Economic Theory (2007) 31: 293–306
DOI 10.1007/s00199-006-0100-z

RESEARCH ARTICLE

P. A. V. B. Swamy · George S. Tavlas

The new Keynesian Phillips curve and inflation expectations: re-specification and interpretation

Received: 10 November 2005 / Accepted: 8 March 2006 / Published online: 19 April 2006

Abstract A theoretical analysis of the new Keynesian Phillips curve (NKPC) is provided, formulating the conditions under which the NKPC coincides with a real-world relation that is not spurious or misspecified. A time-varying-coefficient (TVC) model, involving only observed variables, is shown to exactly represent the underlying "true" NKPC under certain conditions. In contrast, "hybrid" NKPC models, which add lagged-inflation and supply-shock variables, are shown to be spurious and misspecified. We also show how to empirically implement the NKPC under the assumption that expectations are formed rationally.

Keywords Time-varying-coefficient model · Inflation-unemployment trade-off · "Objective" probability · Spurious correlation · Rational expectation · Coefficient driver

JEL Classification Numbers C51 · E31 · E42 · E50

We are grateful to C. D. Aliprantis, Harris Dellas and Arnold Zellner for helpful comments. The questions and comments of an anonymous referee were extremely stimulating. The views expressed are the authors' own and do not constitute policy of their respective institutions.

P. A. V. B. Swamy
Bureau of Labor Statistics,
Postal Square Building,
Room # 4985, 2 Massachusetts Avenue, N.E.,
Washington, DC 20212, USA

G. S. Tavlas (✉)
Economic Research Department, Bank of Greece,
21 El. Venizelou St.,
102 50 Athens, Greece
E-mail: GTavlas@bankofgreece.gr

1 Introduction

A Phillips curve relation of some kind or other has long been at the center of much theoretical work in macroeconomics. In essence, the Phillips curve is a proposition about the effects of monetary policy, whereby changes in monetary policy push inflation and unemployment in opposite directions in the short run (see Clarida et al. 1999). Beginning with Phillips's original paper (1958) on the relation between percentage changes of money wages and the unemployment rate in the United Kingdom, "the" Phillips curve has undergone a number of modifications. The specific model that has received most attention in recent years is the "new" Keynesian Phillips curve (NKPC), which is based on a dynamic extension of static new Keynesian models of price adjustment. A number of researchers (e.g., Mankiw 2001; Gali 2003; Walsh 2003) have noted, however, that while the NKPC is appealing from a theoretical standpoint, empirical estimates of the NKPC have, by-and-large, not been successful in explaining the standard stylized facts about the dynamic effects of monetary policy, whereby monetary shocks are thought to first affect output, followed by a delayed and gradual effect on inflation. Mankiw (2001, p. C59), for example, characterized the state of empirical estimates of the NKPC as "ultimately a failure."[1] As discussed below, a consequence of this situation has been that recent empirical applications have modified the specification of the "pure" NKPC (that is, the version based on theory), yielding a "hybrid" Phillips relation.

This paper provides a theoretical analysis of the reasons why empirical estimates of the NKPC that can replicate the stylized facts have proved elusive. We show that the "pure" NKPC can be formulated in terms of a relationship that is not spurious or misspecified. In contrast, "hybrid" versions that augment the "pure" NKPC with the addition of (1) lagged inflation involved in traditional backward-looking models of inflation–unemployment dynamics and (2) a supply-shock variable, in an attempt to explain the standard stylized facts about the dynamic effects of monetary policy, are shown to be spurious and misspecified. Testing of the assumed NKPC employing a broad range of data is also discussed.

The reminder of this paper is divided into four sections. Section 2 establishes the connection between the NKPC and the underlying "true" model, stating the conditions needed for the existence of the "true" model. It uses this connection to derive the *explicit expressions* for the omitted-variable and measurement-error biases contained in the coefficients of certain operational versions (i.e., those based on certain proxies for expected future inflation and the natural rate of unemployment) of the NKPC. The section also shows that (1) there is very little role for lagged inflation in the NKPC once all the determinants of current inflation are included in the "true" model, and (2) forecasting future inflation with past inflation can be far from rational. Sections 3 and 4 provide a method of implementing the NKPC empirically, under the assumption that expectations are formed rationally. Specifically, section 3 defines "objective" probabilities required to define rational expectations of future inflation, since expectations of future inflation appear as an explanatory variable in the NKPC, while Section 4 presents a statistically efficient (in the sense of Lehmann and Casella 1998, p. 439) method of correcting for omitted-variable and measurement-error biases contained in the coefficients of the operational versions of the NKPC relation. Section 5 concludes.

[1] See, also, Walsh (2003, p. 241).

2 The new Keynesian Phillips curve

Unlike earlier versions of the Phillips curve, an attractive feature of the NKPC is that it is derived explicitly from a model of optimizing behavior on the part of price setters, conditional on the assumed economic environment (e.g., monopolistic competition, constant elasticity demand curves, and randomly-arriving opportunities to adjust prices) (see Walsh 2003, pp. 263–268). Perhaps the most popular formulation of nominal rigidities used in the derivation of the NKPC is due to Calvo (1983). In the Calvo model, firms follow time-contingent rules whereby price adjustment follows a random process. In any given period, a firm has a fixed probability that it will keep its price unchanged during that period, and, hence, one minus that probability that it will adjust prices. Each firm is assumed to have the same probability of being one of the firms to adjust price regardless of when it last adjusted its price.

