University *of* Leicester

DEPARTMENT OF ECONOMICS

TIME-VARYING COEFFICIENT ESTIMATION IN THE PRESENCE OF NON-STATIONARITY

Stephen G. Hall, University of Cambridge, UK

P. A. V. B. Swamy, Retired from Federal Reserve Board, USA

George S. Tavlas, Bank of Greece, Greece

George Hondroyiannis, Harokopio University, Athens

Working Paper No. 09/13

September 2009

Time-Varying Coefficient Estimation in the Presence of Non-Stationarity

Stephen G. Hall , P. A. V. B. Swamy, and George S. Tavlas*

and

George Hondroyiannis

Abstract Time-varying coefficient (TVC) estimation is a technique that has been developed to produce consistent estimates of parameters in the simultaneous face of measurement errors, unknown functional form and omitted variables. Previous work on the technique has not paid explicit attention to the issue of non-stationarity. This paper outlines the basic stages of the technique and discusses in detail how the issue of non-stationarity and cointegration affect each stage of the TVC estimation procedure.

Keywords Specification Problem · Correct interpretation of coefficients · Appropriate assumption · Time-varying coefficient model · Coefficient driver

JEL Classification Numbers C130 · C190 · C220

The views expressed in this paper are the authors' own and do not necessarily represent those of their respective institutions.

P. A. V. B. Swamy
Retired from Federal Reserve Board,
Washington, DC,
6333 Brocketts Crossing
Kingstowne, VA 22315
e-mail: swamyparavastu@hotmail.com

* G. S. Tavlas (Corresponding author)
Economic Research Department,
Bank of Greece, 21 El. Venizelos
Ave. 102 50. Athens, Greece
e-mail: GTavlas@bankofgreece.gr

Stephen G. Hall
Leicester University and Bank of Greece and
NIESR, Room: Astley Clarke 116,
University Road, Leicester,
LEI 7RH, UK,
e-mail: s.g.hall@le.ac.uk

George Hondroyiannis
Bank of Greece and Harokopio
University, 21 E. Venizelos Ave.
102 50 Athens, e-mail:
ghondroyiannis@bankofgreece.gr

1 Introduction

The term, 'Time-Varying Coefficient (TVC) estimation' as used in this paper refers to a specific estimation technique that has been developed to provide consistent parameter estimates in the simultaneous presence of measurement errors, omitted variables, and unknown functional form[1] Previous studies using this technique implicitly allow for non-stationarity in the data, but do not explicitly discuss how each stage of the estimation strategy needs to be undertaken when the data exhibits the common property of non-stationarity.[2] Non-stationarities other than unit-root non-stationarities of variables and the nonlineartities of the relationships among these variables are closely related, as will be clear from what follows. The objective of this paper is to discuss how each stage of the estimation strategy would be affected by non-stationary variables that form a nonlinear structural economic relationship and which therefore co-integrate.

Rarely, if ever, are econometric relationships free from specification errors arising from the following four problems: (i) an unknown true functional form, (ii) the correlation between the error terms of econometric models and their explanatory variables, (iii) data on economic variables contain measurement errors and (iv) omitted variables. Consequently, misspecification of models is difficult to avoid.

The following exposition of the TVC approach begins with a coefficient-decomposition that illustrates how the various forms of misspecification interact; it then goes on to propose a specific form of time variation for the parameters that allows a decomposition of the parameters that match the coefficient decomposition. This allows us to recover consistent estimates of the parameters of interest. We argue that each of these stages is valid in the presence of non-stationary variables, but the non- statonarity has implications for how the coefficients will behave and the procedure for carrying-out the decomposition.

---

[1] See Swamy and Tinsley (1980), Havenner and Swamy (1981), Chang, Hallahan and Swamy (1992), Yokum, Wildt and Swamy (1998), Chang, Swamy, Hallahan and Tavlas (2000), Swamy and Tavlas (2001), Swamy, Yagi, Mehta and Chang (2007), Hall, Hondroyiannis, Swamy and Tavlas (2008), Hall, Hondroyiannis, Swamy and Tavlas (2009).

