the forecasting performance or profitability of a certain strategy.

Fundamentalists base their expectations on economic theory about future asset prices and their trading strategy upon market fundamentals. They believe that the market price will revert to the intrinsic value of an asset and therefore bases expectations on the deviation of the market price from the fundamental economic value. In contrast, technical traders, or chartists, base their expectations on past price behaviour and try to extrapolate the trend in the most recent period(s). They expect trends to continue in the same direction and exploit these historical patterns in their investment decisions. Fundamentalist behaviour is generally found to have a stabilising effect on prices, while chartists tend to have a destabilising effect driving asset prices away from the intrinsic value of the asset.

3 Early Contributions and Supporting Evidence

3.1 Early Contributions

One of the earliest examples of a heterogeneous agent model that we can find in the literature is Zeeman (1974). He recognises and distinguishes two types of agents in the stock market, similar to the ones used in the 'modern-day' heterogeneous agent models. One group, chartists, chases trends, therefore buying when prices go up and selling when prices go down. The other group, fundamentalists, is aware of the true fundamental value, and buys (sells) when the stock is currently undervalued (overvalued). Zeeman explains the slow feedback flow observed in the stock market by the fact that the rate of change of stock market indices responds to chartist and fundamentalist demand faster than their demand responds to the return changes of these indices. In other words, while chartists and fundamentalists demand has a direct effect on returns, fundamentalists may only start selling when a stock is overvalued by a certain amount, thereby causing bull (chartists driving the price up) and bear (both chartists and fundamentalists selling stocks) markets. Although Zeeman's model is very similar in terms of set-up and implications to the heterogeneous agent models as we know them now, it lacked clear micro-foundations (Hommes, 2006) and his theory was not picked up at the time.

Another important contribution came from Beja and Goldman (1980). According to them, it is obvious that a man-made market where people interact and respond to each other cannot be fully efficient. Therefore, discrepancies will exist and human beings will naturally respond to these discrepancies by speculating on their expected direction of the market. Since this is bound to lead to different price dynamics than would occur under the efficient markets hypothesis, they propose an alternative theory. In line with Zeeman (1974), Beja and Goldman (1980) assume a mechanism where the speed of price changes and the speed of demand changes are not in line. Furthermore, they propose a market which consists of fundamental demand (based on the expectations of future equilibrium prices) and speculative demand (based on the state of the market). Dynamics in the aggregate demand especially occur due to relative sizes of the fundamental and speculative demand (which becomes larger if the price change is larger than expected) and the flexibility of the trend followers. The market will be stable if the impact of the fundamental demand is sufficiently high or if the impact of the trend followers is sufficiently low.

The heterogeneous agents literature has thereafter benefitted a lot from contributions from, among others, Frankel and Froot (1987, 1990a,b) and Brock and Hommes (1997, 1998). Frankel and Froot showed, by using survey data, that expectations could be classified as extrapolative, regressive and adaptive (1987), or as chartist and fundamentalist (1990a). Brock and Hommes (1997, 1998) introduced an intuitive switching rule, effectively implying that investors would switch to the rule with the best recent performance. HAM have been very well able to explain and replicate certain stylised facts of financial markets (Lux, 2009), such as volatility clustering, fat tails, and bull and bear markets. For comprehensive overviews of the (theoretical) HAM literature, see, for example, Hommes (2006), Chiarella et al. (2009) and LeBaron (2000).[3]

3.2 Supporting Evidence on the Micro-Level

Over the years, studies have collected empirical evidence in favour of the chartist–fundamentalist approach in various ways. In this section, we will discuss some of the evidence collected on the micro-level, of which the majority comes from laboratory experiments and survey studies.

Schmalensee (1976) was one of the first to use experimental methods to reveal expectation formation processes for time series, in particular with respect to technical rules. Smith et al. (1988) are able to replicate bubbles and crashes in a laboratory environment. De Bondt (1993) and Bloomfield and Hales (2002) use classroom experiments and find evidence of trend-following behaviour, where the latter also find support for the assumption in Barberis et al. (1998) that investors perceive past trend reversals as an indicator for the probability of future reversals

[3] Not all papers on HAM estimation are positive about the use and appropriateness of such models. Amilon (2008) uses maximum likelihood and efficient method of moments and finds that the models generally have a poor fit and do not generate all the stylised facts that some of the simulation studies are able to match.

even though they are aware of the random walk character. A laboratory experiment is used by Hommes et al. (2005) to evaluate how subjects form expectations when all they know is dividend yield, interest rates and past realised prices. The authors find that participants make use of very similar linear rules, such as autoregressive or adaptive strategies, in forming expectations. Assenza et al. (2014) provide an excellent summary of the relevant experimental work in this field.

As (laboratory) experiments are, in general, not fully able to replicate the real world situation, and their generalisability has therefore been questioned, attempts have been made to directly measure investor expectations and expectation formation rules. To this end, both quantitative and qualitative surveys have been conducted. Taylor and Allen (1992) show, based on a questionnaire survey, that 90% of the foreign exchange dealers based in London use some form of technical analysis in forming expectations about future exchange rates, particularly for short-term horizons. The foreign exchange dealers further stated that they see fundamental and technical analyses as complementary strategies for making forecasts and that technical analysis can serve as a self-fulfilling mechanism. Menkhoff (2010) gathered similar data from fund managers in five different countries. In line with the findings of Taylor and Allen, he finds that 87% of the fund managers surveyed use technical analysis. About 20% of the fund managers consider technical analysis as more important than fundamental analysis. Various quantitative surveys have been evaluated as well. For a more extensive overview, see Jongen et al. (2008). Frankel and Froot (1987, 1990a,b) have had a substantial impact on the foreign exchange literature and the further development of heterogeneous agent models. They were among the first to show that survey data reveals non-rationality and heterogeneity of investors. They also find evidence for the chartist–fundamentalist approach employed in many of the heterogeneous agent models. Others have confirmed these findings in later years, and with various datasets. Dick and Menkhoff (2013) use forecasters' self-assessment to classify themselves as chartists, fundamentalists or a mix. They find that forecasters who classify their forecasting tools as chartist use trend-following strategies and who classify as fundamentalist have a stronger preference for purchasing power parity (PPP). They also find that chartists update their forecasts more frequently than fundamentalists.

Ter Ellen et al. (2013) are among the first to estimate a full dynamic heterogeneous agent model (HAM) on survey data, meaning that the expectations of investors can be dynamic in various ways. They find that three forecasting rules fit the survey data very well: a PPP rule (fundamentalist), a momentum rule (chartist) and an interest parity rule. They confirm the earlier finding from Frankel and Froot (1990a,b) that investors use more speculative strategies for shorter horizons (1 month) and more fundamental strategies for longer horizons (12 months). Moreover, investors switch between forecasting rules depending on the past performance of these rules. Goldbaum and Zwinkels (2014) find that a model with fundamentalists and chartists can explain the survey data well. As in Ter Ellen et al. (2013), they find that fundamentalists are mean reverting and that this model is increasingly used for longer horizons. Chartists have contrarian expectations. A model with time-varying weights on the different strategies outperforms a static version of this model. Jongen et al. (2012) also allow the weights on different strategies to vary depending on market circumstances. However, instead of directly explaining the survey expectations, they analyse the dispersion between forecasts. They find that the dispersion is caused by investors using heterogeneous forecasting rules and having private information. This is in line with the earlier findings of Menkhoff et al. (2009) for a dataset on German financial market professionals.

Zwinkels and co-authors have collected evidence for heterogeneous beliefs from data on fund managers' exposure. Verschoor and Zwinkels (2013) show that foreign exchange fund managers behave like heterogeneous agents. They find that fund managers allocate capital to a momentum, carry and value strategy depending on the past performance of these strategies. They make money by employing a negative feedback strategy: shifting money from recent winning strategies to recent losing strategies. Schauten et al. (2015) apply a heterogeneous agent model to hedge fund risk exposure. Because of the non-linear trading strategies that hedge fund managers employ, a non-linear model with dynamic weights seems to be appropriate to capture the hedge fund risk exposure. The heterogeneity of the hedge funds lies in the dynamic weighting of exposure to different risk factors.

3.3 An Example

We will now provide an example of a heterogeneous agent model with chartists, fundamentalists, and dynamic weighting of the two groups. Many of the models employed can be simplified to this model. The form of the model we show here is mostly related to some of our own applications of HAM (e.g. De Jong et al., 2010; Ter Ellen and Zwinkels, 2010; Chiarella et al., 2014), which are largely based on the functional form from Brock and Hommes (1997, 1998) and Boswijk et al. (2007).

The base of the model is the price of an asset. The price of an asset tomorrow, P_{t+1}, equals the price of today, P_t, and the weighted demand of different types of agents, typically chartists and fundamentalists[4]:

$$P_{t+1} = P_t + W_t D_t^c + (1 - W_t) D_t^f \tag{1}$$

Here, W_t is the chartist weight in the market, D_t^c is the chartist demand, $(1 - W_t)$ is the weight of fundamentalists in the market, and D_t^f is the demand function of fundamentalists. The demand functions can be specified as the difference between the current asset price and the expected asset price under chartist ($E_t^c[P_{t+1}]$) or fundamentalist ($E_t^f[P_{t+1}]$) expectations:

$$D_t^c = a^c (E_t^c[P_{t+1}] - P_t) \tag{2}$$

[4]Note that this simple linear function can follow from mean-variance optimising agents and zero outside supply, see Brock and Hommes (1998) and Hommes (2001), for example.

$$D_t^f = a^f(E_t^f[P_{t+1}] - P_t) \tag{3}$$

The demand is naturally positively related to the expected price change for both chartists and fundamentalists. In other words, when agents expect the price to increase in the coming period, they will increase their demand for that asset today. However, chartists and fundamentalists differ in the way they form expectations about future prices. Chartists form their expectations based on some form of technical analysis. Commonly used rules are moving average (MA) rules and AR(n) rules. For simplicity, we will focus on a simple AR(1) rule for chartists:

$$E_t^c[P_{t+1}] = P_t + \beta_c(P_t - P_{t-1}). \tag{4}$$

According to this rule, chartists expect price movements to continue if $\beta_c > 0$ or to reverse if $\beta_c < 0$. This often depends on the time horizon, i.e. whether t denotes a week, month or year, for example. Fundamentalists form their expectations based on their perception of a fundamental value of the asset, $(\overline{P_t})$, and the current price deviation thereof:

$$E_t^f[P_{t+1}] = P_t + \beta_f(\overline{P_t} - P_t). \tag{5}$$

Often, fundamentalists are a stabilising force, which means that they expect prices to revert to their fundamental levels. In such a case, $\beta_f > 0$. Computing a fundamental value as input for the model is one of the most challenging tasks of estimating a HAM. For some markets, there are multiple competing models, for example, in the foreign exchange market (PPP, UIP, monetary model, etc.), at other times there are no obvious candidates at all (for example, in commodity markets).

In many applications, the dynamics of the market can be best explained with time-varying weights for chartists and fundamentalists (in other words, when agents can 'switch' between the strategies). Switching functions may vary. For an evaluation of different switching functions, see Baur and Glover (2014). The example we show is an adapted multinomial logit rule from Brock and Hommes (1997, 1998) and similar to Ter Ellen and Zwinkels (2010). In this case, the weight of the chartists depends on the recent forecasting accuracy of the chartist forecasting rule, Π_t^c, relative to the recent forecasting accuracy of the fundamentalist rule, Π_t^f:

$$W_t = \left[1 + \exp\left(\gamma\left[\frac{\Pi_t^c - \Pi_t^f}{\Pi_t^c + \Pi_t^f}\right]\right)\right]^{-1} \tag{6}$$

In this setup, W_t is the proportion of chartists in the market (or the weight put on the chartist forecasting rule), and $1 - W_t$ is the proportion of fundamentalists. The forecasting accuracy of chartists (fundamentalists) is measured as the mean squared error of the chartists (fundamentalists) over the past period. Note that it is also possible that the agents evaluate the rule over more than one period.

$$\Pi_t^c = [(E_{t-1}^c[P_t] - P_{t-1}) - \Delta P_t]^2 \tag{7}$$

$$\Pi_t^f = [(E_{t-1}^f[P_t] - P_{t-1}) - \Delta P_t]^2 \tag{8}$$

As in Ter Ellen and Zwinkels (2010), Eq. (6) differs slightly from the weighting mechanism originally proposed by Brock and Hommes (1997). Instead of using the absolute difference in forecasting accuracy of the two rules, $\Pi_t^c - \Pi_t^f$, weights are calculated by using the relative forecasting (in)accuracy $\left(\frac{\Pi_t^c - \Pi_t^f}{\Pi_t^c + \Pi_t^f}\right)$. Ter Ellen and Zwinkels 2010 and Ter Ellen et al. 2017 argue that this method has the advantages of ease of estimation and comparability between different markets. The coefficient γ is called the intensity of choice and represents the investors' speed of switching. If $\gamma = 0$, investors do not adapt the importance given to the two rules and $W_t = 0.5$. The other extreme is when $\gamma = \infty$ where investors are perfectly adaptive and immediately adjust all weights to the rule with the smallest forecast error. A small positive γ can be an indication of status quo bias, introduced by Kahneman et al. (1982). If investors suffer from this bias, they are reluctant to change their status quo belief, which results in a slower updating of beliefs.

4 Estimation

Due to the complex and nonlinear nature of the bounded rationality heterogeneous agent models, most of the early papers in this field were restricted to theoretical explanations and simulations of these models. These simulations produced interesting results and were able to reproduce many of the stylised facts observed in (financial) markets. Therefore, direct confrontation of the model with real financial data was desirable. Vigfusson (1997) was the first to make an attempt to estimate the parameters of a model with chartists and fundamentalists to financial data.

Given that the dynamic weighting of the two strategies is unobserved, Vigfusson applied the Markov regime switching approach to the foreign exchange market, where chartist and fundamentalist behaviour can be seen as different states. After him, several other authors used this approach for the foreign exchange market (Ahrens and Reitz, 2003) and the stock market (Alfarano et al., 2006; Chiarella et al., 2012). Baak (1999) and Chavas (2000) suggested an approach with General Method of Moments (GMM) and Kalman filtering to estimate a chartist–fundamentalist model for the beef market. Not much later, Winker and Gilli (2001) and Gilli and Winker (2003) used a simulation-based indirect estimation approach by minimising loss functions based on the simulated moments and the realised moments from foreign exchange data. Westerhoff, Reitz and Manzan use a STAR-GARCH approach in several papers. An important characteristic of this estimation technique is that only one type of agents can have a deterministic time-varying weight. Westerhoff and Reitz (2003, 2005) incorporate dynamic weighting in one of the two types of agents by means of a STAR-GARCH estimation for the foreign exchange market (2003, time-varying fundamentalist impact) and the commodity market (2005, time-varying chartist impact). Manzan and Westerhoff (2007) also apply this method with time-varying weights on the chartist impact for the foreign exchange market, whereas Reitz and Slopek (2009) apply it to the oil market.

An important contribution in the estimation of heterogeneous agents models came from Boswijk et al. (2007). They use nonlinear least squares estimation combined with a multinomial logit switching rule to empirically validate a heterogeneous agents model for the S&P500. The main improvements of their method over estimating based on Markov switching are the smaller number of parameters to be estimated and the deterministic nature of their switching process, in contrast to a stochastic Markov process. Many empirical papers on heterogeneous agents models have successfully used, and sometimes adapted, the techniques from Boswijk et al. (2007) for stock markets (De Jong et al., 2009; Chiarella et al., 2014) and foreign exchange markets (De Jong et al., 2010), but also for less obvious asset classes, such as oil (Ter Ellen and Zwinkels, 2010), housing (Kouwenberg and Zwinkels, 2014), gold (Baur and Glover, 2014), options (Frijns et al., 2010), hedge funds (Schauten et al., 2015) and credit markets (Chiarella et al., 2015).

A recent survey study by Lux and Zwinkels (2018) extensively covers various techniques for estimating agent-based models. Here, we rather focus on the results from estimating heterogeneous agent models.

4.1 Results

Most empirical studies on heterogeneous agent models use the classification of chartists and fundamentalists as found in the theoretical literature, where chartists base their expectations either on an autoregressive or on a moving average rule, and fundamentalists choose a fundamental value that is appropriate for the asset class under consideration. According to the theory on chartists and fundamentalists, chartists generally play a destabilising role by extrapolating and enforcing trends, whereas fundamentalists have a stabilising impact on the asset price due to their mean reverting expectations. This presumption is confirmed by many empirical validations of the model (Table 1).

4.1.1 Stock Market

One of the most widely used methods for estimating a heterogeneous agents model (HAM) is with nonlinear least squares or maximum likelihood, combined with a multinomial logit switching rule which is inspired by the work of Brock and Hommes (1997, 1998). This method was introduced by Boswijk et al. (2007),

Table 1 Overview of empirical validations of heterogeneous agent models

Study	Market	Sample	Freq	Estimation	Fundamental	Agents
Equity						
Alfarano et al. (2006)	ASX & AUDUSD	1980(/1983)–2004	D	ML	Implied	Noise & fund
Boswijk et al. (2007)	S&P500	1871–2003	A	NLS	GG	Chart & fund
Amilon (2008)	S&P500	1980–2000	D	EMM / ML	GG	Chart & fund
De Jong et al. (2009)	Hang S. & B-SET	1980–2007	Q	ML	GG	Chart & fund & int
Chiarella et al. (2012)	S&P500	2000–2010	M	Markov RS	GG	Chart & fund & noise
Chiarella et al. (2014)	S&P500	1970–2012	M	ML	GG	Chart & fund & noise
Lof (2014)	S&P500	1871–2011	A	NLS	GG	fund & rat spec & cont spec
Frijns and Zwinkels (2016b)	Can. firms	2010–2011	HF	VECM	–	Chart & arb
Hommes and in ’t Veld (2017)	S&P500	1950–2012	Q	NLS	GG	Chart & fund
Huang and Tsao (2018)	Taiwan SE	2010–2011	HF	ML	CAPM	Chart & fund & liq
Forex						
Vigfusson (1997)	CAD	1983–1992	D	Markov RS	PPP & TOT	Chart & fund
Winker and Gilli (2001)	DM	1991–2000	D	SMM	–	Chart & fund
Gilli and Winker (2003)	DM	1991–2000	D	SMM	–	Chart & fund
Westerhoff and Reitz (2003)	GBP, DM, JPY	1980–1996	D	STAR-GARCH	PPP	Chart & fund
Manzan and Westerhoff (2007)	DM, JPY, CAD, FF, GBP	1974–1998	M	OLS	PPP	Chart & fund
Menkhoff et al. (2009)	EUR, GBP, JPY	1992–2006	M	OLS	PPP & MM	Chart & fund
De Jong et al. (2010)	7 EMS currencies	1979–1998	M	ML	EMS parity	Chart & fund
Jongen et al. (2012)	EUR, JPY, GBP	1989–2009	M	NLS with panel	PPP	Chart & fund & carry
Ter Ellen et al. (2013)	JPY, GBP, EUR	2003–2008	W	NLS	PPP	Chart & fund & IP
Dick and Menkhoff (2013)	EUR	1991–2011	M	OLS	PPP	Chart & fund & inter
Goldbaum and Zwinkels (2014)	JPY	1995–2007	M	OLS	MM	Chart & fund

Commodities						
Westerhoff and Reitz (2005)	Corn	1973–2003	M	STAR-GARCH ML	LRA	Chart & fund
Reitz and Westerhoff (2007)	6 commodities	1973–2003	M	STAR-GARCH ML	LRA	Chart & fund
Reitz and Slopek (2009)	Oil (WTI)	1986–2006	M	STAR-GARCH	Demands	Chart & fund
Ter Ellen and Zwinkels (2010)	Oil (Brent & WTI)	1983–2009	M	ML	2Y MA	Chart & fund
Baur and Glover (2014)	Gold	1970–2012	M	NLS	EWMA	Chart & fund
Housing						
Kouwenberg and Zwinkels (2014)	US	1960–2012	Q	ML	PV of rents	Chart & fund
Kouwenberg and Zwinkels (2015)	US	1960–2014	Q	ML	PV of rents	Chart & fund
Eichholtz et al. (2015)	Amsterdam	1649–2005	A	ML	Inflation	Chart & fund
Bolt et al. (2014)	US, UK, NL, JP, CH, ES, SE, BE	1970–2013	Q	NLS	PV of rents	Chart & fund
Credit						
Chiarella et al. (2015)	13 European CDS	2004–2013	W	ML	Hazard rate	Chart & fund
Frijns and Zwinkels (2016a)	European bonds & CDS	2008–2015	D	ML	Latent factor	Arb & chart & liq
Other						
Frijns et al. (2010)	DAX 30 Volatility	2000–2000	D	GJR-GARCH	LRA	Chart & fund
Frijns et al. (2013)	US equity MF	1998–2004	D	ML	–	–
Verschoor and Zwinkels (2013)	FX FM	2000–2009	M	ML	–	Chart & fund & carry
Schauten et al. (2015)	HF exposure	1996–2009	M	ML	–	–
Cornea-Madeira et al. (2017)	US inflation	1968–2015	Q	NLS	Real MC	Fund & RW

Notes: Estimation methods refer to Maximum Likelihood (ML), Nonlinear Least Squares (NLS), Method of Moments (MM), and Markov regime shifting (Markov RS). Agent types refer to chartists (chart), fundamentalists (fund), arbitrageurs (arb), internationalists (int), liquidity traders (liq), noise traders (noise), rational speculators (rat) and contrarian speculators (cont). More detailed description of these papers can be found in Sect. 4.1

who directly estimate a HAM on stock returns (S&P500). In their model, there are heterogeneous agents with access to the fundamental value of a risky asset, but with different beliefs about the persistence of the deviation between the spot price and the fundamental price of the asset. Switching between the different beliefs takes place based on the relative past profitability of that strategy. Chiarella et al. (2014) estimate a heterogeneous agents model for the S&P500 with three types of agents: fundamentalists, chartist and noise traders. Consistent with most of the other empirical studies, fundamentalists play a stabilising role with respect to the fundamental value of the asset. Chartists trade based on a moving average rule given by a geometric decay process, while most empirical studies rely on an AR(1) rule. While the relative weight of fundamentalists and chartists in the market changes over time based on the relative performance of these rules, the impact of noise traders is assumed to be constant. Noise traders have no specific expectations of future returns, and their demand is driven by a noisy signal that depends on volatility. Both Boswijk et al. (2007) and Chiarella et al. (2014) find support for mean reversion in fundamentalists' expectations and trend extrapolation in chartists' expectations of the S&P500. The model with time-varying weights has a significantly better fit than the static model.

Lof (2014) also estimates a heterogeneous agent model on S&P500 data. The types of agents he distinguishes are fundamentalists, rational speculators and contrarian speculators. The latter two types have exactly opposing beliefs to one another. He finds that the existence of contrarians can explain some of the most volatile episodes of the S&P500. De Jong et al. (2009) also distinguish three types of agents, to shed light on the Asian crisis in the context of heterogeneous agents. Besides chartists and fundamentalists, they distinguish internationalists, who condition their expectations on foreign market conditions. In a two-country model (with Hong Kong and Thailand) for the stock market, chartists and fundamentalists base their expectations on past price changes and the price deviation from the fundamental value, respectively, whereas internationalists base their expectations on the past price changes of the foreign market. Market dynamics occur due to switching between the different groups conditional on their past forecasting performance. Their estimation method is in many ways comparable to the one in Boswijk et al. (2007), yet De Jong et al. (2009) use maximum likelihood techniques instead of nonlinear least squares. All these studies compute a fundamental stock price by taking the discounted value of expected future dividends, which comes down to a simple Gordon growth model when a constant growth rate of dividends is assumed. Given the earlier critique on the use of a benchmark fundamental value with constant risk premium, Hommes and in 't Veld (2017) also calculate a fundamental value based on the Campbell–Cochrane consumption-habit model that allows for variation in the risk premium. Even with this model as a benchmark, they find substantial behavioural heterogeneity for the S&P500.

Alfarano et al. (2006) use Markov switching to estimate a HAM for Australian stock and FX data. They recognise the complexity of the agent-based models and the fact that this makes it difficult to directly estimate all the underlying parameters. They simplify the model to a closed-form solution for returns to overcome this problem. Although their model is highly simplified compared to some of the earlier agent-based models for financial markets, the authors are still able to reproduce some of the stylised features of stock returns. The two groups of traders are labelled as fundamentalists and noise traders, and switching between the two groups occurs based on asymmetric switching probabilities, inspired by Kirman's herding mechanism. The switching is asymmetric because the transition probability of an agent switching from the group of noise traders to the group of fundamentalists differs from the transition probability of a switch in the opposite direction. Chiarella et al. (2012) use Markov regime switching to explain the market dynamics of the S&P500. In their model, investors' beliefs about returns are regime dependent, and regimes (a bull state of the market with positive returns and low volatility or a bust state of the market with negative returns and high volatility) are generated by a stochastic process.

Recent contributions have used the heterogeneous agent framework to explain very high frequency stock price movements. Frijns and Zwinkels (2016b) look at cross-listed Canadian firms to find out where price discovery takes place. The model shows time variation in price discovery that is driven by agents switching between an arbitrage and a speculative strategy. Huang and Tsao (2018) use intraday data on three stocks listed on the Taiwan Stock Exchange to investigate whether there is evidence of heterogeneity of beliefs. They find that fundamentalists are stabilising, given that they expect mispricing to reduce in the next period. Chartists (technical analysts) behave as contrarians, but extrapolate buyer-initiated trades as a sign that prices will rise, and seller-initiated trades as a sign that prices will decline. Interestingly, they also find that chartists perform slightly better than fundamentalists.

4.1.2 Foreign Exchange Market

Vigfusson (1997) is the first to empirically test the chartist–fundamentalist approach for the foreign exchange market, and does this by means of a Markov switching approach. He tests two different specifications for fundamentalists and two for chartists. He finds that more important than the functional form of the types of agents is the different variances in the two regimes. He concludes that the USDCAD market is certainly characterised by quite regular regime shifts, but that it is not straightforward to conclude that this directly stems from the presence of chartists and fundamentalists in the market.

De Jong et al. (2010) estimate a full heterogeneous agents model with switching on exchange rates. By estimating the chartist–fundamentalist model on EMS rates, they circumvent the problem of having to choose a fundamental rate. Instead, they can use the 'parity' rate. With a survey dataset from Consensus Economics London, Goldbaum and Zwinkels (2014) directly test investor heterogeneity and expectation formation for the Japanese yen and the euro against the US dollar. The authors estimate three different models with chartists and fundamentalists. In the first model, both rules are estimated for the full sample of respondents and time.

In the second model, every forecaster is labelled as being either fundamentalist or chartist, based on the sum of the relative difference between the forecast and the outcome of the respective forecasting strategy. Finally, the respondents are allowed to switch their strategy. Every single forecast is labelled as resulting from either the fundamentalist or chartist strategy. The authors use the monetary model to compute a fundamental value for the exchange rates. Another paper that evaluates investor expectations for the foreign exchange market with survey data comes from Ter Ellen et al. (2013). They estimate a full heterogeneous agent model with dynamic weights of PPP traders (fundamentalists), momentum traders (chartists) and interest parity traders on forecasts for the euro, pound sterling and Japanese yen against the US dollar and the Japanese yen against the euro. One of their main findings is that they find forecasters to use rather 'speculative' models, such as momentum and carry, to predict exchange rates for short horizons, and rather 'fundamental' models, such as PPP and UIP, to predict exchange rates for longer horizons. The same strategies are identified by Verschoor and Zwinkels (2013) by looking at currency trader indices. They further find that FX fund managers apply a negative feedback strategy, moving capital from strategies with high past performance to low past performance.

Winker and Gilli (2001) and Gilli and Winker (2003) use a simulation-based indirect estimation approach to find the parameter values of a HAM applied to the US dollar–German mark exchange rate. The parameter values of the model are obtained by minimising a loss function based on the model simulated moments and the moments from the real data. The 2001 paper serves as an introduction of this method and therefore only focuses on two moments: kurtosis and ARCH-effects. The authors only estimate the random switching probability parameter and the probability that an agent will switch after interacting with another agent. In the 2003 paper, the optimisation algorithm is improved and a third parameter, the standard deviation of noise in the majority assessment, is estimated.

Westerhoff and Reitz (2003) estimate a STAR-GARCH model where the impact of fundamentalists depends on the strength of their belief in fundamental analysis. If the misalignment of the exchange rate with the fundamental value increases, fundamentalists lose their faith in fundamental analysis and leave the market. Therefore, the dynamics in the fundamentalists' behaviour further destabilise the exchange rate. This is in stark contrast to the findings in Manzan and Westerhoff (2007). They find that fundamentalists play an increasingly stabilising role in the event of a larger misalignment of the exchange rate. However, chartists play a destabilising role only within a certain range. When the past appreciation or depreciation of the exchange rate is larger than the threshold value, their behaviour becomes stabilising. De Jong et al. (2010) find evidence of stabilising behaviour of all types of agents for EMS rates, a result they assign to the investors' trust in the monetary authorities.

Finally, rather than explaining price movements or expectations directly, a few papers explain the dispersion of beliefs by a model with chartists and fundamentalists (Menkhoff et al., 2009; Jongen et al., 2012). They provide further evidence that agents in the foreign exchange market are heterogeneous due to the use of these different forecasting approaches.

4.1.3 Commodities

Prat and Uctum (2011) describe the expectation formation process for WTI oil prices as a combination of the extrapolative, regressive and adaptive expectation formation processes, based on survey data obtained from Consensus Economics. Reitz and Slopek (2009) explain the large price swings observed in the oil market by stabilising fundamentalists, who have a larger impact the larger the misalignment of the oil price is, and chartists, who are dominant and play a destabilising role when the price of oil is close to its fundamental value. While Reitz and Slopek (2009) take a STAR-GARCH approach with heterogeneous agents to explain large oil price swings, Ter Ellen and Zwinkels (2010) employ maximum likelihood with a multinomial logit switching rule. In their approach, the market impact of trend-extrapolating chartists and mean-reversion fundamentalists is time varying, based on the relative past forecasting accuracy of the strategies. Fundamentalists believe in mean reversion of the WTI and Brent price of crude oil to a long-term moving average of the oil price, whereas chartists extrapolate the price movement from the previous period. Considering that there is no consensus on the fundamental value of oil and computing one can be costly, the authors use a 2-year moving average as a proxy for the fundamental value. They confirm the destabilising (stabilising) effect of chartists (fundamentalists) and additionally find asymmetry in the responses of both chartists and fundamentalists. Furthermore, high weights for the chartist strategy coincide with different price spikes in the sample period, suggesting that they contributed to an oil price bubble in these periods. The model has a good out-of-sample fit. The authors show that the heterogeneous agent model outperforms the random walk model and a VAR(1,1) model.

Baur and Glover (2014) find that investors in the gold market are heterogeneous. They find that whereas both chartists and fundamentalists help to explain the price of gold, it was mostly the extrapolative behaviour of chartists that contributed to the large and persistent increase in the price of gold in the early 2000s. However, the coefficients they obtain for chartist and fundamentalist behaviour are somewhat different from what is commonly found in other financial markets. One such surprising results is that in some specifications, fundamentalists in the market for gold play a destabilising role, i.e. they behave more like the chartists in the original model of Brock and Hommes (1997).

Westerhoff and Reitz (2005) estimate a model for the US corn market with constant stabilising fundamentalist behaviour and dynamic technical trading activity, which is time varying depending on the misalignment of the corn price. They find that chartists play a highly destabilising role, and that this effect becomes stronger the further the price of corn is away from its fundamental, or long-run equilibrium, price. They estimate a similar model, but with time variation in fundamentalists beliefs, in Reitz and Westerhoff (2007) for cotton, lead, rice, soybeans, sugar and zinc, and find that for these commodities, fundamentalists play a stabilising role when the misalignment is sizable enough.

4.1.4 Credit

Chiarella et al. (2015) analyse the large deviations from fundamental levels of credit risk for some European countries during the European sovereign debt crisis and find that these can be partly explained by a combination of increased global risk aversion and the dynamics between momentum traders (chartists) and fundamentalists. Although the increase in credit spreads for peripheral European countries during the sovereign debt crisis was initially caused by deteriorating fundamentals, a large part of the surge can be explained by momentum traders further extrapolating these trends of higher CDS spreads. Frijns and Zwinkels (2016a) jointly model the bond and CDS market for a very similar sample. Rather than calculating the underlying fundamental value, they treat the fundamental process as an unobservable factor driving both markets. They find that, on average, only 5.5% of spread variation can be explained by speculation, but that the effect varies over time.

4.1.5 Housing

Kouwenberg and Zwinkels (2014, 2015) show that even the price movements in the US housing market can be well explained by a dynamic heterogeneous agent model. The model is estimated with maximum likelihood, including fundamentalists who believe in mean reversion of house prices to a rents-based fundamental value and chartists who destabilise the market by extrapolating trends. Agents switch between strategies based on the past forecasting accuracy of the respective strategies. They further find that the dominance of chartists in the housing market from 1992 to 2005 can explain the bubble-like behaviour of house prices in that period. Their model with time-varying impact of fundamentalists, who believe in mean reversion to a fundamental value based on rents, and chartists, who extrapolate past price trends, explains the house price for the in-sample period, and is also able to predict the decline in house prices from 2006 onwards.

Bolt et al. (2014) estimate a heterogeneous agent model on housing data for eight countries, including the USA. Different from Kouwenberg and Zwinkels, Bolt et al. (2014) include (the possibility of) a risk premium in the fundamental value calculation. Also, their chartists extrapolate price misalignments rather than price trends. Overall, they find that the housing markets in all countries studied are prone to behavioural bubbles. They also suggest some policies that can help stabilise prices.

Whereas the aforementioned studies start their samples in the 1960s and 1970s, Eichholtz et al. (2015) study house prices in Amsterdam, the Netherlands, from the seventeenth century onwards. They find that expectation formation depends on the stage of the economic cycle: during economic slowdowns, agents focus more on fundamentals, whereas they are more prone to follow trends during booms.

4.1.6 Other Asset Classes

The evidence in favour of heterogeneous agents extends more and more to other (financial) markets. Frijns et al. (2010) propose a way to model heterogeneous expectations of volatility by applying a heterogeneous agent model to the option market, where volatility is priced and traded. Fundamentalists believe that conditional volatility will revert to the level of the unconditional volatility and chartists trade based on recently observed unexpected shocks. Their heterogeneous agent model simplifies to a GJR-Garch(1,1) model with time-varying coefficients, which depend on the time-varying market impact of chartists and fundamentalists.

Frijns et al. (2013) estimate a switching model on 400 US equity mutual funds where investors can switch between cash and stocks depending on the expected relative performance of stocks or cash, and evaluate the market timing ability of these funds. Strikingly, they find that less than 5% of the mutual funds in their study have positive market timing skills, versus more than 40% with negative timing skills.

Schauten et al. (2015) consider style investing hedge funds, and find that there is time variation in their exposure to certain investment styles. The time variation depends on the recent relative performance of the styles, as is common in the heterogeneous agent literature. Hedge funds display positive feedback trading, but could do better by doing this more aggressively.

As it turns out, housing is not the only macro-variable that can be explained by heterogeneous agents. Cornea-Madeira et al. (2017) estimate a HAM on the US inflation data. Fundamentalists expect inflation to revert back to a fundamental value, which is based on the relation between inflation and real marginal costs. The other group of firms, which they call random walk believers, have naive expectations, and are thus backward-looking. They find that the majority of firms follows such a backward-looking strategy when forming inflation expectations, but that there are also occurrences of the dominance of fundamentalists.

5 Conclusion

Although the rational paradigm has been at the forefront of financial markets research since the seventies, rejections of this paradigm and attempts to model investor behaviour in a different way are gaining ground. Boundedly rational heterogeneous agent models (HAM) are an example of such models. In these models, agents are allowed to form expectations using relatively simple rules of thumb. In the empirical applications, this often boils down to two to four different agent types: fundamentalists, who expect market prices to revert to the fundamental value of the respective assets, chartists, who extrapolate price trends, and third and fourth types that often differ among various applications. In this contribution, we have provided an overview of papers estimating such models and their main results.

We have learned from this literature that investors are not only heterogeneous, they also do not use stable, unconditional, forecasting rules to form their expectation on future movements of exchange rates. Instead, they may change the way they form expectations based on various factors, such as the past performance of different forecasting rules or the horizon for which they form their expectations. The dynamics between the different types of investors can cause periods of severe mispricing and disruption of financial markets.

There is ample micro-evidence that agents indeed do not form rational expectations but use rules of thumb to forecast (financial) variables. Survey datasets that contain analysts' forecast are an important tool to unravel investor expectation mechanisms and dynamics that can otherwise not always be directly observed in the data. Studies based on such data have shown that expectations are not unbiased and do sometimes not even incorporate all available public information. Furthermore, the expectation formation rules that are found to explain the data well can be summarised by extrapolative, adaptive, and regressive rules, much in line with the rules chartists and fundamentalists use in heterogeneous agent models.

More micro-evidence on the behaviour of economic agents has come from experimental studies. Although a common critique of such studies is often the potential lack of external validity, many experimental studies have confirmed the behavioural rules found in survey responses. These rules are very much in line with behavioural rules in heterogeneous agent models: economic agents use (approximate) linear forecasting rules, such as autoregressive, mean reverting or adaptive strategies.

As surveyed in this chapter, heterogeneous agent models typically explain the stylised facts of financial markets well, and they are able to replicate important episodes of turmoil. However, empirically obtained results for various asset markets are often hard to compare, due to the researcher's choice of sample, fundamental value, set of behavioural rules and functional form of the switching function. Some efforts have been made to increase comparability by estimating a generic model on several (asset) prices, based on the same sample, switching function and behavioural rules, and based on a similar model for the fundamental value. In more general terms though, the degrees of freedom of behavioural (asset pricing) models needs to be taken seriously. It is the reason that the models can produce a very good fit of the data, but it can also lead to ad hoc modelling decisions that lack micro-foundations. One reason that the rational expectations paradigm is and has been the dominant one for so long is that there is only one way to be rational (and thus to model rationality), while there are infinite ways to deviate from rationality. When deviating from the rational expectations paradigm, it is important to keep in mind that there needs to be clear evidence on the micro-level for the way expectations are modelled.

Finally, one needs to keep in mind that models based on the heterogeneous beliefs of agents still abstract from reality in many other respects. In reality, it is very likely that agents do not only differ in the way they form beliefs but also in the preferences they have, the shocks that they are hit by and the information set they have access to. Especially on a macro-level, it is very hard to pin down whether people behave different from our model because they are irrational, or because we do not capture their preferences well. Currently, there is ample evidence that heterogeneous agent models beat a random walk model in forecasting financial variables. However, as of yet there is very little work that compares the performance of these models to other deviations of the efficient markets hypothesis, such as full versus limited information/attention, heterogeneous preferences or financial (market) frictions. This can be a promising line of future research.

References

Ahrens, R., & Reitz, S. (2003). Heterogenous expectations in the foreign exchange market: Evidence from daily dollar/DM exchange rates. *Journal of Evolutionary Economics, 15*(1), 65–82.

Alfarano, S., Lux, T., & Wagner, F. (2006). Estimation of a simple agent-based model of financial markets: An application to Australian stock and foreign exchange data. *Physica: Statistical Mechanics and Its Applications, 370*(1), 38–42.

Amilon, H. (2008). Estimation of an adaptive stock market model with heterogeneous agents. *Journal of Empirical Finance, 15*(2), 342–362.

Assenza, T., Bao, T., Hommes, C. H., & Massaro, D. (2014). Experiments on expectations in macroeconomics and finance. In J. Duffy (Ed.), *Experiments in macroeconomics* (pp. 11–70). Bingley: Emerald Group Publishing Limited.

Baak, S. J. (1999). Tests for bounded rationality with a linear dynamic model distorted by hetereogeneous expectations. *Journal of Economics Dynamics and Control, 32*(9), 1517–1573.