To explain the use of the Calvo model in the derivation of the NKPC, we follow the approach proposed by Mankiw (2001), whose derivation is based on three relationships.[2] The first relationship concerns a firm's desired price, which is the price that would maximize profit at a particular point in time; the desired price depends on the overall price level and the deviation of unemployment from its natural rate. Price adjustment, however, is assumed to be infrequent and so firms generally do not set prices equal to desired prices. According to Mankiw's second relationship, when a firm has the opportunity to change its price, its adjustment price equals a weighted average of the current price and all future desired prices.[3] The third relationship concerns the current overall price level, which is assumed to be a weighted average of all the prices firms have set in the past. Given these three relationships, the following equation can be derived:

$$\dot{p}_t = \beta E_t \dot{p}_{t+1} + \eta(U_t - U_t^{\mathrm{n}}), \tag{1}$$

where $\dot{p}_t$ is the inflation rate, $E_t \dot{p}_{t+1}$ is the inflation rate expected in the current period for the next period, U_t is the unemployment rate, U_t^{n} is the natural rate of unemployment, and t indexes time.[4] Equation (1) is considered a "pure" NKPC.[5]

Because of the above-noted failure of the estimated versions of (1), recent empirical work on the NKPC has involved several modifications to the basic, or "pure", specification (see, e.g., Staiger et al. 1997; Gordon 1998; Mankiw 2001). The main variations include the following. First, a measure of real marginal cost or

[2] We provide only an intuitive description of these relationships; for a formal statement, see Mankiw (2001).

[3] In a personal communication, A. Zellner wrote that, based on his experience studying pricing problems in many industries, a more sophisticated theory of pricing is needed to improve Phillips curves. Such an improved Phillips-curve formulation would take account of possible entry and exit of firms, firm interactions, possible actions by industrial regulators, appropriate formation of expectations, etc. The issue of expectations formation is dealt with later.

[4] In Mankiw's theoretical derivarion of equation (1), $\beta = 1$. Other theoretical formulations are such that $0 < \beta < 1$ (see, e.g., Gali 2003). O.J. Blanchard and J. Gali (unpublished manuscript) derived a formulation based on the difference between current employment and potential employment, which they refer to as the "employment gap".

[5] In contrast to the Lucas (1973) imperfect information model, in the new Keynesian model firms set nominal price based on the expectations of future marginal costs, where the variable, $U_t - U_t^n$, captures movements in marginal costs associated with variations in excess demand.

a measure of de-trended output is typically used in place of the unemployment gap, $U_t - U_t^{\mathrm{n}}$, used in (1). Second, to capture such unexpected shocks as the oil-price hikes of the 1970s, a supply shock variable, denoted in what follows by ν_t, is often included in the specification. Third, to empirically implement the NKPC something needs to be assumed about how expectations of future inflation are formed. Much econometric work assumes that expectations are formed backwardly, on the basis of previous information. For example, some writers (see, e.g., Staiger et al. 1997; Ball and Mankiw 2002) use the previous period's inflation rate as a proxy for expected inflation based on the supposition that in the United States over the past four decades inflation has been a random walk. In such conditions, it is argued that forecasting future inflation with past inflation is not far from rational (see Ball and Mankiw 2002, p. 119). Fourth, in order to capture the inflation persistence, which some authors believe is contained in the data, and apart from the particular inflation-expectations mechanism assumed, it is common to augment the basic forward-looking inflation adjustment equation with the addition of lagged inflation. Incorporating the second and fourth modifications described above yields a hybrid Phillips curve of the following form (see Gali 2003; Walsh 2003, p. 242):

$$\dot{p}_t = \theta\beta E_t \dot{p}_{t+1} + \eta(U_t - U_t^{\mathrm{n}}) + (1-\theta)\dot{p}_{t-1} + \nu_t, \tag{2}$$

where the parameter θ is typically described to be a measure of the degree of forward-looking behavior, so that $(1-\theta)$ is a measure of backward-looking behavior.

2.1 Effects of nonlinearities

Beginning with Phillips's original estimates, a good deal of empirical work has found that the inflation–excess–demand relation is nonlinear. The curve is convex with respect to the origin under certain conditions. These conditions are as follows: increases in demand lead to diminishing marginal returns; successive uniform declines in the unemployment rate require larger increments in excess demand and, thus, in inflation rates, to achieve a given decline in the unemployment rate. This convexity can itself lead to shifts in the Phillips relation. Therefore, it is important that the NKPC relation has the correct functional form. Otherwise, shifts in the relation may be due to the effects of an incorrectly-specified functional form.

Because causal relationships are not spurious, our aim is to reformulate the NKPC in terms of a causal relationship. To assess whether the NKPC in (1) and the hybrid Phillips curve in (2) are causal or spurious, we first address the issue of functional forms. A straightforward way of capturing the unknown nonlinear functional form of (1) is to allow its coefficients to vary freely. That is, a purpose of allowing the coefficients to vary is to capture the unknown functional form (see Swamy and Tavlas 2001). Following this approach, we write

$$\dot{p}_t = \gamma_{0t} + \gamma_{1t}x_{1t} + \gamma_{2t}x_{2t}, \tag{3}$$

where x_{1t} is a proxy for $E_t\dot{p}_{t+1}$, x_{2t} is a proxy for $U_t - U_t^{\mathrm{n}}$, and the errors in these proxies as well as the correct definitions of γ's are explicitly dealt with below.[6] Equation (3) is referred to as "the time-varying coefficients (TVC) model." It is not necessarily linear, since the TVCs permit the equation to pass through every data point. Thus, with TVCs, the equation can be nonlinear.