[2] We are using here the term 'non-stationarity' to refer to the processes that are not either weakly or strongly stationarity defined in Anderson (1971, pp. 373-374).

The remainder of this paper is divided into three sections. Section 2 presents the interpretation of the coefficients of a misspecified econometric model and discusses the implications of non-statonarity. The next section discusses the formulation of a specific form of time variation of the coefficients that allows identification of the underlying coefficients. Section 3 presents the methods of consistently estimating the unknown quantities of the underlying coefficients. Section 4 concludes.

2 The Interpretations of Model Coefficients and Appropriate Assumptions

Conventional econometrics is to a large extent the study of individual causes of biased parameter estimates: 'non-sufficient sets' of omitted variables, measurement errors, an incorrect functional form, etc. These problems are usually dealt with one at a time in a textbook context, but of course practical work is plagued by all these problems at once. In this section, we outline (1) the basic problem of interpreting coefficients when these problems are present and (2) our proposed procedure for dealing with these problems simultaneously. In particular, we are concerned with the case in which the dependent variable of an economic relationship is non-stationary and at least two sets of its determinants are also non-stationary and where there is a (possibly) nonlinear relationship between these variables which produces a combination of them with constant coefficients, thus a general nonlinear form of cointegration. We restrict ourselves here to the case of two sets of non-stationary variables simply because this allows for all the cases we believe are of interest. These two sets of variables could be equally thought of as two individual non-stationary variables. We also allow for measurement error and the omitted variables may be either stationary or one or both of the two sets of non-stationary independent variables.

Denote the dependent variable by y_t^*; it is related to a hypothesized set of $K - 1$ of its determinants, denoted by x_{1t}^*, …, $x_{K-1,t}^*$, where K-1 may be only a subset of the complete set of determinates of y_t^*, in which case the relation of y_t^* to x_{1t}^*, …, $x_{K-1,t}^*$ may be subject to omitted-variable biases. Any specific functional form may be incorrect and may lead to specification errors. In addition to these problems, the available data on y_t^*, x_{1t}^*, …, $x_{K-1,t}^*$ may not be perfect measures of the underlying true variables, causing an errors-in-variables problem.

Suppose that T measurements on y_t^*, x_{1t}^*, …, $x_{K-1,t}^*$ are made and these measurements are actually the sums of "true" values and measurement errors: $y_t = y_t^* + \mathrm{v}_{0t}$, $x_{jt} = x_{jt}^* + \mathrm{v}_{jt}$, j = 1, …, K-1, t = 1, …, T, where the variables y_t, x_{1t}, …, x_{Kt} without an asterisk are the observable variables, the variables with an asterisk are the unobservable "true" values, and the v's are measurement errors. Given the possibility that the true functional form we are estimating may be unknown and that there may be some important variables missing from x_{1t}, …, $x_{K-1,t}$; we need a model which will capture all these potential problems.

It is useful at this point to clarify what we believe to be the main objective of econometric estimation. In our view the objective is to obtain consistent estimates of the bias-free effect on a correctly measured dependent variable of changing one of its correctly measured determinants holding all of its other correctly measured determinants constant. That is to say, we aim to find an estimate of the partial derivative of y_t^* with respect to any x_{jt}^* if y_t^* is a continuous function of x_{jt}^* and the bias-free effect of any x_{jt}^* on y_t^* otherwise. This of course is the interpretation, which is usually placed on the coefficients of a standard econometric model, but this interpretation depends crucially on the assumption that the conventional model has bias-free coefficients, which is, of course, not the case in the presence of model misspecification. Note that the term "bias-free" here means without both omitted-variable and measurement-error bias components.