Ball, R., & Brown, P. (1968). An empirical evaluation of accounting income numbers. *Journal of Accounting Research, 6*(2), 159–178.

Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics, 49*(3), 307–345.

Baur, D. G., & Glover, K. J. (2014). Heterogeneous expectation in the gold market: Specification and estimation. *Journal of Economic Dynamics and Control, 40*(C), 116–133.

Beja, A., & Goldman, M. B. (1980). On the dynamic behaviour of prices in disequilibrium. *Journal of Finance, 35*(2), 235–248.

Bilson, J. (1981). The speculative efficiency hypothesis. *Journal of Business, 54*, 435–451.

Black, F., Jensen, M. C., & Scholes, M. (1972). The capital asset pricing model: Some empirical tests. In M. C. Jensen (Ed.), *Studies in the theory of capital markets* (pp. 79–124). New York: Praeger.

Blake, D., Beenstock, M., & Brasse, V. (1972). The performance of UK exchange rate forecasters. *The Economic Journal, 96*(384), 986–999.

Bloomfield, R., & Hales, J. (2002). Predicting the next step of a random walk: Experimental evidence of regime-switching beliefs. *Journal of Financial Economics, 65*(3), 397–414.

Bolt, W., Demertzis, M., Diks, C., Hommes, C., & van der Leij, M. (2014). *Identifying Booms and Busts in House Prices Under Heterogeneous Expectations*. De Nederlandsche Bank Working Paper (450).

Boswijk, H. P., Hommes, C. H., & Manzan, S. (2007). Behavioral hetreogeneity in stock prices. *Journal of Economic Dynamics and Control, 31*(6), 1938–1970.

Brock, W. A., & Hommes, C. H. (1997). A rational route to randomness. *Econometrica, 65*(5), 1059–1095.

Brock, W. A., & Hommes, C. H. (1998). Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic Dynamics and Control, 22*(8), 1235–1274.

Cavaglia, S., Verschoor, W. F. C., & Wolff, C. C. P. (1993). Further evidence on exchange rate expectations. *Journal of International Money and Finance, 12*(1), 78–98.

Cavaglia, S., Verschoor, W. F. C., & Wolff, C. C. P. (1994). On the biasedness of forward foreign exchange rates: Irrationality or risk premia? *The Journal of Business, 67*(3), 321–343.

Chavas, J.-P. (2000). On information and market dynamics: The case of the US beef market. *Journal of Economic Dynamics and Control, 25*(5), 833–853.

Chiarella, C. (1992). The dynamics of speculative behavior. *Annals of Operational Research, 37*(1), 101–123.

Chiarella, C., Dieci, R., & He, X.-Z. (2009). Heterogeneity, market mechanisms and asset price dynamics. In T. Hens & K. R. Schenk-Hoppe (Eds.), *The handbook of financial markets: Dynamics and evolution* (pp. 277–344). Burlington: Elsevier.

Chiarella, C., He, X.-Z., Huang, W., & Zheng, H. (2012). Estimating behavioural heterogeneity under regime switching. *Journal of Economic Behaviour and Organization, 83*(3), 446–460.

Chiarella, C., He, X.-Z., & Zwinkels, R. C. J. (2014). Heterogeneous expectations in asset pricing: Empirical evidence from the S&P500. *Journal of Economic Behaviour and Organization, 105*, 1–16.

Chiarella, C., Ter Ellen, S., He, X.-Z., & Wu, E. (2015). Fear or fundamentals? Heterogeneous beliefs in the European sovereign CDS market. *Journal of Empirical Finance, 32*, 19–34.

Cornea-Madeira, A., Hommes, C., & Massaro, D. (2017, forthcoming). Behavioral heterogeneity in US inflation dynamics. Journal of Business & Economic Statistics.

Cumby, R., & Obstfeld, M. (1984). International interest rate and price level linkages under flexible exchange rates. In: J. F. O. Bilson & R. C. Marston (Eds.), *Exchange rate theory and practice* (pp. 121–152). Chicago: University of Chicago Press.

Day, R. H., & Huang, W. (1990). Bulls, bears and market sheep. *Journal of Economic Behaviour and Organization, 14*(3), 299–329.

De Bondt, W. F. M. (1993). Betting on trends: Intuitive forecasts of financial risk and return. *International Journal of Forecasting, 9*(3), 355–371.

De Bondt, W. F. M., & Thaler, R. (1985). Does the stock market overreact? *The Journal of Finance, 40*(3), 793–805.

De Grauwe, P., Dewachter, H., & Embrechts, M. (1993). *Exchange rate theory: Chaotic models of foreign exchange markets*. Oxford: Blackwell.

De Grauwe, P., & Grimaldi, M. (2006). Exchange rate puzzles: A tale of switching attractors. *European Economic Review, 50*(1), 1–33.

De Jong, E., Verschoor, W. F. C., & Zwinkels, R. C. J. (2009). Behavioural heterogeneity and shift-contagion: Evidence from the Asian crisis. *Journal of Economic Dynamics and Control, 33*(11), 1929–1944.

De Jong, E., Verschoor, W. F. C., & Zwinkels, R. C. J. (2010). Heterogeneity of agents and exchange rate dynamics: Evidence from the EMS. *Journal of International Money and Finance, 29*(8), 1652–1669.

De Long, J. B., Shleifer, A., Summers, L., & Waldmann, R. (1990). Noise trader risk in financial markets. *Journal of Political Economy, 98*(4), 703–738.

Dick, C. D., & Menkhoff, L. (2013). Exchange rate expectations of chartists and fundamentalists. *Journal of Economic Dynamics and Control, 37*(7), 1362–1383.

Dominguez, K. (1986). Are foreign exchange rate forecasts rational? New evidence from survey data. *Economics Letters, 21*(3), 277–281.

Eichholtz, P., Huisman, R., & Zwinkels, R. C. J. (2015). Fundamentals or trends? A long-term perspective on house prices. *Applied Economics, 47*(10), 1050–1059.

Engel, C. (1996). The forward discount anomaly and the risk premium: A survey of recent evidence. *Journal of Empirical Finance, 3*(2), 123–192.

Fama, E. F. (1965). Random walks in stock market prices. *Financial Analysts Journal, 51*(1), 55–59.

Fama, E. F. (1984). Forward and spot exchange rates. *Journal of Monetary Economics, 14*, 319–338.

Frankel, J. A., & Froot, K. A. (1987). *Understanding the US Dollar in the Eighties: The Expectations of Chartists and Fundamentalists*. NBER Working Paper, No. R0957.

Frankel, J. A., & Froot, K. A. (1990a). Chartists, fundamentalists, and trading in the foreign exchange market. *American Economic Review, 80*(2), 181–185.

Frankel, J. A., & Froot, K. A. (1990b). *Exchange Rate Forecasting Techniques, Survey Data, and Implications for the Foreign Exchange Market*. IMF Working Paper.

Friedman, M. (1953). The case for flexible exchange rates. In *Essays in positive economics* (pp. 157–203). Chicago: University of Chicago Press.

Frijns, B., Gilbert, A., & Zwinkels, R. C. (2013). Market timing ability and mutual funds: A heterogeneous agent approach. *Quantitative Finance, 13*(10), 1613–1620.

Frijns, B., Lehnert, T., & Zwinkels, R. C. J. (2010). Behavioral heterogeneity in the option market. *Journal of Economic Dynamics and Control, 34*(11), 2273–2287.

Frijns, B., & Zwinkels, R. C. (2016a). *Speculation in European Sovereign Debt Markets*. Working paper.

Frijns, B., & Zwinkels, R. C. (2016b). *Time-Varying Arbitrage and Dynamic Price Discovery*. Working paper.

Froot, K. A. (1989). New hope for the expectations hypothesis of the term structure of interest rates. *Journal of Finance, 44*(2), 283–305.

Froot, K. A., & Thaler, R. H. (1990). Anomalies: Foreign exchange. *Journal of Economic Perspectives, 4*(3), 179–192.

Gilli, M., & Winker, P. (2003). A global optimization heuristic for estimating agent based models. *Computational Statistics and Data Analysis, 42*(3), 299–312.

Goldbaum, D., & Zwinkels, R. C. J. (2014). An empirical examination of heterogeneity and switching in foreign exchange markets. *Journal of Economic Behavior & Organization, 107*, 667–684.

Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *American Economic Review, 70*(3), 393–408.

Hommes, C. H. (2001). Financial markets as nonlinear adaptive evolutionary systems. *Quantitative Finance, 1*(1), 149–167.

Hommes, C. H. (2006). Heterogeneous agent models in economics and finance. In L. Tesfatsion & K. L. Judd (Eds.), *Handbook of computational economics* (Vol. 2, pp. 1109–1186). Amsterdam: North-Holland.

Hommes, C. H., & in 't Veld, D. (2017). Booms, busts and behavioural heterogeneity in stock prices. *Journal of Economic Dynamics and Control, 80*, 101–124.

Hommes, C. H., Sonnemans, J., Tuinstra, J., & van de Velden, H. (2005). Coordination of expectations in asset pricing experiments. *Review of Financial Studies, 18*(3), 955–980.

Hsieh, D. A. (2017). Tests of rational expectations and no risk premium in forward exchange markets. *Journal of International Economics, 17*, 173–184.

Huang, Y.-C., & Tsao, C.-Y. (2018). Discovering traders' heterogeneous behavior in high-frequency financial data. *Computational Economics, 51*(4), 821–846.

Ito, T. (1990). Foreign exchange rate expectations: Micro survey data. *The American Economic Review, 80*(3), 434–449.

Jegadeesh, N. (1990). Evidence of predictable behavior of security returns. *The Journal of Finance, 45*(3), 881–898.

Jongen, R., Verschoor, W. F. C., & Wolff, C. C. P. (2008). Foreign exchange rate expectations: Survey and synthesis. *Journal of Economic Surveys, 22*(1), 140–165.

Jongen, R., Verschoor, W. F. C., Wolff, C. C. P., & Zwinkels, R. C. J. (2012). Explaining dispersion in foreign exchange expectations: A heterogeneous agent approach. *Journal of Economic Dynamics and Control, 36*(5), 719–735.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.

Kouwenberg, R., & Zwinkels, R. C. (2015). Endogenous price bubbles in a multi-agent system of the housing market. *PloS One, 10*(6), 1–10.

Kouwenberg, R., & Zwinkels, R. C. J. (2014). Forecasting the U.S housing market. *International Journal of Forecasting, 30*(3), 415–425.

LeBaron, B. (2000). Agent-based computational finance: Suggested readings and early research. *Journal of Economic Dynamics and Control, 24*(5–7), 679–702.

Lof, M. (2014). Rational speculators, contrarians and excess volatility. *Management Science, 61*(8), 1889–1901.

Lux, T. (1998). The socio-economic dynamics of speculative markets: Interacting agents, chaos and fat tails of return distributions. *Journal of Economic Behavior and Organization, 33*(2), 143–165.

Lux, T. (2009). Stochastic behavioral asset pricing models and the stylized facts. In T. Hens & K. R. Schenk-Hoppe (Eds.), *The handbook of financial markets: Dynamics and evolution* (pp. 161–215). Burlington: Elsevier.

Lux, T., & Zwinkels, R. C. J. (2018). Empirical validation of agent-based models. In *Handbook of Computational Economics*. Amsterdam: Elsevier.

MacDonald, R. (1990a). Are foreign exchange market forecasters ‘rational’?: Some survey-based tests. *The Manchester School, 58*(3), 229–241.

MacDonald, R. (1990b). Expectations formation and risk in three financial markets: Surveying what the surveys say. *Journal of Economic Surveys, 14*(1), 69–100.

MacDonald, R., & Marsh, I. W. (1996). Currency forecasters are heterogeneous: Confirmation and consequences. *Journal of International Money and Finance, 15*(5), 665–685.

Manzan, S., & Westerhoff, F. H. (2007). Heterogeneous expectations, exchange rate dynamics and predictability. *Journal of Economic Behavior and Organization, 64*(1), 111–128.

Menkhoff, L. (2010). The use of technical analysis by fund managers: International evidence. *Journal of Banking and Finance, 34*(11), 2573–2586.

Menkhoff, L., Rebitsky, R. R., & Schröder, M. (2009). Heterogeneity in exchange rate expectations: Evidence on the chartist-fundamentalist approach. *Journal of Economic Behavior and Organization, 70*(1), 241–252.

Milgrom, P., & Stokey, N. (1982). Information, trade and common knowledge. *Journal of Economic Theory, 26*(1), 17–27.

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica, 29*(3), 315–335.

Pojarliev, M., & Levich, R. (2008). Do professional currency managers beat the benchmark?. *Financial Analyst Journal, 64*(5), 18–30.

Prat, G., & Uctum, R. (2007). Switching between expectation processes in the foreign exchange market: A probabilistic approach using survey data. *Review of International Economics, 15*(4), 700–719.

Prat, G., & Uctum, R. (2011). Modelling oil price expectations: Evidence from survey data. *The Quarterly Review of Economics and Finance, 51*(3), 236–247.

Prat, G., & Uctum, R. (2015). Expectation formation in the foreign exchange market: A time-varying heterogeneity approach using survey data. *Applied Economics, 47*(34–35), 3673–3695.

Reitz, S. & Slopek, U. (2009). Non-linear oil price dynamics: A tale of heterogeneous speculators?. *German Economic Review, 10*(3), 270–283.

Reitz, S., & Westerhoff, F. (2007). Commodity price cycles and heterogeneous speculators: A STAR-GARCH model. *Empirical Economics, 33*, 231–244.

Schauten, M., Willemstein, R., & Zwinkels, R. C. J. (2015). A tale of feedback trading by hedge funds. *Journal of Empirical Finance, 34*, 239–259.

Schmalensee, R. (1976). An experimental study of expectation formation. *Econometrica, 44*(1), 17–41.

Shiller, R. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review, 71*(3), 421–436.

Simon, H. (1957). *Models of man*. New York: Wiley.

Smith, V. L., Suchanek, G. L., & Williams, A. W. (1988). Bubbles, crashes, and endogenous expectations in experimental spot asset markets. *Econometrica, 56*(5), 1119–1151.

Spronk, R., Verschoor, W. F., & Zwinkels, R. C. (2013). Carry trade and foreign exchange rate puzzles. *European Economic Review, 60*(C), 17–31.

Taylor, M. P., & Allen, H. (1992). The use of technical analysis in the foreign exchange market. *Journal of International Money and Finance, 11*(3), 304–314.

Ter Ellen, S., Hommes, C. H., & Zwinkels, R. C. J. (2017). *Comparing Behavioural Heterogeneity Across Asset Classes*. Norges Bank working paper, 12/2017.

Ter Ellen, S., Verschoor, W. F. C., & Zwinkels, R. C. J. (2013). Dynamic expectation formation in the foreign exchange market. *Journal of International Money and Finance, 37*, 75–97.

Ter Ellen, S., & Zwinkels, R. C. J. (2010). Oil price dynamics: A behavioural finance approach with heterogeneous agents. *Energy Economics, 32*(6), 1427–1434.

Verschoor, W. F. C., & Zwinkels, R. C. J. (2013). Do foreign exchange fund managers behave like heterogeneous agents? *Quantitative Finance, 13*(7), 1125–1134.

Vigfusson, R. J. (1997). Switching between chartists and fundamentalists: A Markov regime-switching approach. *International Journal of Economics and Finance, 2*(4), 291–305.

Westerhoff, F., & Reitz, S. (2003). Nonlinearities and cyclical behaviour: The role of chartists and fundamentalists. *Studies in Nonlinear Dynamics and Econometrics, 7*(4), 1558–3708.

Westerhoff, F., & Reitz, S. (2005). Commodity price dynamics and the nonlinear market impact of technical traders: Empirical evidence for the US corn market. *Physica A: Statistical Mechanics and its Applications, 349*(3), 641–648.

Winker, P., & Gilli, M. (2001). *Indirect Estimation of the Parameters of Agent Based Models of Financial Markets*. FAME Research Paper, 38.

Wolff, C. C. (1987). Forward foreign exchange rates, expected spot rates, and premia: A signal-extraction approach. *Journal of Finance, 42*(2), 395–406.

Zeeman, E. C. (1974). The unstable behavior of stock exchange. *Journal of Mathematical Economics, 1*(1), 39–49.

High Frequency Trading in the Equity Markets During US Treasury POMO

Cheng Gao and Bruce Mizrach

Abstract We analyze high frequency trading (HFT) activity in equities during US Treasury permanent open market (POMO) purchases by the Federal Reserve. We construct a model to study HFT quote and trade behavior when private information is released and confirm it empirically. We estimate that HFT firms reduce their inside quote participation by up to 8% during POMO auctions. HFT firms trade more aggressively, and they supply less passive liquidity to non-HFT firms. Market impact also rises during Treasury POMO. Aggressive HFT trading becomes more consistently profitable, and HFT firms earn a higher return per share. We also estimate that HFT firms earn profits of over $105 million during US Treasury POMO events.

JEL Classification G12, G21, G24

1 Introduction

High frequency trading (HFT) has grown since the adoption of the Regulation National Market System in 2005, and now represents the majority of equity trading volume in the USA. The impact of HFT on the equity markets has become a central question in the policy debate about market structure and in the academic literature on market microstructure.

HFT firms engage in a variety of strategies. Hagstromer and Norden (2013) divide these approaches into market making and opportunistic. Our work considers both aspects by looking at both liquidity provision and aggressive trading profits. Menkveld (2013) analyzes the arrival of the Chi-X high frequency platform in Europe and concludes that HFT firms act as market makers in the new market. Hasbrouck and Saar (2013) suggest that HFT activity improves traditional market

C. Gao · B. Mizrach (✉)
Department of Economics, Rutgers University, New Brunswick, NJ, USA
http://snde.rutgers.edu

F. Jawadi (ed.), *Uncertainty, Expectations and Asset Price Dynamics*,
Dynamic Modeling and Econometrics in Economics and Finance 24,
https://doi.org/10.1007/978-3-319-98714-9_4

quality measures such as short-term volatility, spreads, and displayed depth in the limit order book. Carrion (2013) studies a data set from Nasdaq that identifies HFT firms. He finds that HFT participants supply liquidity when it is low and take liquidity when it is high. Brogaard et al. (2014) analyze the same data set and argue that HFT increases price efficiency.

While HFT firms are often passive liquidity providers, this contribution asks whether their role changes during periods of market turbulence. This question is motivated in part by the "Flash Crash" of May 6, 2010 when over 200 stocks traded down to a penny bids before the market quickly rebounded. The U.S. Commodity Futures Trading Commission and Securities and Exchange Commission (2010) task force report analyzed HFT activity from the 12 largest firms during the crash. Half significantly curtailed their trading activity during the crash including two firms that stopped trading for the rest of the day. The "Flash Crash" helps to clarify why reporting the average effect of HFT firms on the market may provide a misleading portrait of their contribution to market quality. Analyzing their impact when the market is under stress or reacting to news needs to be isolated from their contribution during less turbulent periods.

Benos and Wetherilt (2012) note that HFT firms are in competition with designated market makers (DMMs). They emphasize that the HFT firms have no affirmative quoting obligations. This allows them to "compete with DMMs when market-making is profitable but withdraw altogether from the market when it is not…"

We examine this claim by looking at periods of potential market stress, the US Treasury purchases made by the Fed beginning in late 2008 as part of its quantitative easing program. The Federal Reserve's asset purchase program began in November 2008 with \$600 billion of GSE debt and mortgage backed securities. In March 2009, the Federal Open Market Committee enlarged these programs and authorized purchases of \$300 billion in long-term Treasuries. We examine 57 auctions between 2008 and 2010.

We develop a theoretical model in which HFT firms receive valuable private information before other market participants. This information will lead HFT firms to temporarily abandon their role as liquidity providers and trade aggressively in the direction of the news. Consistent with our model's prediction that HFT firms reduce their inside quoting activity, we find that during Treasury POMO auctions HFT firms reduce their inside bid participation by 8%. The model also predicts that HFT firms would trade more aggressively when obtaining valuable information. We do find that HFT firms buy more frequently in good news and sell more often in bad news using aggressive orders. We also find that HFT firms are less likely to supply liquidity to non-HFT firms that trade in the direction of news. These results are even stronger once we control for microstructure effects.

The ability of HFT firms to receive private information may create additional price impact. Zhang (2010) observes that HFT is positively correlated with stock price volatility and hinders the ability of the market prices to reflect fundamental information. Cvitanic and Kirilenko (2010) provide a theoretical perspective and show that HFT activity effects volume and the distribution of transaction prices.

Martinez and Roşu (2013) model HFT participants as informed traders who observe news stream and trade quickly. They find that HFT generates volatility and decreases liquidity.

Consistent with these studies, we find that the release of auction bidding information raises market impact. A 1000 share order from a HFT firm moves the market on average \$0.0318, but on POMO days this rises to \$0.0341. This is evidence that high frequency traders appear to have superior information.

Whether they are at the active or passive side, HFT trades are more profitable when the counterpart is a non-HFT firm rather than a HFT firm. High frequency traders are able to generate the most profit from private information because of their ability to trade quickly. Baron et al. (2017) study the profitability of HFT in the E-mini futures contract and find that HFT firms make high and persistent profits from all categories of non-HFT participants. Hirschey (2013) provides evidence that HFT firms anticipate the order flow from non-HFT investors and their aggressive trades are highly correlated with future returns.

We find that HFT firms are consistently profitable trading during POMO events. They are profitable 88% of the time on aggressive trades and 100% of the time on passive trades. We estimate a daily average profit per stock of \$1300.35 which rises to \$1895.37 on POMO days. The profits per share from aggressive trading rise 300%. Extrapolating these results to the market as a whole, we estimate profits of more than \$105 million.

The chapter is organized as follows. We develop a theoretical model in Sect. 2 to study HFT behavior when private information is released. Section 3 describes the HFT data set. Section 4 describes the POMO purchases by the Federal Reserve. In Sect. 5, we analyze HFT quote and trade activities during POMO and provide empirical support for our model. Additional empirical results on market impact and profits of HFT are presented in Sect. 6. We perform robustness checks in Sect. 7, and Sect. 8 concludes.

2 The Model

2.1 Model Setup

Consider a risky security, with the terminal value V, that changes from its initial value V_0 based on random innovations $\varepsilon \sim N(0, \sigma^2)$ and fundamental information,

$$V = V_0 + \varepsilon + \eta. \tag{1}$$

η is the expected change in the fundamental value due to the information arrival of POMO auctions. η is assumed to be independent from ε and can take three values: $\eta = +\delta > 0$ if the news is positive, $\eta = -\delta$ if the news is negative, and $\eta = 0$ if no information arrives.

There are three types of traders in our model: noise traders (NTs), limit order traders (LOTs), and high frequency traders (HFTs). NTs submit orders for liquidity reasons and use market orders that hit the bid or offer on the limit order book. We assume that a noise trader arrives exogenously with probability θ, and will submit a buy order with probability γ or a sell order with probability $1-\gamma$.

LOTs provide liquidity by placing bid and offer quotes competitively. HFTs are profit maximizing. They trade either passively to earn the bid-ask spread by submitting limit orders or aggressively to realize a positioning profit using marketable limit orders. We assume in our model that HFTs trade faster than NTs and LOTs in the sense that they are more quickly informed of the value of η than noise traders and limit order traders.

To simplify the analysis we assume that all orders by each type of traders are for one unit of the security. Because the order flow of noise traders is exogenous, we only need to focus on two players: LOTs and HFTs. We then analyze their decision problems under different market conditions.

2.2 Limit Order Traders

LOTs do not observe the value of η, but they infer its value based on trading activity. Their conjecture about the probability distribution of η: $\eta = +\delta$ with probability α, $\eta = -\delta$ with probability β, and therefore $\eta = 0$ with probability $1-\alpha-\beta$. The unconditional expectation of η is calculated as $\overline{\eta} = \delta(\alpha-\beta)$. LOTs post bid and offer quotes at B and A respectively, and are aware that HFTs have superior information about η. Following Glosten and Milgrom (1985), we consider buys and sells separately. Given that other traders buy, the expected profit of LOTs is

$$\mathrm{E}\left[\pi_{\mathrm{LOT}}|\mathrm{Buy}\right] = A - \mathrm{E}\left[V|\mathrm{Buy}\right] = A - V_0 - \mathrm{E}\left[\eta|\mathrm{Buy}\right]. \tag{2}$$

To calculate the expectation of η given that the trade is a buy, we first compute the conditional probabilities of η.

$$\begin{aligned}\Pr(\eta = +\delta|\mathrm{Buy}) &= \frac{\Pr(\eta = +\delta, \mathrm{Buy})}{\Pr(\mathrm{Buy})} = \frac{\alpha(1+\theta\gamma)}{\alpha+\theta\gamma}, \\ \Pr(\eta = -\delta|\mathrm{Buy}) &= \frac{\Pr(\eta = -\delta, \mathrm{Buy})}{\Pr(\mathrm{Buy})} = \frac{\beta\theta\gamma}{\alpha+\theta\gamma}.\end{aligned} \tag{3}$$

We assume that competition among LOTs drives their expected profit to a positive amount c_{LOT}. Therefore, the best offer is set as

$$A = V_0 + \mathrm{E}\left[\eta|\mathrm{Buy}\right] + c_{\mathrm{LOT}} = V_0 + \overline{\eta} + \frac{\delta\alpha(1-\alpha+\beta)}{\alpha+\theta\gamma} + c_{\mathrm{LOT}}. \tag{4}$$

Similarly, we can obtain the best bid by LOTs. The conditional probabilities of η given that other traders sell is

$$\Pr(\eta = +\delta|\text{Sell}) = \frac{\Pr(\eta = +\delta, \text{Sell})}{\Pr(\text{Sell})} = \frac{\alpha\theta(1-\gamma)}{\beta + \theta(1-\gamma)}, \tag{5}$$
$$\Pr(\eta = -\delta|\text{Sell}) = \frac{\Pr(\eta = -\delta, \text{Sell})}{\Pr(\text{Sell})} = \frac{\beta(1 + \theta(1-\gamma))}{\beta + \theta(1-\gamma)}.$$

If the expected profit is driven to c_{LOT} by competition, the best bid is set as

$$B = V_0 + \text{E}[\eta|\text{Sell}] - c_{\text{LOT}} = V_0 + \overline{\eta} - \frac{\delta\beta(1 + \alpha - \beta)}{\beta + \theta(1-\gamma)} - c_{\text{LOT}}. \tag{6}$$

The bid-ask spread is

$$A - B = \frac{\delta[2\alpha\beta + \theta(\alpha(1-\gamma)(1+\beta-\alpha) + \beta\gamma(1+\alpha-\beta))]}{(\alpha + \theta\gamma)(\beta + \theta(1-\gamma))} + 2c_{\text{LOT}}. \tag{7}$$

It is not hard to show that the spread is always positive. As seen in (7), the bid-ask spread can be decomposed into two components for LOTs. The first term captures the adverse selection risk and the second one covers the inventory cost.

In the symmetric case that LOTs' conjecture on positive or negative news has equal probability, i.e. $0 \leq \alpha = \beta \leq \frac{1}{2}$, the bid-ask spread reduces to

$$A - B = \frac{\delta\alpha(2\alpha + \theta)}{(\alpha + \theta\gamma)(\alpha + \theta(1-\gamma))} + 2c_{\text{LOT}}. \tag{8}$$

2.3 High Frequency Traders

HFTs maximize their profit using either limit orders or marketable limit orders. They can expect to earn the bid-ask spread on passive trades by placing limit orders. By using marketable limit orders HFTs must pay the spread. A trader may want to do so because valuable limit orders can disappear quickly given the competition from other HFTs and the cancellation of limit orders. In this way they expect to gain trading profits. We assume that HFTs are informed of the value of η under news release and have the same conjecture as LOTs about the distribution of η when no information arrives. We then study the optimal order placement decision of HFTs based on different market conditions.

When there is no news expected on POMO auctions, HFTs' conjecture about the terminal security value is

$$\text{E}[V|\text{non}-\text{POMO}] = V_0 + \overline{\eta}. \tag{9}$$

Since it lies between the best bid and offer quotes by LOTs, they could expect a loss if they submit marketable limit orders by crossing the spread. For example, the expected profit for a buy marketable limit order would be

$$\mathrm{E}\left[V|\text{non}-\text{POMO}\right]-A=-\frac{\delta\alpha\left(1-\alpha+\beta\right)}{\alpha+\theta\gamma}-c_{\mathrm{LOT}}<0. \tag{10}$$

Instead, HFTs are better off under no expected news if they provide liquidity by posting bid and offer quotes and earn the bid-ask spread on passive trades.

HFTs place their quotes at the same bid and offer prices as the limit order traders. They are not adversely selected by other traders, so they would earn a higher expected profit than LOTs at the bid and offer quotes specified in (6) and (4). At the bid side HFTs expect to have a profit of

$$c_{\mathrm{HFT}}^{B}=\frac{\delta\beta\left(1+\alpha-\beta\right)}{\beta+\theta(1-\gamma)}+c_{\mathrm{LOT}}, \tag{11}$$

and their expected profit at the offer side would be

$$c_{\mathrm{HFT}}^{A}=\frac{\delta\alpha\left(1-\alpha+\beta\right)}{\alpha+\theta\gamma}+c_{\mathrm{LOT}}. \tag{12}$$

When a positive information of POMO auctions is expected, HFTs' conjecture about the terminal security value is

$$\mathrm{E}\left[V|\text{positive}\right]=V_0+\delta. \tag{13}$$

For a marketable limit order purchase, their expected profit is

$$\mathrm{E}\left[\pi_{\mathrm{HFT}}|\text{positive}\right]=V_0+\delta-A=\delta\left(1-\alpha+\beta\right)-c_{\mathrm{HFT}}^{A}. \tag{14}$$

It is positive when $\delta > c_{\mathrm{HFT}}^{A}/\left(1-\alpha+\beta\right)$. This suggests that HFTs would take the profitable opportunity to buy the security at the offer quote A by LOTs when they are informed of a good news with a relatively big rise of the equity value. Although they pay the spread in this way, HFTs can earn trading profits because of their superior information about the news.

It also indicates that in this situation HFTs would withdraw their liquidity provision at the inside offer and then post a higher offer quote at $V_0+\delta+c_{\mathrm{HFT}}^{A}$.

The analysis for HFTs' order strategy with a negative expected information is similar. Their expected profit for a sell marketable limit order is

$$\mathrm{E}\left[\pi_{\mathrm{HFT}}|\text{negative}\right]=B-(V_0-\delta)=\delta\left(1+\alpha-\beta\right)-c_{\mathrm{HFT}}^{B}, \tag{15}$$

which is greater than zero when $\delta > c_{\mathrm{HFT}}^{B}/\left(1+\alpha-\beta\right)$. It suggests that HFTs would sell the security to LOTs at the bid quote B when they expect a bid drop of the equity value. In this case they would choose to scale back from the inside bid and then post a lower bid at $V_0 - \delta - c^B_{\mathrm{HFT}}$. We then validate these theoretical implications by analyzing their quoting and trading activities during the period of Treasury POMO auctions.

3 HFT Data Set

We utilize a data set from Nasdaq that identifies HFT firms. This is the same data set used by Carrion (2013) and Brogaard et al. (2014). The data tracks 120 stocks, listed in Table 1, and has information at different intervals and samples about quotes and trades from 26 HFT firms.

The trade information is most complete. It includes all trades on the Nasdaq exchange during regular market hours, apart from the opening and closing crosses, from January 2008 to December 2009, plus the week of February 22–26, 2010. We begin our analysis in December 2008 with the onset of POMO activity by the Federal Reserve. This sample covers the entire first round of asset purchases by the Federal Reserve. The data set tells whether a HFT firm initiated or filled a trade. These 26 firms are involved in 76% of all the trading activity during the period January 2008 through February 2010.

There are detailed Nasdaq order book data snapshots sampled from the first week of each quarter from January 2008 to December 2009, and then February 22–26, 2010. We observe whether a HFT firm is providing liquidity at each tier of the order book. To supplement the HFT data for our market impact analysis, we make

Table 1 Stocks in HFT database

AA	AZZ	CDR	CSE	FFIC	IMGN	MANT	PFE
AAPL	BARE	CELG	CSL	FL	INTC	MDCO	PG
ABD	BAS	CETV	CTRN	FMER	IPAR	MELI	PNC
ADBE	BHI	CHTT	CTSH	FPO	ISIL	MFB	PNY
AGN	BIIB	CKH	DCOM	FRED	ISRG	MIG	PPD
AINV	BRCM	CMCSA	DELL	FULT	JKHY	MMM	PTP
AMAT	BRE	CNQR	DIS	GAS	KMB	MOD	RIGL
AMED	BW	COO	DK	GE	KNOL	MOS	ROC
AMGN	BXS	COST	DOW	GENZ	KR	MRTN	ROCK
AMZN	BZ	CPSI	EBAY	GILD	KTII	MXWL	ROG
ANGO	CB	CPWR	EBF	GLW	LANC	NC	RVI
APOG	CBEY	CR	ERIE	GOOG	LECO	NSR	SF
ARCC	CBT	CRI	ESRX	GPS	LPNT	NUS	SFG
AXP	CBZ	CRVL	EWBC	HON	LSTR	NXTM	SJW
AYI	CCO	CSCO	FCN	HPQ	MAKO	PBH	SWN

The table lists the 120 ticker symbols in the HFT database provided by Nasdaq

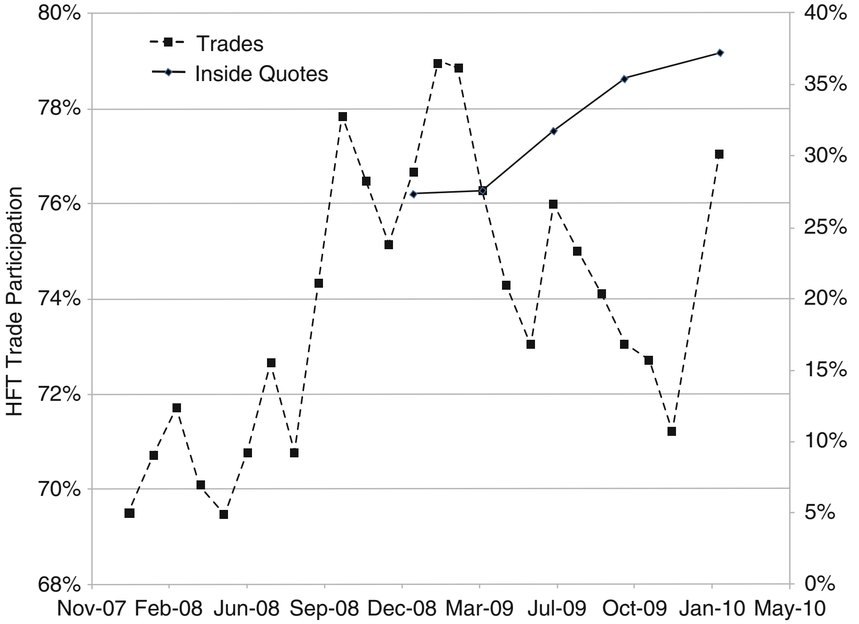


Fig. 1 Trade and inside quote activity due to HFT firms. This figure presents the percentage of trades and inside quote participation by HFT firms

use of the ITCH data set. ITCH provides full order book level detail for the Nasdaq market, but it does not provide any HFT information. We only analyze inside quote activity in both data sets though.

Market participants[1] and regulators[2] have been concerned about the size and scope of HFT activity in recent years. Our data set documents a growing role for HFT activity. Figure 1 plots the monthly average percentage of HFT trades. HFT trading activity appears to trend up in 2008, back down in 2009, before stabilizing in early 2010.

Another measure of HFT liquidity is the extent to which HFT firms make up the inside quote. We also graph this frequency in Fig. 1. Inside quote activity continues to uptrend in 2009, unlike the trade series. We will control for these trends in our analysis of the Federal Reserve auctions.

[1] See, e.g., Christopher Matthews, "High Frequency Trading: Wall Street's Doomsday Machine?", *Time Magazine*, August 8, 2012.

[2] SEC Chairman Mary Jo White, in testimony before the Senate Banking Committee on March 13, 2013, noted "..high frequency trading, complex trading algorithms, dark pools, and intricate new order types raise many questions and concerns."

4 Permanent Open Market Operations (POMO)

After the federal funds rate reached the zero lower bound in December 2008, the Federal Reserve began its large scale asset purchases in March 2009. The Fed increased reserve bank credit from $893 billion on September 4, 2008 to $2298 billion on March 25, 2010 during (what turned out to be) the first round of quantitative easing (QE1). The Federal Reserve not only purchased US Treasuries as it normally would, it also bought GSE mortgage backed securities and debt. Because these assets were intended to remain on the balance sheet for an extended period, they were called "permanent" open market operations (POMO).[3]

We focus on US Treasury purchases because Treasury securities play a unique role in the asset markets. We are motivated by the work of Lou, Yan and Zhang (LYZ, 2013) who find that regularly scheduled Treasury auctions have price impacts on the Treasury, repo and equity markets. They link these effects to the limited risk bearing capacity of the primary dealers and fund flows influenced by temporary price distortions. Krishnamurthy and Vissing-Jorgensen (2012) highlight the unique role of Treasuries and note that changes in their supply effect the equilibrium price of liquidity and safety.

4.1 Announcement Effects

Gagnon et al. (2011) isolate five critical days in the evolution of the Fed's unconventional monetary policy. The days include the announcement of the program on November 25, 2008, the extension of asset purchases to the Treasury market on December 16, 2008, and the enlargement of the program on March 18, 2009. Gagnon et al. (2011) note that the Treasury market fell a cumulative 107 basis points on those 5 days.

The stock market reactions are in Table 2. On these 5 days, the CRSP value weighted index rose 5.03%. Between November 24, 2008 and March 24, 2010, the S&P 500 index rose from 851.81 to 1167.72, a gain of 37%. The belief that QE benefited equities is widely held and led to the view that Bernanke has placed a "put" under the market.[4]

4.2 Details of the Auctions

US Treasury security purchases began in March 2009. The Fed purchased $295.4 billion in 57 auctions in which dealers offered $1137 billion for sale. Maturities

[3]The history and motivation of the Federal Reserve program is analyzed in Gagnon et al. (2011).

[4]See, e.g., David Tepper, manager of the Apaloosa Hedge Fund, in the *Financial Times* of October 8, 2010.

Table 2 Stocks and bonds on POMO announcement days

Date	Stocks	10Y bond	Event
25-Nov-2008	4.83%	−36	LSAP announced
01-Dec-2008	−5.21%	−25	Bernanke speech
16-Dec-2008	3.90%	−33	LSAP extended to treasuries
01-Jan-2009	0.09%	28	Fed stands ready to buy more
18-Mar-2009	1.42%	−41	LSAPs enlarged
Event sum	5.03%	−107	

The table reports the 2-day changes ($t - 1$ to $t + 1$) in the CRSP value weighted stock index and the 10-year bond yield on the five event days

ranged from 2 to 30 years for 160 different CUSIPs. This represented about 3% of the outstanding Treasury debt, and about 8% of the available Treasury supply.[5]

D'Amico and King (2013) provide details on the implementation of the Treasury purchases. On every other Wednesday, the Open Market Desk at the Federal Reserve Bank of New York would announce the range of the yield curve they were purchasing and the dates on which bids could be submitted. At 10:15 AM on each auction day, the Fed would publish a list of CUSIPs that it would consider purchasing. Most days, the bidding would commence at 10:30 AM. Shortly after bidding closed at 11:00 AM, the Fed used a confidential algorithm to determine which bids to accept.

Table 3 provides details on the first Treasury purchase on March 25, 2009. The Federal Reserve announced that it would consider purchasing securities with maturities between February 29, 2016 and February 15, 2019.

It listed 18 CUSIPs in this maturity range, but excluded one security, the 5.125% note, maturing on June 15, 2016. On March 25, they accepted bids on 13 of the securities, buying $7.5 billion overall. This was 31% of the $21.9 billion submitted. We find below that the ratio of accepted bids to those submitted captures the liquidity effect of the auctions on the equity market.