[6] The proxies, x_{1t} and x_{2t}, are defined in section 4 below.

2.2 A real-world relation

We treat the observed measurements, $\dot{p}_t = \dot{p}_t^* + \upsilon_{0t}$ and $x_{jt} = x_{jt}^* + \upsilon_{jt}$, $j = 1, 2$, as the sums of (unobserved) "true" values and (unknown) measurement errors. The symbols with an asterisk denote "true" values: $\dot{p}_t^*$ = "true" value of the inflation rate, $x_{1t}^* = E_t \dot{p}_{t+1}^*$ and $x_{2t}^* = U_t - U_t^{\mathrm{n}}$.[7] The symbols, $\dot{p}_t$ and x_{jt}, $j = 1, 2$, without an asterisk denote observable variables. The symbols, υ_{jt}, $j = 0, 1, 2$, denote measurement errors.

Theorem 1 *The sufficient conditions for the TVC model to be an exact representation of the "true" model,*

$$\dot{p}_t^* = \alpha_{0t}^* + \alpha_{1t}^* x_{1t}^* + \alpha_{2t}^* x_{2t}^* + \sum_{g=3}^{m_t} \alpha_{gt}^* x_{gt}^*, \tag{4}$$

linking the "true" variables involving the "true" coefficients, are that

$$\gamma_{0t} = \alpha_{0t}^* + \sum_{g=3}^{m_t} \alpha_{gt}^* \lambda_{0gt}^* + \upsilon_{0t} \tag{5}$$

and

$$\gamma_{jt} = \left(\alpha_{jt}^* + \sum_{g=3}^{m_t} \alpha_{gt}^* \lambda_{jgt}^*\right)\left(1 - \frac{\upsilon_{jt}}{x_{jt}}\right) \quad (j = 1, 2) \tag{6}$$

for all t, where λ_{jgt}^*, $j = 0, 1, 2$, are the "true" coefficients of the "auxiliary" regressions of excluded variables on the included explanatory variables, x_{1t} and x_{2t},

$$x_{gt}^* = \lambda_{0gt}^* + \sum_{j=1}^{2} \lambda_{jgt}^* x_{jt}^* \quad (g = 3, \ldots, m_t) \tag{7}$$

Proof See Chang et al. (2000) and Swamy and Tavlas (2005). □

The following remarks (i)–(iv) clarify equations (3)–(7):

(i) The TVC model in (3) is based on observed dependent and explanatory variables, while the "true" model in (4) is based on "true", but unobserved, dependent and explanatory variables.[8] We treat $\dot{p}_t$ and x_{jt}, $j = 1, 2$, in the TVC model as the included (dependent and explanatory) variables and x_{gt}^*, $g = 3, \ldots, m_t$, in the "true" model as excluded variables for the simple reason they are included in, and excluded from, the TVC model, respectively. There can be infinitely many excluded variables. (The supply shock, ν_t, in (2) is an example of such an excluded variable. According to the pure NKPC, past values of x_{1t}^*, x_{2t}^*, and ν_t included in (2) are not the determinants of $\dot{p}_t^*$ and hence cannot be considered as excluded variables.) Equation (7) is a regression of an excluded variable on the "true" values of all the included explanatory variables, allowing the coefficients (the λ^*'s) to change

[7] As will become apparent, the switch in symbols is made to economize on notation.

[8] Possible objections to our use of the term "true model" are addressed below.

over time, so that any nonlinearities of the regression are captured. Assumption (7) is made because, contrary to much of the Phillips-curve literature, the included explanatory variables cannot be considered independent of 'the' excluded variables, such as ν_t.[9]

(ii) We assume that for an unspecified value of m_t, the explanatory variables on the right-hand side of (4) are all the determinants of $\dot{p}_t^*$ for all t. In other words, there are no determinants of $\dot{p}_t^*$ excluded from (4). Also, the number, m_t, may depend on time, indicating that the number of the determinants of $\dot{p}_t^*$ may change over time. Equation (4) is purely conceptual. This conceptual model (4) is "true" because, by construction, it is correctly specified since none of the determinants of $\dot{p}_t^*$ is excluded, none of the included determinants of $\dot{p}_t^*$ is mismeasured, and the model has the correct functional form (because its coefficients are assumed to have the correct, but unknown, time profiles). Thus, although not much may be known about the "true" model, it can be used as an algebraic device to derive the mapping, (5) and (6), between the coefficients of the TVC and "true" models, without making an incorrect assumption about the functional form of the "true" model. By allowing the coefficients (α^*'s) of the "true" model to follow the correct, but unspecified, time profiles, these coefficients are used to express our ignorance – and what we would like to know – about the "true" functional form. The coefficients, α_{1t}^* and α_{2t}^*, are the correct forms of the coefficients, β and η, respectively, of equation (1) because they appear in the "true" model with the correct, but unknown, time profiles.

(iii) Clearly, unless it exists, the "true" model cannot generate our data on the included variables and cannot be called "the data-generating model". In essence, we need to impose the conditions for the existence of the "true" model. To formulate an existence condition, in what follows we define 'potential values'.