We begin by specifying a set of time-varying coefficients, which provide a complete explanation of the dependent variable y.

$$y_t = \gamma_{0t} + \gamma_{1t}x_{1t} + \cdots + \gamma_{K-1,t}x_{K-1,t} \qquad (t = 1, \ldots, T) \tag{1}$$

which we call "the time-varying coefficient (TVC) model". The explanatory variables of this model are called the included variables. As this model provides a complete explanation of y, all the misspecification in the model, as well as the true coefficients must be captured by the time-varying coefficients. Note that if the true functional form is non-linear the time-varying coefficients may be thought of as being the partial derivatives of the true non-linear structure and so they are able to capture any possible function. These coefficients will also capture the effects of measurement error and omitted variables.

Because we are dealing with a potentially non-linear true model, which is assumed to be unknown, we need a slightly more general definition of non-stationarity and cointegration than is usually used. Normally, we focus on the order of integration of a variable; however in the presence of general non-linearity, variables may not be integrated at all. A variable is integrated of order *d* if it becomes stationary after being first differenced *d* times. When $d = 0$, such a variable is (weakly or strongly) stationary, and when $d > 0$, it is unit-root non-stationary. There are also non-stationary variables that are not unit-root non-stationary. Let $x_t = (1, x_{1t}, \ldots, x_{K-1,t})'$ and $\gamma_t = (\gamma_{0t}, \gamma_{1t}, \ldots, \gamma_{K-1,t})'$ is time dependent. Then equation (1) is nonlinear and shows that the first difference of y_t,

$$\Delta y_t = y_t - y_{t-1} = x_t'\gamma_t - x_{t-1}'\gamma_{t-1} + x_{t-1}'\gamma_t - x_{t-1}'\gamma_t = \Delta x_t'\gamma_t + x_{t-1}'\Delta\gamma_t,$$

is , in general, neither stationary nor unit-root non-stationary and hence y_t is non unit-root non-stationary and is not integrated. Also, Δy_t does not possess a finite unconditional mean if x_t and/or γ_t follow a random walk processes. Thus, each time equation (1) is differenced additional terms enter into it giving a non-parsimonious form unless equation (1) is linear or its intercept and slopes (excluding its error term) are constant, which will not generally be true.

There are a number of possible definitions of cointegration. A rather restricted one says that a set of integrated variables is said to be cointegrated if these variables follow a linear model in which (i) the error term is integrated of order zero with mean zero such that it is mean independent of the included explanatory variables and (ii) the coefficients are free of specification biases (see Greene 2008, p. 756). This is a highly specialized definition of cointegration that rarely if ever applies to practical situations. A more general definition that allows for non-linearity and omitted variables is as follows: The variables y_t and x_t are cointegrated if the bias-free components of the coefficients of x_t in the relation of y_t to x_t are nonzero. This is very much in keeping with the original idea of cointegration. In general cointegration should only arise if there is a (possibly non-linear) structural relationship holding a set of variables together. If there is such a relationship then this implies that the bias free effect of x on y will be non zero.

Equation (1) is called the observation equation and its coefficients are called the state variables if it is embedded in a state-space model. We now apply a formal decomposition of these time-varying coefficients which illustrates the various components they contain.

Notation and Assumptions Let m_t denote the total number of the determinants of y_t^*. The exact value of m_t is usually unknown at any time. We assume that m_t is larger than *K*-1 (that is, the number of determinants is greater than the determinants for which we have observations) and possibly varies over time. This assumption means that there are determinants of y_t^* that are excluded from equation (1). Let x_{gt}^*, $g = K, \ldots, m_t$, denote these excluded determinants. Let α_{0t}^* denote the intercept and let both α_{jt}^*, $j = 1, \ldots, K\text{-}1$, and α_{gt}^*, $g = K, \ldots, m_t$, denote the other coefficients of the regression of y_t^* on all of its determinants. The true functional form of this regression determines the time profiles of α^*'s. These time profiles are unknown, since the true functional form is unknown. For $g = K, \ldots, m_t$, let $x_{gt}^* = \lambda_{0gt}^* + \lambda_{1gt}^* x_{1t}^* + \cdots + \lambda_{K-1,gt}^* x_{K-1,t}^*$. The true functional forms of these regressions determine the time profiles of λ^*'s.