4.3 Effect of POMO Flows on the Equity Market

Lou, Yan and Zhang's (LYZ, 2013) work suggests that dealers have a limited risk bearing capacity, and that following US Treasury auctions, capital returns to other markets, raising equities by nearly 50 basis points. LYZ suggest that this effect on the stock market operates through asset reallocations by hedge funds and mutual funds.

[5] There are three purchases of Treasury Inflation Protected securities (TIPS) totaling $4.5 billion, but we did not include them in our analysis.

Table 3 US Treasury purchase detail for March 25, 2009

Release time:		10:30
Close time:		11:00
Settlement date:		March 26, 2009
Maturity/call date range:		02/29/2016–02/15/2019
Total par amt $	Submitted	21,937,000,000
	Accepted	7,500,000,000

CUSIP	Coupon	Maturity	Par amt accepted ($)
912828KS8	2.6250	2/29/2016	2,836,000,000
912810DW5	7.2500	5/15/2016	115,000,000
912828FQ8	4.8750	8/15/2016	1,031,000,000
912828FY1	4.6250	11/15/2016	739,000,000
912810DX3	7.5000	11/15/2016	147,000,000
912828GH7	4.6250	2/15/2017	35,000,000
912828GS3	4.5000	5/15/2017	950,000,000
912810DY1	8.7500	5/15/2017	238,000,000
912828HA1	4.7500	8/15/2017	702,000,000
912810DZ8	8.8750	8/15/2017	159,000,000
912828HH6	4.2500	11/15/2017	0
912828HR4	3.5000	2/15/2018	0
912828HZ6	3.8750	5/15/2018	0
912810EA2	9.1250	5/15/2018	23,000,000
912828JH4	4.0000	8/15/2018	0
912828JR2	3.7500	11/15/2018	0
912810EB0	9.0000	11/15/2018	193,000,000
912828KD1	2.7500	2/15/2019	0
912810EC8	8.8750	2/15/2019	332,000,000
Exclusions			
912828FF2	5.125	6/15/2016	0

This is the first of 214 Treasury purchases between March 2009 and June 2011. Details can be found on the New York Federal Reserve web site, http://www.newyorkfed.org/markets/pomo/display/index.cfm

We are able to identify a channel from the POMO auctions into equities at an intra-daily frequency. We regress the 15-min CRSP value weighted return[6] on the accepted/submitted ratio of bids in the POMO auctions. A high ratio here indicates that firms may be freeing up more capital to redeploy elsewhere.

The empirical estimates in Table 4 support a POMO liquidity channel into stocks. The average accepted/submitted ratio in the sample is 27.62%. When this ratio rises to 43.75%, stocks rise 0.50%. More than 13% of 15-min returns are explained by

[6]This effect is present from 1-min up to 30-min after the auction. The peak impact on equity returns is at the 15-min horizon.

Table 4 POMO stock liquidity model

Dep. variable: 15-min value weighted return	
Intercept	−0.0028 (0.001)
Accepted/submitted bid ratio	0.0114 (3.097)
$\overline{R}^2$	0.1330

t-statistics in parentheses
The table reports the regression model estimate for the 15-min value weighted CRSP return on the 57 Treasury POMO days between November 2008 and March 2010

this ratio. This use of weighted averages is similar to the results in Abou and Prat (2000).

These auctions results show that while the total amount of assets to be purchased was known prior to the auction, the amount of buying and selling interest did provide news to the market. We then try to model in the next section how HFT firms might alter their trading activity upon receipt of this news.

The goal of the large-scale asset purchases was "an effort to drive down private borrowing rates, particularly at longer maturities." Using an event study, Gagnon et al. (2011) conclude that 10-year Treasury bond yields fell 91 basis points and that 10-year agency debt yields declined 156 basis points. Joyce et al. (2011) find that a program of similar scale in the UK lowered gilt yields by 100 basis points.

Central bank asset purchases can also have impact on related asset markets. Neely (2010) and Joyce et al. (2011) have both emphasized the portfolio balance channel in which declining exposure to Treasury also raises other asset prices. Neely (2010) shows that announcements related to the US asset purchase program also lowered 10-year government bond yields in Australia, Canada, Germany, Japan and the UK between 19 and 78 basis points. Krishnamurthy and Vissing-Jorgensen (2011) estimate that US corporate bond yields fell between 43 and 130 basis points. Neely (2010) also finds evidence for reallocation into the US stock market: the S&P 500 index rises a cumulative 3.42%.

Even though the size of Fed's overall program was largely known by to the market by March 2009, the specific securities they would buy and the bids they would accept at each auction were not. The participation levels and prices paid, just like any auction, reveal information to the markets. Lou et al. (2013) suggest that the bid-to-cover ratio is likely to be an informative signal. We find that a closely related variable, the accepted-to-submitted ratio explains up to 13% of market returns in the period following the auction.

5 Empirical Results

We analyze HFT quote and trade activities during US Treasury POMO auctions. We find that (1) HFT firms pull back as market makers during periods of information release; (2) HFT firms use information to trade aggressively in the direction of the news; (3) HFT firms provide less passive liquidity on the opposite side of the news; (4) market impact rises during US Treasury POMO auctions; (5) HFT profits rise during POMO events. The empirical evidence provides support to our theoretical model.

5.1 HFT Firms Pull Back from the Inside Quote

Our model implies that, around the release of news, market makers should become more cautious. Admati and Pfleiderer (1988) have emphasized that the risk of trading against valuable private information is higher, and market makers should widen spreads and reduce their depth.

To examine this empirically, we estimate how frequently HFT firms participate in the inside quote on the Nasdaq market. Our order book overlaps with the US Treasury POMO auctions on 5 trading days.

We calculate the percentage of ticks in which the HFT firms is at the inside bid or offer. Interpreting this raw percentage requires some care. First, we have to account for the trend in the data that we noted in Fig. 1. We find that a quadratic trend fits the data well.

The data are also seasonal intra-daily. The vast majority of POMO auctions occur between 10:30 and 11:00 AM. This is a relatively quiet time during the day in which HFT participation tends to fall off. Therefore, we include a time dummy for the period from 10:30 to 11:00 AM in the model.

We also need to control for the typical microstructure factors that influence order aggressiveness. These include realized volatility, which we measure as a ten-tick moving average, trade volume, and the order imbalance of buyer less seller initiated trades. These variables are all lagged one period.

We estimate the model in the probit form with robust standard errors on a pooled cross-section of the 120 stocks, using maximum likelihood. We report the result for the bid and offer side, respectively. Table 5 shows that the participation rate of HFT firms in the inside quote falls significantly during Treasury purchases, whether it is a positive or negative news. The result is consistent with our model.

HFT firms are almost 8% less likely to quote on the inside bid and 5% less frequently on the inside offer during US Treasury POMO auctions.

Table 5 HFT inside quote frequency during US treasury POMO

Variable	Bid	Offer
Trend	0.0022 (62.38)	0.0017 (50.34)
$\text{Trend}^2/1000$	−0.0012 (−16.00)	−0.0008 (−10.93)
Returns_{t-1}	0.0206 (1.92)	−0.0368 (−3.55)
Realized vol_{t-1}	−0.0146 (87.94)	−0.0121 (−77.25)
$\text{Volume}_{t-1}/1000$	0.0745 (221.70)	0.0401 (158.12)
Order $\text{imbalance}_{t-1}/1000$	1.6756 (13.36)	−1.1129 (−10.29)
S10:30–11:00	−0.0266 (−4.62)	−0.0335 (−5.92)
Positive UST news	−0.0763 (−4.00)	−0.0450 (−2.43)
Negative UST news	−0.0772 (−6.87)	−0.0458 (−4.16)
Constant	−0.4383 (−107.41)	−0.3251 (−82.64)
$\overline{R}^2$	0.1450	0.1155

t-statistics in parentheses

The table reports estimates of a model for the inside quote participation of the 26 HFT trading firms. We control for the growth in HFT activity using a linear and quadratic trend. We also include standard regressors for market making aggressiveness, past returns, volatility, volume, and order imbalance. We also include a time dummy for the quiet period from 10:30 to 11:00. Finally, we measure the effect of US Treasury POMO auctions using two dummy variables, one for positive news and the other for negative

5.2 *HFT Firms Trade More Aggressively in the Direction of News*

Given the fact that the HFT firms tend to withdraw liquidity from the inside quotes during POMO auctions, the other question to ask is whether they demand more liquidity from other non-HFT market participants. The trade data set tells us whether traders are HFT or non-HFT firms at both sides of a trade. We treat the HFT firms as a group and focus particularly on HN and NH trades, where the first letter refers to the liquidity seeker and the second to the liquidity provider. We study the trading behavior of the HFT firms when they expect a positive and a negative news, respectively.

Table 6 Unconditional HFT net buy counts

	Positive UST news		Negative UST news		Non-POMO	
	HN	NH	HN	NH	HN	NH
Avg.	16.08	−9.11	−11.51	7.63	2.34	−1.71
SD	46.46	36.70	36.76	33.66	35.57	33.91
H_0: $\overline{c}_{\text{POMO}} = \overline{c}_{\text{Non}}$						
t-stat	1.94	−1.48	−2.10	1.68		

This table reports the average number of HFT net buys at a 1-min frequency from 10:30 to 11:00. HFT net buy is defined as the difference between the number of HFT buyer and seller initiated trades. We calculate it on days with positive and negative news from US Treasury auctions and on non-POMO days, and in aggressive HN and passive NH trades respectively

We report, in Table 6, the average difference of the number of HFT buyer and seller initiated trades between 10:30 and 11:00 AM and the rest of the day, at a 1-min frequency.

We divide US Treasury purchases into positive and negative news events based on the 1-h equity return after the start of the auction. The event is treated as positive news if the average return across the 120 stocks is greater than zero, and a negative one otherwise. Among 57 Treasury purchases, there are 32 positive and 25 negative news events.

We compute the average difference on non-POMO days and on days with positive and negative news from US Treasury auctions. We then test the differences in these net buy counts during event and non-event periods. We find a statistically significant reduction in buyer initiated trades on negative news days, with a reduction of 345 net buy trades during the POMO period. There is an increase of 482 net buy trades on positive news days, although this result is only significant at the 10% level.

POMO announcements days are volatile periods for the market, and this should lead to a less aggressive trading posture for HFT firms. To confirm and perhaps strengthen the results in Table 6, we need to then control for microstructure factors. We add lagged returns, realized volatility, volume and order imbalances as before, as well as a seasonal time dummy. We also include two dummy variables for positive and negative news.

The dependent variable is the 1-min net differential between buyer and seller initiated trades. We then estimate a least squares model in Table 7 for HFT net trade counts in aggressive HN and passive NH trades, respectively.

We estimate a significantly positive effect of US Treasury POMO events on HFT net trades, indicating the more aggressive stance of the HFT firms during the auctions. Once we control for microstructure factors, HFT firms increase their net buying by 600 trades on good news days and decrease their net buying by 891 trades when there is bad news. This result is similar to Brogaard et al. (2014) who find that, marketwide, HFT firms trade in the direction of the news flow.

Table 7 HFT net buy counts during US treasury POMO

Variable	HN	NH
Returns_{t-1}	−80.3836	−54.0185
	(−30.40)	(−18.46)
$\text{Realized vol}_{t-1}$	−0.0332	−0.0526
	(−2.90)	(−3.04)
$\text{Volume}_{t-1}/1000$	−0.0020	0.0021
	(−2.21)	(1.86)
$\text{Order imbalance}_{t-1}/1000$	0.0553	−0.0674
	(33.03)	(−32.03)
S10:30–11:00	0.0471	−0.0317
	(3.61)	(−2.04)
Positive UST news	0.1667	−0.1328
	(5.40)	(−3.65)
Negative UST news	−0.2476	0.1360
	(−7.59)	(3.50)
Constant	−0.0235	0.0075
	(−4.64)	(1.31)
$\overline{R}^2$	0.0024	0.0031

t-statistics in parentheses
The table reports estimates of models for HFT net trade counts in aggressive HN and passive NH trades, respectively. We include standard regressors for trading aggressiveness, past returns, volatility, volume, and order imbalance. We also include a time dummy for the quiet period from 10:30 to 11:00. Finally, we measure the effects of positive and negative US Treasury auctions using two dummy variables, respectively

5.3 HFT Firms Reduce Their Passive Liquidity Supply

We then do the same comparison in Table 6 for NH trades in which HFT firms are the passive liquidity suppliers. We find that non-HFT firms reduce their net buys by 273 trades on positive news days and increase their net buys by 229 trades on bad news days. This indicates that HFT firms have become more reluctant to supply passive liquidity to noise traders trading in the direction of the news. Neither of these results is significant at the 10% level though.

Introducing microstructure variable controls helps to isolate the effects predicted by our model. When we regress NH net buyers counts, the effect of the POMO auctions becomes much more strongly significant. Non-HFT firms decrease their net buying by 478 trades on good news days and increase their net buying by 490 trades with bad news.

We have now confirmed three of the primary predictions of the model. HFT firms become less active participants in the inside market on either the bid or offer. HFT firms increase their net buying activity on good news days and decrease on bad news days. Finally, we show that non-HFT firms are not able to trade as aggressively as HFT firms in the direction of the news because the HFT firms reduce their passive liquidity supply.

6 Additional Effects of HFT Activity

In this section, we study market impact and profits during the POMO auctions. Our model anticipates that the release of private information during the POMO events should raise market impact. Wider spreads and more informed trading should also lead to higher trading profits.

6.1 Market Impact of Trades by HFT Firms Becomes Higher

Another measure of liquidity is the market impact of trades. This is a dynamic indicator which incorporates the bid-ask spread, market depth, the persistence in order flow, and the resiliency of the order book.

Let $r_{i,t}$ be the change in the midpoint of the bid-ask spread, $(p^b_{i,t} + p^a_{i,t})/2 - (p^b_{i,t-1}+p^a_{i,t-1})/2$. $x_{i,t} \in \{-1,+1\}$ is an indicator variable which measures the trade direction. It is assigned as $+1(-1)$ if the transaction is a buy(sell). Let $V_{i,t}$ denote the size of the trade. We follow Hasbrouck (1991) using a vector autoregressive (VAR) model of their dynamic interaction. We also use Hasbrouck's identifying assumption that the current trade can effect the current quote, but not vice versa,

$$r_{i,t} = a_{r,0} + \sum_{j=1}^{10} a_{r,j} r_{i,t-j} + \sum_{j=0}^{10} b_{r,j} x_{i,t-j} V_{i,t-j} + \varepsilon_{r,t}, \tag{16}$$

$$x_{i,t} V_{i,t} = a_{x,0} + \sum_{j=1}^{10} a_{x,j} r_{i,t-j} + \sum_{j=1}^{10} b_{x,j} x_{i,t-j} V_{i,t-j} + \varepsilon_{x,t}. \tag{17}$$

We use 10 lags in the VAR. The estimates are not sensitive to this choice.

Market impact is a dynamic process

$$\partial r_{i,t+j} / \partial x_t V_t \tag{18}$$

which we will now compute during POMO and non-POMO intervals. We sum the aggregate effect

$$\overline{\Lambda} = \frac{1}{120} \sum_{i=1}^{120} \sum_{j=1}^{50} \partial r_{i,t+j} / \partial x_{i,t} V_{i,t} \tag{19}$$

arbitrarily after 50 trades, filtering out negative impacts.

Table 8 HFT market impact

	10:30–11:00		11:00–11:30		13:45–14:15	
	POMO	Non-POMO	POMO	Non-POMO	FOMC	Non-FOMC
Avg. (10^{-5})	3.4138	3.1761	3.0856	3.0129	2.8937	2.8927
SD (10^{-5})	0.3007	0.3037	0.4289	0.2251	0.2317	0.3630
H_0: $\overline{\Lambda}_{\text{POMO}} = \overline{\Lambda}_{\text{Non}}$						
t-stat	2.54		1.10		0.71	

This table reports the average market impact calculated by (19). We use the eight FOMC announcements in 2009: January 29, March 18, April 29, June 24, August 12, September 23, November 4, and December 16. For the FOMC results we use the 60 Nasdaq stocks in the sample with quotes from the ITCH feed

The number of POMO days we can include is limited by the availability of our Nasdaq inside quote data. We have only 5 US Treasury POMO days to contrast with 14 non-POMO days. To do reasonable comparisons, we expand the sample to 14 POMO days using Nasdaq ITCH data.

Our HFT data set classifies trades into four categories. The trade has an aggressor and a passive supplier. Either can be a HFT or not. We report the comparison of average market impact by HFT trades, $x_t \in \{x_t^{HH}, x_t^{HN}, x_t^{NH}\}$, across the 120 stocks in Table 8.

We find, as our quote analysis indicated, that market impact from HFT is significantly higher during the US Treasury POMO auctions than the corresponding period on non-POMO days. A 1000 share order moves the market on average \$0.0318, but on POMO days this rises to \$0.0341. The rise in market impact of trades could make the trading costs of non-HFT firms even higher. These nonlinear market impacts are consistent with the empirical findings in Jawadi and Prat (2012).

6.2 *HFT Firms Make More Profits During POMO*

Menkveld (2013) makes a useful division of trading profits for a HFT firm. On passive trades, designated NH in our sample, they can expect to earn the bid-ask spread. On aggressive trades, designated HN, they must pay the spread, hoping to realize a positioning profit.

Under some assumptions, we can estimate the profitability of the HFT firms as a group using our trade data. We assume that HFT firms try to end the day flat and assess their profits by valuing any position at the day's average price. By construction, we consider only HN and NH trades.

The HFT daily profits for stock i are estimated as

$$\pi_i^{\text{HFT}} = \sum_{t=1}^{T} \left[D_{i,t}^S p_{i,t} q_{i,t} - D_{i,t}^B p_{i,t} q_{i,t}\right] + \frac{\sum_{t=1}^{T} p_{i,t} q_{i,t}}{\sum_{t=1}^{T} q_{i,t}} \sum_{t=1}^{T} \left[D_{i,t}^B q_{i,t} - D_{i,t}^S q_{i,t}\right], \tag{20}$$

Table 9 HFT daily profits per stock

	UST POMO	Non-POMO
Mean	\$1895.37	\$1300.35
SD	2736.70	3642.63
H_0: $\overline{\pi}^{\text{HFT}}_{\text{POMO}} = \overline{\pi}^{\text{HFT}}_{\text{Non}}$		
t-stat	1.65	

This table reports estimated HFT daily profits per stock by (20) during US Treasury POMO and non-POMO days

Table 10 HFT daily profits per share

	Total		HN		NH	
	UST	Non-POMO	UST	Non-POMO	UST	Non-POMO
Mean	\$0.0178	\$0.0129	\$0.0099	\$0.0032	\$0.0341	\$0.0298
SD	0.0085	0.0105	0.0132	0.0158	0.0149	0.0177
Min	−\$0.0036	−\$0.0254	−\$0.0235	−\$0.0594	\$0.0064	−\$0.0084
Max	\$0.0385	\$0.0422	\$0.0550	\$0.0484	\$0.0897	\$0.1002
% Days>0	96.49%	88.13%	87.72%	60.63%	100.00%	96.88%
H_0: $\overline{\pi}^{\text{HFT}}_{ps,\text{POMO}} = \overline{\pi}^{\text{HFT}}_{ps,\text{Non}}$						
t-stat	3.86		3.42		2.11	

This table reports estimated HFT daily profits per share by (21) during US Treasury POMO and non-POMO days. HFT profits are also calculated in aggressive HN and passive NH trades, respectively

where D^B and D^S are buy and sell indicators, respectively, $p_{i,t}$ is the price of stock i at time t, and $q_{i,t}$ is the quantity. It closes out open positions at the end of the day using daily average trade prices. The method of calculating profits is similar to Brogaard et al. (2014) and Baron et al. (2017).

Profits per stock for POMO US Treasury days are compared to profits on non-POMO days in Table 9.

On non-POMO days, we estimate profits of \$1300.35 per stock for the entire sample of trading days between December 2008 and February 2010. This compares to Brogaard et al.'s (2014) estimate of \$2284.89 for the entire trading sample back to January 2008. We find that HFT firms increase their average daily profits by 46% on US Treasury POMO days. The increase in profit of \$595.02 per stock is marginally significant at 10% level.

To approximate returns from HFT activity, we also estimate in Table 10 the profits per share $\pi^{\text{HFT}}_{i,ps}$ from their aggressive and passive trades,

$$\pi^{\text{HFT}}_{i,ps} = \frac{\pi^{\text{HFT}}_{i}}{\sum_{t=1}^{T} q_{i,t}/2}. \tag{21}$$

Given an average share price of around \$30, the returns are quite modest. The trades, however, are very short term and rarely lose money. On US Treasury POMO days, profits per share are positive 96.49% of the time. On the 2 days where the HFT firms lose money, they lose only 2/10 and 3.6/10 of one cent per share, respectively, compared with the largest gain of nearly \$0.04 per share on April 30, 2009.

In HN trades, HFT firms also rarely lose money on Treasury POMO days. They make profits 87.72% of the time. Crossing the spread on non-POMO days is much more risky. Profits are positive on only 60.63% of non-POMO trading days. The average profit per share when crossing the spread is typically small, only \$0.0032 per share, but this rises by a statistically significant 300% during POMO auctions. The wider spreads on POMO days, while helping their passive profits, should reduce.

In NH trades where HFT firms are passive liquidity providers, profits per share are always positive on US Treasury POMO days. Compared to their performance on non-POMO days, HFT firms increase the average profits per share by 38% on US Treasury POMO days, and the effect is statistically significant.

On POMO days, HFT firms became more aggressive. While this should raise the profits on their passive activity, it should actually reduce their profits on aggressive trades unless their positioning profits are higher. This is evidence that HFT firms receive valuable private information during the POMO auctions because their profits per share rise despite the wider spreads.

Extrapolating the daily profit estimates to the broader market requires an estimate of the percentage of high frequency trading in the market captured by our sample. We present an estimate here in based on the 12.3% of total market capitalization represented by the firms in our sample. We sum daily profits across the 120 stocks, the 57 US Treasury POMO days, and we assume similar activity in the sample we do not observe. We estimate profits of over \$105 million during US Treasury POMO auctions.

7 Robustness Checks

7.1 Time Window

We analyze the HFT firm behavior in the half-hour after US Treasury POMO as a robustness check. In terms of inside quote frequency by the HFT firms, we use a similar model described in Sect. 5.1 but replace the variable US Treasury Purchase with a dummy variable indicating the half-hour after purchases. The results for the HFT inside bid and offer participations are reported in Table 11. The effect is not statistically significant for either bids or offers during the half-hour after US Treasury POMO auctions.

Table 11 HFT inside quote frequency after US treasury POMO

Variable	Bid	Offer
Trend	0.0013	0.0011
	(32.88)	(28.59)
$\text{Trend}^2/1000$	−0.0007	−0.0006
	(−8.26)	(−7.06)
Returns_{t-1}	0.0348	−0.0561
	(2.95)	(−4.83)
Realized vol_{t-1}	−0.0088	−0.0078
	(48.18)	(−44.28)
$\text{Volume}_{t-1}/1000$	0.0394	0.0323
	(116.47)	(109.12)
Order $\text{imbalance}_{t-1}/1000$	1.3817	−1.1281
	(10.45)	(−8.93)
Inside quote_{t-1}	1.6012	1.6150
	(465.57)	(473.96)
S11:00–11:30	0.0006	−0.0035
	(0.08)	(−0.52)
After UST purchases	−0.0235	0.0076
	(−1.67)	(0.54)
Constant	−1.0536	−1.0210
	(−220.77)	(−217.52)
$\overline{R}^2$	0.4211	0.4118

t-statistics in parentheses
The table reports estimates of a model for the inside quote participation of the 26 HFT firms in the half-hour after US Treasury purchases. We use a similar model presented in Table 5 but replace UST news variables with a dummy variable indicating the half-hour after purchases

The market impact of HFT trades in the half-hour after US Treasury POMO is not significantly different from the same period of non-POMO days either. The result is shown in the second set of columns in Table 8.

7.2 *FOMC Days*

We also contrast our results with the behavior of HFT firms on the eight Federal Open Market Committee (FOMC) dates in our sample listed in the third set of columns in Table 8.

We compare the market impact of HFT trades during the period from 13:45 to 14:15, the half-hour before the Fed announces its policy intentions. We felt this period was analogous to our half-hour before the release of POMO Treasury purchases. We used trades from the HFT database and quotes from the Nasdaq ITCH feed. This limits our analysis to the 60 Nasdaq stocks in the sample. For the 60 stocks, the market impact on FOMC and non-FOMC days is little changed. From this, we conclude that the POMO auctions were more important events for the market.

8 Conclusion

HFT firms perform a dual role as market makers. During the POMO auctions though, our model predicts that they may shift their focus from being liquidity providers to trading aggressively. We find that HFT firms reduce their presence at the inside quote and less frequently provide liquidity to non-HFT firms. Studying HFT activity in event windows like POMO may give us a better indication of how HFT firms will perform in stressful market conditions.

Acknowledgements We would like to thank Nasdaq OMX for providing the high frequency dataset, an anonymous referee, Michael Fleming and seminar participants at National Chiao Tung, Nankai, the Modeling High Frequency Data in Finance Conference, and the Second International Symposium in Computational Economics and Finance.

References

Abou, A., & Prat, G. (2000). Modelling stock price expectations: Lessons from microdata. In F. Gardes, & G.Prat (Eds.) *Price expectations in goods and financial markets* (pp. 313–46). Cheltenham: Edward Elgar.

Admati, A., & Pfleiderer, P. (1988). A theory of intraday patterns: Volume and price variability. *Review of Financial Studies, 1*, 3–40.

Baron, M., Brogaard, J., Hagstromer, B., & Kirilenko, A. (2017). Risk and return in high frequency trading. *Journal of Financial and Quantitative Analysis, forthcoming*. Available at http://ssrn.com/abstract=2433118

Benos, E., & Wetherilt, A. (2012). The role of designated market makers in the new trading landscape. *Bank of England Quarterly Review Q2, 52*, 343–353.

Brogaard, J., Hendershott, T., & Riordan, R. (2014). High frequency trading and price discovery. *Review of Financial Studies, 27*, 2267–2306.

Carrion, A. (2013). Very fast money: High-frequency trading on the NASDAQ. *Journal of Financial Markets, 16*, 680–711.

Cvitanic, J., & Kirilenko, A. (2010). High frequency traders and asset prices. California Institute of Technology working paper. Available at http://ssrn.com/abstract=1569067

D'Amico, S., & King, T. B. (2013). Flow and stock effects of large-scale treasury purchases: Evidence on the importance of local supply. *Journal of Financial Economics, 108*, 425–448.

Gagnon, J., Raskin, M., Remache, J., & Sack, B. (2011). The financial market effects of the Federal Reserve's large-scale asset purchases. *International Journal of Central Banking, 7*, 3–44.

Glosten, L. R., & Milgrom, P. R. (1985). Bid, ask, and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics, 14*, 71–100.

Hagströmer, B., & Norden, L. (2013). The diversity of high-frequency traders. *Journal of Financial Markets, 16*, 741–770.

Hasbrouck, J. (1991). Measuring the information content of stock trades. *Journal of Finance, 46*, 179–207.

Hasbrouck, J., & Saar, G. (2013). Low-latency trading. *Journal of Financial Markets, 16*, 646–679.

Hirschey, N. H. (2013). Do high-frequency traders anticipate buying and selling pressure? London Business School working paper. Available at: http://ssrn.com/abstract=2238516

Jawadi, F., & Prat, G. (2012). Arbitrage costs and nonlinear stock price adjustment in the G7 countries. *Applied Economics, 44*, 1561–1582.

Joyce, M., Lasaosa, A., Stevens, I., & Tong, M. (2011). The financial market impact of quantitative easing in the United Kingdom. *International Journal of Central Banking, 7*, 113–161.

Krishnamurthy, A., & Vissing-Jorgensen, A. (2011). The effects of quantitative easing on interest rates: Channels and implications for policy. *Brookings Papers on Economic Activity, 42*(Fall), 215–287.

Krishnamurthy, A., & Vissing-Jorgensen, A. (2012). The aggregate demand for treasury debt. *Journal of Political Economy, 120*, 233–267.

Lou, D., Yan, H., & Zhang, J. (2013). Anticipated and repeated shocks in liquid markets. *Review of financial studies, 26*, 1891–1912.

Martinez, V. H., & Roşu, I. (2013). High frequency traders, news and volatility. Working paper, HEC Paris.

Menkveld, A. J. (2013). High frequency trading and the new-market makers. *Journal of Financial Markets, 16*, 712–740.

Neely, C. (2010). The large-scale asset purchases had large international effects. Federal Reserve Bank of St. Louis working paper no. 2010-018A.

U.S. Commodity Futures Trading Commission and Securities and Exchange Commission. (2010). Findings regarding the market events of May 6, 2010, Washington, D.C.

Zhang, F. (2010). High-frequency trading, stock volatility, and price discovery. Yale University working paper. Available at http://ssrn.com/abstract=1691679

Part III
Transmission and Market Integration

Crude Oil and Biofuel Agricultural Commodity Prices

Semei Coronado, Omar Rojas, Rafael Romero-Meza, Apostolos Serletis, and Leslie Verteramo Chiu

Abstract Crop prices in the United States (USA), and especially corn prices, have been displaying important changes in the last 10 years, after the ethanol mandate in 2005. Motivated by these significant price changes, there has been a growing interest in the study of price transmission from oil prices to agricultural commodity prices. In this contribution, we concentrate on the relationship between the price of oil and the prices of three agricultural commodities that are used for biofuels production: corn, soybeans, and sugar. In doing so, we apply linear Granger causality tests, the nonlinear causality test of Diks and Panchenko (J Econ Dyn Control 30:1647–1669, 2006), and the Brooks and Hinich (J Empir Financ 6:385–404) cross-bicorrelation test to daily data over the period from 1990 to 2016.

Coherent with the previous studies, we find weak linear Granger causality, but strong bidirectional nonlinear causality, especially for the period from 2006 to 2016. Using the Brooks and Hinich test, we also identify the number of epochs (nonoverlapped windows) where there is nonlinear dependence between each pair of series. We find that most cross-bicorrelation windows coincide from 2006 to 2016, indicating that the nonlinear dynamics between the series studied have changed in recent years in the aftermath of the ethanol mandate. Our results provide hints in

S. Coronado
Departamento de Métodos Cuantitativos, Universidad de Guadalajara, Guadalajara, Mexico

O. Rojas
Escuela de Ciencias Económicas y Empresariales, Universidad Panamericana, Zapopan, Jalisco, Mexico

R. Romero-Meza
Facultad de Economia y Negocios, Universidad Alberto Hurtado, Santiago, Chile

A. Serletis (✉)
Department of Economics, University of Calgary, Calgary, AB, Canada
http://econ.ucalgary.ca/profiles/apostolos-serletis

L. Verteramo Chiu
Charles H. Dyson School of Applied Economics and Management, Cornell University, Ithaca, NY, USA

F. Jawadi (ed.), *Uncertainty, Expectations and Asset Price Dynamics*, Dynamic Modeling and Econometrics in Economics and Finance 24,
https://doi.org/10.1007/978-3-319-98714-9_5

order to improve our understanding of the effects of the implemented policies in the energy sector on agricultural commodities.

JEL Classification C32, G15, O13, Q13, Q43

1 Introduction

There have been significant changes in the dynamics between crude oil and biofuel agricultural commodity prices in recent years. This is particularly evident in the case of corn after the ethanol mandate of 2005. The effect of the increased demand for corn to meet the ethanol mandate spurred corn prices to record levels also increasing their volatility. This mandate created a direct link between ethanol production and corn demand, which in turn strengthened an indirect link between the crude oil and corn markets. This mandate requires a certain amount of biofuel (ethanol) to be mixed with gasoline sold in the US market. Bioethanol, accounting for more than 90% of biofuels, is produced by the fermentation and distillation of sugars or starch of biomass, usually from grains, cereals, sugar cane, and sugar beets. Starch and cellulose from crops need to be converted into sugars through the use of enzymes. This process is called saccharification. Once the sugars are obtained, they can be fermented, with the addition of yeast, and distilled. Bioethanol from sugar cane is a simpler process since it does not require saccharification. In the United States (USA), almost all bioethanol production comes from feedstocks, corn being the main crop. In order to guarantee supply of corn, and other feedstocks, for other markets, the USA sets a limit in the amount of bioethanol produced from starch-based feedstocks, like corn. Sugar cane is used at a lower scale for bioethanol production in the USA; however, it is the leading crop in bioethanol production in Brazil.

The relationship between oil and crop prices is critical to energy and agricultural policymakers and researchers. However, the current nature of this relationship is becoming more complex. Understanding it would allow policymakers to have the tools to make better policy decisions anticipating possible unintended consequences that would carry significant costs. Unanticipated changes in crop prices in the USA not only affect the corn producers and food consumers domestically but also the developing countries, particularly when agriculture plays a significant role in their economy. It is, therefore, important from the US domestic policy point of view to fully understand this dependency.

Agricultural commodity prices and their volatility affects farmers' welfare. While high prices are beneficial to producers, a concomitant increase in volatility makes risk management more difficult. Exposure to price volatility may even be detrimental to international agricultural trade—see Cho et al. (2002). The social costs and benefits of biofuel policies, which are aimed at reducing dependence on oil, among other reasons, are also analyzed by De Gorter and Just (2010); they find that the current ethanol policy can increase the inefficiencies of farm subsidies, and vice versa and also improve the international terms of trade in both corn exports and oil imports. However, as they argue, "the effects of each biofuel policy and their interaction with other policies (biofuel or otherwise) are very complex, the economics of which can seem impenetrable. This is due to the intricate interrelationships between energy and commodity markets and the varied environmental consequences." The crude oil and food price linkage due to biofuel production can also upset the relationship between food producers and consumers, with potential implications on food security—see Ford and Senauer (2007). Given the importance of having a better understanding of the joint movement between crude oil prices and biofuel agricultural commodity prices, and among themselves, this work analyzes the dependence structure of international crude oil prices and the US corn, soybeans, and sugar prices. We investigate whether their linkage changed after the ethanol mandate of 2005, using two influential techniques of nonlinear analysis, that to the best of our knowledge have not been applied together.

There are a large number of recent papers looking for modelling the behavior of oil prices and the market interactions between oil and agricultural commodity prices. Prat and Uctum (2011) attempt to model oil price expectations using survey data. Serra et al. (2011a) analyze the dynamics among corn, ethanol, gasoline, and oil prices from 2005 to 2007 in the USA. Using a smooth transition vector error-correction model, they find evidence of a linkage between corn and gasoline and oil markets created by the ethanol market. Looking at the volatility among these markets, including soybeans, based on weekly series from 1989 to 2007, Zhang et al. (2009), using cointegration, vector error corrections, and multivariate generalized autoregressive conditional heteroskedasticity models, find that ethanol and oil prices are affected by gasoline prices, while the ethanol price positively affects corn and soybeans prices. Looking at the volatility transmission in ethanol, oil, and sugar markets in Brazil, Serra et al. (2011b), using error-correction and multivariate generalized autoregressive conditional heteroscedasticity models, estimated in a single step by maximum likelihood, find a positive volatility transmission from sugar and oil prices to ethanol prices, but not a significant transfer in the opposite direction. They also find evidence of linkages between ethanol and sugar prices in Brazil, and that oil and sugar prices precede ethanol prices, but not the opposite.

There has also been an increased interest in analyzing the price dynamics between corn and other agricultural commodities, especially after the ethanol mandate. The ethanol mandate began when the first Renewable Fuels Standard (RFS) became law in 2005. It established the amount of biofuel to be blended for domestically sold gasoline and diesel. This energy policy aimed at using renewable biofuels as an environmentally friendly octane component in gasoline and diesel, and as a way to decrease dependence in foreign oil. The amounts of biofuels set by the RFS are calculated as a percentage of the yearly estimated nonrenewable gasoline and diesel supply. The resulting estimated biofuel amount, known as Renewable Volume Obligations (RVO), is then distributed among refiners and importers of gasoline and diesel. Transactions of biofuels used in gasoline and diesel blend are governed by Renewable Identification Numbers (RIN). RIN are tradable renewable fuel credits generated by biofuel producers and importers.

The target amount of biofuel used under the mandate is planned to increase yearly. The mandate established a target of 4 billion gallons of biofuel to be blended with gasoline and diesel for 2006. This target is set to 36 billion gallons by 2022; however, in 2007 the RFS established that only 15 billion gallons of biofuel should come from corn, and the rest from "advanced biofuels." Some sources of advanced biofuels include ethanol from cellulose and sugar, and biodiesel from soybean oil. It is estimated that 10% of motor gasoline sold in the USA is ethanol produced from corn (U.S. Energy Information Administration, 2013); this percentage was 3% in 2005. It is estimated that in 2011, 40% of the US corn was used to produce ethanol. Worldwide, this figure is 15% (Carter et al. 2012). De Gorter and Just (2010) find that the ethanol policy has significant impacts on corn prices, and is creating a linkage between the ethanol, grain, and oilseed markets. Specifically, De Gorter et al. (2015) suggest that the jump in agricultural commodity prices in 2006 was caused by the increased demand for crops for biofuel production, effectively modifying the market fundamentals for these commodities. In fact, according to Verteramo and Tomek (2016), corn prices seem to have entered a new level in 2006. Prior to that, the average farm price of corn in the USA from 1973 to 2005 was $2.36 per bushel, while the average price from 2006 to 2015 increased to $4.49 per bushel. The same authors show that after 2006 the corn demand curve has been shifting outwards, with the shift being traced back to the ethanol mandate.

Other analyses also consider the effects of oil prices on local agricultural commodity prices. Using the average nominal prices for the period, analyzing the relationship between local prices of corn in east Africa and international oil prices, Dillon and Barrett (2015), using error-correction models, find that local agricultural price variation depends on the market distance from the coast, since international oil prices affect transportation costs directly. However, they do not find a causal relationship through biofuel or production cost channels. They find that international oil prices pass on faster to cost of farm inputs and then to local corn prices than do global corn prices. Finally, they present evidence that for markets utmost inland, changes in international oil prices have larger effects on local corn prices than do changes in world corn prices. In a study of local prices, Nazlioglu and Soytas (2011) investigate the short- and long-run linkage between international oil prices, the lira–dollar exchange rate, and individual agricultural commodity prices (wheat, maize, cotton, soybeans, and sunflower) in Turkey. They use autoregressive vectors and impulse-response analysis to monthly data. They find no evidence of both direct and indirect effects of oil prices on agricultural commodity prices. Also, they find no indirect effects of oil prices through exchange rates. However, Cho et al. (2002), applying the gravity model, find evidence that for some commodities and in some countries, exchange rate fluctuations affect agricultural commodity trade. Mensi et al. (2014), applying the models MGARCH, VAR-BEKK-GARCH, and VAR-DCC-GRACH, find that there is a linkage between energy and cereal markets; in particular, they find that OPEC's announcements affect international cereal markets.