The existence of a model means that it is a real-world relation. As Basmann (1988, p. 99) has pointed out, causation is a real-world relation between events rather than a mere property of its linguistic representation. Only when the "true" model exists, can it be considered a causal law. We say that $\dot{p}_t^*$ – the "true" value – is related to the "true" values of its determinants, x_{jt}^*, $j = 1, 2$, and x_{gt}^*, $g = 3, \ldots, m_t$, by a law if for every vector of values of the determinants there exists a vector of values of the coefficients, α_{jt}^*, $j = 1, 2$, and α_{gt}^*, $g = 3, \ldots, m_t$, which on the tth observation associates with every vector of possible values of the determinants a value of $\dot{p}_t^*$. We call these values of $\dot{p}_t^*$, defined for every vector of values of its determinants on every observation t, 'potential' values if they are unobserved. The only value of $\dot{p}_t^*$ that will be realized on any one observation t is that corresponding to the one vector of the realized values of its determinants. It is only when the potential values exist that the "true" model can be considered

[9] The argument leading up to the adoption of (7) is due to Pratt and Schlaifer (1988, p. 34) who show that "the condition [that the included explanatory variables be independent of 'the' excluded variables themselves] is meaningless unless the definite article is deleted and can then be satisfied only for certain 'sufficient sets' of excluded variables some if not all of which must be defined in a way that makes them unobservable as well as unobserved". Pratt and Schlaifer (1988, p. 47) caution that because the value of $\dot{p}_{t-1}$ was determined before the value of the current joint effect ν_t of excluded variables, it must not be assumed that $\dot{p}_{t-1}$ necessarily satisfies the condition that it was independent of ν_t. It may well have been influenced by a forecast of an excluded variable represented in ν_t, or both $\dot{p}_{t-1}$ and ν_t may have been affected by some third variable – in common parlance, a 'common cause'. We heed this caution in this paper.

to be a causal law.[10] As shown by Pratt and Schlaifer (1988, p. 28), it is the existence of potential values that distinguishes a law from a statistical association. In other words, if the potential values of $\dot{p}_t^*$ do not exist, the observed relationship between $\dot{p}_t$ and the included explanatory variables is a pure statistical artifact (i.e., a spurious regression).[11] With the existence condition, the coefficients of the "true" model – i.e., equation (4) – can be interpreted as follows: For $j > 0$, the quantity, $\alpha_{jt}^* + (\partial\alpha_{jt}^*/\partial x_{jt}^*)x_{jt}^*$, measures the causal – or, what we call the "direct" – effect of x_{jt}^* on $\dot{p}_t^*$, with all the determinants of $\dot{p}_t^*$ other than x_{jt}^* held constant.[12]

(iv) Having the causal law, i.e., the "true" model satisfying the existence condition, how should we use it? This issue has been addressed by Zellner and Basmann. Building on the work of Feigl (1953), Zellner (1979, 1988) defined causality in terms of 'predictability according to a law or set of laws'. However, Basmann (1988, p. 91) argued that 'causality' is not definable in the explicit sense, but that it is an *open* term, only partially characterized by the domain-specific *interpretative systems* in which it is used. We take account of both Feigl's definition and Basmann's view – i.e., causality is not definable, but can only be partially represented by an empirical model – so that causality implied by the NKPC is predictability according to, or is fully characterized by, the "true" (or correctly specified) model, whenever the "true" model exists.[13]

Remarks (i)–(iv) prove the following:

Theorem 2 *Causality is only partially characterized by the NKPC in (1) if (i) the "true" model in (4) exists and (ii) the coefficients,* α_{1t}^* *and* α_{2t}^**, on* $E\dot{p}_{t+1}^*$ *and* $U_t - U_t^{\mathrm{n}}$*, in the "true" model satisfy the restrictions,* $0 < \alpha_{1t}^* \leq 1$ *and* $-1 < \alpha_{2t}^* < 0$*, respectively, for all t. In the alternative case where* $\alpha_{1t}^* = 0$ *and* $\alpha_{2t}^* = 0$ *for all t, the NKPC is spurious.*

The theoretical literature on the NKPC considers the intervals of the values of α_{1t}^* and α_{2t}^* given in Theorem 2 to be theoretically correct (see, e.g., Ball and Mankiw 2002; Walsh 2003, p. 241).

[10] The extension of Neyman's potential outcome notation to define causal effects in both non-randomized and randomized studies is due to Rubin (2005).

[11] The idea is that we do not know whether the potential values exist. To form a causal law, we need to assume that they do exist.

[12] Zellner drew our attention to a widely accepted view that there is no such thing as a "true" model. An implication of this view is that the "true" model in (4) never exists and is, therefore, fictitious. Our use of the terminology "true model" does not run counter to this view, since we allow for the possibility that the "true" model may never satisfy the existence condition. Furthermore, the "true" model is not empirically implementable. This is what makes it non-falsifiable. The widely accepted view notwithstanding, the problems of spurious correlations and specification biases due to functional form misspecifications, omitted variables, and measurement errors are considered in the econometric literature. As will become apparent, these problems cannot be solved satisfactorily without considering the "true" models of the type (4).

[13] Some elaboration may be helpful. Remark (iii) above imposes the existence condition on the "true" model so that the "true" model is a real-world relation. This real-world relation is a causal law because, as Basmann has pointed out, causation is a real-world relation between events. Basmann also has pointed out that (i) causation is not definable in the explicit sense, and (ii) it can only be characterized by the real-world relation in which it is used. Following both Zellner and Basmann we state that causality between the dependent variable ($\dot{p}_t^*$) and its determinants in equation (4) is predictability according to, or is fully characterized by the causal law or the real-world relation which equals the "true" (or correctly specified) model (4), when the "true" model exists.