Theorem 1 *The intercept of (1) satisfies the equation,*

$$\gamma_{0t} = \alpha^*_{0t} + \sum_{g=K}^{m_t} \alpha^*_{gt}\lambda^*_{0gt} + \mathrm{v}_{0t}, \tag{2}$$

and the coefficients of (1) other than the intercept satisfy the equations,

$$\gamma_{jt} = \alpha^*_{jt} + \sum_{g=K}^{m_t} \alpha^*_{gt}\lambda^*_{jgt} - \left(\alpha^*_{jt} + \sum_{g=K}^{m_t} \alpha^*_{gt}\lambda^*_{jgt}\right)\left(\frac{\mathrm{v}_{\mathrm{jt}}}{\mathrm{x}_{\mathrm{jt}}}\right) \quad (j = 1, \ldots, K\text{-}1). \tag{3}$$

where the λ^*_{0gt} are a 'sufficient set' of excluded variables in the sense that they in conjunction with the x^*_{jt} are at least sufficient to determine y^*_t.

Proof see Swamy and Tavlas (2001, 2007).

Thus, we interpret the TVC's of (1) in terms of the underlying correct coefficients, a 'sufficient set' of excluded variables, the observed explanatory variables and their measurement errors. By assuming that the α^*'s and λ^*'s are possibly time varying, we do not *a priori* rule out the possibility that the relationship of y^*_t with all of its determinants and the regressions of the determinants of y^*_t excluded from (1) on the determinants of y^*_t included in (1) are non-linear.

In terms of non-stationarity and nonconstancy we can consider 3 cases, assuming that x^*_{jt}, j = 1, …, K – 1, and λ^*_{0gt}, $g = K$, …, m_t, are the two sets of the determinants of y^*_t .

1. Both the 'sufficient set' of excluded variables $\sum_{g=K}^{m_t} \alpha^*_{gt}\lambda^*_{0gt}$ and the measurement error v_{0t} are white noise with mean zero and the true intercept α^*_{0t} is constant for all t, both the components α^*_{jt} and $\sum_{g=K}^{m_t} \alpha^*_{gt}\lambda^*_{jgt}$ of the coefficient of x^*_{jt} are constant and the measurement error v_{jt} = 0 for all j and t . In which case the non-stationarity will be confined to the mean of y_t, as in the standard regression models. .
2. If x^*_{jt}, j = 1, …, K-1, are non-stationary but the other non-stationary variables are λ^*_{0gt}, $g = K$, …, m_t, a sufficient set of excluded variables, then even if the

α^*_{jt} are constant for all *j* and *t*, the γ_{jt} will be nonconstant if their other components are nonconstant,

3. Both the sets of variables x^*_{jt}'s and λ^*_{0gt}'s may be non-stationary, in which case again the γ_{jt} will be nonconstant if any or all of their components in (2) and (3) are nonconstant.

Theorem 2 *For* $j = 1, \ldots, K\text{-}1$, *the component* α^*_{jt} *of* γ_{jt} *in (3) is the partial derivative of* y^*_t *with respect to* x^*_{jt} *if* y^*_t *is a continuous function of* x^*_{jt} *and is the direct or bias-free effect of* x^*_{jt} *on* y^*_t *with all the other determinants of* y^*_t *held constant otherwise and is unique.*

Proof It can be seen from equation (3) that the component α^*_{jt} of γ_{jt} is free of omitted-variables bias ($= \sum_{g=K}^{m_t} \alpha^*_{gt}\lambda^*_{jgt}$), measurement-error bias ($= -\left(\alpha^*_{jt} + \sum_{g=K}^{m_t} \alpha^*_{gt}\lambda^*_{jgt}\right) \times \left(\mathrm{v}_{jt} / x_{jt}\right)$), and of functional-form bias, since we allow the α^*s and λ^*s to have the correct time profiles. These biases are not unique being dependent on what determinants of y^*_t are excluded from (1) and the v_{jt}. Only α^*_{jt} is unique being the coefficient of x^*_{jt} in the correctly specified relation of y^*_t to all of its determinants. The component α^*_{jt} represents the direct, or bias-free, effect of x^*_{jt} on y^*_t with all the other determinants of y^*_t held constant. The nonzero direct effect is unique because it represents a property of the real world that remains invariant against mere changes in the language we use to describe it (see Basmann 1988, p. 73; Pratt and Schlaifer 1984, p. 13; Zellner 1979, 1988).