Dependence structures between markets have traditionally been analyzed using linear models. In recent years, however, nonlinear models are being used more frequently to examine market interactions. Looking at monthly prices to estimate a linear relationship among corn, rice, soybeans, sugar, wheat, and ethanol, gasoline, and oil from 1981 to 2007, Zhang et al. (2010), applying a model of error correction and Granger causality, find an effect from corn prices to ethanol prices. Nazlioglu (2011) uses both the Toda-Yamamoto linear Granger causality test and the Diks-Panchenko nonlinear Granger causality test to study the relationship between oil and agricultural commodity weekly prices. He finds no indication of causality under the linear analysis, but nonlinear causality is revealed through nonlinear tests. Thus, he concludes that on the one hand there are nonlinear feedbacks between the oil and the agricultural prices, and on the other hand that there is a persistent unidirectional nonlinear causality running from the oil prices to the corn and to the soybeans prices. Beckmann and Czudaj (2014), using a nonlinear smooth transition model to analyze futures prices of two nearby contracts, state that simplistic linear models are no longer reliable for agricultural price analysis. Using newly developed tests on causality in variance, Nazlioglu et al. (2013) analyze volatility transmission between oil prices and wheat, corn, soybeans, and sugar prices by applying a causality test in the variance. They find no evidence of volatility transmission between those markets before 2006; however, after that year there is evidence of volatility transmission, indicating that the relationship between oil and agricultural commodities is dynamic. The nonlinearities found in agricultural commodity prices may also be caused by volatility transmissions among commodities, see Beckmann and Czudaj (2014). Using data for the prices of nearby futures contracts for corn, cotton, and wheat and estimating GARCH-in-mean VAR models, they conclude that there is short-run volatility transmission process in agricultural futures markets. Alternative models to capture nonlinear dependency between agricultural and energy markets include structural VAR models, see Du and McPhail (2012); in their paper, they find that after 2008, ethanol, gasoline, and corn prices became more linked together due to the ethanol mandate. Moreover, variance decomposition indicates that ethanol can explain about 23% of corn price variation, and a big portion (27%) of ethanol price variations is determined by corn price oscillation. Using a nonlinear vector error correction model, Balcombe and Rapsomanikis (2008) analyze the long-run equilibrium among the sugar–ethanol–oil markets; they find the oil price to be a long-run component of sugar prices in Brazil. Moreover, in the price adjustment processes there exist nonlinearities of sugar and ethanol to oil. There is evidence of interdependence also in stock markets. For example, Jawadi and Prat (2012) studying arbitrage costs and nonlinear adjustment in the G7 stock markets found that a two-regime Smooth Transition Error Correction Model is appropriate to reproduce the dynamics of stock price deviations from fundamentals for these markets. The results from these studies warrant the use of nonlinear models to estimate risk transmission between different markets.

Nonlinear behavior in agricultural commodity prices has also been observed in older studies. One of the earliest accounts is the analysis of futures spreads and crop stocks by Working (1949), who observed a nonlinear relationship between futures prices and total crop stocks. In a more recent paper, Deaton and Laroque (1992) use a rational expectations competitive storage model to study the behavior of commodity prices, and find nonlinearities arising from the impossibility of the market to carry negative inventories. Other works applying nonlinear techniques on commodity prices include Deaton and Laroque (1995) and Mackey (1989), among others. Deaton and Laroque (1995) deal with the estimation of a model where a stochastic variable associated to an agricultural commodity generates a competitive price. Because storage cannot be negative, the relationship between prices and the stochastic variable is inherently nonlinear. Mackey (1989) develops a continuous time model for the price adjustment of a single commodity market, formulated as a delay differential equation. The paper considers the nonlinearities in supply and demand, and also production plus storage that may depend on the market price.

We apply the Diks and Panchenko (2006) nonparametric Granger causality test and complement it with the Brooks and Hinich (1999) nonlinear cross-bicorrelation test, to daily prices for crude oil, corn, soybeans, and sugar. These agricultural commodities are selected because corn represents the most important input in the production of biofuels in the international market. Soybeans are a close substitute of corn for both producers and feed buyers. Sugar is included to test if similar effects as those found in Brazil can be found for the USA. All these agricultural commodities have been studied before in relation to their effects with respect to biofuels and oil markets. Regarding our methodology, it is to be noted that the Diks and Panchenko (2006) test has been applied to similar data, but with monthly observations, by Nazlioglu (2011). However, it is the first time that the Brooks and Hinich (1999) test is applied to agricultural commodity prices, and the test provides new information about the underlying nonlinear dynamics. Previous applications of the Brooks and Hinich (1999) test include the study of exchange rate dynamics by Serletis et al. (2012) and economic activity by Romero-Meza et al. (2014).

The structure of this chapter is as follows. Section 2 describes the data, provides summary statistics, and examines the univariate time series properties of the variables. Section 3 presents the empirical results of linear causality, nonlinearity, nonlinear causality, and a cross-bicorrelation analysis. The final section briefly concludes the contribution.

2 Data

We consider daily prices for crude oil and three agricultural commodities—corn, soybeans, and sugar. The crude oil price used is the West Texas Intermediate (WTI) spot price, measured in dollars per barrel. We use futures prices for corn, soybeans, and sugar, obtained from *Bloomberg*. Corn and soybeans are traded at the Chicago Board of Trade (CBT) and sugar is traded at the Intercontinental Exchange (ICE). Futures contracts are quoted in cents per bushel for corn and soybeans, and in cents per pound for sugar. Each corn futures contract represents 5000 bushels

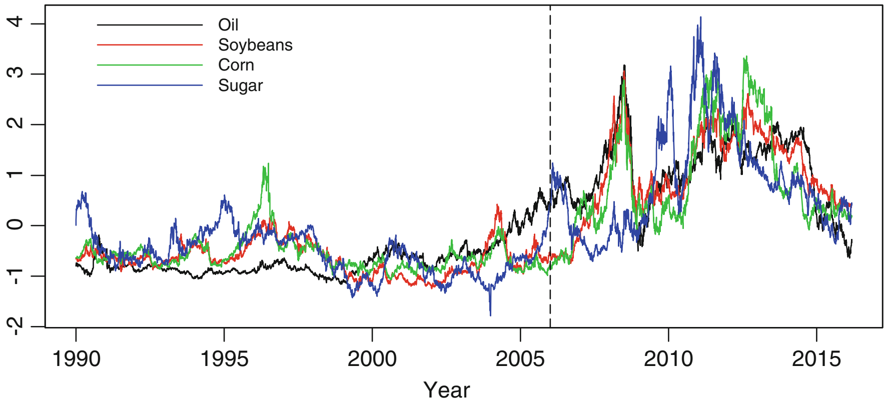


Fig. 1 Time series plots of standardized prices. The dashed vertical line is drawn at the split point between the two subsamples, December 31, 2005

(about 127 metric tons) with delivery months in March, May, July, September, and December. Similarly, each soybeans contract represents 5000 bushels (about 136 metric tons) with delivery months in January, March, May, July, August, September, and November. The size of the sugar contract is 112,000 pounds and the delivery months are March, May, July, and October. We use the price quotes for the nearest futures contracts. The sample period starts in January 1, 1990 and ends in March 10, 2016, for a total of 6834 daily observations; it is long enough to capture price interactions when the linkage between the markets of these agricultural commodities and the oil market was not as strong as in the period after the ethanol mandate.

Figure 1 presents the standardized prices of all four time series (standardized by subtracting the mean and dividing by the standard deviation), for comparison purposes. From 1990 to 2005, these prices seem to follow similar dynamics; however, from 2006 higher volatility is apparent and perhaps more co-movement in the series. To take into account the ethanol mandate and how this might have affected the price dynamics of the agricultural commodities under study, we partition the sample into two subperiods: a pre-mandate period that goes from January 1, 1990 to December 30, 2005 (subsample-1) with a total of $n_1 = 4175$ observations, and a post-mandate period from January 2, 2006 to March 10, 2016 (subsample-2) with a total of $n_2 = 2659$ observations.

The price series, p_t, are transformed into series of continuously compounded percentage returns, by taking (natural) logarithmic first differences, i.e., $r_t = 100(\ln p_t - \ln p_{t-1})$, and are shown in Fig. 2. In what follows, we work with the return series, since stationarity of the time series is required for the methods employed. Summary statistics for the return series are shown in Table 1. Both subsamples exhibit statistics in accordance with the common stylized facts of financial time series, see Cont (2001). All return series, for both subsamples, are platykurtic and non-normally distributed, as the Jarque and Bera (1987) test shows. They are also stationary, according to the Augmented Dickey–Fuller (ADF) unit

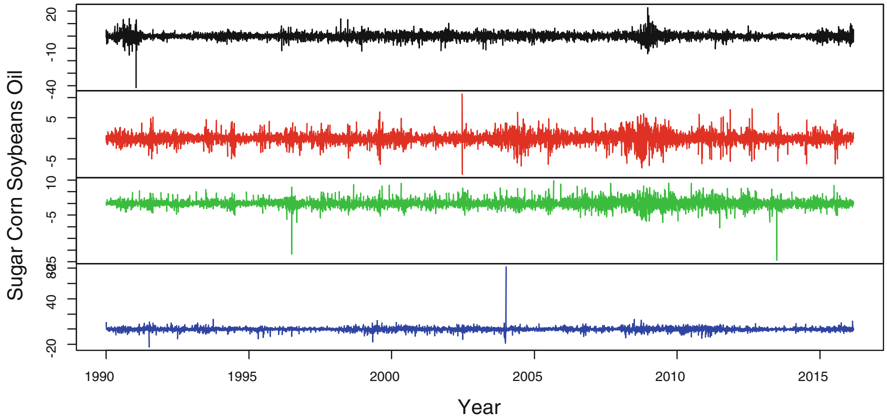


Fig. 2 Time series plots of returns

root test [see Dickey and Fuller (1981)] and the "residual augmented least squares" (RALS) unit root test by Im et al. (2014). The latter test has the advantage of being appropriate for data derived from non-normal distributions. All null hypotheses of normality and of the presence of a unit root are rejected at the 1% level (as an asterisk in Table 1 indicates).

Correlations between the return series are presented in Table 2, where Pearson correlation coefficients are reported. The upper diagonal elements in Table 2 correspond to subsample-1, and the lower diagonal elements to subsample-2. For subsample-1, all correlations are relatively small, except for the corn–soybeans correlation, which remains almost the same in subsample-2. However, all other correlations increase greatly from subsample-1 to subsample-2, which might be an indication of co-movement and perhaps causality after the ethanol mandate. The high correlation between corn and soybeans derives from their high degree of substitution on both the supply and demand sides. Producers decide on a combination of acreage of both crops that maximize expected returns while reducing production risk. These decisions are based on expected returns, where relative prices are a major factor in acreage decisions, and associated risks, see Chavas and Holt (1990). On the demand side, both crops are considered substitutes and are extensively used as feed crops. This degree of substitution is not observed among the other commodities in the study, hence the low correlations in subsample-1. The increase in correlation observed in subsample-2 corresponds to the increase in demand for crops derived from biofuels production after the ethanol mandate. This change in correlation is suggestive of a new market structure, and warrants further investigation of the price dynamics.

Table 1 Descriptive statistics for the return series

Statistic	Oil	Soybeans	Corn	Sugar
Panel A. subsample-1				
Mean	0.02	0.00	−0.00	0.00
Minimum	−41.55	−8.68	−21.65	−23.49
Maximum	14.24	10.80	9.80	81.62
Standard deviation	2.38	1.13	1.40	2.48
Skewness	−1.42	−0.04	−0.37	7.94
Kurtosis	58.79	9.17	20.30	289.12
Jarque–Bera	117073.38	6623.47	52158.54	14281930.92
ADF	−40.28*	−66.59*	−61.82*	−47.55*
RALS	−41.75*	−70.04*	−65.58*	−58.24*
Panel B. subsample-2				
Mean	−0.02	0.01	0.02	0.00
Minimum	−14.25	−7.12	−24.53	−12.37
Maximum	22.92	7.26	8.66	13.06
Standard deviation	2.37	1.37	2.00	2.22
Skewness	0.37	−0.31	−0.69	−0.11
Kurtosis	10.24	7.29	13.07	6.20
Jarque–Bera	5861.61	2081.34	11456.84	1137.78
ADF	−8.71*	−52.09*	−30.03*	−29.79*
RALS	−48.27*	−268.12*	−105.29*	−110.42*

Note: An asterisk indicates that the unit root null hypothesis is rejected at the 1% level

Table 2 Correlation matrix for returns

	Oil	Soybeans	Corn	Sugar
Oil	1.00	0.03	0.04	0.03
Soybeans	0.34	1.00	0.59	0.05
Corn	0.26	0.58	1.00	0.05
Sugar	0.21	0.25	0.24	1.00

Note: Upper triangular elements correspond to subsample-1 and lower triangular ones to subsample-2

3 Causality Dynamics

In this section, we present the results of linear Granger causality tests between the time series under study. We also perform a nonlinear test in order to determine if the dynamics of the series exhibit nonlinear behavior in which case nonlinear causality tests would be relevant.

3.1 Testing for Linear Granger Causality

According to Granger (1969), for a bivariate process $\{x_t, y_t\}$, it is said that $\{x_t\}$ causes $\{y_t\}$ if lagged values of $\{x_t\}$ provide additional information on future values of $\{y_t\}$ compared to the information already contained in $\{y_t\}$. The null hypothesis of non-causality, using an F-test, assumes linearity of the bivariate process, an assumption that might be quite restrictive and give way to misleading results. Results of the linear Granger causality tests, for lagged variables of order one, according to the Bayesian Information Criterion (BIC), are presented in Table 3. For subsample-1, causality is found from oil to corn at the 10% level and from corn to sugar at the 1% level. The causality found may reflect the effect of oil used as agricultural input in the form of gas and diesel, but most importantly as fertilizer. Fertilizer use for corn production is the second largest expense, after cash rent, with $148/acre. The causality from corn to sugar prices may be derived from the production of high fructose corn syrup, a widely used sweetener in the food industry. Since its price depends on the corn price, it may also affect the demand for sugar, a substitute. Approximately, 6% of corn production in the USA is used to produce high fructose corn syrup (see http://www.ers.usda.gov). For subsample-2, causality is found only from oil to soybeans and at the 10% level. However, there might be nonlinear causal relationships between the series, as we shall see below.

Table 3 F -statistics for linear Granger causality

Causality	Subsample-1	Subsample-2
Oil $\nrightarrow$ Soybeans	2.008 (0.157)	1.883 (0.094)
Soybeans $\nrightarrow$ Oil	0.153 (0.695)	1.176 (0.318)
Oil $\nrightarrow$ Corn	3.193 (0.074)	1.470 (0.196)
Corn $\nrightarrow$ Oil	0.016 (0.899)	0.990 (0.320)
Oil $\nrightarrow$ Sugar	0.881 (0.348)	0.888 (0.346)
Sugar $\nrightarrow$ Oil	1.850 (0.174)	0.433 (0.511)
Soybeans $\nrightarrow$ Corn	0.009 (0.923)	0.014 (0.906)
Corn $\nrightarrow$ Soybeans	1.086 (0.298)	0.673 (0.412)
Soybeans $\nrightarrow$ Sugar	0.567 (0.452)	1.459 (0.227)
Sugar $\nrightarrow$ Soybeans	0.122 (0.727)	1.020 (0.313)
Corn $\nrightarrow$ Sugar	9.674 (0.001)	0.000 (0.999)
Sugar $\nrightarrow$ Corn	0.000 (0.988)	0.951 (0.330)

Note: $x_t \nrightarrow y_t$ denotes the null hypothesis that x_t does not cause y_t. Numbers in parentheses are p-values. Low p-values reject the null

3.2 Testing for Nonlinearity

In this subsection, we apply the Brock et al. (1996) BDS test to test for nonlinear dependence in each of the time series. We choose the BDS test, because it has proved more reliable than other nonlinearity tests—see Patterson and Ashley (2000) and Zivot and Wang (2006). It is based on the concept of the correlation integral at embedding dimension m, that measures the frequency of temporal patterns occurring in the data. In order to filter out any possible linear dependence in the time series, an AR(p) is fitted to the data and the test is applied to the residuals. The lag p is chosen according to the Bayesian Information Criterion (BIC). Rejection of the null hypothesis implies a nonlinear dependence structure in the series. We perform the BDS test on the residuals of the return time series for dimension $m = 2$ and values of ϵ equal to $0.5\sigma_x$, σ_x, and $1.5\sigma_x$, where σ_x denotes the standard deviation of the return series r_t. The tests were performed for other embedding dimensions as well, and gave similarly conclusive results, rejecting the null hypothesis in all cases (see Table 4; the corresponding p-values are all less than 0.001), thus providing evidence of nonlinear behavior.

Given the evidence for nonlinearity in the return time series under study, in the next subsection we investigate causality relationships from a nonlinear point of view. As Granger (2014) put it, "univariate and multivariate nonlinear models represent the proper way to model a real world that is almost certainly nonlinear."

3.3 Testing for Nonlinear Causality

Given that linear Granger causality tests might fail to detect nonlinear causal relations, and the nonlinear nature of the time series under study, in this subsection we examine the dynamic relationship between the series using the nonlinear

Table 4 BDS test statistics for $m = 2$

Series	0.5σ	σ	1.5σ
Panel A. subsample-1			
Oil	6.071	7.101	8.558
Soybeans	4.427	6.066	7.421
Corn	7.065	8.142	8.497
Sugar	8.385	9.538	10.448
Panel B. subsample-2			
Oil	10.941	12.830	13.941
Soybeans	6.608	8.643	10.369
Corn	6.326	7.414	8.029
Sugar	4.906	4.550	4.947

Note: The null hypothesis is that of *iid* residuals causality test of Diks and Panchenko (2006)—see also Serletis and Istiak (2018) for a detailed description of the test. This test is an extension of the causality test by Baek and Brock (1992), which is based, as the BDS test, on the correlation integral. However, the Diks and Panchenko (2006) test does not rely on any assumptions of the time series being mutually and individually independent and identically distributed. It has also been shown to be more robust than the Hiemstra and Jones (1994) test, reducing over-rejection of the null hypothesis—see Bekiros and Diks (2008). The test statistic T_{n,ϵ_n}, with lag lengths of order one, i.e., $l_x = l_y = 1$ and bandwidth $\epsilon_n = Cn^{-\beta}$ for $C > 0$ and $1/4 < \beta < 1/3$ has a standard normal distribution as its limiting distribution.

We apply the Diks and Panchenko (2006) test after we first remove any linear dependence by fitting a VAR(p), where the lag length p is chosen according to the BIC and using the residuals as input for the test. The resulting optimal bandwidth for subsample-1 is $\epsilon = 0.68$ and for subsample-2 is $\epsilon = 0.90$. The results of the Diks and Panchenko (2006) test are presented in Table 5. In subsample-1, bidirectional causality is found between corn and soybeans and between oil and soybeans. For subsample-2, bidirectional causality is uncovered, at different significance levels, for all but the sugar–soybeans and oil–soybeans relationships. The nonlinearities between oil and agricultural markets, and among agricultural markets, especially in the latter period may be indicating the new relationship caused by the new biofuels policy. Nonlinear effects are harder to measure and predict than linear relationships. Consequently, hedging strategies may become more difficult to establish for agricultural producers and developing countries. The effects of oil price changes to agricultural commodity prices may become larger and more difficult to anticipate.

Table 5 Nonlinear causality tests

Causality	Subsample-1	Subsample-2
Oil $\nrightarrow$ Soybeans	2.149 (0.0158)	1.131 (0.1291)
Soybeans $\nrightarrow$ Oil	2.516 (0.0059)	2.689 (0.0036)
Oil $\nrightarrow$ Corn	0.730 (0.2327)	2.040 (0.0207)
Corn $\nrightarrow$ Oil	1.509 (0.0656)	1.540 (0.0618)
Oil $\nrightarrow$ Sugar	0.408 (0.6585)	3.656 (0.0001)
Sugar $\nrightarrow$ Oil	1.728 (0.0420)	3.012 (0.0013)
Soybeans $\nrightarrow$ Corn	3.407 (0.0003)	3.383 (0.0004)
Corn $\nrightarrow$ Soybeans	3.632 (0.0001)	1.735 (0.0414)
Soybeans $\nrightarrow$ Sugar	0.909 (0.1817)	1.416 (0.0784)
Sugar $\nrightarrow$ Soybeans	0.570 (0.2845)	0.623 (0.2665)
Corn $\nrightarrow$ Sugar	0.293 (0.3848)	1.888 (0.0295)
Sugar $\nrightarrow$ Corn	1.706 (0.4400)	2.925 (0.0017)

Note: $x_t \nrightarrow y_t$ denotes the null hypothesis that x_t does not cause y_t
Numbers in parentheses are p-values. Low p-values reject the null

3.4 The Cross-Bicorrelation Test

Motivated by our results based on the Diks and Panchenko (2006) nonlinear causality tests, we are now interested in measuring the degree of nonlinear correlation between the time series under study. In order to do so, we apply the nonlinear cross-bicorrelation test of Brooks and Hinich (1999), which is a multivariate extension of the Hinich (1996) test and does not require knowledge of the type of dynamics, stochastic or chaotic, of the series. The Brooks and Hinich (1999) test is based on the third-order statistics between two variables and the cross-bicorrelation indicates a correlation between the current value of a variable and the value of previous cross-correlations between the two variables. The test statistic H_{xxy} is asymptotically distributed as a χ^2 with $L(2L-1)$ degrees of freedom, where $L = N^c$ for $0 < c < 1/2$, N being the sample size. Rejection of the null hypothesis of the two time series being independent pure white noise processes implies that the cross-bicovariances $C^{rs}_{xxy} = E(x_t x_{t+r} y_{t+s})$ are nonzero, and therefore causality in the Granger sense between the series is inferred. As with the Diks and Panchenko (2006) test, the Brooks and Hinich (1999) test is applied to the residuals of a VAR(p) to remove linear dependence. The series are partitioned into equal length nonoverlapping moving frames and the test indicates if there is causality in each of such windows or frames. For our sample data, we use the optimal window length of size 32 and the results are significant at the 5% level.

Cross-bicorrelations between the studied subsamples are presented in Table 6. The upper diagonal elements correspond to subsample-1 and the lower diagonal elements to subsample-2 (in the same fashion as in Table 2). Cross-bicorrelations are higher than the correlations presented in Table 2, which confirms our suspicion of higher nonlinear dependence between the variables. All bidirectional relationships increase from the first subsample to the second one. The cross-bicorrelation between all agricultural commodity prices increases during the latter period. This result may be due to the new market integration between crops and energy, and to the substitutability among crops in the production of biofuels.

Correlograms of the correlation (Table 2) and cross-bicorrelation (Table 6) matrices are shown in Figs. 3 and 4, respectively. As can be seen, the cross-bicorrelations indicate more correlation between the variables under study than the simple Pearson correlations. Figure 5 shows the relative frequencies of the

Table 6 Cross-bicorrelation matrix for returns

	Oil	Soybeans	Corn	Sugar
Oil	1.00	0.40	0.47	0.63
Soybeans	0.64	1.00	0.80	0.46
Corn	0.77	0.73	1.00	0.55
Sugar	0.65	0.60	0.64	1.00

Note: Upper triangular elements correspond to subsample-1 and lower triangular elements to subsample-2

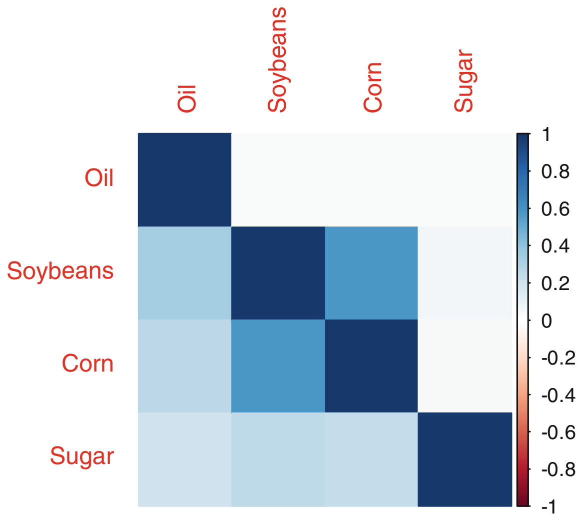


Fig. 3 Correlogram of the correlation matrix (Table 2). Upper triangular elements correspond to subsample-1 and lower triangular ones to subsample-2

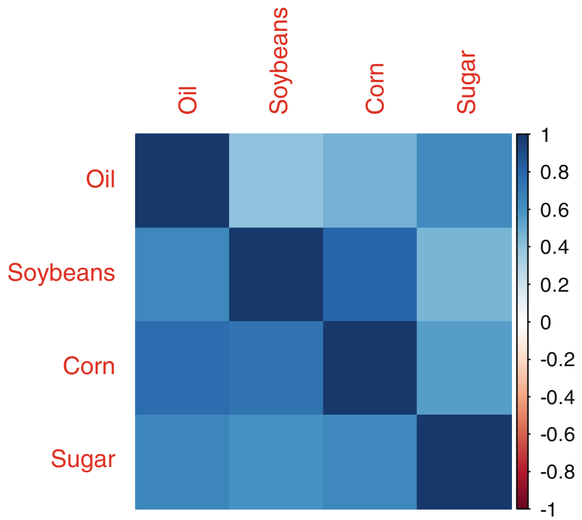


Fig. 4 Correlogram of the cross-bicorrelation matrix (Table 6). Upper triangular elements correspond to subsample-1 and lower triangular ones to subsample-2

significant windows of Brooks and Hinich (1999) causality by year. It is clear that after 2005 the number of windows that exhibit cross-bicorrelations increased, indicating a change in the nonlinear dynamics after the ethanol mandate. However, there are some years after 2005, that the relative frequency drops, potentially

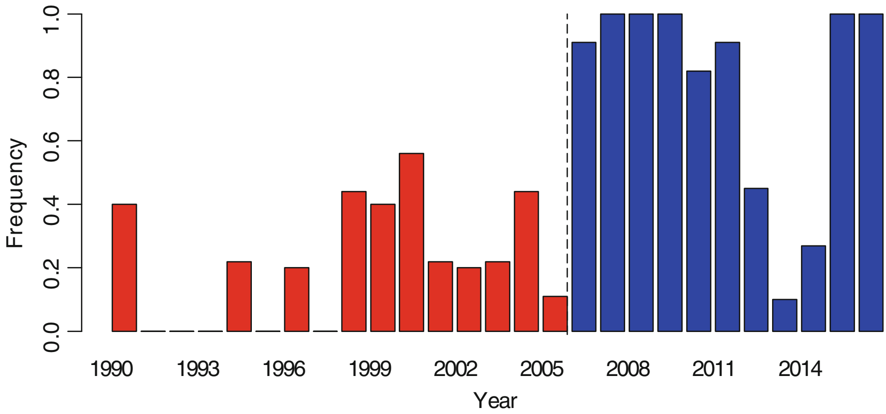


Fig. 5 Relative frequency of significant windows of Brooks and Hinich causality by year for causality from oil to soybeans and corn to sugar, for subsample-1 (left) and subsample-2 (right)

indicating asymmetric spillover effects between the oil price and the agricultural commodity prices.

4 Conclusion

This contribution applies two influential techniques of nonlinear analysis—the Diks and Panchenko (2006) test and the Brooks and Hinich (1999) test—to daily data of oil prices and three agricultural commodity prices, corn, soybeans, and sugar, over the period from 1990 to 2016, in order to understand the dynamics of price and volatility transmission from the oil market to the biofuels agricultural commodity markets. We find evidence of nonlinear dependence, and also certain periods of higher nonlinear interactions between these markets, especially after 2005. We conclude that the ethanol mandate might have changed the dynamics of dependence between the oil market and the biofuel agricultural commodity markets. Our results, in addition to providing a deeper understanding of the dynamic interactions between these markets, can also help policymakers in better assessing the welfare impacts and unintended consequences of energy policies and food prices. This is particularly relevant today as oil and corn prices have dropped to their lowest levels in 5 years.

For example, understanding the current nonlinear dynamic linkages among these markets can help market participants and governments in the design and implementation of better risk management strategies. Models not accounting for this nonlinearity may create biased projections of agricultural commodity prices. Moreover, if due to the ethanol mandate of 2005, agricultural commodity prices are more sensitive to energy prices, then food subsidy programs should be considered in the short run and investments in the agricultural sector should be increased in the long run, in order to mitigate the effects of rising and volatile commodity prices that mostly affect the poor.

Acknowledgements Rafael Romero-Meza and Semei Coronado are grateful for the support of FONDECYT (Project 1111034) and Universidad de Guadalajara for funding this research.

References

Baek, E., & Brock, W. (1992). A general test for nonlinear Granger causality: Bivariate model. Iowa State University and University of Wisconsin at Madison, Working Paper.

Balcombe, K., & Rapsomanikis, G. (2008). Bayesian estimation and selection of nonlinear vector error correction models: The case of the sugar-ethanol-oil nexus in Brazil. *American Journal of Agricultural Economics, 90*, 658–668.

Beckmann, J., & Czudaj, R. (2014). Volatility transmission in agricultural futures markets. *Economic Modelling, 36*, 541–546.

Bekiros, S. D., & Diks, C. G. H. (2008). The relationship between crude oil spot and futures prices: Cointegration, linear and nonlinear causality. *Energy Economics, 30*, 2673–2685.

Brock, W. A., Scheinkman, J. A., Dechert, W. D., & LeBaron, B. (1996). A test for independence based on the correlation dimension. *Econometric Reviews, 15*, 197–235.

Brooks, C., & Hinich, M. J. (1999). Cross-correlations and cross-bicorrelations in Sterling exchange rates. *Journal of Empirical Finance, 6*, 385–404.

Carter, C., Rausser, G., & Smith, A. (2012). *The effect of the US ethanol mandate on corn prices*. Unpublished manuscript. http://www.ourenergypolicy.org/wp-content/uploads/2013/07/The-Effect-oftheUS-Ethanol-Mandate-on-Corn-Prices-.pdf

Chavas, J.-P., & Holt, M. T. (1990). Acreage decisions under risk: The case of corn and soybeans. *American Journal of Agricultural Economics,72*, 529–538.

Cho, G., Sheldon, I. M., & McCorriston, S. (2002). Exchange rate uncertainty and agricultural trade. *American Journal of Agricultural Economics, 84*, 931–942.

Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance, 1*, 223–236.

De Gorter, H., Drabik, D., & Just, D. R. (2015). *The economics of biofuel policies: Impacts on price volatility in grain and oilseed markets*. New York: Springer.

De Gorter, H., & Just, D. R. (2010). The social costs and benefit of biofuels: The intersection of environmental, energy and agricultural policy. *Applied Economic Perspectives and Policy, 32*, 4–32.

Deaton, A., & Laroque, G. (1992). On the behaviour of commodity prices. *Review of Economic Studies, 59*, 1–23.

Deaton, A., & Laroque. G. (1995). Estimating a nonlinear rational expectations commodity price model with unobservable state variables. *Journal of Applied Econometrics, 10*, S9–S40.

Dickey, D. A., & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica, 49*, 1057–1072.

Diks, C., & Panchenko, V. (2006). A new statistic and practical guidelines for nonparametric Granger causality testing. *Journal of Economic Dynamics and Control, 30*, 1647–1669.

Dillon, B. M., & Barrett, C. B. (2015). Global oil prices and local food prices: Evidence from east Africa. *American Journal of Agricultural Economics, 98*, 154–171.

Du, X., & McPhail, L. L. (2012). Inside the black box: The price linkage and transmission between energy and agricultural markets. *The Energy Journal, 33*, 171–194.

Ford, C., & Senauer, B. (2007). How biofuels could starve the poor. *Foreign Affairs, 86*, 41–53.

Granger, C. (1969). Investigating causal relations by econometric models and cross spectral methods. *Econometrica, 37*, 424–438.

Granger, C. (2014). *Forecasting in business and economics*. London: Academic Press.
Hiemstra, C., & Jones, J. D. (1994). Testing for linear and nonlinear Granger causality in the stock price-volume relation. *Journal of Finance, 49*, 1639–1664.
Hinich, M. J. (1996). Testing for dependence in the input to a linear time series model. *Journal of Nonparametric Statistics, 6*, 205–221
Im, K. S., Lee, J., & Tieslau, M. A. (2014). More powerful unit root tests with non-normal errors. In *Festschrift honor Peter Schmidt* (pp. 315–342). New York: Springer.
Jarque, C., & Bera, A. (1987). A test for normality of observations and regression residuals. *International Statistical Review, 55*, 163–172.
Jawadi, F., & Prat, G. (2012). Arbitrage costs and nonlinear adjustment in the G7 stock markets. *Applied Economics, 44*, 1561–1582.
Mackey, M. C. (1989). Commodity price fluctuations: Price dependent delays and nonlinearities as explanatory factors. *Journal of Economic Theory, 48*, 497–509.
Mensi, W., Hammoudeh, S., Nguyen, D. K., & Yoon, S.-M. (2014). Dynamic spillovers among major energy and cereal commodity prices. *Energy Economics, 43*, 225–243.
Nazlioglu, S. (2011). World oil and agricultural commodity prices: Evidence from nonlinear causality. *Energy Policy, 39*, 2935–2943.
Nazlioglu, S., Erdem, C., & Soytas, U. (2013). Volatility spillover between oil and agricultural commodity markets. *Energy Economics, 36*, 658–665.
Nazlioglu, S., & Soytas, U. (2011). World oil prices and agricultural commodity prices: Evidence from an emerging market. *Energy Economics, 33*, 488–496.
Patterson, D. M., & Ashley, R. A. (2000). *A nonlinear time series workshop: A toolkit for detecting and identifying nonlinear serial dependence*, vol. 2. New York: Springer.
Prat, G., & Uctum, R. (2011). Modelling oil price expectations: Evidence from survey data. *The Quarterly Review of Economics and Finance, 51*, 236–247.
Romero-Meza, R., Coronado, S., & Serletis, A. (2014). Oil and the economy: A cross bicorrelation perspective. *Journal of Economic Asymmetries, 11*, 91–95.
Serletis, A., & Istiak, K. (2018). Broker-dealer leverage and the stock market. *Open Economies Review, 29*(2), 215–222.
Serletis, A., Malliaris, A. G., Hinich, M. J., & Gogas, P. (2012). Episodic nonlinearity in leading global currencies. *Open Economies Review, 23*, 337–357.
Serra, T., Zilberman, D., & Gil, J. (2011a). Price volatility in ethanol markets. *European Review of Agricultural Economics, 38*, 259–280.
Serra, T., Zilberman, D., Gil, J. M., & Goodwin, B. K. (2011b). Nonlinearities in the U.S. corn-ethanol-oil price system. *Agricultural Economics, 38*, 259–280.
U.S. Energy Information Administration. (2013). Accessed Dec 10, 2016 from http://www.eia.gov/todayinenergy/detail.php?id=11511
Verteramo, L. J., & Tomek, W. G. (2016). Anticipatory signals of structural changes in U.S. corn demand. In *Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management, St. Louis, MO*.
Working, H. (1949). The theory of price of storage. *American Economic Review, 39*, 1254–1262.
Zhang, Z., Lohr, L., Escalante, C., & Wetzstein, M. (2009). Ethanol, corn, and soybean price relations in a volatile vehicle-fuels market. *Energies, 2*, 320–339.
Zhang, Z., Lohr, L., Escalante, C., & Wetzstein, M. (2010). Food versus fuel: What do prices tell us? *Energy Policy, 38*, 445–451.
Zivot, E., & Wang, J. (2006). *Modeling financial time series with S-PLUS*. New York: Springer.

Financial Integration and Business Cycle Synchronization in Sub-Saharan Africa

Julien Acalin, Bruno Cabrillac, Gilles Dufrénot, Luc Jacolin, and Samuel Diop

Abstract This contribution studies the relationship between financial integration and the correlation of business cycles in sub-Saharan African countries. We consider asymmetric dynamics during expansions and recessions and desynchronized fluctuations that capture the costs and benefits of financial integration. Our study suggests that the effect of financial integration is heterogeneous across groups of countries. In the CEMAC, we find a positive relationship between financial integration and business cycles, while for WAEMU and SADC, financial integration increases the dephasing of business cycles. Further, reserve pooling does not play a substantial role in smoothing idiosyncratic shocks.

1 Introduction

This chapter is written in the honor of Georges Prat. The topic is in the field of financial analysis to which he devoted a significant part of his scientific activity. We discuss issues related to the link between financial integration and business cycles as an illustration of the way in which markets relate prices and economic fundamentals. This issue is central in Georges' works. Market integration can be captured in several ways but is usually referred to as an illustration of the law of one price. The convergence of prices in financial markets is conditioned by numerous factors, among which market expectations, economic fundamentals, the horizon of pricing, and arbitrage costs, but also by animal spirits as originally suggested by Keynes or Minsky (see for instance, Abou and Prat 2000, 2010; Jawadi and Prat 2012, 2017; Prat and Uctum 2007, 2011, 2013, 2015; Uctum et al. 2017).

J. Acalin · B. Cabrillac · L. Jacolin · S. Diop
Banque de France, Paris, France

G. Dufrénot (✉)
Aix-Marseille School of Economics, Marseille, France

CEPII, Paris, France

F. Jawadi (ed.), *Uncertainty, Expectations and Asset Price Dynamics*,
Dynamic Modeling and Econometrics in Economics and Finance 24,
https://doi.org/10.1007/978-3-319-98714-9_6

This chapter discusses the links between financial integration and economic cycles in sub-Saharan African countries. We investigate whether a higher financial integration has contributed to synchronizing business cycles over time. This issue is of interest for both academics and policymakers, since many countries want to reinforce their monetary, financial, and economic integration through the creation of new monetary and economic unions (see Table 5 in Appendix).

A usual necessary condition for currency unions to be a success is that the shocks affecting the economies are symmetric. This facilitates the efficiency of the common monetary policy in absorbing the effects of negative shocks on the real sector. According to the theories of optimum currency areas, this can be achieved through several channels: (1) trade deepening (by reducing transportation costs and eliminating regulatory barriers to trade), (2) increased mobility of factors of production, as well as price and wage flexibility, (3) a coordination of fiscal and macroeconomic policies, or (4) financial integration. Our contribution focuses on the role of financial integration as a stimulating factor to enhance the synchronization of business cycles in sub-Saharan Africa. We consider two types of financial integration indicators: (1) bank-based indicators and (2) risk-sharing indicators.

Our motivations for choosing a bank-sector-based indicator are threefold. Firstly, the financial sector in Africa is dominated by banks, and the empirical evidence suggests that the development of the banking sector promotes financial development in sub-Saharan Africa.[1] Secondly, the costs of interbank cross-border payments have decreased over time. For instance, the WAEMU has adopted in June 2004 a real-time gross settlement system (STAR-UEMOA) which allows interbank transfers, an interbank clearing system (SICA-UEMOA), and an interbank payment system for bank cards (GIM-UEMOA). The WAMZ is currently finalizing a project of upgrading the existing RTGS system between Ghana and Nigeria to other three member countries (the Gambia, Guinea, and Sierra-Leone) to facilitate the use of banking services across countries. The medium-term outcome is a deeper integration of the financial system. In Eastern Africa, South Africa and the EAC (East African Community) have harmonized their accounting standards to facilitate reporting principles when investing cross border.

A third motivation for adopting bank-based indicators is the rapid expansion of pan-African bank groups and, in parallel, the volume of cross-border transactions between African banks.[2] As an illustration, four African banks account for one-third percent of total deposits in at least 13 countries (Bank of Africa, Ecobank, Stanbic, and United Bank of Africa). The implication is narrowed interest rate spreads in the credit markets. In this work, we propose indicators of the convergence of commercial banks' intermediation margins (net intermediation margin) and financial performance (ROE, ROA), based on time-varying sigma convergence.

The second type of indicator relevant for our study is derived from risk-sharing models of consumption smoothing widely used in the literature. A major

---

[1] See Kablan (2010).

[2] See Monfort et al. (2013).

difference with the traditional literature is our assumption that the contributions of the different factors in smoothing shocks to GDP are time-varying. Indeed, depending upon the changing linkages of the balance of payments across time, the smoothing can operate through different factors of unequal importance (factor income, capital depreciation, saving, etc.). Moreover, shock smoothing is not a continuous phenomenon and can alternate with periods of substantial volatility.

Our aim is to see whether these variables have fostered business cycle synchronization. Our analysis is based on a panel of 31 sub-Saharan African countries over the period from 2000 to 2012 (see Table 1).

We find that financial integration does not necessarily imply co-movements in output phases (be they recessions or expansions). Another finding is the heterogeneity in the financial integration-output correlation link. SADC and WAEMU countries behave in a different manner than CAEMC and a subgroup of countries including ECOWAS and EAC countries. Our finding of GDP outphasing in WAEMU and SADC suggests that some other factors reduce the beneficial effect of risk sharing, especially during expansion. We finally discuss the relevance of examining the effects of higher financial integration on the dephasing of the cycles. Indeed, this provides some information about whether transferring resources to or receiving resources from the neighbors implies a cost or a gain for individual countries.