2.3 Interpretation of the intercept of the TVC model

Equations (4)–(7) can be used to give the correct interpretations of the coefficients of the TVC model, equation (3), as follows.[14]

Proposition 1 *As shown by the mapping in (5), the intercept, γ_{0t}, of the TVC model is the sum of (i) the intercept, α^*_{0t}, of the "true" model (or the "true" component), (ii) the joint effect, $\sum_{g=3}^{m_t} \alpha^*_{gt}\lambda^*_{0gt}$, on $\dot{p}^*_t$ of the portions of the "true" values of excluded variables remaining after the effects of the "true" values of included explanatory variables have been removed, and (iii) the measurement error (i.e., υ_{0t}) in $\dot{p}_t$.*

Comparing the TVC model in (3) with the NKPC in (1) shows that the term, ν_t, which appears in equation (2) in order to capture supply shocks that are omitted from equation (1), is equal to the term, $\sum_{g=3}^{m_t} \alpha^*_{gt}\lambda^*_{0gt}$, of γ_{0t}. As Proposition 1 shows, this term of γ_{0t} represents the joint effect on $\dot{p}^*_t$ of the portions of the "true" values of excluded variables remaining after the effects of the "true" values of included explanatory variables have been removed and not the net effect of explanatory variables omitted from (1). Therefore, the claim that the term, ν_t, in equation (2) can capture supply shocks is incorrect.

2.4 Omitted-variable and measurement-error biases

Proposition 2 *As shown by the mapping in (6), for $j > 0$, the jth coefficient, γ_{jt}, of the TVC model is the sum of (i) the jth coefficient, α^*_{jt}, of the "true" model (or the "true" component), (ii) a term, $\sum_{g=3}^{m_t} \alpha^*_{gt}\lambda^*_{jgt}$, capturing omitted-variables bias due to excluded variables, and (iii) a measurement-error bias, $-\left(\alpha^*_{jt} + \sum_{g=3}^{m_t} \alpha^*_{gt}\lambda^*_{jgt}\right)(\upsilon_{jt}/x_{jt})$, due to mismeasuring the jth included explanatory variable.*

This proposition shows that the coefficients of (3), the TVC model, are biased, in general, because some determinants of $\dot{p}^*_t$ are omitted from (3) or because the determinants of $\dot{p}^*_t$ included in (3) are measured with error. Estimates of the coefficients of the NKPC in (1) or of the hybrid Phillips curve in (2) can have both incorrect signs and magnitudes unless the omitted-variable and measurement-error biases which they contain are completely removed. We refer to the coefficients of the "true" model, equation (4), as "bias-free" because they do not contain any biases.

2.5 Uniqueness of the coefficients of the TVC model

Proposition 3 *Suppose that the coefficients of the TVC model satisfy mappings (5) and (6). Then rewriting the "true" model in terms of the included explanatory variables and a function of excluded and included explanatory variables, leaves the coefficients of the TVC model invariant and hence these coefficients are unique.*

[14] For the derivation of Propositions 1 and 2 below, see Chang et al. (2000).

Proof See Swamy et al. (1996). □

The coefficients and error terms in (1) and (2) are not unique because they do not satisfy equations (5) and (6). The coefficients of the TVC model are non-unique if their correct interpretations given in Propositions 1 and 2 are contradicted by the assumptions made about them. For example, the coefficients of the TVC model are non-unique when omitted-variable and measurement-error biases, which they contain, are not dealt with. The true assumptions needed to ascertain non-unique coefficients may not exist.

If the correct functional form of the "true" model is nonlinear, then its coefficients are time-varying, in which case the following results hold: (1) The first difference of $\dot{p}_t^*$ is not stationary and therefore, $\dot{p}_t^*$ cannot be a random walk.[15] (2) The supposition that inflation in the United States over the past four decades has been a random walk is incorrect. (3) Forecasting future inflation with past inflation is far from rational. (4) Autoregressive integrated moving average models of inflation that require the assumption that inflation becomes stationary after being first differenced d times are inconsistent with the "true" model.[16]

2.6 Spuriousness of the hybrid Phillips curve

Lagging equation (4), the "true" model, by one period yields $\dot{p}_{t-1}^* \equiv \alpha_{0,t-1}^* + \sum_{j=1}^{2} \alpha_{j,t-1}^* x_{j,t-1}^* + \sum_{g=3}^{m_{t-1}} \alpha_{g,t-1}^* x_{g,t-1}^*$. Also, recall that according to the "true" model, past values of x_{1t}^* $(= E_t \dot{p}_{t+1}^*)$, x_{2t}^* $(= U_t - U_t^n)$, and v_t $(= x_{gt}^*)$, are not the determinants of current inflation. Under such conditions, $\dot{p}_{t-1}^*$ can be highly correlated with $\dot{p}_t^*$ *without being a determinant of* $\dot{p}_t^*$. Such correlations, however, do not imply causality leading to the following results: (1) The "true" component (defined in Proposition 2 as α_{jt}^*) of the coefficient on the lagged inflation in the hybrid Phillips curve in (2) is zero. (2) An efficient estimate of this coefficient should be insignificant when it is appropriately corrected for omitted-variables and measurement-error bias. (3) There is very little role for lagged inflation once all the determinants of current inflation are included in the "true" model in their correct form. These results together with Proposition 2 prove the following:

Theorem 3 *Any nonzero value of the coefficient on the lagged inflation in the hybrid Phillips curve in (2) arises due to a spurious correlation if the "true" model in (4) exists. The hybrid Phillips curve in (2) is misspecified if the specification biases due to omitted variables, measurement errors, and incorrect functional forms contained in its coefficients are not taken into account.*

3 Rational expectations

This section and the next show how the NKPC might be estimated under the assumption that expectations are formed rationally. If we assume that individual

[15] Taking the first differences of both sides of (4) with time-varying coefficients demonstrates this point.