This is true irrespective of the non-stationarity of the variables under consideration as the 'sufficient set' of omitted variables can fully reflect the omitted non-stationary variables. The real issue, however, is correctly identifying the omitted variable biases in the case of non-stationarity. We turn to this issue in the next section.

3 Identification and Consistent Estimation of Time-Varying Coefficient Model

3.1 Identification

As noted above, we believe that empirical researchers are interested in the bias-free or direct effects (or the partial derivatives) α^*'s, not in the omitted-variable and measurement-error biases. That is, they are not interested in the γ_{jt}, which are contaminated by omitted-variable and measurement-error biases. To obtain accurate estimates of the α_{jt}^* using the observations in (1), we need to first decompose each γ_{jt} with $j > 0$ into its components in (3). Our method of identifying these components and performing the decomposition is based on the following assumptions.

Assumption 1 (Auxiliary information) *Each coefficient of (1) is linearly related to certain drivers plus a random error,*

$$\gamma_{jt} = \pi_{j0} + \sum_{d=1}^{p-1} \pi_{jd} z_{dt} + \varepsilon_{jt} \quad (j = 0, 1, \ldots, K\text{-}1), \tag{4}$$

where the π s *are fixed parameters, the* z_{dt} *are what are called the coefficient drivers, and different coefficients of (1) can be functions of different sets of coefficient drivers.*

Here, the issue of identification is quite important. If both γ_{jt} and α_{jt}^* are constant then clearly they are not identifiable. . Regardless of whether α_{jt}^* is constant or not, if γ_{jt} is nonconstant (due to its nonconstant omitted-variables and measurement-error bias components) we need to include a set of coefficient drivers to identify its components.

Assumption 2 *For j* = 1, …, *K*-1, *the set of p-1 coefficient drivers and the constant term in (4) divides into two disjoint subsets* S_1 *and* S_2 *so that* $\sum_{d \in S_1} \pi_{jd} z_{dt}$ *has the same pattern of time variation as* α^*_{jt} *and* $\sum_{d \in S_2} \pi_{jd} z_{dt} + \varepsilon_{jt}$ *has the same pattern of time variation as the sum of the last two terms on the right-hand side of equation (3) over the relevant estimation and forecasting periods. The definition that* $\alpha^*_{jt} = \sum_{d \in S_1} \pi_{jd} z_{dt}$ *is correct.*

Here we are assuming that the drivers in the set S_1 separate the direct effect α^*_{jt} from the specification biases in the model. Here again we can draw out some important implications for the division between S_1 and S_2. If the component α^*_{jt} is constant while γ_{jt} is non-stationary (or nonconstant) then in general the variables in S_1 should also be constant (although there is the possibility of a case where the variables in S_1 are non-stationary (or nonconstant) but cointegrate over the sample so that in combination all the variables in S_1 are constant. In this case we know that all non-stationary (or nonconstant) drivers should be in the set S_2. Of course it is possible that α^*_{jt} itself is non-stationary (or nonconstant), due to the unknown non-linear functional form, in which case we have a difficult problem of splitting out just the correct amount of non-stationarity (or nonconstancy) between the sets S_1 and S_2. However the assumption that α^*_{jt} is constant is not a very strong one if y^*_t is linearly related to all of its determinants.