The remainder of the chapter is organized as follows. Section 2 presents the variables and the empirical model. Section 3 contains our results. Finally, Sect. 4 concludes.

2 The Variables and the Empirical Model

2.1 Defining an Indicator of Business Cycle Synchronization

The measurement of business cycles for African economies raises several problems. First, unlike the industrialized countries, it is difficult to find some reliable advanced indicators of the economic activity. Secondly, though industrialization has been on the rise over the last decade, many economies owe their economic growth to that of the primary and services sectors. Agricultural output is quite volatile and services are, for a significant part, informal. Therefore, extracting cycles from series capturing the activity in both the primary and tertiary sectors could lead to measurement errors. Thirdly, some series capturing the economic activity contain lots of breaks caused by exogenous “shocks” like social and political crises.

With regard to these difficulties, we construct indicators of business cycles in sub-Saharan Africa using statistical filters applied to national account data. Some widely used techniques include calculating the residuals of GDP series from different filters (HP, Baxter-King, unobserved components), using innovations from VAR models, considering lagged differences of GDP series, or adopting frequency

Table 1 Countries in our sample

South Africa (S, SADC)	Tanzania (S, others)	Togo (W, WAEMU)	Sierra Leone (W, others)
Angola (S, SADC)	Uganda (S, others)	Cameroun (W, CAEMC)	
Botswana (S, SADC)	Zambia (S, SADC)	Central Afr. Republic (W, CAEMC)	
Kenya (S, others)	Zimbabwe (S, SADC)	Chad (W, CAEMC)	
Lesotho (S, SADC)	Benin (W, WAEMU)	Congo (W, CAEMC)	
Malawi (S, SADC)	Burkina Faso (W, WAEMU)	Equatorial Guinea (W, CAEMC)	
Mauritius (S,SADC)	Côte d'Ivoire (W, WAEMU)	Gabon (W, CAEMC)	
Mozambique (S,SADC)	Mali (W, WAEMU)	The Gambia (W, others)	
Namibia (S, SADC)	Niger (W, WAEMU)	Ghana (W, others)	
Swaziland (S, SADC)	Senegal (W, WAEMU)	Nigeria (W, others)	

Note: *S*, *W*, Southern and Western; *SADC*, Southern African Development Community; *WAEMU*, West African Economic and Monetary Union; *CAEMC*, Central African Economic and Monetary Community; others, Uganda, Tanzania, Kenya, the Gambia, Nigeria, Ghana, Sierra Leone
Due to data unavailability and sample size, no EAC group was created

domain approaches.[3] A limitation in the use of most of these techniques is that they are based on parametric models, the assumptions of which may affect GDP estimates significantly.

We therefore choose a method of dating the major peaks and troughs of domestic business cycles using the Bry and Boschan algorithm, which has the advantage of being nonparametric and thereby not requiring any assumption on the underlying data generation process of the GDP series. Peaks and troughs are defined as local extrema. Using the following pair of binary variables($\bigwedge_t, \bigvee_t$), where $\bigwedge_t = 1$ indicates a peak and $\bigvee_t = 1$ indicates a trough, we retain the following definitions:

$$\begin{aligned}\bigwedge_t = \mathbf{1}\{&(\mathrm{GDP}_t - \mathrm{GDP}_{t-2}) > 0, (\mathrm{GDP}_t - \mathrm{GDP}_{t-1}) > 0,\\ &(\mathrm{GDP}_{t+2} - \mathrm{GDP}_t) < 0, (\mathrm{GDP}_{t+1} - \mathrm{GDP}_t) < 0\}\end{aligned} \tag{1a}$$

$$\begin{aligned}\bigvee_t = \mathbf{1}\{&(\mathrm{GDP}_t - \mathrm{GDP}_{t-2}) < 0, (\mathrm{GDP}_t - \mathrm{GDP}_{t-1}) < 0,\\ &(\mathrm{GDP}_{t+2} - \mathrm{GDP}_t) > 0, (\mathrm{GDP}_{t+1} - \mathrm{GDP}_t) > 0\}\end{aligned} \tag{1b}$$

where $\mathbf{1}(x)$ equals 1 if x is true and zero otherwise. This allows identifying the periods between a peak and a trough with a recession (or contraction) and periods between troughs and peaks as expansion phases.

The next question is how to measure business cycle synchronization. In the African context, the authors who have investigated this issue primarily rely on dynamic factor models.[4] However, as noted by Kemegue and Seck (2014), these approaches suffer the criticism of being benchmarked by backward-looking aggregate factors. To avoid this problem, a nonparametric approach can be retained that also captures changing dynamics in the co-movements of the economic activities. We define an index of concordance of the domestic business cycle phases based on Harding and Pagan (2002). However, instead of considering only the fraction of time the countries are in the same cycle phase, we distinguish between situations in which the co-movements in GDPs occur during expansions and recessions. Moreover, we calculate indices of business cycle desynchronization. In the following formula, a domestic country is represented by the index i, while j refers to the other countries in the group. For the latter, we adopt the following convention. If, in a given subgroup of countries, a majority of member countries is experiencing a recession (resp. an expansion), then the group is considered as being in recession (resp. in expansion). For each time t, we thus define the following index:

$$C_t^{ij} = S_t^i S_t^j + \left(1 - S_t^i\right)\left(1 - S_t^j\right) \tag{2}$$

[3]For recent applications to sub-Saharan African countries, see Coleman (2011), Hitaj et al. (2013), Mafusire and Brixiova (2013), and Ssozi (2011).

[4]See Carmignani (2009) and Houssa (2008).

Table 2 Number of coincidences, synchronization, and idiosyncratic shocks

Domestic country	Expansion	Recession	Expansion	Recession
SADC				
# coincidences	73	46	24	39
Fraction of total observations	0.40	0.25	0.13	0.21
WAEMU				
# coincidences	41	19	9	22
Fraction of total observations	0.39	0.18	0.09	0.21
CAEMC				
# coincidence	28	26	10	14
Fraction of total observations	0.36	0.33	0.13	0.18
Others				
# coincidence	14	27	49	10
Fraction of total observations	0.15	0.30	0.44	0.11

To capture business cycle synchronization, S_t^k is defined, respectively, as $\wedge_t^k$ and $\vee_t^k$ $(k = i, j)$. For business cycle desynchronization, for a couple of variables $\left(S_t^i, S_t^j\right)$, we define $S_t^i = \wedge_t^i$ (or $\vee_t^i$) and $S_t^j = 1 - \wedge_t^j$ (or $1 - \vee_t^j$).

Table 2 displays the number of coincidences found over the period 2000–2012 based on Eq. (2). These numbers are expressed in absolute value and as share of total observations. In the first two columns, the latter can be thought of as the degree of synchronization between the expansion and recession phases. In the other two columns, it is relevant to see it as an indicator of desynchronization.

Inspection of the values in the table indicates a co-movement of the domestic countries' fluctuations vis-à-vis the others' during expansions and recessions. In SADC and WAEMU, the countries are more similar in times of expansions than during recessions. In CAEMC they seem to be synchronized with equal strength whatever the business cycle phases, reflecting the large weight of oil exports in their GDP. It is also noteworthy that the business cycle phases are less disparate (desynchronized) when a subregion evolves in an expansion phase (compared with a phase of recession). Indeed, we see that the fractions of coincidences in the fifth column are generally higher than in the fourth column. The asynchronous connections between the business cycle phases thus suggest that the positive shocks hitting a group of countries have been more idiosyncratic than the negative shocks occurring in bad times (during recessions).

2.2 Variables of Financial Integration

2.2.1 Risk-Sharing Variables

It has become common wisdom in the literature to examine financial integration by investigating cross-country risk sharing within a group of countries. Such a methodology, originally proposed for the industrialized countries,[5] has been applied to sub-Saharan African countries. The idea is the following. Idiosyncratic shock smoothing can be achieved by risk sharing among countries through several channels of international income insurance: a repatriation of income earned on capital assets owned abroad or a reduction of capital held abroad, remittances from residents living abroad (both account for net factor income from abroad), international transfers, and interregional credits.

Formally, the risk-sharing assumption implies a decorrelation between consumption and GDP, which amounts to estimating the following equation where $(1 - \rho^{\mathrm{NP}})$ captures the degree of risk sharing:

$$\Delta \log C_{it} = \delta^{\mathrm{NP}} + \rho^{\mathrm{NP}} \Delta \mathrm{logGDP}_{it} + \varepsilon_{it}, \quad 0 \leq \rho^{\mathrm{NP}} \leq 1 \tag{3}$$

where C_{it} is aggregate consumption in country i at time t and ε_{it} is an error term. Using a national account approach, total risk-sharing can be decomposed in four components representing, respectively, the insurance of income through net factor income from abroad, net fiscal transfers, interregional credits, and capital appreciation/depreciation. Let us consider the following national accounts identities: GNI = GDP + net income from abroad, national income (NI) = GNI – capital depreciation, disposable national income (DNI) = NI + international transfers, and consumption (C) = DNI – net saving. Based on these identities, the different components of risk sharing are measured by estimating the systems of equations that consist of Eq. (3) and the following equations:

$$\Delta \log \mathrm{GDP}_{it} - \Delta \log \mathrm{GNI}_{it} = \delta^{\mathrm{FR}} + \rho^{\mathrm{FR}} \Delta \mathrm{logGDP}_{it} + \varepsilon_{it}^{\mathrm{FR}}, \tag{4}$$

$$\Delta \log \mathrm{GNI}_{it} - \Delta \log \mathrm{NI}_{it} = \delta^{D} + \rho^{D} \Delta \mathrm{logGDP}_{it} + \varepsilon_{it}^{D}, \tag{5}$$

$$\Delta \log \mathrm{NIP}_{it} - \Delta \log \mathrm{DNI}_{it} = \delta^{T} + \rho^{T} \Delta \mathrm{logGDP}_{it} + \varepsilon_{it}^{T}, \tag{6}$$

$$\Delta \log \mathrm{DNI}_{it} - \Delta \log C_{it} = \delta^{S} + \rho^{S} \Delta \mathrm{logGDP}_{it} + \varepsilon_{it}^{S}, \tag{7}$$

The upper indices on the coefficients and the residual terms refer to factor revenue (FR), capital depreciation (D), international transfers (T), and saving (S). The coefficients ρ^{FR}, ρ^{D}, ρ^{T}, and ρ^{S} measure the fraction of asymmetric shocks that are smoothed by factor revenues, capital depreciation, international transfers, and saving. ρ^{NP} is a measure of the proportion of shocks unsmoothed. The five coefficients sum to one. They are not necessarily positive, since the idiosyncratic shocks on the GDP can amplify the volatility of consumption (de-smoothing).

[5]See Asdrubli et al. (1996), Brennan and Solnik (1989), and Sorensen and Yosha (1998) for the seminal papers.

In the literature risk-sharing indicators are often considered as a measure of the degree of financial integration within a group of countries, though they are based on data capturing financial flows with the rest of the world. This interpretation presents a limitation when applied to the African countries because financial flows do not occur as in a "gravity" model and income smoothing occurs mainly with the rest of the world and not necessarily between the African countries. Risk sharing more likely tells us something about whether these countries smooth the shocks on consumption through an easier borrowing from the bilateral/multilateral donors, through remittances from diasporas leaving in African and non-African countries or through domestic saving.

Unlike previous studies,[6] we assume that the coefficients are time-varying, and we estimate them using a rolling window approach to assess their possible instability over time. The reason is that the GDP series from sub-Saharan African countries exhibit a highest variability than those of the industrialized countries because their economic activities are driven by shocks that are more volatile. Accordingly, there may be some biases in the estimation under the assumption that the variance of the GDP is constant over time. To take this into account, we consider a rolling window estimate with a fixed width of 10 years beginning in 1970 until 2012. Thus the estimated risk-sharing coefficients for 2000 are obtained through an estimate of the system over the period 1990–2000; the coefficients for 2001 are obtained by the estimation over 1991–2001 and so forth until 2000–2012. We use seemingly unrelated regressions (SUR) with iterated weighted least squares. Common shocks are captured by time fixed effects in the regressions.

Tables 3a, b, c, and d report the estimates of the smoothing components related to factor income, transfers, and saving, together with the degree of smoothing. The estimates suggest that the smoothing effects have been changing over time and are heterogeneous across subgroups of countries. In WAEMU (Table 3a) risk sharing has been little at work until the period 1996–2006 with only 5% of the shocks absorbed through factor income. Then, saving has played a heavy influence accounting for roughly 30–40% on average of shock smoothing. This portrays a situation of non-integration of the WAEMU with the rest of the world. In the remainder of the survey, we ask the following question: Does this lack of risk-sharing dynamics explain a weak co-movement between the WAEMU countries' growth rates?

The conclusion seems to be different for SADC (Table 3b) where the degree of risk sharing is stronger than in WAEMU and for a longer time period (since 1992). Aggregate risk is shared equally through income factor and saving. In our opinion, the role of factor income in smoothing the asymmetric shocks can be explained by the fact that the group of countries that are members of SADC consists of middle-income countries and some of them play a key role in explaining factor and capital mobility within the subregion (Botswana, South Africa, and Mauritius). Moreover

[6] See Nnyanzi (2013), Tapsoba (2009), and Yehoué (2011).

Table 3 Income smoothing by national account components: 2000–2012

	Degree of smoothing	Factor income		Transfers		Saving	
	Coeff.	Coeff.	T-ratio	Coeff.	T-ratio	Coeff.	T-ratio
(a) WAEMU							
1990–2012	0.25	0.03	1.1	−0.02	−0.35	0.25**	3.66
1990–2000	0	0.04	0.94	−0.04	−0.48	0.13	1.5
1991–2001	0.05	0.05**	2.02	−0.08	−0.95	0.13	1.55
1992–2002	0.05	0.05**	2.01	−0.07	−0.83	0.07	0.82
1993–2003	0.05	0.05*	1.65	−0.07	−0.84	0.16*	1.8
1994–2004	0.07	0.07**	2.38	−0.03	−0.33	0.13	1.39
1995–2005	0	0.04	1.44	−0.04	−0.48	0.09	0.97
1996–2006	0.04	0.04*	1.77	−0.04	−0.52	0.13	1.36
1997–2007	0.26	0.05*	1.98	0.03	0.53	0.21**	2.3
1998–2008	0.31	0.04	0.82	−0.02	−0.45	0.31***	3.42
1999–2009	0.41	0.04	0.84	−0.01	−0.2	0.41***	4.21
2000–2010	0.49	0.03	0.68	−0.02	−0.21	0.49***	4.4
2001–2011	0.39	0.03	0.65	−0.02	−0.24	0.39***	3.18
2002–2012	0.44	−0.03	−0.53	0.06	0.75	0.44***	3.47
(b) SADC							
1990–2012	0.37	0.22***	4.57	−0.009	−0.84	0.15**	2.34
1990–2000	0.16	0.16*	1.72	−0.029	−1.29	−0.03	−0.32
1991–2001	0	−0.03	−0.34	−0.036	−1.55	0.128	1.13
1992–2002	0.18	−0.05	−0.61	−0.04*	−1.75	0.182*	1.69
1993–2003	0.15	0.15**	2.11	−0.02	−1.18	0.04	0.42
1994–2004	0.37	0.22***	2.89	−0.03	−1.5	0.17*	1.66
1995–2005	0.47	0.23***	3.18	−0.02	−0.84	0.24**	2.47
1996–2006	0.58	0.29***	5.43	0.0086	0.56	0.29***	3.05
1997–2007	0.43	0.27***	5.14	0.006	0.44	0.165*	1.75
1998–2008	0.47	0.23***	4.67	0.006	0.66	0.24***	2.93
1999–2009	0.52	0.21***	4.29	0.003	0.34	0.31***	3.66
2000–2010	0.54	0.26***	5.47	0.0038	0.39	0.28***	3.5
2001–2011	0.52	0.25***	5.48	0.004	0.52	0.27***	3.52
2002–2012	0.55	0.26***	5.7	−0.004	−0.55	0.29***	3.98
(c) CAEMC							
1990–2012	0.3	0.24**	2.04	0.06**	2	0.30*	1.79
1990–2000	0.28	0.28**	2.13	0.06	1.61	0.37	1.62
1991–2001	0.9	0.41***	3.52	0.11***	3.17	0.49**	2.48
1992–2002	0.54	0.43***	2.79	0.11***	2.6	0.36	1.63
1993–2003	0.9	0.38**	2.17	0.10**	2.17	0.49**	2.12
1994–2004	0.41	0.32*	1.88	0.09**	2.06	0.34	1.51
1995–2005	0.39	0.30*	1.76	0.09*	2.28	0.35	1.52
1996–2006	0.41	0.31*	1.88	0.10**	2.42	0.35	1.52
1997–2007	0.37	0.28*	1.67	0.09**	2.33	0.35	1.52

(continued)

Table 3 (continued)

	Degree of smoothing	Factor income		Transfers		Saving	
	Coeff.	Coeff.	T-ratio	Coeff.	T-ratio	Coeff.	T-ratio
(c) CAEMC							
1998–2008	0.52	0.39**	1.99	0.13**	2.54	0.32	1.16
1999–2009	0.58	0.44**	2.14	0.14**	2.57	0.45	1.61
2000–2010	0.58	0.08	0.4	0.05	0.85	0.58*	1.86
2001–2011	0	0.18	0.88	0.07	1.26	0.21	0.89
2002–2012	0	−0.27	−1.04	−0.05	−0.64	−0.29	−0.95
(d) Others							
1990–2012	0	0.04	0.77	0.009	0.21	0.05	0.36
1990–2000	0	0.07	0.74	0.03	0.74	−0.05	−0.22
1991–2001	0	0.11	1.08	0.04	1.29	−0.004	−0.016
1992–2002	0	0.05	0.65	0.03	0.92	0.03	0.14
1993–2003	0	0.16	1.63	0.04	1.14	−0.05	−0.21
1994–2004	0.05	0.12	1.51	0.05*	1.86	−0.004	−0.02
1995–2005	0.05	0.13*	1.66	0.05*	1.82	−0.00014	−0.0006
1996–2006	0	0.03	0.35	0.04	1.46	−0.07	−0.347
1997–2007	0	0.03	0.38	0.04	1.32	−0.11	−0.49
1998–2008	0.05	0.04	0.46	0.05**	2.56	−0.13	−0.56
1999–2009	0.06	0.07	0.86	0.06***	2.82	−0.12	−0.52
2000–2010	0	0.04	0.55	0.05	0.57	−0.09	−0.43
2001–2011	0	0.03	0.39	0.04	0.46	0.12	0.63
2002–2012	0	−0.02	−0.37	−0.03	−0.31	0.25	1.45

Note: *, **, *** statistically significant at 10%, 5%, and 1%

the higher risk sharing could be explained by the fact that the "core" countries in SADC have a high integration with the rest of the world.

In the CAEMC (Table 3c), the smoothing of the negative impact of the idiosyncratic shocks occurs through factor income and transfers with an intensity which is as high as in the SADC (though shock smoothing has ceased since 2009). Since the beginning of 1990s, the main smoothing factor has been a factor income, in spite of a low degree of labor mobility within the region. One may wonder whether there is an "oil effect" here, since income transfers are mainly related to oil income transfers (net transfers from oil companies).

For the remainder of our sample (which includes countries from ECOWAS and EAC), no evidence of risk-sharing smoothing is found.

2.2.2 Indicators Related to the Banking Sector

We consider a second set of financial integration indicators based on data from the banking sector. We propose indicators of the convergence of commercial banks'

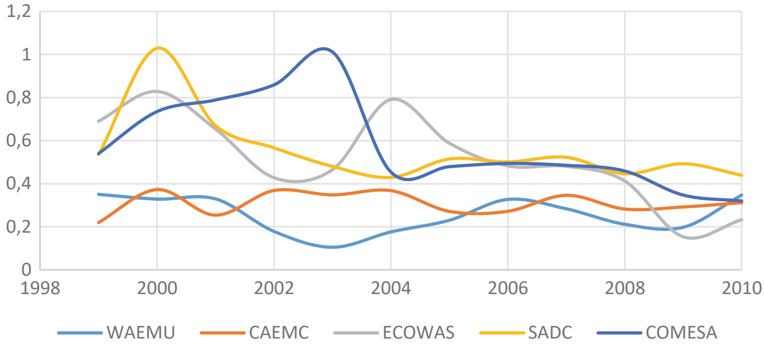


Fig. 1 Sigma convergence: net interest margins (standard error/average)

intermediation margins (net intermediation margin) and financial performance (ROE), based on time-varying sigma convergence. Net intermediation margins are defined as the accounting value of a bank's net interest revenue as a share of its interest-bearing (total earning) assets, while bank ROE is the average return on assets (net income/total equity). Data are taken from the World Bank financial structure dataset. We assume that convergence of these indicators over time could be seen as reflecting a greater integration.

Following the methodology proposed by Sy (2007) to estimate the degree of integration, we build a time series of the cross-sectional dispersion in net interest margins and ROEs. At each period t, we calculate the standard deviation of these variables across countries belonging to the same group. The formula is the following:

$$\sigma_t = \left[\frac{1}{n-1}\sum\left(X_{i,t} - \overline{X}_t\right)^2\right]^{\frac{1}{2}} \tag{8}$$

where i represents an individual country and X our variable of interest (NIM or alternatively the ROE). This variable is normalized by the time average. Therefore, we compute the dispersion coefficients.

The main advantage of using cross-sectional dispersions is that, contrary to correlations, they can be calculated at each point in time. When series are highly correlated, as they should be in integrated markets, variables generally move in the same direction, and cross-sectional dispersion is low. Alternatively, dispersion is high when the variables in different countries drift apart.

In Fig. 1, for purpose of illustration, we present the cross-sectional and time-series patterns of NIMs for the different groups of our sample. The graph suggests that net interest margins have globally converged over the period in ECOWAS, SADC, and COMESA, while the evolution is more contrasted in the franc zone.

The evolution in ECOWAS, SADC, and COMESA may be related to an increased competition between more efficient banks. Indeed, a reduced market concentration has led to reduce interest margins in sub-Saharan Africa (Ahokpossi 2013).

2.3 Control Variables

While examining the role of financial integration, other factors explain that the economies are more or less exposed to asymmetric shocks. Most studies investigating how the African economies could absorb idiosyncratic shocks suggest that, in addition to deeper financial integration, this could come through higher trade linkages, a convergence of macroeconomic policies, or a reduction of structural asymmetries.[7]

We consider the following control variables which capture several channels through which the countries' cycles affect each other: terms of trade, real effective exchange rates, financial stress in the industrialized countries, changes in M2 over GDP ratio, inflation rate, and public debt as share of GDP.

These variables have been shown to be robust determinants of growth accelerations and decelerations in the African countries.[8] Our question is whether they are equally effective as financial integration in explaining business cycle co-movements.

The dependence of the African countries to globalization is captured here by an index of financial stress indicator proposed by Balakrishnan et al. (2009). This control variable is important because a high vulnerability of the African countries to international financial shocks could be a motivation to improve their financial integration. Inflation and debt over GDP ratio are considered as proxies of monetary and fiscal policies. M2 as share of GDP accounts for the degree of development of the banking sector, which is a prerequisite for financial integration and could influence the effects of the bank-based indicator (defined above as the sigma-convergence of banks' ROE). Terms of trade and the real exchange rate are considered to reflect possible asymmetric shocks to the trade balance.

All the data are taken from WDI (World Bank Development Indicators database) over the period 2000–2012. In the probit regressions below, some of the control variables are measured in terms of their deviations to the average of the subgroup: terms of trade, real effective exchange rate, inflation rate, the ratio of debt over GDP, and changes in the ratio of M2 over GDP.

[7]See Tapsoba (2009) and Wang et al. (2007).

[8]For recent papers, see Tsangarides (2012).

2.4 The Econometric Model

The econometric methodology is based on a probit model. Assume that y_i^{*E} and y_j^{*E} represent the "true" unobserved values of the GDP during expansion in countries i and j and that the computed series using our filter are y_i^E and y_j^E. A synchronization of two business cycles is described by the following notation (where C_t^{ij} is defined by Eq. 2, when $S_t^i = \wedge_t^i$ and $S_t^j = \wedge_t^j\Big)$:

$$C_t^{ij} = \begin{cases} 1, \ if \ y_{it}^{*E} > 0 \text{ and } y_{jt}^{*E} > 0 \\ 0, \text{ otherwise} \end{cases} \tag{9}$$

We proceed in a similar way to characterize synchronization during recession and asynchronous phases. Since our endogenous variables consist of zeros and ones, they can be interpreted as probabilities.

We want to assess the probability that the countries' GDP co-move or evolve in opposite directions during the expansion and recession phases of their business cycles. Specifically, we want to see whether the information content of our financial integration variables has predictive power in forecasting business cycle synchronization or desynchronization.

The estimated equation is as follows:

$$\Pr\left(C_t^{ij} = 1/X\right) = \alpha + \sum\nolimits_{j=1}^{4} \beta_j \text{Fin}_{it}^j + \sum\nolimits_{k=1}^{2} \gamma_k \text{Res}_{it}^k + \delta \text{control}_{it} + \varepsilon_{it}^1 \tag{10}$$

where Φ is the standard normal CDF (probit). X is the matrix of observations of the independent variables. Using the estimate of β for each subgroup of countries, we compute the average partial effect of a higher degree of financial integration on the probability that a country evolves in a similar business cycle phases than the majority of the others in the same subgroup. The estimation of the coefficient is performed using the Mundlak-Chamberlain estimator on random effects model. This allows for a serial correlation in the unobserved variables determining the endogenous variable and relies on the assumption of a correlation between the fixed effects and the explanatory variables.

3 The Results: Is Financial Integration Large Enough to Affect Business Cycle Synchronization?

Table 4a, b, c, and d present the results from estimating Eq. (10) when we distinguish between groups of countries according to their membership to economic unions: SADC, WAEMU, CAEMC, and others. We report the average partial effects estimated using the Chamberlain-Mundlak estimator.

Table 4 Financial integration and business cycle synchronization

			Net interest margin		ROA	
	Risk sharing		Sigma convergence		Sigma convergence	
	Coefficient	T-ratio	Coefficient	T-ratio	Coefficient	T-ratio
(a) During expansions						
Financial integration						
SADC	0.154	0.154	0.101	1.57	0.032	0.67
WAEMU	0.475***	4.92	0.314***	13.86	0.098	1.59
CAEMC	0.167***	∞	0.078***	∞	0.05	0
Others	−52.92***	∞	−1.95***	−31.59	−6.62***	−95.03
Control variables						
Terms of trade—subregion(terms of trade)	−0.012	−0.595	−0.00024	−0.013	−0.02	−0.99
Financial stress	0.04	0.583	0.0316	0.421	0.04	0.478
Real exchange rate—subregion REER	−0.013	−0.167	−0.046	−0.646	−0.01	−0.143
Changes in M2—regional changes in M2 (%GDP)	−0.01***	−175.3	−0.022***	−224.08	−0.03***	−359.41
Inflation rate—subregional inflation	−0.00002	−0.034	−0.000014	−0.02	−0.000026	−0.036
Debt/GDP—subregional debt/GDP	−0.000715	−0.006	−0.00052	−0.032	−0.000577	−0.023
Prediction	$y = 0$	$y = 1$	$y = 0$	$y = 1$	$y = 0$	$y = 1$
$P(y = 1) \leq 0.5$	248	127	221	96	239	125
$P(y = 1) > 0.5$	13	15	32	46	22	17
% correct	95.01	10.56	87.74	32.39	91.57	11.97
% incorrect	4.98	89.44	12.26	67.61	8.43	88.03

(b) During recessions						
Financial integration						
SADC	−0.02	−0.174	−0.119***	−3.61	−0.037	−0.76
WAEMU	−0.291***	−4.24	−0.394***	−38.57	−0.17***	−3.94
CAEMC	−0.056	−0.04	−0.08***	−6.23	−0.04**	−2.69
Others	0.445***	10.38	−0.08**	−2.21	0.009	0.21
Control variables						
Terms of trade—subregion(terms of trade)	−0.004	−0.309	−0.028**	−2.33	−0.009	−0.67
Financial stress	−0.03	−0.63	−0.015	−0.325	−0.03	−0.56
Real exchange rate—subregion REER	0.05	0.837	0.07**	2.03	0.05	0.87
Changes in M2—regional changes in M2 (%GDP)	−0.004***	−121.5	0.03**	736.32	0.017***	3.99
Inflation rate—subregional inflation	0.000015	0.04	0.000005	0.01	0.000016	0.03
Debt/GDP—subregional debt/GDP	−0.00043	−0.006	−0.00051	−0.05	−0.00045	−0.027
Prediction	$y = 0$	$y = 1$	$y = 0$	$y = 1$	$y = 0$	$y = 1$
$P(y = 1) \leq 0.5$	293	109	281	97	293	109
$P(y = 1) > 0.5$	0.0	1.0	12	13	0.0	1.0
% correct	100	0.01	95.90	11.82	100	0.9
% incorrect	0.0	0.99	4.09	88.18	0.0	99.1

(continued)

Table 4 (continued)

	Risk sharing		Net interest margin		ROA	
			Sigma convergence		Sigma convergence	
	Coefficient	T-ratio	Coefficient	T-ratio	Coefficient	T-ratio
(c) Country, expansion; others, recession						
Financial integration						
SADC	−0.12***	−5.21	−0.065***	−52.11	−0.048***	−6.91
WAEMU	−0.63***	−65.15	−0.22***	−618.72	−0.21***	−37.88
CAEMC	−0.148	−0.95	−0.058***	−147.81	−0.172***	−142.80
Others	3.92***	∞	0.03***	26.22	0.056***	9.00
Control variables						
Terms of trade—subregion(terms of trade)	0.029***	15.26	0.078***	183.89	0.095	48.27
Financial stress	−0.003	−0.326	0.007***	4.13	0.007	1.03
Real exchange rate—subregion REER	−0.126***	−10.79	−0.151***	−51.80	−0.16***	−12.71
Changes in M2—regional changes in M2 (%GDP)	−0.136***	∞	−0.131***	∞	−0.14***	∞
Inflation rate—subregional inflation	−0.0003***	−5.08	−0.0001***	−8.33	−0.00017	−4.02
Debt/GDP—subregional debt/GDP	0.000026	0.0022	−0.000021	−0.076	−0.00004	−0.026
Prediction	$y = 0$	$y = 1$	$y = 0$	$y = 1$	$y = 0$	$y = 1$
$P(y = 1) \leq 0.5$	323	70	308	45	319	62
$P(y = 1) > 0.5$	4	6	19	31	8	14
% correct	98.77	7.89	94.19	40.79	97.55	18.42
% incorrect	1.23	92.11	5.81	59.21	2.45	81.58

(d) Country, recession; others, recession						
Financial integration						
SADC	0.097**	2.42	0.042***	6.18	0.004	0.30
WAEMU	0.26***	10.97	0.146***	53.76	0.037**	2.23
CAEMC	0.09***	∞	0.033	∞	0.004	∞
Others	−36.73	∞	−1.46***	−78.27	−5.05***	−272.53
Control variables						
Terms of trade—subregion(terms of trade)	−0.03***	−7.53	−0.02***	−3.66	−0.03***	−4.83
Financial stress	−0.011	−0.652	−0.015	−0.65	−0.01	−0.526
Real exchange rate—subregion REER	0.07***	3.95	0.04*	1.91	0.06**	2.75
Changes in M2—regional changes in M2 (%GDP)	0.06***	∞	0.05	∞	0.05***	∞
Inflation rate—subrégional inflation	−0.00004	−0.281	−0.00003	−0.19	−0.00004	−0.174
Debt/GDP—subregional debt/GDP	0.0008	0.026	0.0009	0.18	0.0009	0.136
Prediction	$y = 0$	$y = 1$	$y = 0$	$y = 1$	$y = 0$	$y = 1$
$P(y = 1) \leq 0.5$	328	74	328	74	328	74
$P(y = 1) > 0.5$	0.0	1	0.0	1	0.0	1.0
% correct	100	1.33	100	1.33	100	1.33
% incorrect	0.0	98.66	0.0	98.66	0.0	98.66

Probit estimation (random effect: Chamberlain-Mundlak)—average partial effect
Note : **,*** : statistically significant at 5% and 1%

The financial integration variables and the control variables have a very low power in predicting the strength of recessions across subgroups of countries. Indeed, based on the computation of the percentage of correct predictions of the endogenous variable in Table 4b and d, we obtain a goodness of fit which is extremely small (in the best case, 12% and very often only 1% and 2%). The model appears to predict better the occurrence of co-movement in the GDP growth during expansions (as suggested by Table 4a). The predictability accuracy is also improved when the endogenous variable is the degree of desynchronization between a country's expansion and the other countries' recession (Table 4c). We therefore focus our comments on the results of Table 4a and c.

In Table 4a, the regression with net interest margin as the financial integration variable outperforms the other two models with risk sharing and ROA variables, since 32% of the endogenous variable is correctly predicted against, respectively, 11% and 12% for the other two models. The same feature is observed in Table 4c where the number of correct predictions jumps to 41%. Comparing Table 4a and c, the regressions with net interest margin and ROA suggest that financial deepening weakens the links between outputs during expansions and increases the desynchronization between the business cycles when a country experiences an expansion while the neighbors are in recession (the coefficients for SADC, WAEMU, and CAEMC are positively signed in Table 4a and carry a negative sign in Table 4c). This finding plays in favor of the assumption of "polarization within the region." For instance, the heterogeneity of the economic fundamentals between countries leads to lending activities that are more risky in some countries than in others.

When financial integration is measured by risk-sharing variables, during expansions a higher risk sharing does feed back into higher output correlation in WAEMU and CAEMC (the coefficients are positively signed and statistically significant in Table 4a). However, this happens only if the subregion as a whole is also experiencing an expansion. When a country experiencing an expansion is out of phase with the neighbors, the coefficients turn to be negative (see Table 4c). This observation is also in line with the hypothesis of polarization of economic activities within the two subregions. For WAEMU and CAEMC, one explanation may be that the domestic financial markets are shallow. For SADC, in line with conventional findings in frontier and emerging economies, what seems to be at play is called "asset substitutability" with a concentration of capital flows to the richest countries of the subgroup (South Africa, Botswana, and Mauritius). Since financial investment is likely to lead higher returns in these countries, this makes the subgroup more exposed to asymmetric shocks.

All in all, our main and new findings regarding the role of financial integration on business cycle synchronization point to a higher "polarization" during a process of higher financial integration. This is in line with the theoretical literature providing an argument suggesting that risk sharing across countries induces productive specialization, thereby increasing the vulnerability to their own idiosyncratic shocks. This conclusion, usually obtained for developed countries, also holds for the African countries where the production activities have often been substitutable rather than complements (even though countries are members of economic unions), particularly when the risk sharing is analyzed vis-à-vis the rest of the world.

Concerning the control variables, the results suggest that smoothing through positive shocks on the terms of trade is not substantial (nonsignificant coefficients in Table 4a) but can sometimes de-correlate the cycles (positive and statistically significant estimates in Table 4c). Changes in M2 to GDP ratio (compared to the average changes in the neighbor countries) are found to exert a systematically significant role, both on business cycle synchronization and desynchronization.

The other control variables seem to be informative of the dynamics of the endogenous variables in Table 4c. Shocks on the real exchange rate increases the desynchronization of the business cycle phases.

4 Conclusion

The question as whether financial integration could be a channel for higher economic integration in sub-Saharan Africa (as measured by the synchronization of the business cycles) is an important issue in the policy circles. No studies so far have examined this question empirically, probably due to the observation that the development of the African financial markets is unequal across the continent (though progress is on the rise). This contribution tries to fill this gap. Our innovations consist in (1) tackling this issue by considering different countries groupings (to account for observed heterogeneity), (2) building intuitions by disentangling situations where synchronization occurs during periods of recession and expansion and by also taking into account the impact of financial integration on the desynchronization of the business cycles, and (3) quantifying the potential for time-varying risk-sharing and price-based indicators of convergence.

Our study suggests that the effect of financial integration is heterogeneous across groups of countries. In the CEMAC, we find a positive relationship between financial integration and business cycle correlations, but it is hard to disentangle the impact of financial integration per se from the paramount importance of oil output and exports in determining business cycles. We also find that, in the case of regional recessions, the expanding country may incur significant risk-sharing costs when transferring resources to other group members. This calls for more regional financial integration to mitigate these individual costs. Regarding WAEMU and SADC, we find that financial integration increases the dephasing of business cycles. This could be the result of weak absorptive capacity of financial due to the lack of financial and economic development in WAEMU. In SADC, this undesirable effect of risk sharing could be explained by gaps of economic development between member countries (with some countries acting as “hub” of development) and economic specialization. Further, reserve pooling does not play a substantial role in smoothing idiosyncratic shocks.

The weak impact of financial integration on business cycle synchronization mirrors other studies in the literature focusing on the link between financial development and economic growth, which shows that this correlation is nonlinear according to financial development. In emerging countries with intermediate levels of financial development, this link is usually found to be positive. In countries with low levels of financial development, the relationship is not significant at all.[9] Given the rapid progress of economic and financial growth in sub-Saharan Africa, one might expect more robust and more homogenous relationship between financial integration and business cycle correlations to emerge gradually. This calls for more studies on financial integration in developing countries, in parallel with those on financial development, as financial integration carries its own set of financial risks (contagion, etc.) and policy recommendations.

Appendix

Table 5 Convergence criteria in monetary unions projects in sub-Saharan Africa

African union projects	First rank convergence criteria in 2013				
	Inflation (%)	Global fiscal deficit/GDP (%)	International reserves (in months of import)	Public debt/GDP (%)	Government financing from the Central Bank
EAC	$\leq$8	$\leq$3	$\geq$4, 5	$\leq$50	–
WAMZ	$\leq$5	$\leq$4	$\geq$6	–	$\leq$10% of tax receipts of year $n-1$
AU	$\leq$3	$\leq$3	$\geq$6	–	0%
SADC (2012–2018)	$\leq$5	$\leq$3	$\geq$6	$\leq$60	$\leq$5% of tax receipts of year -1
COMESA (2011–2015)	$\leq$3	$\leq$4	$\geq$5	–	0%

Monetary Union Projects
EAC East African Community, with five member countries (Burundi, Kenya, Uganda, Tanzania, and Rwanda)
WAMZ West African Monetary Zone, six member countries (Gambia, Sierra Leone, Guinéa, Ghana, Nigeria, and Liberia). Initially set for January 2003, then postponed to July 2005, and then January 2010, the merger between West African Monetary zone and WAEMU is now envisaged for 2020
AU African Union, 53 African countries
SADC Southern African Development Community, 15 member countries, of which South Africa, Lesotho, Namibia, and Swaziland, which has a common currency, the south African Rand, and based on the use of currency board mechanisms
COMESA Common Market for Eastern and Southern Africa, 19 member countries, of which, Burundi, Comoros, Djibouti, Egypt, Libya, and Rwanda

[9] The interested reader can refer to recently published papers on this topic (see, among others, Eggoh (2010), Fu-Sheng (2009), Hong-Chuan and Shu-Lin (2009), Maher (2012), and Masten et al. (2008).

References

Abou, A., & Prat, G. (2000). Modelling stock price expectations: Lessons from microdata. In F. Gardes & G. Prat (Eds.), *Price expectations in goods and financial markets* (pp. 313–346). Cheltenham: Edward Elgar.

Abou, A., & Prat, G. (2010). The dynamics of U.S. equity risk premia: Lessons from professional's view. *Bankers, Market and Investors, 104*, 4–20.

Ahokpossi, C. (2013, January). *Determinants of bank interest margins in Sub-Saharan Africa*. IMF working paper n° 13/34.

Asdrubli, P., Sorensen, B., & Yosha, O. (1996). Channels of inter-state risk-sharing: United States 1963–1990. *Quarterly Journal of Economics, 111*, 1081–1110.