[16] Forecasts from inconsistent models are incoherent in a Bayesian sense (see de Finetti 1974).

expectations about future inflation involved in the NKPC in (1) are rational, then we should not assume that they depend on recently observed inflation. The reason is that forecasting future inflation with past inflation can be far from rational, as we have shown above.

Definition 1 (Rational expectations) *Individual expectations about future inflation are rational if they agree with the forecasts generated from the "true" model.*

This definition means that rational expectations cannot be formed unless one knows how to generate forecasts from the "true" model with all the ignorance he or she has about the "true" model. To generate forecasts from the "true" model, we need the conditional probability distribution of $\dot{p}_t^*$, given all the determinants of $\dot{p}_t^*$. This conditional distribution is "objective" if it can be derived from the "true" model without the aid of any subjective priors and restrictions (see Swamy and Tavlas 2006).

Definition 2 ("Objective" probabilities) *Let χ denote the set of all (realizable and potential) values of $\dot{p}_t^*$ obeying the "true" model and let A be a σ-field of subsets of χ. Let P be a probability measure defined over the measure space (χ, A). Determinations of the functions, $P(\dot{p}_t^* \in A|x_{jt}^*, j = 1, 2, x_{gt}^*, g = 3, \ldots, m_t)$, exist which, for each set of fixed x_{jt}^*, $j = 1, 2, x_{gt}^*$, $g = 3, \ldots, m_t$, define a conditional probability when χ is Euclidean (see Lehmann and Casella 1998, p. 35).*

Any additional restrictions on the "true" model may make these conditional probabilities subjective. If the variance of the "objective" conditional distribution of $\dot{p}_t^*$ in Definition 2 is finite, then its mean is the minimum average mean square error (or best) predictor of $\dot{p}_t^*$ (see Rao 1973, p. 264). Economic agents cannot use this mean if they do not know the "objective" conditional distribution of $\dot{p}_t^*$. However, they can use the TVC model under certain assumptions about its coefficients. The forecasts from the TVC model can agree with the best forecasts from the "objective" conditional distribution of $\dot{p}_t^*$. In the next section, we show how such a situation might arise.

4 Efficient estimation of the TVC model

Let x_{1t} in (3) represent an irrational expectation about $\dot{p}_{t+1}$. Examples of such expectations are: (1) last period's inflation, $\dot{p}_{t-1}$, (2) the Michigan and Livingston survey measures of expected inflation, and (3) the Federal Open Market Committee's (FOMC) inflation forecasts in Greenbook. What makes any of these x_{1t} an irrational expectation is the measurement error, υ_{1t}, it contains. We will show in this section that irrational expectations about $\dot{p}_{t+1}$ can be used to estimate the TVC model efficiently, provided the appropriate omitted-variable and measurement-error biases are taken into account.

Let x_{2t} be a measure of de-trended U_t, i.e., deviations of U_t from a smooth trend. Examples of estimates of the trend are: (1) a fitted quadratic function of time, (2) the Hodrick-Prescott filter with a smoothing parameter, and (3) a moving average of the U_t's.

It is important that the TVC model is estimated under the assumptions that are consistent with the correct interpretations of its coefficients; otherwise inconsistencies arise. The coefficients of the TVC model cannot be constant if their components on the right-hand side of (5) and (6) vary over time. The real-world sources of variation in these components are: (1) the nonlinearities of the "true" model causing variation in the α^*'s, (2) the nonlinearities of the "auxiliary" regressions of excluded variables on the included explanatory variables causing variation in the λ^*'s, (3) changes in m_t, (4) changes in the variables that constitute m_t, and (5) variations in υ_{0t} and (υ_{jt}/x_{jt}) with $j > 0$. In the TVC model, all the γ_{jt}'s involve the coefficients (α^*_{gt}) on excluded variables and for $j > 0$, γ_{jt} involves the jth included explanatory variable, x_{jt}, and the measurement error υ_{jt}. The implications of these interpretations are that in the TVC model, the coefficients including the intercept are correlated with each other and the included explanatory variables are correlated with their own coefficients. Any explanatory variable that is correlated with its own coefficient cannot be exogenous. These are the implications of the correct interpretations of the coefficients of the TVC model. They are the prime considerations guiding the selection of the features of the TVC model that ought to be treated as constant parameters.

Assumption 1 *The coefficients of the TVC model satisfy the equation*

$$\gamma_{jt} = \pi_{j0} + \sum_{d=1}^{p-1} \pi_{jd} z_{dt} + \varepsilon_{jt} \quad (j = 0, 1, 2), \tag{8}$$

where for all j, d, and t, $z_{dt} \neq 1$ and ε_{jt} and x_{jt} are conditionally independent given z_{dt} and the mean of ε_{jt} is zero. The ε_{jt} may be contemporaneously and serially correlated.

The z's have been called "the coefficient drivers" (see Swamy and Tavlas, 2006). We use these coefficient drivers to decompose each coefficient of the TVC model into its components. For example, we assume that for $j > 0$, the sum of $p_1 (< p)$ specific terms of $\pi_{j0} + \sum_{d=1}^{p-1} \pi_{jd} z_{dt}$ is equal to the "true" component, α^*_{jt}, of γ_{jt} and the sum of the remaining terms on the right-hand side of equation (8) is equal to the sum of omitted-variables and measurement-error bias components of γ_{jt}. From this assumption it follows that only those coefficients of the "true" model that appear as the "true" components of the coefficients on the included explanatory variables in the TVC model are identifiable – subject to the restrictions implied by (8) – on the basis of the available data on the included variables, whereas the "true" coefficients on excluded variables in the "true" model are not identifiable.