Assumption 3 *The K-vector* $\varepsilon_t = (\varepsilon_{0t}, \varepsilon_{1t}, ..., \varepsilon_{K-1,t})'$ *of errors in (4) follows the stochastic equation,*

$$\varepsilon_t = \Phi \varepsilon_{t-1} + u_t, \tag{5}$$

where Φ *is a* $K \times K$ *(not necessarily diagonal) matrix whose eigenvalues are less than 1 in absolute value, the K-vector* $u_t = (u_{0t}, u_{1t}, ..., u_{K-1,t})'$ *is distributed with* $E(u_t | z_{1t}, \ldots, z_{p-1,t}) = 0$ *and*

$$E(u_t u'_{t'} \mid z_{1t}, \ldots, z_{p-1,t}) = \begin{cases} \sigma_u^2 \Delta_u & \text{if } t=t' \\ 0 & \text{if } t \neq t' \end{cases}, \tag{6}$$

where Δ_u *may not be diagonal.*

This assumption considerably generalizes (4). If we assumed that the errors in (4) were independent, this would imply a very simple dynamic structure. By making the assumption that the errors in fact have a serial correlation structure we are allowing a much richer dynamic structure although we are imposing some common factors in this structure to keep the model tractable.

In terms of non-stationarity by assuming that all the eigenvalues are less than 1 in absolute terms we are ruling out the possibility that non-stationarity in γ_{jt} is generated by the error process ε_t. This then isolates the non-stationarity as coming from the coefficient drivers.

Assumption 4 *The regressor* x_{jt} *of (1) is conditionally independent of its coefficient* γ_{jt} *given the coefficient drivers in (4) for all j and t.*

A vector formulation of model (1) is

$$y_t = x'_t \gamma_t, \tag{7}$$

where x_t and γ_t are as defined below equation (1). A matrix formulation of (4) is

$$\gamma_t = \Pi z_t + \varepsilon_t, \tag{8}$$

where $\Pi = \left[\pi_{jd}\right]_{0 \le j \le K-1, 0 \le d \le p-1}$ is a $K \times p$ matrix having π_{jd} as its $(j+1, d+1)$-th element and $z_t = (1, z_{1t}, \ldots, z_{p-1,t})'$. Substituting (8) into (7) gives

$$y_t = (z'_t \otimes x'_t)\pi^{Long} + x'_t \varepsilon_t, \tag{9}$$

where $\otimes$ denotes a Kronecker product, and π^{Long} is a *Kp*-vector, denoting a column stack of Π. The observations in (1) can be displayed in a matrix form as

$$y = X_z \pi^{Long} + D_x \varepsilon, \tag{10}$$

where $y = (y_1, \ldots, y_T)'$ is a *T*-vector, $X_z = (z_1 \otimes x_1, \ldots, z_T \otimes x_T)'$ is $T \times Kp$, $D_x = diag_{1 \le t \le T}(x'_t)$ is $T \times KT$, and $\varepsilon = (\varepsilon'_1, \ldots, \varepsilon'_T)'$ is a *TK*-vector.

Theorem 3 *Under Assumptions 1-4,* $E(y \mid X_z) = X_z\pi^{Long}$ *and* $\mathrm{Var}(y \mid X_z) = D_x\sigma_u^2\Sigma_\varepsilon D_x'$ *where* $\sigma_u^2\Sigma_\varepsilon$ *is the covariance matrix of* ε.

Proof See Swamy, Yaghi, Mehta and Chang (2007, p. 3386). □

Under Assumptions 1 and 3, the variance of γ_{jt} is finite for all *j* and *t*. The Chebychev inequality shows that if γ_{jt} has a small variance, then its distribution is tightly concentrated about its mean implied by Assumptions 1 and 3 (see Lehmann 1999, p. 52). Assumptions 2 and 4 provide a prime consideration guiding the selection of coefficient drivers, especially in the presence of non-stationarity. The magnitude of ε_{jt} gets reduced as the number of correct coefficient drivers in (4) increases. The larger the number of correct coefficient drivers in (4), the smaller the magnitude of ε_{jt} and the smaller the variance of γ_{jt}. Including many correct coefficient drivers in (4) may imply that the errors of equation (4) are white-noise variables or the matrix Φ in equation (5) is null. If Assumption 3 is replaced by the assumption that ε_t follows a random walk for all *t*, then the unconditional variance of γ_{jt} is not finite.