Balakrishnan, R., Danninga, S., Elekdag, S., & Tytell, I. (2009). *The transmission of financial stress from advanced to emerging economies*. IMF working paper #09/133.

Brennan, B., & Solnik, M. J. (1989). International risk-sharing and capital mobility. *Journal of International Money and Finance, 8*(3), 359–373.

Carmignani, F. (2009). *Endogenous optimal currency areas: The Central African economic and monetary community*. Discussion paper # 390, School of Economics, University of Queensland.

Coleman, S. (2011). *Investigating business cycle synchronization in West Africa*. Discussion paper in economics, Nottingham: Trent University.

Eggoh, J. (2010). Financial development and growth: A panel smooth regression approach. *Journal of Economic Development, 35*(1), 15–33.

Fu-Shen, H. (2009). Explaining the nonlinear effects of financial development on economic growth. *Journal of Economics, 97*(1), 41–65.

Harding, D., & Pagan, A. R. (2002). Dissecting the cycle: A methodological investigation. *Journal of Monetary Economics, 49*(2), 365–381.

Hitaj, E., Kolerus, C., Shapiro, D., & Zdzienicka, A. (2013). *Responding to shocks and maintaining stability in the West African Monetary Unions*. Washington, DC: African Department, International Monetary Fund.

Ho-Chuan, H., & Shu-Lin, L. (2009). Nonlinear finance-growth nexus. A threshold with instrumental variable approach. *Economics of Transition, 17*(3), 439–466.

Houssa, R. (2008). Monetary union in West Africa and asymmetric shocks: A dynamic structural factor model. *Journal of Development Economics, 85*(1–2), 319–347.

Jawadi, F., & Prat, G. (2012). Arbitrage costs and nonlinear stock price adjustment in the G7 Countries. *Applied Economics, 44*(12), 1561–1582.

Jawadi, F., & Prat, G. (2017). Equity prices and fundamentals: A DDM-APT mixed approach. *Journal of Quantitative Finance and Accounting, 49*(3), 661–695.

Kablan, S. (2010). *Banking efficiency and financial development in sub-Saharan Africa*. IMF working paper #10/136.

Kemegue, F., & Seck, O. (2014). Do African monetary arrangements make sense? Evidence based on structural symmetry. In D. Seck (Ed.), *Regional economic integration in West Africa*. Heidelberg: Springer.

Mafusire, A., & Brixiova, Z. (2013). Macroeconomic shock synchronization in the East African Community. *Global Economy Journal, 13*(2), 261–280.

Maher, A. (2012). Nonlinear growth impacts of financial development in Euro area. *International Journal of Economics and Finance, 12*, 23–38.

Masten, A., Coricelli, F., & Mastar, J. (2008). Non-linear effects of financial development: Does financial integration matter? *Journal of International Money and Finance, 27*(2), 295–317.

Monfort, M., Seok, G. P., & Masafumi, Y. (2013). *Banking in sub-Saharan Africa: The macroeconomic context*. African Department Publications, International Monetary Fund.

Nnyanzi, J. (2013). Remittances and risk-sharing in Africa. *Issues in Business Management and Economics, 5*, 107–114.

Prat, G., & Uctum, R. (2007). Switching between expectational processes in the foreign exchange market: A probabilistic approach using survey data. *Review of International Economics, 5*(4), 700–719.

Prat, G., & Uctum, R. (2011). Modelling oil price expectations: Evidence from survey data. *Quarterly Review of Economics and Finance, 51*(3), 236–247.

Prat, G., & Uctum, R. (2013). Modeling the Horizon-dependent risk premium in the Forex: Evidence from survey data. *Journal of International Financial Markets, Institutions and Money, 33*, 33–54.

Prat, G., & Uctum, R. (2015). Expectation formation in the foreign exchange market: A time-varying heterogeneity approach using survey data. *Applied Economics, 47*(34–35), 3673–3695.

Sorensen, B., & Yosha, O. (1998). International risk-sharing and European Monetary Unification. *Journal of International Economics, 45*, 211–238.

Ssozi, J. (2011). Business cycle synchronization in the proposed East African Monetary Union: An unobserved component approach. *Review of Development Economics, 15*(4), 664–675.

Sy, A. (2007). Financial integration in the West African Monetary Union. *Journal of Financial Transformation, 19*, 91–103.

Tapsoba, S. (2009). Trade intensity and business cycle synchronicity in Africa. *Journal of African Economies, 18*(2), 287–318.

Tsangarides, C. (2012). *Determinants of growth spells: Is Africa different?* IMF working paper #12/227.

Uctum, R., Renou, P., Prat, G., & Lecarpentier, S. (2017). Persistence of announcement effects on the intraday volatility of stock returns: Evidence from individual data. *Review of Financial Economics, 35*, 43–56.

Wang, J.-Y., Masha, I. Shirono, K., & Leighton, H. (2007). *The common monetary area in Southern Africa: Shocks, adjustment, and policy challenges*. IMF working paper # 07/158.

Yehoué, E. (2011). International risk-sharing and currency unions: The CFA zones. *Journal of International Development, 23*(7), 936–958.

Part IV
Fundamentals and Bubbles

Informational Efficiency and Endogenous Rational Bubbles

George A. Waters

Abstract In a model where rational bubbles form and collapse endogenously, properly specified tests of return predictability have little power to reject deviations from the efficient markets model. A weighted replicator dynamic describes how agents switch between a forecast based on fundamentals, a rational bubble forecast, and a weighted average of the two. A significant portion of the population may adopt the rational bubble forecast, which is inconsistent with the efficient markets model but satisfies informational efficiency. Tests on simulated data show excess variance in the price and unpredictable returns.

JEL Classification C22, C73, G12, D84

1 Introduction

Eugene Fama received a Nobel Prize for his work providing evidence that stock prices embody all available information. Robert Shiller received a Nobel Prize for the work providing evidence such as the excess variance in stock prices that asset price fluctuations could not be fully explained by fundamentals. These views represent weak and strong versions of the efficient markets hypothesis and are not mutually exclusive. Evidence that returns are not forecastable, informational efficiency (IE), does not imply that asset prices are solely determined by fundamentals, as in the standard efficient markets model (EMM).[1]

This contribution describes an asset pricing model that is informationally efficient, but the asset price could depend on extraneous information. A representative

[1]This terminology is used by Shiller (1981). Others refer to the model as representing the strong efficient markets hypothesis.

G. A. Waters
Department of Economics, Illinois State University, Normal, IL, USA

F. Jawadi (ed.), *Uncertainty, Expectations and Asset Price Dynamics*, Dynamic Modeling and Econometrics in Economics and Finance 24,
https://doi.org/10.1007/978-3-319-98714-9_7

agent rational bubble model such as Blanchard (1979) has these features, but it is unclear how agents would coordinate on such forecasts. The present model includes heterogeneous forecasting strategies where rational bubbles endogenously form and collapse, representing a deviation from the efficient markets model. However, properly specified tests of IE have very little power to reject such deviations. Therefore, tests of return forecastability give evidence about the forecastability of returns and little else.

The endogenous rational bubbles produce excess variance in the returns as found in financial data.[2] Under the standard EMM, where the price is determined by discounted expected future dividends, Cochrane (2008) shows that the variance of the price–dividend ratio can be decomposed into the unforecastable components of the dividend innovations and the returns. Under the present model, that relationship is broken, since the asset price could depend on extraneous data. Hence, one cannot accept the conclusion, as in Cochrane (2008), that a limited degree of return forecastability rules out bubbles.

An evolutionary game theory mechanism describes how agents switch between three forecasting strategies. The *fundamental* forecast is based on the EMM. The *mystic* forecast is formed using a model of a rational bubble where extraneous information affects the asset price. The third *reflective* forecast is a weighted average of the two and represents the rational forecast in an environment with heterogeneity. Payoffs to these forecasting strategies are based on past forecast errors. A weighted replicator dynamic describes the evolution of the fractions of agents using the different strategies and allows for the parameterization of how fast agents switch to better performing strategies.

For some parameter choices, the fundamental forecast dominates, and the model behaves according to the EMM. For other plausible settings, however, mysticism can gain a temporary following, leading to deviations from the EMM, in that non-fundamental information affects the price. Such deviations are detected as excess variance in the simulated data. Mystic outbreaks require that the magnitude of the shock to the extraneous information is similar to that of the fundamental information and that agents are sufficiently aggressive about switching to better performing strategies. Mysticism cannot outperform reflectivism indefinitely so such mystic bubbles form and collapse endogenously, in contrast to other models of rational bubbles.

The work of Georges Prat provides empirical support for the consideration of dynamic, heterogeneous forecasting strategies. Using survey data for stock prices (Abou and Prat, 2000), oil prices (Prat and Uctum, 2011), and foreign exchange rates (Prat and Uctum 2015), these authors reject models with a single rational forecast and models with fixed fractions of the population using different forecasts. Only models with dynamically switching, heterogeneous forecasts are able to properly describe the survey data.

[2] Shiller (1981) and LeRoy and Porter (1981) show evidence of excess volatility.

Many studies of IE regress asset returns on a lagged predictor variable such as the price–dividend ratio and conduct a standard t-test to determine if the coefficient on that variable is significantly different than zero. The coefficients are usually different than zero, though the level of significance and R^2 are often unimpressive (Cochrane, 2008). Campbell and Yogo (2006) point out that such tests do not properly account for the persistence of the predictor variable, and they construct a test that does. Their results on IE are mixed in that they depend on the sample, a reflection of the related literature.

Applying the test of Campbell and Yogo (2006) on simulated data from the model that allows for endogenous rational bubbles shows that returns are not predictable using both dividends and the price–dividend ratio as predictors. For a range of parameter choices, the simulated data shows excess variance but not return predictability. Hence, even if the null of IE cannot be rejected, endogenous rational bubbles could be present, violating the EMM.

The behavior of all agents in the present work satisfies the *cognitive consistency principle,* described in Evans and Honkapohja (2011), which says that agents in a model are as smart as economists. More precisely, agents form expectations using reasonable models according to economic theory. The model also explains observed features of financial markets data such as excess kurtosis and GARCH effects in the returns, see Parke and Waters (2007) for details. Parke and Waters (2014) conduct a formal stability analysis on a related model. The primary aim of the present work is to study the implications for the interpretation of econometric tests of IE. Another contribution is the formal analysis of the endogenous collapse of the bubbles.

There are a number of interesting alternative approaches to asset pricing that involve deviations from the EMM, though the implications for return predictability need to be examined in detail. Adam et al. (2016) and Lansing (2010) are able to match a number of the features of the US stock market data. In the model in the former paper, a representative agent updates its estimate of the long-run growth rate of the asset price, which is used for forecasting. In Lansing (2010), the forecasting model (perceived law of motion) includes a geometric random walk, making bubbles a possibility, and agents update a parameter in the forecasting model that determines the impact of the bubble. The agents in Branch and Evans' (2011) model of bubbles update an estimate of the conditional variance of the return using a linear model. The time series implications of this approach have yet to be explored in detail.[3]

There are a number of asset pricing models with heterogeneous forecasting strategies. In LeBaron (2012), some agents use a "buy and hold" strategy, which has intuitive appeal but may not satisfy cognitive consistency. The cognitive consistency of agents in the asset pricing models of Brock and Hommes (1998) and Branch

[3]Similarly, the noise trader model of DeLong et al. (1991) has deviations from the strong EMH, but the relationship to return predictability is discussed informally.

and Evans (2007) is open to interpretation.[4] In the former paper, some agents have perfect foresight but must pay a cost. In Branch and Evans (2007), some agents use underparameterized models, which exclude information that affects the asset price. In contrast, in the present model, all the forecasting strategies can be described as rational though they differ on beliefs about which information is important. The fundamental and mystic forecasts can be found in the asset pricing literature, and the reflective forecast is the unbiased forecast in the presence of heterogeneity.

The chapter is organized as follows. The asset pricing model, forecasting strategies, and their payoffs are specified in Sect. 2. Section three has a discussion of the intuition behind the formation and collapse of mystic bubbles and the implications for return predictability. Section 4 describes the evolutionary game theory mechanism and the requirements for bubbles to arise. Section 5 gives details about the econometric tests on the simulated data and Sect. 6 concludes. Versions of the model are described in Sects. 2 and 4 can be found in Parke and Waters (2007, 2014). The formal analysis in Sects. 3, 5, and 6 forms the primary contributions, which develop the intuition and implications of the model of mysticism.

2 Asset Pricing

This section specifies the three forecasts and the resulting realization of the asset price, which thereby determine the forecast errors for each strategy. The underlying motivation is the standard asset pricing equation:

$$p_t = \alpha p^e_{t+1} + d_t, \tag{1}$$

where the asset price is p_t, the dividend is d_t, and the parameter α is the discount factor. This model is not fully sufficient for our purpose, since there is a representative forecast of the price. Brock and Hommes (1997) develop a model with mean-variance optimization where risk-neutral investors choose between a riskless and risky asset in constant supply. With risk-neutral agents who have a common belief about the variance of the returns, the model with heterogeneous forecasts can be written as:

$$p_t = \alpha \sum_{h=1}^{n} x_{h,t} e_{h,t} + d_t + RP \tag{2}$$

where the vectors $e_t = \left(e_{1,t}, \ldots, e_{n,t}\right)$ and $x_t = \left(x_{i,t}, \ldots, x_{n,t}\right)$ are the different forecasts of p_{t+1} and the fractions of agents using the forecasts, respectively. The

[4]These papers are part of a large literature using the multinomial logit dynamics to describe the evolution of heterogeneous forecasts. See Hommes (2006) for a survey.

constant RP is a risk premium, which is set to zero in the following to simplify the presentation.

The forecasts considered are motivated by the multiplicity of solutions to the model (1) in the homogeneous case. According to the efficient markets model (EMM), the price is given by the discounted expected future dividends as in the following solution to the model (1).

$$p_t^* = d_t + \sum_{j=1}^{\infty} \alpha^j E_t(d_{t+j}) \tag{3}$$

Agents referred to as *fundamentalists* adopt the forecast $e_{2,t}$ determined by the above solution.

$$e_{2,t} = E_t(p_{t+1}^*) = \sum_{j=1}^{\infty} \alpha^{j-1} E_t(d_{t+j}) \tag{4}$$

However, this solution is not unique. As discussed in the rational bubbles literature, Evans (1991), for example, there is a continuum of solutions to (1) of the form:

$$p_t^m = p_t^* + \alpha^{-t} m_t$$

where the stochastic variable m_t is a martingale such that $m_t = m_{t-1} + \eta_t$, for *iid*, mean zero shocks η_t. Though the information contained in the martingale m_t may be extraneous, if agents believe that information is important, it does affect the asset price. Agents that adopt the forecast $e_{3,t}$ based on the rational bubble solution above are called *mystics*, and their forecast is as follows:

$$e_{3,t} = E_t\left(p_{t+1}^m\right) = E_t(p_{t+1}^*) + \alpha^{-t-1} m_t \tag{5}$$

A primary objection to such a solution is that it violates a transversality condition, see Lundqvist and Sargent (2004, section 13.6). As discussed in Lansing (2010), an agent could profitably short the risky asset if the prices follow such a path. However, this hypothetical agent would need to be infinitely lived with unlimited resources or ability to borrow. Furthermore, agents in the present model can adopt or abandon the forecast at any time so this objection to the mystic forecast is not a concern.

Both the mystic and fundamental forecasts satisfy rational expectations in that they are unbiased in the homogeneous case. However, our goal is to allow for possible heterogeneity in forecasting strategies, so we introduce the *reflective* forecast, which satisfies rational expectations even in the presence of heterogeneity. Inclusion of the reflective forecast also ensures the endogenous collapse of any bubble, see Proposition 3. The reflective forecast $e_{1,t}$ is an average of the alternative forecasts used in the population weighted according to the relative popularity, such that

$$e_{1,t} = (1 - n_t)\, e_{2,t} + n_t e_{3,t} \tag{6}$$

where

$$n_t = \frac{x_{3,t}}{x_{2,t} + x_{3,t}}$$

The variable n_t is the relative popularity of mysticism among agents using mysticism or fundamentalism.

Reflectivism depends on alternative strategies, so to ensure its existence, we make the following assumption.

Assumption The fraction of fundamentalists $x_{2,t}$ never falls below some minimum $\delta_2 > 0$.

This assumption is not particularly restrictive, considering that in most asset pricing models, all investors are fundamentalists.

Given these three forecasting strategies (4), (5), and (6) and the asset pricing model allowing for heterogeneity (2), the realization of the asset price is

$$p_t = p_t^* + \alpha^{-t} n_t m_t. \tag{7}$$

One can verify that the reflective forecast has the same form as the realization of the price such that $e_{1,t} = E_t p_{t+1}$. The reflective forecast embodies the "beauty contest" characterization (Keynes, 1935) of asset markets in that agents use the martingale in their forecast only to the extent that other agents use it, not because they regard it as inherently important. See Parke and Waters (2014) for a detailed discussion.

Remark 1 If mysticism is present in the population, the efficient markets model is violated.

If the fraction n_t is greater than zero, then the extraneous martingale affects the asset price. Whether such an outbreak of agents adopting the mystic forecast is possible and quantitatively significant is a primary issue in the simulation results.

An evolutionary game theory mechanism describes how agents choose from the above forecasting strategies. Agents evaluate the performance of the forecasting strategies by comparing payoffs based on squared forecast errors. Hommes (2001) shows that the mean-variance optimization underpinning the model (2) is equivalent to minimizing squared forecast errors. Payoffs are defined as follows:

$$\pi_{i,t} = -(p_t - e_{i,t-1})^2 \tag{8}$$

The reflective forecast error U_t plays an important role in the payoffs to all three forecasting strategies, and is comprised of two terms.

$$U_t = (p_t^* - E_{t-1}(p_t^*)) + \alpha^{-t}(n_t m_t - n_{t-1} m_{t-1}) \tag{9}$$

The first term is the innovation to the current period dividend payment, which is the new fundamental information. The second term embodies the new information about the martingale's impact on the asset price.

In a model with a representative forecast, the fundamental and mystic forecasts are unbiased, but their forecast errors are affected by the level of the martingale in the presence of heterogeneity. A key term in the payoffs is the weighted martingale $A_{t-1} = \alpha^{-t} m_{t-1}$. The reflective forecast depends only on U_t and, using (8) and (9), has payoff

$$\pi_{1,t} = -U_t^2. \tag{10}$$

Fundamentalism has forecast error $U_t + n_{t-1} A_{t-1}$, so its payoff is

$$\pi_{2,t} = -U_t^2 - 2n_{t-1} U_t A_{t-1} - n_{t-1}^2 A_{t-1}^2. \tag{11}$$

Similarly, the payoff to mysticism is as follows:

$$\pi_{3,t} = -U_t^2 + 2(1 - n_{t-1}) U_t A_{t-1} - (1 - n_{t-1})^2 A_{t-1}^2 \tag{12}$$

In some studies with heterogeneous expectations, such as Brock and Hommes (1997), Evans and Ramey (1992), and Waters (2009), a fixed cost is often included in the payoffs to represent differences in the information content or forecasting sophistication of a strategy. In the present environment where all information is freely available, no such costs are introduced. Even though the reflective forecast incorporates more information, giving the other strategies and inherent advantage is not necessary for heterogeneity in the forecasting strategies to arise endogenously.

3 Intuition

Much of the intuition behind the potential for endogenous formation and collapse of rational bubbles and the implications for IE can be observed in the above three payoffs, assuming a reasonable evolutionary dynamic where agents switch to strategies with observed superior payoffs. The third terms in the payoffs to mysticism (12) and fundamentalism (11) are unambiguously damaging to those payoffs in comparison with the payoff to reflectivism (10). If there is heterogeneity in the choice of forecasting strategies ($0 < n_{t-1} < 1$), then mysticism and fundamentalism over- and underreact to the martingale. The unconditional expectation of the "covariance" term $U_t A_{t-1}$ is zero, so reflectivism outperforms the other two strategies.[5]

However, mysticism can outperform the other strategies in a given period. If the realization of the covariance $U_t A_{t-1}$ is positive and sufficiently large, the second term in (12) may outweigh the third term so that $\pi_{3,t} > \pi_{1,t} > \pi_{2,t}$. Such a positive covariance corresponds to a fortunate (for the mystic) correlation between

[5]The term UA is not a covariance strictly speaking. The word is used descriptively, since the term depends on the covariances between the shocks and the level of martingale.

the martingale and the innovations in the model. Similarly, fundamentalism could outperform reflectivism for a negative covariance $U_t A_{t-1}$ that outweighs the third term in the fundamentalist payoff (11).

In distribution, dividend innovations are uncorrelated with the martingale, but over a number of periods, such correlations are likely to occur. For mysticism to have a chance of success, the level of A_t must be large enough to that the covariance is significant, but not so large that the martingale terms dominate. Intuitively, a forecast like "Dow 36 thousand" might attract a significant following (as it did, see Glassman and Hassett, 1999), but an absurdly large forecast such as "Dow 36 billion" would be dismissed.

Some formal results clarify the implications for the endogeneity of bubbles and return predictability. Saying that expected excess returns are constant is an alternative way of saying that returns are not forecastable, as in Ohlson (1977) for one example. Interpreting the price p_t and the dividend d_t as logs, excess returns given by:

$$Z_t = p_t + d_t - \alpha^{-1} p_{t-1}. \tag{13}$$

If asset prices embody all available information, then only news should cause a change in the price, so informational efficiency IE is equivalent to $E_{t-1} Z_t = \bar{Z}$ for a constant $\bar{Z}$. If all agents are fundamentalists, this condition can be derived immediately.

In the present model with heterogeneity, excess returns are equivalent to the reflective forecast error up to a constant $\overline{Z}$ such that $Z_t = \overline{Z} + U_t$, given the underlying model based on mean-variance optimization. This point is verified using the price realization (7).

Given the reasonable assumptions for the information structure, the reflective forecast is unbiased and returns are unpredictable. The following proposition is not conclusive but does shed light on the underlying issues.

Proposition 2 *Given the following assumptions:*

i) the innovations to the dividend process d_t are iid,
ii) the innovations to the martingale η_t and n_t are uncorrelated,
iii) the martingale m_t and the change Δn_t are uncorrelated,
iv) the fraction n_t is unpredictable, $E_{t-1} \Delta n_t = 0$,
the conditional forecast of excess returns is constant, $E_{t-1} Z_t = \bar{Z}$.

Proof Using the expression for U_t (9), the term $E_{t-1} Z_t$ can be written as the sum of three expectations:

$$E_{t-1} Z_t = \bar{Z} + E_{t-1}(p_t^* - E_{t-1}(p_t^*)) + \alpha^{-t} E_{t-1} \left(n_t \eta_t \right) - \alpha^{-t} E_{t-1} \left(\Delta n_t m_t \right).$$

The four assumptions in Proposition 2 guarantee that the three expectations are zero, since the martingale innovation is mean zero $E_{t-1} \eta_t = 0$. Hence, $E_{t-1} Z_t = \bar{Z}$. ∎

The first two assumptions in Proposition 2 are innocuous as *i*) is satisfied for most specifications of dividends in the finance literature[6] and the martingale innovation in *ii*) is independently distributed. Assumptions *iii*) and *iv*) are plausible, but potentially unjustified. For any level of the martingale m_t, it is equally likely that the fraction n_t rises or falls, since it depends on the covariance $U_t A_{t-1}$. Furthermore, agents do not know the value of n_t when they make their forecast of p_{t+1}. However, if there is persistence in n_t, the inherent persistence in the martingale m_t could lead to correlations between the two.

Proposition 2 implies that the forecastability of returns depends on the forecastability of the populations' choices of forecasting strategies or how well people predict in Keynes' beauty contest. If agents are unable to forecast the fraction n_t, the model with mysticism satisfies IE but not the EMM. If there is a way to use the potential persistence in n_t, which depends on how agents update their choices of forecasting strategy, returns could be predictable. In practice, while there are many sources of extraneous data, forecasting strategy choices would be difficult, but in a model with three strategies, cannot be ruled out. Whether this phenomenon quantitatively impacts the IE of the market is a question to be addressed with the simulation exercises.

Mysticism cannot maintain a following indefinitely. Given the presence of reflectivists and the existence of a minimum fraction of fundamentalists δ_2, bubbles collapse endogenously. If fundamentalism could be eliminated from the population, then the fraction n_t is one, the payoff to mysticism (12) is identical to the payoff to reflectivism (10), and the model collapses to a representative agent rational bubble model. However, the presence of a minimum fraction of fundamentalists implies that $n_t < 1$ and that the reflective and mystic forecasts are not identical.

Proposition 3 *Given that n_t is fixed at its maximum $1 - \delta_2$, for $\delta_2 > 0$, and assumptions i) and ii) from Proposition 2, the expectation of the reflective payoff is strictly greater than the mystic payoffs, i.e., $E_{t-1}\pi_{1,t} > E_{t-1}\pi_{3,t}$, with probability one.*

Proof The difference in the payoffs (10) and (12) is

$$\pi_{1,t} - \pi_{3,t} = -2(1 - n_{t-1})U_t A_{t-1} + (1 - n_{t-1})^2 A_{t-1}^2$$

Given a constant n_t and $\Delta n_t = 0$, then $E_{t-1} U_t A_t = E_{t-1}\left[(p_t^* - E_{t-1}(p_t^*))\alpha^{-t} m_t\right] + \alpha^{-2t} E_{t-1}\left(n_t \eta_t m_t\right)$ using the expression for U_t in Eq. (9) and $A_t = \alpha^{-t} m_{t-1}$ by definition. Since the innovations d_t and η_t are mean zero and uncorrelated with m_t, the expectation $E_{t-1} U_t A_t = 0$. Hence, the unconditional expectation

$$E_{t-1}\left(\pi_{1,t} - \pi_{3,t}\right) = (1 - n_{t-1})^2 A_{t-1}^2.$$

[6]For model with drift, dividends would be defined as deviations from the deterministic model.

Since $\delta_2 > 0$, so is $1 - n_{t-1}$, so the right-hand side is positive as long as $A_{t-1} \neq 0$, which only occurs if $m_t = 0$. This condition is met with probability 1, as is $E\pi_{1,t} > E\pi_{3,t}$. ■

Mysticism cannot dominate indefinitely. Since the expected value of the term $U_t A_{t-1}$ in (12) is zero, reflectivism outperforms mysticism in the long run. Further, the magnitude of A_t (a submartingale) grows over time, so the third term in the payoff to mysticism (11) dominates and the performance of mysticism deteriorates over time. While mysticism can gain a following temporarily, whereby the martingale affects the asset price, eventually agents abandon mysticism in favor of reflectivism, so bubbles endogenously form and collapse. The goal of the simulations is to determine the quantitative importance of such outbreaks of mysticism.

Since it limits the life of bubbles, the minimum fraction of fundamentalists plays a similar role as the *projection facility* used with least squares learning as in Adam et al. (2016). Similarly, Lansing (2010) limits parameters so that agents focus on the one bubble of a continuum of solutions, that leads to stationarity in the first difference of the endogenous variable being forecast. In these models, a representative agent updates the estimate of the parameters in a forecasting rule, but the projection facility limits the acceptable estimates. In the present model, a small fraction of agents rejects extraneous information.

The present model represents a minimal departure from rationality when mystics are introduced into the population. Mysticism appears due to a disagreement about what constitutes fundamental information, but all agents form expectations with a reasonable economic model, i.e., agents meet the *cognitive consistency* principle described in Evans and Honkapohja (2011). Further, both mysticism and fundamentalism satisfy rationality in the homogeneous case, and reflectivism satisfies rationality when there is heterogeneity in the forecasting strategies, and this forecasting strategy is available to agents at all times. When mystics are eliminated from the population, the reflective and fundamental forecasts coincide. Only when mystics are introduced do the mystic and fundamental forecasts deviate from rationality, but mystics believe that the extraneous information in the martingale is relevant to the forecast of the asset price, and that other agents will eventually realize it. All agents believe that they are making efficient use of the available information.

4 Evolutionary Dynamics

A generalization of the replicator dynamic, a workhorse in the evolutionary game theory literature, describes the evolution of the vector x_t of the fractions of agents using the different forecasting strategies. This dynamic allows for the parameterization of agents' aggressiveness in switching to better performing strategies, which is a key determinant for the potential adoption of mysticism. This section also discusses the necessary conditions for the resulting emergence of rational bubbles.

Let the weighting function $w(\pi)$ be a positive, increasing function of the payoffs. The general replicator dynamic is

$$x_{i,t+1} - x_{i,t} = x_{i,t} \frac{w\left(\pi_{i,t}\right) - \overline{w}_t}{\overline{w}_t}, \tag{14}$$

where the expression $\overline{w}_t$ is the weighted population average $\overline{w}_t = x_{1,t} w\left(\pi_{1,t}\right) + \cdots + x_{n,t} w\left(\pi_{n,t}\right)$ and $\sum_{i=1}^{n} x_{i.t} = 1$. A strategy gains followers if its weighted payoff above the weighted population average, i.e., has positive fitness, in the language of evolutionary game theory. Such a dynamic is said to be *imitative* since strategies that are popular today, larger $x_{i,t}$, tend to gain more adherents if they are successful. Such dynamics have the potential to impart persistence to x_t, so assumptions *iii*) and *iv*) in Proposition 2 could be violated, which motivates the simulation exercises.

A general form for the dynamic (14) allows for a range of behavior of the agents. Compared to a linear weighting function $w(\pi)$, under a convex $w(\pi)$, agents switch to better performing strategies more quickly, see Hofbauer and Weibull (1996). A linear weighting function in the dynamic (14) corresponds to the special case of the replicator dynamic studied in Weibull (1998) and Samuelson (1997). Sandholm (2011) gives a thorough comparison of the features of a number of evolutionary dynamics. Waters (2009) discusses discrete time dynamics used in macroeconomic applications.

Using a version of the dynamic (14) with an alternate timing, Parke and Waters (2014) demonstrate that, for bounded dividends, the payoff to reflectivism is always above the population average.[7] Therefore, under the replicator (linear $w(\pi)$), mysticism cannot take followers away from reflectivism. Under linear weighting, the covariance (second) terms in the payoffs to mysticism and fundamentalism, (12) and (11), cancel in the population average payoff,[8] but the third terms with A_{t-1}^2 do not. Since the payoff to reflectivism is unaffected by the martingale, it is larger than the population average, so reflectivism gains followers.

Reflectivism's dominance is weaker in the case of a convex weighting function. Here, a positive covariance term $U_t A_{t-1} > 0$ has greater benefit to mysticism than harm to fundamentalism, so it enters the population average payoff and, if it is large enough, mysticism can gain a following. The model used for simulations focuses on the exponential weighting function:

$$w(\pi) = e^{\theta^2 \pi}, \tag{15}$$

so θ parameterizes the aggressiveness of the agents. An increase in θ means that agents are switching more quickly to the best strategy, but as θ decreases the dynamic approaches the linear weighting case.

[7] The timing for the present work is chosen to avoid complications in the tests for return predictability.

[8] Given the timing of the present version of the model, the covariance terms may not cancel out to zero, but their impact is minimal.

One drawback to imitative dynamics such as the generalized replicator (14) is their lack of inventiveness, see Waters (2009) for a discussion. If a strategy has no followers $(x_i = 0)$, then it cannot gain any. Hence, game theorists usually focus on equilibria that are evolutionarily stable, meaning that they are robust to the introduction of a small fraction of deviating agents. Similarly, the focus of the present class of models is whether the fundamental forecast is robust to the introduction of a small fraction of mystics.[9]

It is possible for mysticism to gain a following given the following conditions. (1) Some agents believe that extraneous information may be important to the value of an asset. (2) In some periods, the extraneous martingale is correlated with fundamentals. (3) Agents must be sufficiently aggressive in switching to superior performing strategies. See Parke and Waters (2014) for a formal analysis.

5 Simulations

Simulation results show that the model satisfies IE but not the EMM. For some parameter choices, the data from the simulations is well represented by the EMM, but if agents are sufficiently aggressive about switching to better performing strategies and shocks to the martingale are of a similar magnitude to the dividend shocks, the data shows bubble-like behavior including excess variance in the asset price. However, returns are not significantly predictable under any circumstances.

The underlying dividend process is calibrated to the annual S&P500 data for the sample 1871–2013 used by Shiller (2005) (updated here), using earnings as a proxy for dividends. Since not all firms pay dividends, earnings are a better measure of market fundamentals. Given the dividend d_t and the martingale m_t, the model is determined by the dynamic (14) along with the exponential weighting function (15), the payoffs (10), (11), and (12), and the realization of the asset price (7). The dividend process is specified as a stationary process with parameter choices below[10]:

$$d_t = \overline{d} + \rho_d \left(d_{t-1} - \overline{d}\right) + v_t \tag{16}$$

$\overline{d}$	ρ_d	σ_v
0.1166	0.465	0.203

The constant $\overline{d}$ is chosen so that for $\alpha = 0.95$, the steady-state price–dividend ratio (log difference) is 2.66, which is close to the long-run average for the S&P500

[9]A referee notes that such minima could be interpreted as arising from hedge funds with fixed strategies and the leverage to support their approach.

[10]LeBaron et al. (1999) and Branch and Evans (2011) use stationary dividends. Adam et al. (2016) and Lansing (2010) both model dividends as a random walk with drift, which would complicate the present model and is left as a possibility for future work.

from the Shiller data. The persistence parameter ρ_d and shocks $v_t \sim N(0, \sigma_v)$ are chosen to match values from the H-P detrended earnings series. They are also close to the linearly detrended series for postwar sample 1945–2013.

The free parameter θ, which measures agent aggressiveness, and σ_η, the standard deviation of the martingale innovations, play are large role in determining the potential for outbreaks of mysticism and bubbles. For such events to occur, agents must be sufficiently aggressive, meaning that θ is sufficiently large, and the magnitude of the martingale innovations must be large enough to have a noticeable impact on the payoffs and the asset price, but not so large so that the third term in the payoff to mysticism (12) dominates.

Figures 1 and 2 demonstrate the effects of varying these two parameters. The simulations are initialized where the fraction of followers of fundamentalism and mysticism are at their minima. Furthermore, the minimum fraction of mysticism 0.001 is much smaller than the minimum fraction of fundamentalism 0.1, so the

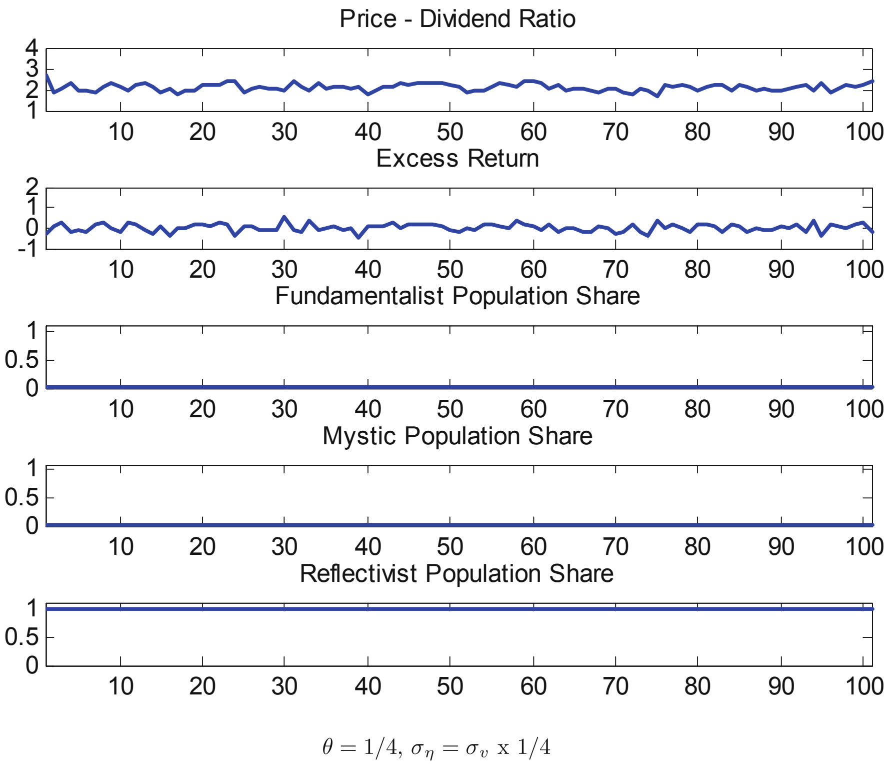
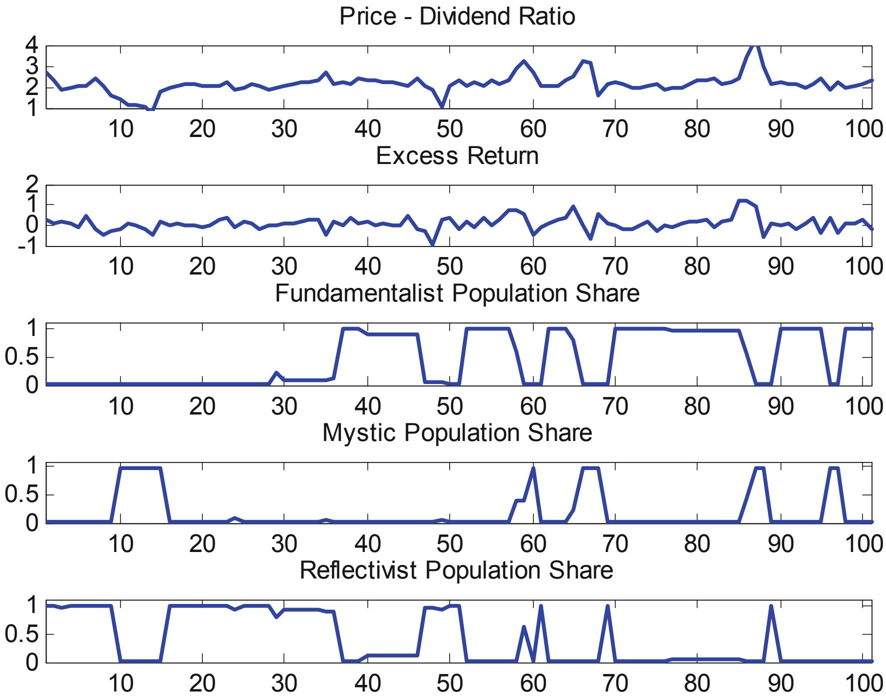


Fig. 1 The price-dividend ratio $p_t - d_t$, the excess returns Z_t and the fraction of reflectivists $x_{1,t}$, fundamentalists $x_{2,t}$ and mystics $x_{3,t}$ are simulations of the model where the aggressiveness parameter in Eq. (15) is $\theta = 1/4$ and the standard deviation of the martingale innovations described in Proposition 2 are $\sigma_\eta = \sigma_v/4$, where the the term δ_v is the standard deviation of the dividend innovations in Eq. (16)



$\theta = 10.0$, $\sigma_\eta = \sigma_\upsilon$ x 2.0

Fig. 2 The price-dividend ratio $p_t - d_t$, the excess returns Z_t and the fraction of reflectivists $x_{1,t}$, fundamentalists $x_{2,t}$ and mystics $x_{3,t}$ are simulations of the model where the aggressiveness parameter in Eq. (15) is $\theta = 1/4$ and the standard deviation of the martingale innovations described in Proposition 2 are $\sigma_\eta = \sigma_\upsilon/4$, where the the term σ_υ is the standard deviation of the dividend innovations in Eq. (16)

fraction n_t is initially small, and the effect of mysticism on the asset price is minimal. If the fraction of followers of mysticism $x_{3,t}$ falls below its minimum, it is reset to the minimum and the martingale m_t is restarted at zero.

In Fig. 1, both parameters θ and σ_η are set to low values. Here, mysticism never gains a following, and the log price–dividend ratio and the excess returns are determined solely by the fundamental price p_t^*, which is determined by the exogenous dividend process under the EMM. However, for the larger values of these parameters in the simulation of Fig. 2, there are occasional outbreaks where mysticism gains a following, and the price–dividend ratio deviates significantly from its steady-state value. Note that the log $(p_t - d_t)$ exceeding 3.23 is equivalent to the level of the price dividend ratio doubling its steady-state value. The endogenous formation and collapse of the mystic bubbles and their effect on the asset price are evident in Fig. 2.