Assumption 1 does not contradict the implications of the correct interpretations of the coefficients of the TVC model if (1) the function, $\pi_{j0} + \sum_{d=1}^{p-1} \pi_{jd} z_{dt}$, completely accounts for the correlation between x_{jt} and γ_{jt} so that the remainder, ε_{jt}, obtained by subtracting $\pi_{j0} + \sum_{d=1}^{p-1} \pi_{jd} z_{dt}$ from γ_{jt} is independent of the x_{jt}, given the z_{dt}, and (2) the right-hand side of (8) is expressible as the sum of two sums, one of which is equal to the "true" component of γ_{jt} and the other of which is equal to the sum of omitted-variables and measurement-error bias components of γ_{jt}. The satisfaction of these conditions should underpin the selection of the coefficient drivers. In what follows, we will call the coefficient drivers that satisfy these conditions "the proper coefficient drivers".[17]

If the parameters of the "decision rules" embodied in the TVC model change when economic policies change, then it is appropriate to use the relevant policy changes as coefficient drivers in (8). With the relevant policy changes entering into (8) as coefficient drivers, the TVC model is not subject to "the Lucas (1976) critique".

Substituting (8) into (3) gives the reduced form

$$\dot{p}_t = \pi_{00} + \sum_{d=1}^{p-1} \pi_{0d} z_{dt} + \sum_{j=1}^{2} \pi_{j0} x_{jt} + \sum_{j=1}^{2} \sum_{d=1}^{p-1} \pi_{jd} z_{dt} x_{jt} + \varepsilon_{0t} + \sum_{j=1}^{2} \varepsilon_{jt} x_{jt}. \quad (9)$$

The right-hand side of this equation with the error terms suppressed gives the conditional expectation of $\dot{p}_t$ as a linear function of the included explanatory variables, the proper coefficient drivers, and their interactions.[18] Equation (9) is our assumed NKPC derived from the TVC and "true" models. We can estimate our assumed model. We can draw inferences about the bias-free coefficients, α^*_{jt}, $j = 0, 1, 2$, of the "true" NKPC model using our assumed model because of the connections between these two models shown in (5), (6), and (8). Whether our assumed model provides a good approximation to the TVC or "true" model will depend critically on the choice of coefficient drivers.[19] Note that the instrumental variables that are highly correlated with the x_{jt} and uncorrelated with the error terms of (9) do not exist because the x_{jt} appear in both the random and systematic parts of (9). Thus, we have demonstrated

[17] Obviously, an exhaustive search of the entire set of possible coefficient drivers is not possible. While programs that search an adequate set of coefficient drivers can be designed, ultimately, the search also depends on the domain-specific (empirical) knowledge and expertise of the researcher.

[18] These conditional expectations differ from those given in econometrics textbooks (see, e.g., Greene 2003) because they explicitly account for biases due to omitted variables, measurement errors, and incorrect functional forms.

[19] In a personal communication on an earlier draft of this paper, A. Zellner made the following comments: (i) It has long been recognized that no model is absolutely true. (ii) There is always the possibility that some other model may perform better over both the past and new ranges of data, and be rationalized by a different theory. We are glad that Zellner made these comments, and would add that any model with excluded variables, mismeasured variables, and incorrect functional forms cannot be absolutely true. The "true" model in (4) is not this type of model and hence Zellner's comments (i) and (ii) do not apply to it. By contrast, Newton's laws and Einstein's more general "laws" have some relevant variables excluded from them. For this reason, physicists have not been able to prove that these laws are absolutely true. To have our assumed model for a dependent variable provide a good approximation to the underlying "true" model whenever the "true" model exists, we do the following: (i) formulate an algebraic form of the "true" model of the dependent variable following the correct definition of the "true" model, (ii) impose the existence condition on the "true" model so that the "true" model represents a real-world relation or a causal law, (iii) state a TVC model that involves only the observable dependent and explanatory variables, (iv) derive the exact algebraic expressions for the omitted-variable and measurement-error biases contained in the coefficients of the TVC model without making an incorrect assumption about the functional form of the "true" model, and (vi) subtract from the estimated coefficients of the TVC model the estimates of the biases contained in them to obtain the estimates of the identifiable coefficients of the "true" model. In contrast to this approach, Rubin's (2005) model-based Bayesian framework for causal inference does not deal with omitted-variable and measurement-error biases and the unknown functional-form problem.

Proposition 4 *After the proper coefficient drivers in (8) account for the correlations between the included explanatory variables and their coefficients in the TVC model, the instrumental variables used in Greene (2003, pp. 75–90) to obtain statistical consistency but not statistical efficiency cease to exist.*[20]

An iteratively rescaled generalized least squares (IRSGLS) method developed in Chang et al. (2000) can be used to estimate (9) under Assumption 1. The IRSGLS estimates of α^*_{jt}, $j = 0, 1, 2$, can be used to validate the NKPC. A demonstration that these estimates remained close to the values, $\alpha^*_{0t} = 0$, $0 < \alpha^*_{1t} \leq 1$, and $-1 < \alpha^*_{2t} < 0$, respectively, when one set of proper coefficient drivers after another is introduced into the analysis raises confidence or reasonable belief in the NKPC.[21]

Under (5) and (6), the TVC model is an exact representation of the "true" model, as Theorem 1 shows. This result implies that the predictions of future inflation from the TVC model agree with those from the "true" model and hence are rational if the coefficients of the TVC model satisfy (5) and (6). If we can set up (8) so that (5) and (6) are satisfied, then the forecasts of future inflation from (9) with proper coefficient drivers are rational.