The fixed coefficient vector π^{Long} in (10) is identified if X_z has full column rank. A necessary condition for X_z to have full column rank is that $T > Kp$. The error vector ε is not identified because the necessary condition $T > TK$ for D_x to have full column rank is false. This result implies that ε is not consistently estimable (see Lehmann and Casella 1998, p. 57). Swamy and Tinsley (1980, p. 117) call this phenomenon "a form of Uncertainty Principle". Correct coefficient drivers should be used in (4) to reduce the unidentifiable portions (the ε_{jt}) of the coefficients of (1). However, $D_x\varepsilon$ being equal to $y - X_z\pi^{Long}$ with identifiable π^{Long} is identifiable, provided D_x has full row rank. The best linear unbiased predictor (BLUP) of $D_x\varepsilon$ can be used to obtain consistent estimators of Φ, Δ_u, and σ_u^2 in (5) and (6), as shown in Chang, Hallahan and Swamy (1992) and Chang, Swamy, Hallahan and Tavlas (2000). Under Assumptions 1-4, the BLUP of $D_x\varepsilon$ exists (see Swamy, Yaghi, Mehta and Chang 2007, p. 3387). So we make

Assumption 5 (i) X_z *has full column rank, (ii)* D_x *has full row rank, and (iii) (assumption 5(iii) is here)* $T \geq Kp$ *+ the number of unknown distinct elements of* Φ *,* Δ_u *, and* σ_u^2 *+ 4 so that the degrees of freedom left unutilized after estimating all the unknown parameters of model (10) is at least 4.*

Assumptions 5(i) and 5(ii) make all the coefficients and $D_x\varepsilon$ of (10) statistically meaningful. Equation (4), which establishes a link between the coefficients of (1) and the coefficients and errors of (10), shows that if the coefficients and $D_x\varepsilon$ of (10) are statistically meaningful, then so are the coefficients of (1). In certain situations specified in Judge, Griffiths, Hill, Lütkepohl and Lee (1985, p. 612), the finite moments of the estimators of the coefficients of (10) exist up to the degrees of freedom that remain unutilized after the estimation of these coefficients. Assumption 5(iii) is made to guarantee the existence of at least finite fourth moments for the estimators of the coefficients of (10) in these situations. Swamy, Mehta and Singamsetti (1996) explain how model (10) might be estimated when X_z has less than full column rank and $D_x\Sigma_\varepsilon D_x'$ is singular.

3.2 Consistent estimation

Under certain conditions, an iteratively rescaled generalized least squares estimators of π^{Long} and $D_x\varepsilon$ in (10) are consistent and asymptotically normal (see Swamy, Tavlas, Hall and Hondroyiannis 2009).

4 Conclusions

We argue here that non-stationarity does not pose any particular problem for TVC estimation. However, as in other cases, the explicit recognition of non-stationarity does offer advantages, in particular, in the identification of the correct set of coefficient drivers to correctly identify bias free component of the time varying coefficient.

Anderson, T.W,: The statistical analysis of time series. New York: John Wiley & Sons (1971).

References

Barndorff-Nielsen, O., Cox, D.R.: Edgeworth and saddlepoint approximations with statistical applications. J Roy Statis Soc. B **41**, 279-312 (1979)

Basmann, R.L.: Causality tests and observationally equivalent representations of econometric models. J Econ Ann **39**, 69-104 (1988)

Brown, L.D.: An ancillarity paradox which appears in multiple linear regression (with discussion). Ann Statist **18**, 471-538 (1990)

Cavanagh, C.L., Rothenberg, T.J.: Generalized least squares with nonnormal errors, pp. 276-290. In: Maddala, G.S., Phillips, P.C.B., Srinivasan, T.N. (eds.) Advances in Econometrics and Quantitative Economics. Cambridge, MA: Blackwell Publishers Inc. (1995)

Chang, I-L., Hallahan, C., Swamy, P.A.V.B.: Efficient computation of stochastic coefficients models, pp. 43-54. In: Amman, H.M., Belsley, D.A., Pau, L.F. (eds.) Computational Economics and Econometrics. Boston, MA: Kluwer Academic Publishers (1992)