5.1 Excess Variance

Studies such as Shiller (1981) demonstrate that asset prices fluctuate more than predicted by the EMM, and endogenous rational bubbles can explain such excess variance. Simulations determine a ratio of the realized variance and the predicted variance based on the variance of the dividends and the EMM embodied in (3). A statistical test of the variance of the price–dividend ratio provides more definitive evidence.

In the absence of mysticism ($n_t = 0$), the asset price behaves according to the efficient markets model and depends only on the dividend process.

$$p_t^* = \overline{d}\left(\frac{\alpha}{1-\alpha} - \frac{\alpha\rho_d}{1-\alpha\rho_d}\right) + d_t\left(\frac{1}{1-\alpha\rho_d}\right) \tag{17}$$

Hence, the variance of the asset price determined by fundamentals is $\sigma_{p^*}^2 = \left(1-\alpha\rho_d\right)^{-2}\sigma_d^2$.

The tables report the results of tests for excess variance and return predictability on simulated data for varying choices of the parameters θ and σ_η, the speed with which agents switch forecasts and the magnitude of the shocks to the martingale, respectively. For all the tests, the simulations are initialized with 50 periods followed by a run of 100, which is similar to the samples used for calibration. The tables report the results for 10,000 runs.

For example, Table 1 reports the average ratio $\sigma_p^2/\sigma_{p^*}^2$ of the variance of the simulated asset prices σ_p^2 and the variance $\sigma_{p^*}^2$ predicted under the EMM, using the sample variance $\widehat{\sigma}_d^2$ of the simulated dividends. The top-left cell of Table 1 shows that for $\theta = \frac{5}{8}$ and $\sigma_\eta = 0.0254 = \frac{1}{8}\cdot\sigma_v$ (see the values under (16)), the ratio is $\sigma_p^2/\sigma_{p^*}^2 = 1.009$, a value close to unity, which corresponds to the EMM and is typical of low levels of θ and σ_η. For higher levels, the ratio rises above one, and, for some pairs of parameter choices, over two. This level is smaller than Shiller's

Table 1 Each cell shows the average over 10,000 runs (each with 100 periods) of the ratio $\sigma_p^2/\sigma_{p^*}^2$ of the variance of p_t to the variance of the fundamental price p_t^* for each choice of θ and σ_η

		$\sigma_\eta = \sigma_v$ x						
		1/8	1/4	1/2	1	2	4	8
θ	5/8	1.009	1.016	1.025	1.034	1.043	1.0642	1.1689
	5/4	1.010	1.016	1.026	1.035	1.041	1.0583	1.148
	5/2	1.046	1.074	1.302	1.679	1.701	1.4557	1.2933
	5	1.713	2.118	2.413	1.904	1.420	1.1937	1.1356
	10	1.110	1.263	1.501	1.599	1.531	1.5396	1.8083
	20	1.014	1.043	1.116	1.221	1.320	1.5088	1.9944
	40	1.003	1.012	1.039	1.091	1.169	1.3216	1.6861

Table 2 Each cell shows the fraction of runs with significant (at the 5% level) excess variance for each choice of θ and σ_η

		$\sigma_\eta = \sigma_\nu$ x						
		1/8	1/4	1/2	1	2	4	8
θ	5/8	0.000	0.000	0.000	0.000	0.000	0	0.195
	5/4	0.000	0.000	0.000	0.000	0.000	0.0002	0.1084
	5/2	0.004	0.009	0.025	0.059	0.088	0.087	0.0829
	5	0.145	0.248	0.326	0.346	0.269	0.1798	0.1269
	10	0.081	0.283	0.587	0.762	0.751	0.743	0.7496
	20	0.000	0.004	0.070	0.338	0.615	0.7922	0.8853
	40	0.000	0.000	0.001	0.026	0.222	0.5558	0.7803

(1981) initial estimate, but other research[11] has found smaller estimated values but still demonstrates excess variance.

Table 2 reports results on a test of the significance of the excess variance.[12] Given normal errors in the dividend innovations v_t for the EMM, the ratio σ_p^2/σ_{p*}^2 has the distribution $\chi^2(n)/n$ where n is the number of periods. Table 2 shows the fraction of runs that exhibit excess variance beyond a 5% significance level. For small shocks to the martingale, represented by σ_η, and low levels of aggressiveness, shown by θ, there is little or no excess variance, but for moderate to high values, large fractions, often well over half, show excess variance, corresponding to mystic outbreaks that lead to bubble-like behavior. The pattern in these results on excess variance qualitatively matches the occurrence of excess kurtosis and GARCH effects found in Parke and Waters (2007). Mysticism can produce significant deviations from the predictions of the EMM.

5.2 *Return Predictability*

Proper tests give evidence that returns are not forecastable, so IE holds, a necessary but not sufficient condition of the efficient markets model. Excess returns Z_t are given by Eq. (13) which is the reflective forecast error (9) plus a constant. Proposition 2 shows that expected excess returns are constant if the fractions of followers of the different strategies are unpredictable. Whether those fractions impart persistence on the excess returns is a question for the simulation exercises.

[11] Some examples are LeRoy and Porter (1981), Campbell and Shiller (1989), and LeRoy and Parke (1992). The issue is complicated since some of these models account for a time-varying interest rate or discount factor.

[12] For all the tests, the simulations are initialized with 50 periods followed by a run of 100, which is similar to the samples used for calibration. The table report the results over 10,000 runs.

To test predictability, the following equation is estimated to test whether lagged data contains information about current returns Z_t, similar to those used in Fama and French (1989), among many others.

$$Z_t = \beta_0 + \beta_1 x_{t-1} + \varepsilon_t, \tag{18}$$

The error term ε_t has standard deviation σ_ε, and there are multiple candidates for the predictor variable x_t. Results are reported for the price–dividend ratio, a common choice for the predictor, and dividends, which is particularly relevant in the present context. A standard approach is to test the null $\widehat{\beta_1} = 0$ and least squares estimates on market data are often significantly different than zero, though the economic significance is often questionable (Cochrane, 2008). Least squares estimates of $\widehat{\beta_1}$ on the simulated data show predictability in many cases, roughly following the pattern in Tables 1 and 2, meaning that for choices of the parameters θ and σ_η that have frequent occurrences of excess variance also have tend to have estimates of $\widehat{\beta_1}$ that are significantly different than zero. However, the persistence in both choices for the predictor variable means that standard least squares t statistics are biased away from zero.

Campbell and Yogo (2006) discuss this issue in detail and develop a consistent statistic Q where the predictor variable x_t could be persistent such that

$$x_t = \chi + \rho x_{t-1} + \xi_t. \tag{19}$$

The covariance $\sigma_{\xi\varepsilon}$ between the innovations ζ_t and ε_t from (18) is used to define the following parameters:

$$\gamma_{\xi\varepsilon} = \frac{\sigma_{\xi\varepsilon}}{\sigma_\xi^2};\ \delta = \frac{\sigma_{\xi\varepsilon}}{\sigma_\varepsilon \sigma_\xi}.$$

The test statistic Q for the null $\widehat{\beta_1} = 0$ is as follows:

$$Q = \frac{\widehat{\beta_1} - \gamma_{\xi\varepsilon}\left(\widehat{\rho} - \rho\right)}{\sigma_\varepsilon\left(1-\delta^2\right)^{\frac{1}{2}}\left(\sum_{t=1}^{T}\widehat{x}_t^2\right)^{-\frac{1}{2}}} \tag{20}$$

The value of $\widehat{\rho}$ is the least squares estimate of the true predictor persistence parameter ρ in (19). Campbell and Yogo (2006) do not have knowledge of ρ for their applied work, and they develop a method for estimating Bonferroni bounds for this parameter, but that is unnecessary with simulated data.

Note that if the persistence parameter equals the true value $\widehat{\rho} = \rho$ and errors v_t and ε_t are uncorrelated so that $\delta = 0$, the Q statistic collapses to the standard least squares t statistic. Lewellen (2004) refers to the $\gamma_{ev}\left(\widehat{\rho} - \rho\right)$ term as the "finite sample bias" correction, see the discussion in Campbell and Yogo (2006).

Table 3 Using d_t as the predictor variable x_t, each cell shows the fraction of runs where the estimate of $\widehat{\beta}_1$ (from (18)) is significantly different than zero at the 5% level, according to the test statistic Q (20), for each choice of θ and σ_η

		$\sigma_\eta = \sigma_\nu$ x						
		1/8	1/4	1/2	1	2	4	8
θ	5/8	0.048	0.048	0.051	0.049	0.049	0.043	0.038
	5/4	0.048	0.049	0.050	0.049	0.046	0.047	0.043
	5/2	0.046	0.051	0.046	0.052	0.046	0.043	0.051
	5	0.042	0.047	0.044	0.042	0.043	0.050	0.054
	10	0.053	0.042	0.044	0.041	0.050	0.052	0.049
	20	0.051	0.052	0.049	0.048	0.042	0.040	0.026
	40	0.050	0.046	0.050	0.053	0.043	0.038	0.024

Table 4 Using $p_t - d_t$ as the predictor variable x_t, each cell shows the fraction of runs where the estimate of $\widehat{\beta}_1$ (from (18)) is significantly different than zero at the 5% level, according to the test statistic Q (20), for each choice of θ and σ_η

		$\sigma_\eta = \sigma_\nu$ x						
		1/8	1/4	1/2	1	2	4	8
θ	5/8	0.054	0.047	0.043	0.047	0.042	0.041	0.030
	5/4	0.050	0.051	0.048	0.052	0.041	0.042	0.034
	5/2	0.049	0.052	0.054	0.060	0.063	0.066	0.055
	5	0.064	0.086	0.095	0.093	0.089	0.068	0.051
	10	0.063	0.074	0.094	0.112	0.073	0.039	0.024
	20	0.050	0.054	0.064	0.058	0.036	0.026	0.020
	40	0.048	0.049	0.048	0.045	0.033	0.026	0.021

These considerations mean that dividends d_t is a natural choice for the predictor variable, since the true value of the persistence parameter is known from the specification (16) so that $\rho = \rho_d$. The most common choice for the predictor variable in applied work is the price–dividend ratio. Using the price–dividend ratio $x_t = p_t - d_t$, the maintained hypothesis that the EMM holds implies that the persistence parameter ρ in (19) is the same as the case with dividends so that $\rho = \rho_d$. So, the chosen value of ρ_d is the appropriate choice for ρ in the Q statistic (20) for both choices of the predictor.

Tables 3 and 4 report the percentage of runs that show predictability of returns with a significance of 5%. Critical values for Q are determined by a Monte Carlo simulation of 10^6 runs of the model without mysticism, meaning the martingale m_t is fixed at zero. Table 3 shows the results using dividends as a predictor, and the fractions are all close to 0.05, the significance level of the test. Therefore, excess returns are not predictable at all. Table 4 shows analogous results using the price–dividend ratio. There is some deviation from 0.05, but the largest fraction is 0.112 which is not nearly as large as 0.762, the fraction showing significant excess variance for that choice of θ and σ_η. In cases where excess variance occurs more than would be predicted under the EMM, the fraction of runs with predictable returns is far smaller.

The model with mysticism can produce excess variance in returns but returns are not predictable. There are clear deviations from the EMM, but they rarely appear in the tests for IE. Therefore, excess variance does not necessarily correspond to predictable returns. Hence, tests for predictability of returns have very little power to reject the deviations from the EMH caused by mystic bubbles.

5.3 Excess Variance Robustness

The robustness of the results is examined using alternative tests for excess variance and a different parameterization. In the spirit of the construction of the Q-statistic, it might be more appropriate to use the least squares estimate for the persistence parameter $\widehat{\rho}$. Table 5 reports the results about the fraction of runs showing excess variance when the variance of the fundamental price $\sigma_{p*}^2 = \left(1 - \alpha\rho_d\right)^{-2} \sigma_d^2$ is estimated using $\dfrac{\sigma_\varepsilon^2}{1 - \widehat{\rho}^2}$ for the variance of the dividends σ_d^2. This test statistic has the proper level of type 1 error in cases without mysticism, and shows the same pattern, in that there are many cases with excess variance for parameter values that allow for mystic outbreaks.

In a study of the yield curve, Flavin (1983) demonstrates that excess variance results can arise due to small sample bias. For the present model, tests on longer samples only strengthen the evidence for excess variance, as was found in the applied work of Shiller (1990).

Tables 6, 7, 8, 9, and 10 show the same results for a different parameterization of the model, and the conclusions are unchanged. Using a linear trend for the earnings series over the full sample 1871–2013 gives the calibrated values, $\rho = 0.69$ and

Table 5 Each cell shows the fraction of runs with significant (at the 5% level) excess variance for each choice of θ and σ_η, using an estimated value of $\widehat{\rho}$ to compute σ_{p*}^2

		$\sigma_\eta = \sigma_\nu$ x						
		1/8	1/4	1/2	1	2	4	8
θ	5/8	0.044	0.049	0.063	0.081	0.114	0.2498	0.6874
	5/4	0.046	0.052	0.055	0.077	0.118	0.2253	0.6005
	5/2	0.046	0.046	0.056	0.088	0.146	0.1964	0.2601
	5	0.071	0.106	0.144	0.175	0.234	0.2607	0.2167
	10	0.039	0.041	0.098	0.406	0.759	0.8369	0.8527
	20	0.043	0.058	0.114	0.435	0.804	0.9169	0.9596
	40	0.044	0.051	0.096	0.281	0.611	0.8159	0.9206

Table 6 Each cell shows the average over 10,000 runs (each with 100 periods) of the ratio $\sigma_p^2/\sigma_{p^*}^2$ of the variance of p_t to the variance of the fundamental price p_t^* for each choice of θ and σ_η. Tables 6, 7, 8, 9, and 10 have results for the same exercises in Tables 1, 2, 3, 4, and 5 for an alternative parameterization of the dividends

		$\sigma_\eta = \sigma_\nu$ x						
		1/8	1/4	1/2	1	2	4	8
θ	5/8	1.006	1.009	1.012	1.015	1.017	1.0206	1.0458
	5/4	1.006	1.009	1.013	1.016	1.016	1.0217	1.0517
	5/2	1.025	1.075	1.184	1.341	1.277	1.1169	1.0594
	5	1.196	1.280	1.321	1.257	1.136	1.068	1.0533
	10	1.034	1.080	1.143	1.171	1.158	1.1632	1.246
	20	1.006	1.015	1.035	1.062	1.086	1.1302	1.249
	40	1.001	1.004	1.011	1.025	1.043	1.079	1.1672

Table 7 Each cell shows the fraction of runs with significant (at the 5% level) excess variance for each choice of θ and σ_η

		$\sigma_\eta = \sigma_\nu$ x						
		1/8	1/4	1/2	1	2	4	8
θ	5/8	0.046	0.051	0.055	0.057	0.059	0.0667	0.0944
	5/4	0.046	0.055	0.057	0.056	0.061	0.0609	0.0796
	5/2	0.062	0.082	0.111	0.173	0.188	0.1519	0.1079
	5	0.204	0.280	0.343	0.372	0.269	0.1697	0.1384
	10	0.076	0.150	0.291	0.369	0.341	0.3629	0.4951
	20	0.047	0.050	0.073	0.122	0.170	0.2752	0.5233
	40	0.040	0.047	0.057	0.067	0.092	0.1622	0.3664

Table 8 Using d_t as the predictor variable x_t, each cell shows the fraction of runs where the estimate of $\widehat{\beta}_1$ (from (18)) is significantly different than zero at the 5% level, according to the test statistic Q (20), for each choice of θ and σ_η

		$\sigma_\eta = \sigma_\nu$ x						
		1/8	1/4	1/2	1	2	4	8
θ	5/8	0.052	0.052	0.050	0.046	0.046	0.0447	0.0432
	5/4	0.047	0.046	0.053	0.046	0.049	0.0483	0.0412
	5/2	0.053	0.046	0.052	0.052	0.044	0.0483	0.0445
	5	0.050	0.046	0.046	0.042	0.046	0.0468	0.0505
	10	0.048	0.053	0.047	0.051	0.057	0.0495	0.0405
	20	0.051	0.048	0.049	0.051	0.051	0.0412	0.0291
	40	0.046	0.046	0.052	0.051	0.047	0.0417	0.0361

Table 9 Using $p_t - d_t$ as the predictor variable x_t, each cell shows the fraction of runs where the estimate of $\widehat{\beta}_1$ (from (18)) is significantly different than zero at the 5% level, according to the test statistic Q (20), for each choice of θ and σ_η

		$\sigma_\eta = \sigma_\nu$ x						
		1/8	1/4	1/2	1	2	4	8
θ	5/8	0.045	0.048	0.046	0.046	0.043	0.0415	0.0357
	5/4	0.050	0.044	0.050	0.046	0.046	0.0428	0.0389
	5/2	0.052	0.050	0.055	0.063	0.063	0.0719	0.0552
	5	0.058	0.066	0.073	0.073	0.077	0.0624	0.0478
	10	0.052	0.057	0.068	0.075	0.066	0.0393	0.0238
	20	0.048	0.045	0.051	0.052	0.038	0.029	0.0197
	40	0.050	0.047	0.050	0.043	0.039	0.0347	0.0227

Table 10 Each cell shows the fraction of runs with significant (at the 5% level) excess variance for each choice of θ and σ_η, using and estimated value of $\widehat{\rho}$ to compute $\sigma^2_{p^*}$

		$\sigma_\eta = \sigma_\nu$ x						
		1/8	1/4	1/2	1	2	4	8
θ	5/8	0.046	0.051	0.055	0.057	0.059	0.0667	0.0944
	5/4	0.046	0.055	0.057	0.056	0.061	0.0609	0.0796
	5/2	0.062	0.082	0.111	0.173	0.188	0.1519	0.1079
	5	0.204	0.280	0.343	0.372	0.269	0.1697	0.1384
	10	0.076	0.150	0.291	0.369	0.341	0.3629	0.4951
	20	0.047	0.050	0.073	0.122	0.170	0.2752	0.5233
	40	0.040	0.047	0.057	0.067	0.092	0.1622	0.3664

$\sigma_v = 0.228$ for the detrended series d_t given in Eq. (16). Simulations using these parameters are reported in Tables 5, 6, 7, 8, 9, and 10. The results for the excess variance are not as dramatic, for only one choice of θ and σ_η do more than half of the runs show significant excess variance, for example. However, the overall message is the same. Excess variance is a common occurrence, while return predictability is not. In fact, using dividends as the predictor, there is no predictability beyond that at the expected level of type 1 error (Table 8).

6 Conclusion

The model of mysticism is a specific example that satisfies informational efficiency but does not conform to the efficient markets model. Returns are unpredictable, but asset prices depend on more that fundamentals. Proper tests of return predictability have very little power to reject mystic bubbles.

The information structure of the model satisfies the cognitive consistency principle, in that agents' forecasts are based on reasonable economic models. Mystic bubbles form endogenously due to fortuitous correlations between extraneous and fundamental data, but they cannot last since the reflectivist forecast, which takes into account the behavior of the other agents, and outperforms the mystic forecast in the long run. The model explains multiple stylized facts about asset markets such as excess variance and GARCH effects.

One might argue that mystic bubbles are not plausible, but heterogeneous forecasts are an observed fact. Since agents do not adopt mysticism indefinitely, any transversality condition argument does not apply. Further, there are many candidates for extraneous information represented by the martingale such as exchange rates, commodity prices, or "expert" forecasts. The interpretation of the model indicates that the only requirement for such data to have an impact is that the innovation in the data be roughly the magnitude of those of the dividends.

The weak power of the tests of return predictability shows the limitations of the interpretations of all such empirical results. The inability to reject unpredictability is not evidence against the presence of extraneous information nor excess variance in the asset price. The exercise could be conducted in a more sophisticated environment, explicitly modeling trends or including behavioral forecasting strategies, for example. However, if the interpretation of tests for return predictability is problematic in a simple environment, they will not be more meaningful with added complications.

References

Abou, A., & Prat, G. (2000). Modelling stock price expectations: Lessons from microdata. In F. Gardes & G. Prat (Eds.), *Price expectations in goods and financial markets* (pp. 313–46). Cheltenham: Edward Elgar.

Adam, K., Marcet, A., & Niccolini, J. P. (2016). Stock market volatility and learning. *Journal of Finance, 71*, 33–82.

Blanchard, O. (1979). Backward and forward solutions for economies with rational expectations. *American Economic Review, 69*(2), 114–118.

Branch, W., & Evans, G. (2007). Model uncertainty and endogenous volatility. *Review of Economic Dynamics, 10*, 207–237.

Branch, W., & Evans, G. (2011). Learning about risk and return: A simple model of bubbles and crashes. *American Economic Review: Macroeconomics, 3*(1), 159–191.

Brock, W. A., & Hommes, C. H. (1997). A rational route to randomness. *Econometrica, 65*(5), 1059–1095.

Brock, W. A., & Hommes, C. H. (1998). Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic Dynamics and Control, 22*, 1235–1274.

Campbell, J. Y., & Shiller, R. J. (1989). The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies 1*(3), 195–228.

Campbell, J. Y., & Yogo, M. (2006). Efficient tests of stock return predictability. *Journal of Financial Economics, 81*(1), 27–60.

Cochrane, J. (2008). A dog that did not bark: a defense of return predictability. *Review of Financial Studies, 21*(4), 1533–1575.

DeLong, B., Schleifer, A., Summers, L. H., & Waldmann, R. J. (1991). The survival of noise traders in financial markets. *Journal of Business, 64*(1), 1–20.

Evans, G. (1991). Pitfalls in testing for explosive bubbles in asset prices. *American Economic Review, 81*(4), 922–930.

Evans, G., & Honkapohja, S. (2011). Learning as a rational foundation for macroeconomics and finance, Bank of Finland Research Discussion Paper No. 8/2011.

Evans, G., & Ramey, G. (1992). Expectation calculation and macroeconomic dynamics. *American Economics Review, 82*(1), 207–224.

Fama, E. F., & French, K. R. (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics, 25*(1), 23–49.

Flavin, M. (1983). Excess volatility in the financial market: A reassessment of the empirical evidence. *Journal of Political Economy, 91*(6), 929–956.

Glassman, J., & Hassett, K. (1999). *DOW 36,000*. New York, NY: Three Rivers Press.

Hofbauer, J., & Weibull, J. (1996). Evolutionary selection against dominated strategies. *Journal of Economic Theory, 71*, 558–573.

Hommes, C. H. (2001). Financial markets as nonlinear adaptive evolutionary systems. *Quantitative Finance, 1*, 149–167.

Hommes, C. H. (2006). Heterogeneous agent models in economics and finance. In L. Tesfatsion & K. I. Judd (Eds.), *Handbook of computational economics*. Agent-based computational economics, vol. 2. Amsterdam: Elsevier Science BV.

Keynes, J. M. (1935). *The general theory of employment, interest and money*. New York, NY: Harcourt Brace.

Lansing, K. (2010). Rational and near-rational bubbles without drift. *Economic Journal, 120*, 1149–1174.

LeBaron, B. (2012). Heterogeneous gain learning and the dynamics of asset prices. *Journal of Economic Behavior and Organization, 83*(3), 424–445.

LeBaron, B., Arthur, B., & Palmer, R. (1999). Time series properties of an artificial stock market. *Journal of Economic Dynamics and Control, 23*, 1487–1516.

LeRoy, S. F., & Parke, W. R. (1992). Stock price volatility: Tests based on the geometric random walk. *American Economic Review, 82*(4), 981–992.

LeRoy, S. F., & Porter, R. D. (1981). The present value relation: Tests based on implied variance bounds. *Econometrica, 49*(3), 555–574.

Lewellen, J. (2004). Predicting returns with financial ratios. *Journal of Financial Economics, 74*(1), 209–235.

Lundqvist, L., & Sargent, T. J. (2004). *Recursive macroeconomics theory*. Cambridge, MA: The MIT Press.

Ohlson, J. A. (1977). Risk-aversion and the martingale property of stock prices: Comments. *International Economic Review, 18*(1), 229–235.

Parke, W. R., & Waters, G. A. (2007). An evolutionary game theory explanation of arch effects. *Journal of Economic Dynamics and Control, 31*(7), 2234–2262.

Parke, W. R., & Waters, G. A. (2014). On the evolutionary stability of rational expectations. *Macroeconomic Dynamics, 18*(7), 1581–1606.

Prat, G., & Uctum, R. (2011). Modelling oil price expectations: Evidence from survey data. *The Quarterly Review of Economics and Finance, 51*(3), 236–247.

Prat G., & Uctum, R. (2015). Expectation formation in the foreign exchange market : A time-varying heterogeneity approach using survey data. *Applied Economics, 47*(34–35), 3673–95.

Samuelson, L. (1997). *Evolutionary games and equilibrium selection*. Cambridge, MA: The MIT Press.

Sandholm, W. (2011). Population games and evolutionary dynamics. Cambridge, MA: The MIT Press.

Shiller, R. J. (1981). Do stock prices move too much to be justified to expected changes in dividends? *American Economics Review, 71*(3), 421–426.

Shiller, R. J. (1990). The term structure of interest rates. In B. M. Friedman & F. H. Hahn (Eds.), *The handbook of monetary economics*. Amsterdam: Elsevier Science BV.

Shiller, R. J. (2005). *Irrational exuberance*. Princeton, NJ: Princeton University Press.

Waters, G. A. (2009). Chaos in the cobweb model with a new learning dynamic. *Journal of Economic Dynamics and Control, 33*(6), 1201–1216.

Weibull, J. (1998). *Evolutionary game theory*. Cambridge, MA: MIT Press.

Stock Market Bubble Migration: From Shanghai to Hong Kong

Eric Girardin, Roselyne Joyeux, and Shuping Shi

Abstract The speculative nature of the stock market in Mainland China has attracted the attention of many observers. However while the degree of integration of the Hong Kong market with its Mainland counterpart has monopolized the interest of researchers, they have neglected the diffusion of bubbles from the latter to the former. We thus propose the first study of such bubble migration. Focusing on the period 2005–2017, we use the Phillips et al. (Int Econ Rev 56:1043–1078, 2015a; Int Econ Rev 52:201–226, 2015b) recursive explosive root test to detect and date speculative episodes in both markets. We then implement the Greenaway-McGrevy and Phillips (NZ Econ Pap 50:88–113, 2016) methodology to detect the presence of migration between the two markets. We detect significant, but dwindling, bubble migration from Shanghai to Hong Kong.

1 Introduction

The 'Stock Connect' allowing Mainland and Hong Kong investors to invest in the other stock market has been advertised as a way to reinforce the links between them. The sharp rise (and fall) in the Shanghai market in 2015, right after the introduction of the Shanghai-Hong Kong Connect in late 2014, thus offers a unique opportunity to determine whether such a Connect enabled the transmission of this movement to the Hong Kong market or if such transmission was already present before.

This chapter proposes a precise dating of bubble episodes in both Mainland China and Hong Kong's stock markets. It uses recently developed recursive explosive root tests with high-frequency data over a 12-year period from the mid-2000s. We use for

E. Girardin (✉)
Aix-Marseille School of Economics, Aix-Marseille University, Marseille, France

CNRS & EHESS, Marseille, France

R. Joyeux · S. Shi
Macquarie University, Sydney, Australia

F. Jawadi (ed.), *Uncertainty, Expectations and Asset Price Dynamics*, Dynamic Modeling and Econometrics in Economics and Finance 24,
https://doi.org/10.1007/978-3-319-98714-9_8

this purpose recently developed recursive explosive versus unit root tests allowing for multiple bubbles, which enable us to detect and date precisely the start and end of near-explosive behaviour. Another major step forward is that we rely on new bubble-migration tests to try and determine to what extent and when bubbles in one market follow the bubbles in the other market.

It is often argued that markets prone to bubbles are characterized by the dominance of unsophisticated investors, short-sale constraints and prohibitively costly international arbitrage. China's stock market provides a unique opportunity to test this hypothesis since it is (or was) characterized by a dominance of inexperienced individual investors, binding short-sale constraints (lifted only in 2011), a small asset float (before the split-share reform of 2005–2006, see Beltratti et al. 2009), heavy share turnover despite high transaction costs, as well as binding capital controls, which may however have been de facto increasingly sidestepped through rising trade openness (Aizenman 2004). Bailey et al. (2009) document that (the dominant) individual investors in China's stock market are less informed and more subject to behavioural biases than institutional investors. In addition the link with the prospective income is (was) often tenuous. Indeed for many years, Chinese listed firms hardly distributed any dividend. All these factors make very likely the presence of active speculative behaviour in the Chinese stock market.

Being internationally open and free of the restrictions typical of its mainland counterpart, Hong Kong's stock market functions more in line with its peers in OECD countries and should thus be less prone to speculative forces. In particular the pricing of dual-listed Chinese mainland firms may be more in line with their (prospective) earnings, as reflected in the well-documented premium of such firms' shares in mainland markets over their Hong Kong counterparts (Fung et al. 2016). It is thus important to test to what extent such priors are vindicated or whether bubbles in the mainland market are matched with a possible lag in the Hong Kong market. This is especially so at a time when the Shanghai-Hong Kong stock market Connect is actively promoting the links between the two financial centres (Kasyap 2016), through the opening of international diversification opportunities to Mainland China investors.

Intermarket price arbitrage should be expected in the presence of dual-listed shares on the Mainland's and Hong Kong's markets. Direct price arbitrage may take the form either of quasi arbitrage or of pure arbitrage. Quasi arbitrage is at work when an investor in one market sells (buys) a dual-listed stock when the price of the stock in the domestic market is higher (lower) than that in the corresponding foreign market (Fung et al. 2016). Similarly quasi arbitrage can take place with respect to differential stock returns in corresponding markets (Peng et al. 2007). Quasi arbitrage is able to generate co-movements in prices but not necessarily price convergence, while the latter can be generated by pure arbitrage in a perfect market setting. Pure arbitrage of dual-listed stocks requires much more stringent conditions. Indeed it requires that both the currency and stock markets are integrated with free market access, no exchange control, no restriction on capital flows and synchronous cross-market settlement. Exchange control by a government can be in the form of price (exchange rate) control and/or quantity (capital flow) control. A rigid dollar peg for the Chinese RMB was in force until late July 2005 and from July 2008 to June 2010 and a quasi de facto peg since that date (Girardin and Salimi Naim 2017), while the Hong Kong dollar has been in a currency board regime with the US dollar for three and a half decades. Restrictions against capital flows hamper pure arbitrage between the China and Hong Kong markets. Moreover, before the late 2014 Stock Connect, mainland investors were prohibited by law to trade stocks listed in Hong Kong (or other overseas markets) with their stock accounts in Mainland China, while Hong Kong or other foreign investors could not trade the A-shares listed on both Mainland exchanges. Further, mainlanders are prohibited from remitting money abroad (over some low limit), and overseas investors, including those in Hong Kong, cannot directly transfer money to a bank account in the Mainland.

In light of these institutional and regulatory constraints, an abundant literature has attempted to estimate the relationships between share prices and/or returns in Mainland and Hong Kong stock markets (see Sect. 2). Such work typically tests for the nonstationarity of the price-earnings (or dividends) ratio à la Campbell and Shiller (1987) or Diba and Grossman (1984, 1988). However, as is known from the work of Evans (1991), such unit root tests over-reject the presence of a bubble with a non-zero probability of collapse. Tests on the (negative) dependence duration of runs of abnormal returns à la McQueen and Thorley (1994) were abundantly used both for Mainland China and Hong Kong, but their conflicting (and often negative) results fully vindicate Harman and Zuehlke's (2004) skepticism on the ability of such tests to detect bubbles. Even more sophisticated tests applied to the Mainland China and Hong Kong markets generally do not allow for the presence of explosive behaviour. There is no existing literature on the relationship between stock bubbles in Mainland China and Hong Kong.

The intent of this study is to determine whether there is diffusion of bubbles from the Mainland China's stock market to the Hong Kong's stock market. We first test for the existence and date bubble periods in those markets. One of the most popular bubble detection techniques is the recursive rolling explosive root (vs. unit root) test of Phillips, Shi and Yu (2015a, b, PSY).[1] This approach enables real-time detection of bubbles and has been shown to outperform several other approaches, including the recursive method of Phillips, Wu and Yu (2011) and the CUSUM strategy of Homm and Breitung (2012), when multiple bubbles occur in the data. The popularity of the PSY procedure is also due to its ease of implementation, relative to the Markov-switching test of Hall et al. (1999) and the regime-switching bubble test (Brooks and Katsaris 2005). Second, we use a non-parametric regression approach proposed by Greenaway-McGrevy and Phillips (2016) to examine the migration of bubbles between the two markets. We use weekly data on the price-earnings ratio in the two markets covering a sample from the mid-2000s, the termination of the

[1] While the PSY procedure is designed for the detection of positive bubbles, it also has the capability of identifying price crashes (Phillips and Shi 2017), which are defined as an abrupt discontinuity between asset prices and their underlying environment that produces large negative movements in asset prices.

half-a-decade-long bear market in the first half of the 2000s, and ends in the late Spring 2017.

The chapter contributes to the literature by providing the first precise dating of bubbles both in the Shanghai and Hong Kong markets and by determining the existence and extent of migration of such bubbles between the two markets. First we document that, while the Shanghai market experienced three major bubbles in less than a decade, the Hong Kong market only experienced the first one, in 2007. During the latter episode, there was clear evidence of bubble migration from Shanghai to the H-share and Hang Seng indices; the bubbles in the last two markets only started when the bubble in the former market was in its dying stage. There is thus no evidence that the Stock Connect led in any way to bubble transmission from Shanghai to Hong Kong. Such transmission had stopped many years before and the Connect did not revive it. Conversely, it does not seem that the northbound flows associated with the Connect were able to mitigate bubbles in the Mainland, possibly due to the low quota that they face.

The rest of the present chapter will be structured as follows. The next section will review existing literature on the relation between the two stock markets as well as bubble detection in each of them. The explosive root detection and the bubble-migration methodologies will be introduced in Sect. 3, which will also present the data. Section 4 will discuss the results of the application of the explosive root detection and bubble-migration frameworks to the price-earnings ratios and provide interpretations. Section 5 will offer some conclusions.

2 Literature on China's and Hong Kong's Stock Bubbles

Speculation is a major candidate to explain the breakdown of the link between asset prices and fundamentals. Scheinkman and Xiong (2003) and Hong et al. (2006) show that, in the presence of both heterogeneous beliefs and short-sale constraints, investors may be induced to overpay for an asset if they expect to sell it to another investor who will be willing to pay even more in the future. Accordingly, asset prices may contain a sizeable speculative component. The theory of asset market bubbles is surveyed by Scherbina and Schlusche (2014). The dominance of unsophisticated individual investors, binding short-sale constraints, and often costly arbitrage has characterized the stock market in Mainland China during most of our sample. Accordingly, it is likely that such a market would be characterized by its speculative nature.

Existing work on China's stock market[2] provides us with evidence in support of its speculative character (Mei et al. 2009). Bailey et al. (2009) document that (the dominant) individual investors in China's stock market are less informed and

[2]Evidence for the 1990s (Girardin and Liu 2003) shows the presence of a speculative regime with very high returns.

more subject to behavioural biases than institutional investors. However, a number of deep reforms, implemented in the last decade, may have lessened this speculative character. First, an expanding number of listed firms in China are increasingly representative of an economy with dominant non-state-owned firms in industrial activity, restructured state-owned firms, as well as reforms associated with entry into the World Trade Organization (WTO) in December 2001. Second, the split-share reform initiated in May 2005 and completed late 2006 (Beltratti et al. 2009) reduced the sharp hiatus between the float and the capitalization of the stock market, and its links with foreign markets have been gradually enhanced via the Qualified Foreign Institutional Investors' (QFII) scheme, the Qualified Domestic Institutional Investors' (QDII) scheme and the Hong Kong-Shanghai (Shenzhen) Stock Connect from late 2014 (from 2016). All these deep changes imply that the behaviour of the Chinese mainland stock market may have become closer to that of its peer in Hong Kong.

It is well documented that developments in Hong Kong's stock market may not be independent from what happens in Mainland China's stock markets. While abundant empirical work has examined for China and Hong Kong the relationship between prices and fundamentals in the stock market, much more limited research has dealt with bubble detection, and no quantitative work has studied the possible links between bubbles in the two markets.

The links between fundamentals and the stock market in China have been examined both for prices and volatility. Bondt et al. (2010) consider the influence of conventional fundamentals in the spirit of Shiller's (1981) present value model and show that equity market reforms and excess liquidity drive episodes of stock price misalignments.[3] A mixed-frequency and fundamentals-based approach used to model Mainland China's stock return volatility was proposed by Girardin and Joyeux (2013), who show the links between fundamentals and long-run volatility.

The work which has attempted to detect bubbles in Mainland China's stock market reports only partially overlapping, and quite contrasted, results in two directions. The first line of research relies on the duration-dependence tests à la McQueen and Thorley (1994) in which the presence of speculative bubbles is inferred from a long-lasting run of positive abnormal returns, associated with negative duration dependence (Maheu and McCurdy 2000). Jirasakuldech et al. (2006) find no evidence of duration dependence on monthly data for Mainland China over 1993–2004, while Zhang (2008) does find evidence of duration dependence over 1991–2001, for both Shanghai and Shenzhen composite indices. Lehkonen (2010) gets conflicting results with weekly and monthly data over 1992–2008. Chen et al. (2011) sidestep the possible misspecification of the hazard function, thanks to a non-parametric testing procedure proposed by Diebold and Rudebusch (1990), and find evidence of duration dependence for the Chinese mainland stock market with weekly data over 1992–2006. Limits of the duration-dependence test have been

[3] See Jawadi and Prat (2017) for a critical review of existing empirical approaches of the relationship between stock prices and fundamentals.

emphasized by Harman and Zuehlke (2004), in as much as evidence of duration dependence is sensitive to the choice of sample periods, the method of correcting for discrete observations of continuous duration and the use of value-weighted versus equally weighted portfolios. Contradictory results were obtained with the use of monthly versus weekly runs of abnormal returns (confirmed for Mainland China by Lehkonen (2010)). Overall this 'calls into question the efficacy of using hazard models to test for speculative bubbles' (Harman and Zuehlke 2004).

A second line of research is able to provide a dating of speculative episodes. Thus Jiang et al. (2010) use a faster-than-exponential (power law with finite-time singularity) increase in stock prices as the main diagnostic of bubbles for Mainland China over the May 2005 to August 2009 sample and detect two bubbles from mid-2005 to October 2007 and from November 2008 to August 2009. In contrast, Asako and Liu (2013), over the 1999–2010 period, only detect significant bubbles in April to May and August to October 2007, and Chang et al. (2016), over 1995–2013, find very short-lived bubbles, only early and late 2007. The approach is extended and updated by Hu and Li (2017).

Three series of studies have tested for the presence of bubbles in the Hong Kong stock market. In a first group, Indiran et al. (2015) try to detect rational speculative bubbles as movements of prices above fundamental value, representing about one-third around the global economic crisis of 2008. The rational speculative bubble started to form and grow from June 15, 2006, to December 10, 2007. Miyakoshi et al. (2014) use Hong Kong stock market's four subindices, rely on unit root tests and find that, over the subperiods of 1986–2002 and 2000–2012, the bubbles in commerce, industry and utilities, but not in finance and properties, are consistent with rational-expectation bubbles. Ahmed et al. (2010), from the early 1990s to 2006, reject the presence of nonlinear speculative bubbles by studying residuals of vector autoregressive-based fundamentals, the Hamilton regime-switching model and the rescaled range analysis of Hurst. Yu and Sze (2012) examine the Hang Seng index and its dividend ratio from July 1974 to May 2002 and find that both the West specification test and the Diba-Grossman co-integration test suggest the possible existence of bubbles.