5 Conclusions

The lack of success of estimated versions of the pure NKPC in explaining the standard stylized facts about the dynamic effects of monetary policy has resulted in a proliferation of "hybrid" NKPC's, which augment the pure relation with lagged-inflation and supply-shock variables. In this paper, we showed that the apparent empirical successes of such "hybrid" specifications, in terms of their ability to yield significant coefficients on the augmented variables, are likely to reflect spurious correlations and specification biases due to (1) incorrect functional forms, (2) omitted variables, and (3) measurement errors. We also show that forecasting future inflation with past inflation can be far from rational. The feasibility of empirically implementing the NKPC relation under the assumption that expectations are fully rational is demonstrated.

References

Ball, L., Mankiw, N. G.: The NAIRU in theory and practice. J Econ Perspect **16**, 115–136 (2002)
Basmann, R. L.: Causality tests and observationally equivalent representations of econometric models. J Econ Ann **39**, 69–104 (1988)
Calvo, G. A.: Staggered prices in a utility maximizing framework. J Monetary Econ **12**, 383–398 (1983)
Chang, I., Swamy, P.A.V.B., Hallahan, C., Tavlas, G.S.: A computational approach to finding causal economic laws. Comput Econ **16**, 105–136 (2000)
Clarida, R., Gali, J., Gertler, M.: The science of monetary policy: a new Keynesian perspective. J Econ Literat **37**, 1661–1707 (1999)

[20] This result obtains if Assumption 1 is true.

[21] For an empirical application of (9) under Assumption 1, see G. Hondroyiannis, P.A.V.B. Swamy, G.S. Tavlas (unpublished manuscript) who provided estimates of the NKPC for four countries: France, Germany, Italy, and United Kingdom.

de Finetti, B.: The theory of probability. New York: Wiley 1974
Feigl, H.: Notes on causality. In: Feigl, H., Brodbeck, M. (eds.) Readings in the philosophy of science. New York: Appleton-Century-Crofts, Inc. 1953
Gali, J.: New perspective on monetary policy, inflation, and the business cycles, pp 151–197. In: Dewatripont, M., Hansen, L., Turnovsky, S. (eds.) Advances in economic theory. Cambridge: Cambridge University Press 2003
Gordon, R.J.: Foundations of the Goldilocks economy: supply shocks and the time-varying NAIRU. Brookings Papers Econ Act **2**, 297–333 (1998)
Greene, W.H.: Econometric analysis, 5th edn. Upper Saddle River: Prentice Hall (2003)
Lehmann, E.L., Casella, G.: Theory of point estimation, 2nd edn. Berlin Heidelberg New York: Springer (1998)
Lucas, R.E., Jr.: Some international evidence on output-inflation trade-off. Am Econ Rev **63**, 326–334 (1973)
Lucas, R.E., Jr.: Econometric policy evaluations: a critique. Carnegie-Rochester Ser Public Policy **1**, 19–46 (1976)
Mankiw, N.G.: The inexorable and mysterious tradeoff between inflation and unemployment. Econ J **111**, C45–C61 (2001)
Phillips, A.W.: The relationship between unemployment and the rate of change of money wages in the United Kingdom, 1861–1957. Economica **25**, 283–299 (1958)
Pratt, J.W., Schlaifer, R.: On the interpretation and observation of laws. J Econ Ann **39**, 23–52 (1988)
Rao, C.R.: Linear statistical inference and its applications, 2nd edn. New York: Wiley (1973)
Rubin, D.B.: Causal inference using potential outcomes: design, modeling, decisions. J Am Stat Assoc **100**, 322–331 (2005)
Staiger, D., Stock, J.H., Watson, M.W.: How precise are estimates of the natural rate of unemployment? pp. 195–246. In: Romer, C.D., Romer, D.H. (eds.) Reducing inflation: motivation and strategy. Chicago: University of Chicago Press 1997
Swamy, P.A.V.B., Mehta, J.S., Singamsetti, R.N.: Circumstances in which different criteria of estimation can be applied to estimate policy effects. J Stat Plan Inference **50**, 121–153 (1996)
Swamy, P.A.V.B., Tavlas, G.S.: Random coefficient models, Chap. 19. In: Baltagi, B.H. (ed.) A companion to theoretical econometrics. Malden: Blackwell 2001
Swamy, P.A.V.B., Tavlas, G.S.: Theoretical conditions under which monetary policies are effective and practical obstacles to their verification. Econ Theory **25**, 999–1005 (2005)
Swamy, P.A.V.B., Tavlas, G.S.: A note on Muth's rational expectations hypothesis: A time-varying coefficient interpretation. Macroecon Dyn **10**, (2006) forthcoming
Walsh, C. E.: Monetary theory and policy, 2nd edn. Cambridge: MIT Press (2003)
Zellner, A.: Causality and econometrics, 9–54. In: Brunner, K., Meltzer, A.H. (eds.) Three aspects of policy and policymaking. Amsterdam: North-Holland 1979
Zellner, A.: Causality and causal laws in economics. J Econ Ann **39**, 7–21 (1988)