Chang, I-L., Swamy, P.A.V.B., Hallahan, C., Tavlas, G.S.: A computational approach to finding causal economic laws. Comput Econ **16**, 105-136 (2000)

DeGroot, M.H.: Comment on 'Lindley's paradox' by G. Shafer. J Amer Statist Assoc **77**, 336-339 (1982)

Durbin, J., Koopman, S.J.: Time series analysis by state space methods. Oxford: Oxford University Press (2001)

Granger, C.W.J., Newbold, P.: Spurious regression in econometrics. J Econ **2**, 111-120 (1974)

Greene, W.H.: Econometric analysis, 6th edn. Upper Saddle River, New Jersey: Pearson, Prentice Hall (2008)

Hall, S.G., Hondroyiannis, G., Swamy, P.A.V.B., Tavlas, G.S.: The new Keynesian Phillips curve and lagged inflation: A case of spurious correlation? Southern Econ J, forthcoming, (2008)

Hall, S.G., Hondroyiannis, G., Swamy, P.A.V.B., Tavlas, G.S.: A portfolio balance approach to euro-area money demand in a time-varying environment. A paper presented at Bank of England/MMF workshop on money and macromodels (2008)

Havenner, A., Swamy, P.A.V.B.: A random coefficient approach to seasonal adjustment of economic time series. J Econ **15**, 177-209 (1981)

Hondroyiannis, G., Swamy, P.A.V.B., Tavlas, G.S.: The new Keynesian Phillips curve in a time-varying coefficient environment: some European evidence. Macroecon Dyn **13**, (2009) forthcoming

Judge, G.G., Griffiths, W.E., Hill, R.C., Lütkepohl, H., Lee, T.: The theory and practice of econometrics, 2nd edn. New York: John Wiley and Sons (1985)

Kariya, T., Kurata, H.: Generalized least squares. Hoboken, NJ: John Wley & Sons, Ltd (2004)

Lehmann, E.L.: Elements of large-sample theory. Berlin Heidelberg New York: Springer

(1999)

Lehmann, E.L., Casella, G.: Theory of point estimation, 2nd edn. Berlin Heidelberg New

York: Springer (1998)

Lucas, R.E., Jr.: Econometric policy evaluations: a critique. Carnegie-Rochester Ser Public Policy **1**, 19-46 (1976)

Pratt, J.W., Schlaifer, R.: On the nature and discovery of structure. J Am Stat Assoc **79**, 9-21, 29-33 (1984)

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: John Wiley & Sons (1973)

Sethuraman, J.: Some limit theorems for joint distributions. Sankya, Series A, 379-386 (1961)

Swamy, P.A.V.B., Tinsley, P.A.: Linear prediction and estimation methods for regression models with stationary stochastic coefficients. J Econ **12**, 103-142 (1980)

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 Statis Plan Inf **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.: A note on Muth's rational expectations hypothesis: A time-varying coefficient interpretation. Macroecon Dyn **10**, 415-425 (2006)

Swamy, P.A.V.B., Tavlas, G.S.: The new Keynesian Phillips curve and inflation expectations: re-specification and interpretation. Econ Theory **31**, 293-306 (2007)

Swamy, P.A.V.B., Tavlas, G.S., Mehta, J.S.: Methods of distinguishing between spurious regressions and causality. J Stat Theory Appl **1**, 83-96 (2007)

Swamy, P.A.V.B., Yaghi, W., Mehta, J.S., Chang, I-L.: Empirical best linear unbiased prediction in misspecified and improved panel data models with an application to gasoline demand. Comput Statist Data Anal **51**, 3381-3392 (2007)

Swamy, P.A.V.B., Mehta, J.S., Chang, I-L., Zimmerman, T.S.: An efficient method of estimating the true value of a population characteristic from its discrepant estimates. Comput Statist Data Anal **52**, (2008) forthcoming

Yokum, J.T., Wildt, A.R., Swamy, P.A.V.B.: Forecasting disjoint data structures using appropriate constant and stochastic coefficient models. J Appl Statist Sci **8**, 29-49 (1998)

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)