A second group of studies for Hong Kong relies on the duration-dependence test developed by McQueen and Thorley (1994). The earliest (Chan et al. 1998), over the 1975–1994 period, and more recent (Yu and Sze 2012) work, extending the sample to 2002, are unable to detect any duration dependence in the Hang Seng index. The same negative conclusion is reached by Gan et al. (2012) both before (1993–1997) and after (1998–2008) the 1997 Asian financial crisis. This result is not sensitive to the choice of different models, monthly versus weekly runs of returns and equally versus value-weighted portfolios in the Hong Kong stock market. Only Chen and Shen (2007), over a very long sample (1965–2005), document duration dependence but in an unintuitive asymmetrical way, i.e. in a bear, but not in a bull, market.

Finally, a third stream of research on the Hong Kong market takes on board Evans' (1991) cautionary recommendations on the over-rejection of bubbles by unit root tests in the presence of a non-zero probability of bubble collapse. Wu and Xiao (2008) thus propose an original test based on the fluctuations in the partial sum process of the residuals of a regression of stock prices on dividends, which should be proportional to the square root of the number of observations in the absence of a bubble. For the Hong Kong market with weekly data, from January 1974 to September 1998, such an absence of bubble is rejected by this test, and there is evidence of three crashes over that period.

Existing work looking at interactions between the Mainland China's stock market and Hong Kong's has adopted either a disaggregated or an aggregated perspective. First, the literature focusing on price interactions granted a lot of attention at the company level to the specific role of the dual listing of Chinese mainland firms with violations of the law of one price. Many studies have shown that foreign shares typically trade at a premium over their corresponding domestic shares (as in Thailand, Switzerland, Mexico, Indonesia, Malaysia, Singapore and Norway; see Fung et al. (2016), for references). However, research on the Chinese dual-listed shares typically finds the opposite phenomenon,[4] where the Chinese domestic shares trade at a premium over the foreign shares, the so-called puzzle of the Chinese stock market (Fernald and Rogers 2002).

At an aggregate level, recent research on interactions among indices, their returns and their volatilities (for references to earlier work, see, e.g. Girardin and Liu, 2007) includes Wang et al. (2012) with daily data from January 2000 to June 2012 and a wavelet coherence model, who find that there are significant co-movements between the Shanghai, Shenzhen and Hong Kong stock markets in the medium and long run. A major result of this work is that the Hong Kong stock market plays a leading role in the long run, but its leader position is threatened by fast-growing Chinese mainland stock markets, especially the Shanghai Stock Exchange.

Many papers have investigated the integration of the Hong Kong and Shanghai markets with co-integration tests. For example, Zhu et al. (2004) and Groenewold et al. (2004) do not uncover co-integration between market returns in Shanghai, Shenzhen and Hong Kong (see also the work surveyed in Wang et al. 2016). However, the possibility of near-explosive behaviour of stock prices calls for caution in assessing the conclusions of the abundant literature testing for integration of the Hong Kong and Shanghai markets with co-integration tests. Such tests should be restricted to subsamples where stock indices have a unit root and no explosive root. In other words a sequential strategy is required to identify such subsamples before testing for integration.

[4]Existing research (reviewed by Fung et al. 2016; and earlier by Fong et al. 2007) suggest that the discount on the Chinese foreign shares can be explained by the following idiosyncratic factors: information asymmetries between domestic and foreign investors, different liquidity conditions, domestic investors' speculative motive, differential risk, market conditions and short-sale restrictions.

Zhang et al. (2009), who examine the dynamics of the linkages between Shanghai and Hong Kong stock indices, considering both the volatility linkage with a multivariate generalized autoregressive conditional heteroscedasticity (MVGARCH) framework and the dependence of returns with a copula approach, find significant dependence of the returns in the two markets. Ho and Zhang (2012), who examine the volatility dynamics of the Greater China stock markets by employing a multivariate framework that incorporates the features of asymmetries, persistence and time-varying correlations, find some evidence of a common degree of persistence among these markets. They document that the Chinese mainland markets were less volatile than the Hong Kong stock exchanges in the late 1990s and early 2000s and that the Shenzhen and Shanghai stock exchanges are weakly correlated with the Hong Kong market.

Ho et al. (2016), who use a regime-switching model to examine the correlation dynamics of the Chinese mainland and Hong Kong stock markets with high-frequency data, find that all correlations are significantly time-varying with various patterns and there is co-persistence in both low- and high-correlation states. Mohammadi and Tan (2015), who examine the dynamics of daily returns and volatility in the stock markets of Hong Kong and Mainland China over January 2, 2001, to February 8, 2013, find no spillover between the two markets and document a medium correlation of 30% between them, with an increase in correlation since the global financial crisis.

3 Methodology and Data

3.1 Methodology

Greenaway-McGrevy and Phillips (2016; GMP) employ the PSY procedure to identify bubble episodes in the New Zealand property market. They also introduce new techniques to detect spillovers across markets. We use this two-stage procedure to detect the migration of bubbles from Mainland China's to Hong Kong's stock market.

3.1.1 Speculative Bubble Detection

The PSY bubble tests are based on testing for explosive roots in normalized asset prices (in our case stock price indices normalized by earnings). The testing algorithm is based on a right-tailed unit root test (Phillips et al. 2014) with a unit root null and an explosive (bubble) alternative. As in the case of unit root tests, the tests are performed by estimating an autoregression such as:

$$\Delta y_t = \alpha + \beta y_{t-1} + \sum_{i=1}^{K} \gamma_i \Delta y_{t-i} + \varepsilon_t, \tag{1}$$

where y_t is the asset price at period t, K is the lag order (set to one in the application), and ε_t is the error term. The null hypothesis is $\beta \leq 0$ (unit root or stationarity), and the alternative is $\beta > 0$ (explosive behaviour).

In order to allow for structural breaks in β, (2) is estimated recursively on subsamples, and the ADF statistic (t-ratio of the OLS estimate of β) is calculated for each subsample. Denoting by r_1 and r_2 the fractional starting and ending points of a subsample, the corresponding ADF statistics are denoted as $ADF_{r_2}^{r_1}$. The ADF statistic is calculated on a sequence of backward expanding samples. The minimum window size required to estimate a regression is denoted by r_0. To lessen the probability of size distortion, Phillips et al. (2015a) recommend setting r_0 to be $\left(0.01 + 1.8/\sqrt{T}\right)$ with T being the total sample size. For observation $\lfloor rT \rfloor$,[5] the ADF statistic sequence is $\left\{ADF_{r_2}^{r_1}\right\}_{r_2=r}^{r_1 \in [0, r-r_0]}$. The sup value of the ADF sequence, denoted by $SADF_r$, is the relevant statistic to test for explosive behaviour. It is defined as:

$$SADF_r = \max\left\{ADF_{r_2}^{r_1} : r_2 = r \text{ and } r_1 \in [0, r - r_0]\right\}.$$

The bubble origination (termination) date is identified using the first crossing-time principle, i.e. the first chronological observation whose SADF statistic is above (below) its corresponding critical value. Phillips et al. (2015b) and Phillips and Shi (2018) show that this strategy can consistently estimate the origination and termination dates of multiple bubbles, if the time span between two bubbles is longer than r_0.

Since we are studying the explosive behaviour of stock price indices, we expect conditional or unconditional heteroscedasticity to be present (Engle 1982; Bollerslev 1986). To reduce the probability of size distortion under such circumstances, critical values are computed with a wild bootstrapping procedure as suggested by Harvey et al. (2016).

3.1.2 Bubble Migration

The analysis of bubble migration (GMP) is based on the estimated coefficient β of Eq. (1). This coefficient summarizes the dynamics of asset prices. It is expected to be (statistically) positive in the presence of speculation and not greater than zero during normal market periods. Let S and R be the source and recipient markets, respectively. The time-varying β coefficients of these two markets are obtained from rolling window regressions of (1) with a window size of w and denoted by $\left\{\widehat{\beta}_{\mathrm{S,t}}\right\}_{t=w}^{T}$ and $\left\{\widehat{\beta}_{\mathrm{R,t}}\right\}_{t=w}^{T}$.

[5] $\lfloor . \rfloor$ signifies the integral part of the argument.

GMP estimate a non-parametric regression between the centred $\tilde{\beta}$ coefficients as:

$$\tilde{\beta}_{\mathrm{R,t}} = \delta_{t,T}\tilde{\beta}_{\mathrm{S,t}-d} + \varepsilon_t \quad (2)$$

where $\tilde{\beta}_{k,\mathrm{t}} = \widehat{\beta}_{k,\mathrm{t}} - \frac{1}{T-w+1}\sum_{t=w}^{T}\widehat{\beta}_{k,\mathrm{t}}$ with $k = \{S,R\}$ are the centred coefficients, d is the lag order, ε_t is the error term and $t = w + d, \ldots, T$. The time-varying coefficient $\delta_{t,T}$ is estimated by a local-level kernel regression. We refer to GMP and Deng et al. (2017) for details of the estimation method. A positive and increasing value of $\delta_{t,T}$ is expected in the event of bubble migration: a positive $\delta_{t,T}$ is an indication that the autoregressive coefficients move in the same direction, i.e. when the S market's autoregressive coefficient increases so does the autoregressive coefficient of the R market. If $\delta_{t,T}$ is positive but small, the speed of increase is slow; if it is large, the speed is fast. If S is in a bubble period (i.e. autoregressive coefficient is found to be significantly larger than (1)) the evidence of bubble migration from S to R is stronger if $\delta_{t,T}$ is 'large' in that period. For example, if $\delta_{t,T} > 1$ the autoregressive coefficient for R is predicted to be larger than the autoregressive coefficient of S in that period, i.e. the R market is predicted to be also in a bubble period.

Ideally, one would provide a confidence band around the estimated coefficient $\widehat{\delta}_{t,T}$ to signify its significance. This is, however, not feasible at this stage as the limiting distribution of $\widehat{\delta}_{t,T}$ has not yet been derived and it is out of the scope of this chapter.

3.2 Data

For both Hong Kong (Hang Seng) and Mainland China's (Shanghai composite) stock markets as well as for the market for Mainland China's firms in Hong Kong (H-shares), we consider a weekly sample from January 3, 2005, to June 16, 2017. The data on the stock market indices and its earnings are extracted from Bloomberg. Weekly data are calculated by taking the week's average from Thursday to Wednesday. Figure 1 shows three periods of rising prices in Shanghai's stock market, from the Summer 2006 to November 2007, from the end of October 2008 to the end of July 2009 and in mid-2015. It is apparent that the Hong Kong market missed the third episode and the H-shares only experienced it mildly.

4 Detecting Bubbles and Their Migration

In this section we initially test for the presence of explosive behaviour in weekly stock price index-earnings ratios separately for Mainland China's, Hong Kong's and H-share markets, in order to date the timeline of bubbles in each of them. Subsequently, we test for the presence of migration of bubbles from one market to one of the others.

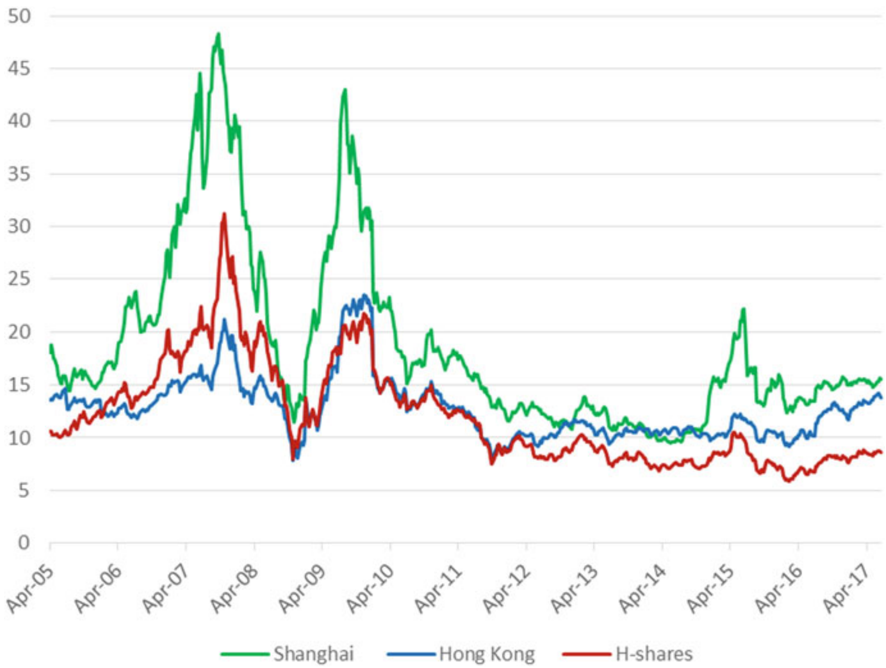


Fig. 1 Stock price index-earnings ratio (PERs) in Hong Kong, Shanghai and H-shares

4.1 Bubble Timeline: Detecting Explosive Behaviour

We plot in Figs. 2, 3 and 4 the identified bubble indicators, along with the log of the stock price-earnings ratios, respectively, for Mainland China, Hong Kong and H-shares.[6] The figures also highlight areas corresponding to periods of explosive behaviour as detected using 90% critical values obtained from wild bootstrapping procedures.

Leaving aside short-lived episodes (shorter than 1 month, such as late 2006), three main bubble episodes can be identified for Shanghai. The first, in 2007, is two-pronged, from mid-January to late June and from early August to early October. The second, in 2009, is short, with only 5 weeks from early July to mid-August. Finally, the third one is two-pronged like the first one, with a 1-month component from early December 2014 to early January 2015, followed by a much longer resumption from mid-March to late June 2015.

[6] We exclude episodes associated with market downturns, as suggested in Phillips and Shi (2017). Such negative movements were detected for the log Shanghai composite PERs and the log H-share PERs from June 18, 2008, to December 3, 2008, and from December 14, 2011, to January 11, 2012.

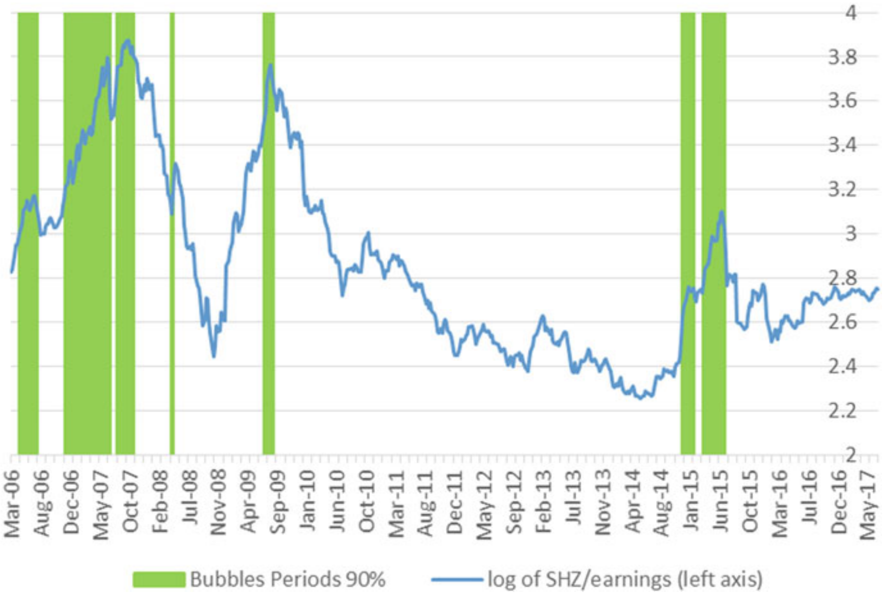


Fig. 2 The identified bubble periods (shaded) and the log Shanghai composite PERs. Note: The bubble periods are identified based on the SADF statistic sequence and the 90% wild bootstrapping critical value sequence

The first bubble in Shanghai follows the full implementation in 2006 of the split-share reform (Beltratti et al. 2009) initiated in the Spring 2005, following a long-lived bear market in the first half of the 2000s—bringing the float closer to capitalization—and accompanied by optimistic investor expectations (Li 2015). However, earnings stopped validating such expectations in the early Autumn 2007 and the bubble burst. The second bubble, which arose and burst in the Summer 2009, marks the end of a long bull market as noted by Jiang et al. (2010). However, the short duration of that bubble is very surprising in the light of the record expansion of bank credit initiated by the Chinese authorities which (successfully) ordered state-owned banks to lend on a large scale (Deng et al. 2017) in the first semester of 2009 in order to try and counter the economic slowdown associated with the global financial crisis.

The third bubble, in 2015, was in part driven by the government, which talked up the market,[7] and in part by the stock market regulator (CSRC) which indirectly fuelled the demand for shares via a relaxation of rules on margin trading (introduced in 2010), lowering thresholds on collateral requirements. Accordingly over its first

[7] Such propping-up was rationalized by some observers by the argument that ‘Higher stock prices would also help China’s state-owned enterprises (SOEs) cut their debt levels because they can sell shares they own to pay back borrowings’ (Source: http://knowledge.wharton.upenn.edu/article/whats-behind-chinas-stock-market-gamble/).

Fig. 3 The identified bubble periods (shaded) and the log Hong Kong PERs. Note: The bubble periods are identified based on the SADF statistic sequence and the 90% wild bootstrapping critical value sequence

5 years of existence, margin trading expanded more than fivefold (from $65 to 355 billion). In addition, from November 2014 monetary easing by China's central bank (PBoC) provided additional macroeconomic liquidity. As a countervailing factor, the (late) actions taken by the CSRC, such as a tightening in margin requirements in January and April 2015 and the widening of short-selling to a larger number of stocks, were unable to dampen stock prices. It is generally accepted that the bubble was pricked when the CSRC announced, on June 12, plans limiting lending for stock trading by brokerages (Frankel 2015).

The Hang Seng market (Fig. 3) participated briefly in the early (from early January to late February) and late (but only for 2 weeks in the second half of June) stages of the first leg of the 2007 Shanghai bubble. In addition a 1-month-long bubble started in the Hang Seng index early October, precisely when the second leg of the 2007 Shanghai bubble ended.

The H-share index participated in part of the last 5 weeks of 2006 to the short-lived warning signals of Shanghai's first bubble. In a similar way to the Hang Seng, a bubble in the H-share index arose (late September 2007) when the second leg of the first Shanghai bubble was exhausted, and similarly lasted until early November. Subsequently there is no evidence of explosive behaviour in the H-share index, which, just as in the case of the Hang Seng index, missed both the second (2009)

Fig. 4 The identified bubble periods (shaded) and the log H-share PERs. Note: The bubble periods are identified based on the SADF statistic sequence and the 90% wild bootstrapping critical value sequence

and the third (2015) Shanghai bubble. Indeed, the 2-week-long episode for H-shares in mid-April 2015 is too short to qualify.

4.2 Bubble Migration

Given that stock market bubbles in Hong Kong (Hang Seng or H-shares) as detected in Fig. 3 never precede bubbles in Shanghai, we rather focus on the migration from Shanghai to either of the Hong Kong indices of which there are some presumptions when confronting Figs. 2 and 3. We estimate rolling autoregressions of the form (2) for each of the series with a fixed window size of $w = 70$. This gives the slope coefficient estimates $\{\widehat{\beta}_{S,i}\}_{i=w}^{T}$ and $\{\widehat{\beta}_{R,i}\}_{i=w}^{T}$ where R stands for one of the Hong Kong price-earnings ratios (recipient) and S for the Shanghai stock price series-earnings ratio (source). We then estimate Eq. (2).

The optimal lag order is estimated to be 0 implying that the migration of bubbles from the Mainland China's market to Hong Kong's stock market is instantaneous. The migration coefficient δ should be interpreted carefully since it is estimated by a smoothing procedure and the β coefficients are estimated over a rolling window. This means that the contagion effect can only change smoothly over time. A rise in δ in a given week indicates that the sensitivity of one market to another is starting to increase.

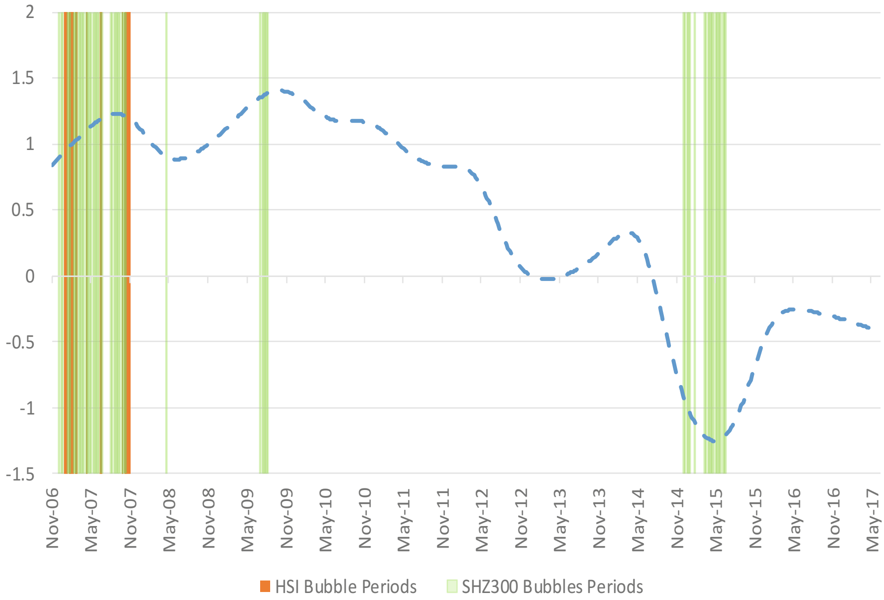


Fig. 5 Bubble migration from Shanghai's stock market to Hong Kong's stock market: the estimated δ coefficient of Eq. (2)

In Fig. 5, the dashed line is the time-varying coefficient of Eq. (2) above. The solid highlighted area shows the periods when there is evidence of a bubble at 10% in the price-earnings ratio in either market. The sensitivity of the Hang Seng to the Mainland China stock market bubble[8] is positive and rises during the period of upward movements in stock prices until September 2009. It falls sharply when the stock market bubble has exploded in the Shanghai market. From mid-2014 the δ coefficient of migration from Shanghai to the Hang Seng becomes negative (Fig. 5), having already reached zero late 2012. This implies that, over that period, the Hang Seng has an adverse reaction relative to the stock market in the Mainland. In the case of the migration from Shanghai to H-shares, the δ coefficient remained positive throughout (Fig. 6) but had already fallen very substantially by late 2013 and kept on falling further until late 2015. Accordingly, it may at first sight be a little surprising that the 2009 Shanghai bubble did not migrate to one of the Hong Kong indices, since δ was still high at that time for both of them, but the short duration of this

[8]Even allowing for this degree of uncertainty in the estimation of the migration coefficient (see methodological section), the finding that this coefficient is stronger from Shanghai to the Hang Seng than to H-shares in 2007–2009 may seem a puzzle, even though it disappeared subsequently. This illustrates the complexity of the links between Greater China's stock markets, which, as argued in the survey section above, existing literature has not helped us understand due to its ignorance of explosive roots.

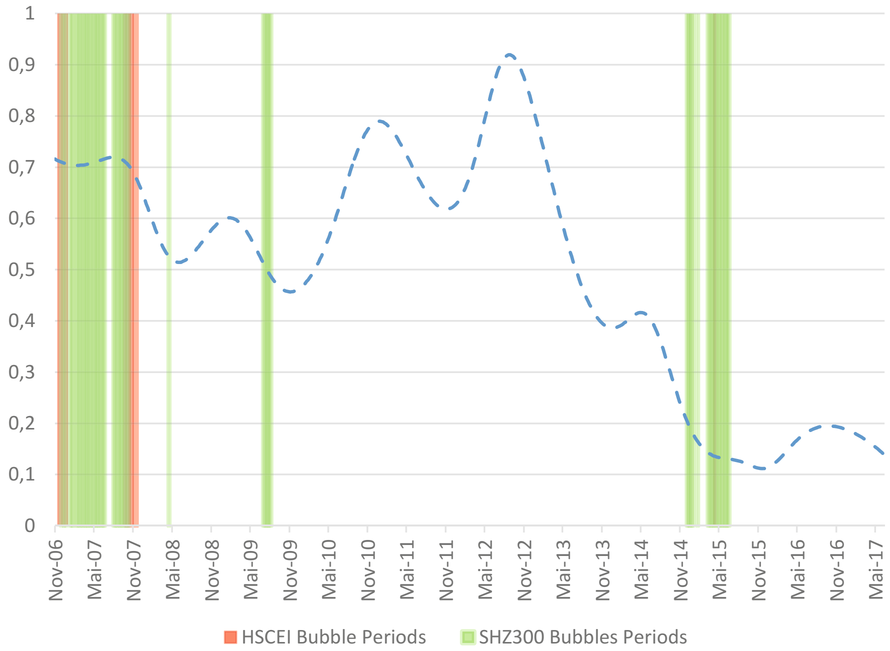


Fig. 6 Bubble migration from Shanghai to H-share stock market: the estimated δ coefficient of Eq. (2)

second Shanghai bubble may partly explain that. By contrast, the low or negative δ_s even prior to the start of the third Shanghai bubble would on an ex ante basis have led to expect no migration to Hong Kong, and this expectation was vindicated by the lack of southern migration of the 2015 Shanghai bubble.

Overall it thus seems that while a Shanghai bubble was able to migrate to the Hong Kong PERs at the time when the former was crashing, passing the baton south as in the early Autumn 2007, this was no more the case subsequently. After the upheavals of the global financial crisis, the international exposure of the Hong Kong market may have then been able to insulate it fully from the speculative waves in the Mainland. The fact that the Shanghai-Hong Kong Stock Connect did not revive the transmission across such markets may be due to a combination of four factors: retail investors not using enough the link, remaining limits on the shorting of A-shares, the prohibition for foreigners to trade in Mainland China's stock index futures and caps on the amounts of cross-border investment allowed under the Connect. Indeed the Connect, launched in late 2014, allows investors from the mainland to buy only a net 10.5 billion yuan ($1.6 billion) of shares a day in Hong Kong, while the reverse flows are capped at 13 billion yuan. It is noteworthy that Wang et al. (2016), using a VAR-GARCH framework, find that the different stages of the Shanghai-Hong Kong Stock Connect boosted returns more in the Shanghai than in the Hang Seng index.

5 Conclusion

The growing concerns of a ‘bubble’ in the Chinese mainland stock markets in the second part of the first decade of the new millennium imply that frequently skyrocketing increases in stock prices may be generated by speculation, in part due to the very unique characteristics of a stock market in Mainland China, which for many years remained relatively closed to overseas investors. By contrast, the Hong Kong stock market, fully open to international investors, may be less prone to bubbles. Existing literature has only studied the possible presence of bubbles for these markets in isolation. The introduction of the Shanghai-Hong Kong Stock Connect late 2014, right before the skyrocketing stock prices in the Mainland, offered us a unique opportunity to examine whether bubble transmission across the border was affected by this opening of cross-border equity investment flows.

We proposed using a weekly data set over 12 years from the mid-2000s, with Shanghai composite, Hang Seng and H-share price-earnings ratios, to address such concerns. To analyse this data, we used a two-step strategy to provide the first timeline of stock market bubbles in the three markets over such a crucial period and bubble migration from one market to one of the others.

Our use of recently developed recursive explosive versus unit root tests implies that concerns about the presence of bubbles in both Mainland China’s and Hong Kong’s stock markets are vindicated, conditioning on fundamentals of the stock markets, i.e. earnings. These results align with the sequential hypothesis concerning bubble creation and collapse of Phillips and Yu (2011) for the US economy. Near-explosive behaviour is detected for three main episodes for the Chinese mainland market. It thus appeared worthwhile to examine the transmission of those bubbles to the Hong Kong markets.

While the Shanghai bubble was able to migrate to the Hong Kong PERs at the time when the former was crashing, passing the baton south as in the early Autumn 2007, this was not a general phenomenon, since bubble migration subsequently stopped. The international exposure of the Hong Kong market may have then been able to insulate it fully from the speculative waves in the Mainland. In as much as the H-share (as well as the broader Hang Seng) market appears now insulated from the collateral damage from the bubbles arising in its Mainland sibling, the further opening to foreign investors and international arbitrage of the Mainland markets would seem to represent a way for the latter to do away with its idiosyncratic speculative bubble episodes. Of course, in return, such an opening may make the Mainland China market more susceptible to imported bubbles. It is of course possible, though much less likely, that further expansion in the quotas allowed under the Stock Connect scheme would enable future speculative pressures in the Mainland markets to spill over to the Hong Kong one.

References

Ahmed, E., Rosser Jr., J. B., & Uppal, J. Y. (2010). Emerging markets and stock market bubbles: Non-linear speculation? *Emerging Markets, Finance and Trade, 46*(4), 23–40.

Aizenman, J. (2004). Financial opening and development: Evidence and policy controversies. *American Economic Review, 94*(2), 65–70.

Asako, K., & Liu, Z. (2013). A statistical model of speculative bubbles with application to the stock market of the United States, Japan and China. *Journal of Banking and Finance, 37*, 2639–2651.

Bailey, W. B., Cai, J., Cheung, Y. L., & Wang, F. (2009). Stock returns, order imbalances, and commonality: Evidence on individual, institutional, and proprietary investors in China. *Journal of Banking and Finance, 33*, 9–19.

Beltratti, A., Bortolotti, B., & Caccavaio, M. (2009). *Stock prices in a speculative market: The Chinese split-share reform*. Bocconi University, Paolo Baffi centre research paper series 2009-09.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics, 31*(3), 307–327.

Brooks, C., & Katsaris, A. (2005). A three-regime model of speculative behaviour: Modelling the evolution of the S&P 500 composite index. *Economic Journal, 115*(505), 767–797.

Campbell, J. Y., & Shiller, R. J. (1987). Cointegration and tests of present-value models. *Journal of Political Economy, 95*(5), 1062–1088.

Chan, K., McQueen, G., & Thorley, S. (1998). Are there rational speculative bubbles in Asian stock market? *Pacific Basin Finance Journal, 6*, 125–151.

Chang, T., Gil-Alana, L., Aye, G. C., Gupta, R., & Ranjbar, O. (2016). Testing for bubbles in BRICS' stock markets. *Journal of Economic Studies, 43*(4), 646–660.

Chen, H., Chong, T.-L., & Li, Z. (2011). Are Chinese stock market cycles duration dependent? *Financial Review, 46*, 151–164.

Chen, S.-W., & Shen, C.-H. (2007). Evidence of duration dependence in the stock markets in the Pacific Rim economies. *Applied Economics, 39*, 1491–1474.

de Bondt, G.J., Peltonen, T.A., & Santabárbara, D. (2010, May). *Booms and busts in China's stock market: estimates based on fundamentals*. ECB working paper N° 1190.

Deng, Y., Girardin, E., Joyeux, R., & Shi, S. (2017). Did bubbles migrate from the stock to the housing market in China between 2005 and 2010? *Pacific Economic Review, 22*(3), 276–292.

Diba, B., & Grossman, H. (1984). *Rational bubbles in the price of gold*. NBER Paper No. 1300.

Diba, B. T., & Grossman, H. I. (1988). Explosive rational bubbles in stock prices? *American Economic Review, 78*(3), 520–530.

Diebold, F. X., & Rudebusch, G. D. (1990). A nonparametric investigation of duration dependence in the American business cycle. *Journal of Political Economy, 98*, 596–616.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica, 50*(4), 987–1007.

Evans, G. (1991). Pitfalls in testing for exploding bubbles in asset prices. *American Economic Review, 76*, 222–230.

Fernald, J., & Rogers, J. H. (2002). Puzzles in the Chinese stock market. *The Review of Economics and Statistics, 84*, 416–432.

Fong, T., Wong, A., & Yong, I. (2007, July). *Share price disparity in Chinese stock markets*. Hong Kong Monetary Authority, Working paper 11/07.

Fung, J., Girardin, E., & Hua, J. (2016, December). *How does an exchange-rate regime change affect dual-listed stock arbitrage? Evidence from China's A- and H-share markets*. Manuscript, Aix Marseille School of Economics, Aix Marseille University.

Gan, C. G., Nartea, V., Ling, D. L., & Hu, B. (2012). Duration dependence test of rational speculative bubbles: A case study of the Hong Kong stock market. *Investment Management and Financial Innovations, 9*(2), 37–48.

Girardin, E., & Joyeux, R. (2013). Market fundamentals as a source of stock market volatility in China: A GARCH-MIDAS approach. *Economic Modelling, 34*, 59–68.

Girardin, E., & Liu, Z. (2003). The Chinese stock market: A casino with buffer zones? *Journal of Chinese Economic and Business Studies, 1*(1), 57–70.

Girardin, E., & Liu, Z. (2007). The financial integration of China: New evidence on temporally aggregated data for the A-share market. *China Economic Review, 18*(3), 354–371.

Girardin, E., & Salimi Naim, F. (2017, July). *Shadowing the dollar or basket pegging in China?* Aix-Marseille School of Economics, Aix Marseille University, Manuscript.

Greenaway-McGrevy, R., & Phillips, P. C. B. (2016). Hot property in New Zealand; Empirical evidence of housing bubbles in the metropolitan centres. *New Zealand Economic Papers, 50*, 88–113.

Groenewold, N., Tang, S., & Wu, Y. (2004). The dynamic interrelationships between the greater China share markets. *China Economic Review, 15*(1), 45–62.

Hall, S. G., Psaradakis, Z., & Sola, M. (1999). Detecting periodically collapsing bubbles: A Markov-switching unit root test. *Journal of Applied Econometrics, 14*(2), 143–154.

Harman, Y. S., & Zuehlke, T. W. (2004). Duration dependence testing for speculative bubbles. *Journal of Economics and Finance, 28*(2), 147–154.

Harvey, D. I., Leybourne, S. J., Sollis, R., & Taylor, A. R. (2016). Tests for explosive financial bubbles in the presence of non-stationary volatility. *Journal of Empirical Finance, 38*, 548–574.

Ho, K.-Y., Shi, Y., & Zhang, Z. (2016). It takes two to tango: A regime-switching analysis of the correlation dynamics between the Mainland Chinese and Hong Kong stock markets. *Scottish Journal of Political Economy, 63*(1), 41–65.

Ho, K.-Y., & Zhang, Z. (2012). Dynamic linkages among financial markets in the greater China region: A multivariate asymmetric approach. *World Economy, 35*(4), 500–523.

Homm, U., & Breitung, J. (2012). Testing for bubbles in stock markets: A comparison of alternative methods. *Journal of Financial Econometrics, 10*(1), 198–231.

Hong, H., Scheinkman, J., & Xiong, W. (2006). Asset float and speculative bubbles. *Journal of Finance, 61*, 1073–1117.

Hu, Z., & Li, C. (2017). New JLS-factor versus standard JLS model: A case study on Chinese stock bubbles. *Discrete Dynamics in Nature and Society*, 8017510.

Indiran, D., Halim, N. A., & Ahmad, W. M. A. (2015). Hong Kong stock market rational speculative bubble periods-2008. *International Journal of Applied Research, 1*(10), 863–865.

Jawadi, F., & Prat, G. (2017). Equity prices and fundamentals: A DDM-APT mixed approach. *Review of Quantitative Finance and Accounting, 49*(3), 661–695.

Jiang, Z., Zhou, W., Sornette, D., Woodard, R., Bastiaensen, K., & Cauwels, P. (2010). Bubble diagnosis and prediction of the 2005–2007 and 2008–2009 Chinese stock market bubbles. *Journal of Economic Behavior and Organization, 74*, 149–162.

Jirasakuldech, B., Emekter, R., & Went, P. (2006). Fundamental value hypothesis and return behavior: Evidence from emerging equity markets. *Review of Pacific Basin Financial Markets and Policies, 29*(1), 97–127.

Kasyap, R. (2016). Hong Kong-Shanghai connect/Hong Kong-Beijing disconnect? Scaling the great wall of Chinese securities trading costs. *Journal of Trading, 11*(3), 81–134.

Lehkonen, H. (2010). Bubbles in China. *International Review of Financial Analysis, 19*, 113–117.

Li, Z. (2015). Emergence of China's 2006-2007 Stock market bubble and its burst. In S. Cheng & Z. Li (Eds.), *The Chinese stock market: Evaluation and prospects* (Vol. 2, pp. 61–124). London: Palgrave.

Maheu, J., & McCurdy, T. (2000). Identifying bull and bear markets in stock returns. *Journal of Business and Economic Statistics, 18*, 100–112.

McQueen, G., & Thorley, S. (1994). Bubbles, stock returns, and duration dependence. *Journal of Financial and Quantitative Analysis, 29*, 186–197.

Mei, J., Scheinkman, J., & Xiong, W. (2009). Speculative trading and stock prices: Evidence from Chinese A-B share premia. *Annals of Economics and Finance, 10*, 225–255.

Miyakoshi, T., Li, K.-W., & Shimada, J. (2014). Rational expectations bubbles: Evidence from Hong Kong subindices. *Applied Economics, 46*(20), 2429–2440.

Mohammadi, H., & Tan, Y. (2015). Return and volatility spillovers across equity markets in Mainland China, Hong Kong and the United States. *Econometrics, 3*(2), 215–232.
Peng, W., Miao, H., & Chow, N. (2007, July). *Price convergence between dual listed A and H shares.* Hong Kong Monetary Authority, China Economic Issues 6/07.
Phillips, P. C., & Shi, S. (2017). *Detecting financial collapse and ballooning sovereign risk*. Cowles Foundation discussion paper no. 3010.
Phillips, P. C., & Shi, S. (2018). Financial bubble implosion and reverse regression. *Econometric Theory, 34*(4), 705–753.
Phillips, P. C., Shi, S., & Yu, J. (2014). Specification sensitivity in right-tailed unit root testing for explosive behaviour. *Oxford Bulletin of Economics and Statistics, 76*(3), 315–333.
Phillips, P. C., Shi, S., & Yu, J. (2015b). Testing for multiple bubbles: Limit theory of realtime detectors. *International Economic Review, 56*(4), 1079–1134.
Phillips, P. C. B., Shi, S., & Yu, J. (2015a). Testing for multiple bubbles: Historical episodes of exuberance and collapse in the S&P 500. *International Economic Review, 56*, 1043–1078.
Phillips, P. C. B., Wu, Y., & Yu, J. (2011). Explosive behaviour in the 1990s Nasdaq: When did exuberance escalate asset values? *International Economic Review, 52*, 201–226.
Phillips, P. C. B., & Yu, J. (2011). Dating the timeline of financial bubbles during the subprime crisis. *Quantitative Economics, 2*, 455–491.
Scheinkman, J., & Xiong, W. (2003). Overconfidence and speculative bubbles. *Journal of Political Economy, 111*, 1183–1219.
Scherbina, A., & Schlusche, B. (2014). Asset price bubbles: A survey. *Quantitative Finance, 14*(4), 589–604.
Shiller, R. J. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review, 71*, 421–436.
Wang, J., Zhu, J., & Dou, F. (2012). Who plays the key role among Shanghai, Shenzhen and Hong Kong stock markets? *China and the World Economy, 20*(6), 102–120.
Wang, Y.-C., Tsai, J.-J., & Lin, Y. (2016). The influence of Shanghai-Hong Kong connect on the Mainland and Hong Kong stock markets. *International Journal of Business and Finance Research, 10*(3), 1–10.
Wu, G., & Xiao, Z. (2008). Are there speculative bubbles in stock markets? Evidence from an alternative approach. *Statistics and its Interface, 1*, 307–320.
Yu, I-W., & Sze, A. (2012, December). *Testing for bubbles in the Hong Kong stock market*. Workshop in financial econometrics, Hong Kong Institute for Monetary research.
Zhang, B. (2008). Duration dependence test for rational bubbles in Chinese stock markets. *Applied Economics Letters, 15*, 635–639.
Zhang, S., Paya, I., & Peel, D. (2009). Linkages between Shanghai and Hong Kong stock indices. *Applied Financial Economics, 19*(23), 1847–1857.
Zhu, H., Lu, Z., & Wang, S. (2004). Causal linkages among Shanghai, Shenzhen, and Hong Kong stock markets. *International Journal of Theoretical and Applied Finance, 7*(2), 135–149.