University of Leicester
Department of Economics

Time Varying Coefficient Models; A Proposal for selecting the Coefficient Driver Sets

Stephen G. Hall, University of Leicester

P. A. V. B. Swamy

George S. Tavlas, Bank of Greece

Working Paper No. 14/18

December 2014

Time Varying Coefficient Models; A Proposal for selecting the Coefficient Driver Sets

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

Abstract

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

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)
Member, Monetary Policy Council
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

1. Introduction

Time-varying-coefficient (TVC) estimation is a way of estimating consistent parameters of a model even when (i) the true functional form is unknown, (ii) there are missing important variables, and (iii) the included variables contain measurement errors.[1] However, an important assumption is needed to make the technique operational. This assumption concerns the choice of what are called "coefficient drivers" (formally defined below) and the separation of these drivers into two sub-sets. This separation allows us to derive estimates of bias-free coefficients. Intuitively, coefficient drivers are a set of variables that feed into the TVC's and explain at least part of their movement. As explained in what follows, this set of drivers is split into two sub-sets, one of which is correlated with the bias-free coefficient that we want to estimate and the other of which is correlated with the misspecification in the model. This split has always been somewhat arbitrary (much as in the case of choosing instrumental variables). The objective of this paper is to put forward a method for producing this split that takes account of the non-linearity that may be in the original data. As we argue below, this method provides a natural split in the driver set.

The remainder of this paper is divided into four sections. Section 2 presents a summary of the TVC approach and formally defines the concept of coefficient drivers and the need to split the drivers into two sets. Section 3 then proposes a method for determining this split. Section 4 provides an example of the practical use of the technique by applying it to modeling the effect of rating agencies ratings on sovereign bond spreads. Section 5 concludes.

2. TVC estimation

Here, we summarize the approach to TVC estimation that has been formalized in Swamy, Hall, Hondroyiannis and Tavlas (2010). TVC estimation proceeds from an important theorem that was first established by Swamy and Mehta (1975), and, which has subsequently been confirmed by Granger (2008). This theorem states that any nonlinear functional form can be exactly represented by a model that is linear in variables, but which

[1] The development of TVC estimation is due to Swamy (1969, 1974). See also, Swamy and Tavlas (2001, 2007) and Swamy, Hall, Hondroyiannis, and Tavlas (2010).

has time-varying coefficients. The implication of this result is that, even if we do not know the correct functional form of a relationship, we can always represent this relationship as a TVC relationship and, thus, estimate it. Hence, any nonlinear relationship may be stated as;

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

Where k-1 is the number of included variables in our model and L> [Steve: something is missing] Consequently, this theorem leads to the result that, if we have the complete set of relevant variables with no measurement error, then by estimating a TVC model we will get consistent estimates of the true partial derivatives of the dependent variable with respect to each of the independent variables given the unknown, non-linear functional form. If we then allow for the fact that we do *not* know the full set of independent variables and that some, or perhaps all, of them may be measured with error, then the TVCs become biased (for the usual reasons). What we would like to have is some way to decompose the full set of biased TVCs into two parts -- the biased component and the remaining part; the latter would be a consistent estimate of the true parameter. While this is asking a great deal of an estimation technique, it is precisely what TVC estimation aims to provide (Swamy, Tavlas, Hall and Hondroyianis, 2010). This technique builds from the Swamy and Mehta theorem, mentioned above, to produce such a decomposition[2].

Swamy, Tavlas Hall and Hondroyianis (2010) show what happens to the TVCs as other forms of misspecification are added to the model. If we omit some relevant variables from the model, then the true TVCs get contaminated by a term that involves the relationship between the omitted and included variables. If we also allow for measurement error, then the TVCs get further contaminated by a term that allows for the relationship between the exogenous variables and the error terms. Thus, as one might expect, the estimated TVCs are no longer consistent estimates of the true partial derivatives of the non-linear function. Instead, they are biased due to the effects of omitted variables and measurement error. There are exact mathematical proofs provided for our statements up to this point.

[2] Mathematically this model may appear to be a state space one. However, the interpretation of the coefficients is quite different from the standard state space representation. Omitted-variable biases, measurement-error biases and the correct functions of certain ‘sufficient sets of excluded variables’ are not considered parts of the coefficients of the observation equations of state-space models. This is the major difference between (1) and the observation equations of a standard state-space model.

To make TVC estimation fully operational, we need to make two key parametric assumptions; first, we assume that the time-varying coefficients themselves are determined by a set of stochastic linear equations which makes them a function of a set of variables we call driver (or coefficient-driver) variables. This is a relatively uncontroversial assumption. Second, we assume that some of these drivers are correlated with the misspecification in the model and some of them are correlated with the time-variation coming from the non-linear (true) functional form. Having made this assumption we can then simply remove the bias from the time-varying coefficients by removing the effect of the set of coefficient drivers which are correlated with the misspecification. This procedure, then, yields a consistent set of estimates of the true partial derivatives of the unknown nonlinear function, which may then be tested by constructing 't' tests in the usual way. An important difference between coefficient drivers and instrumental variables is that for a valid instrument we require a variable that is uncorrelated with the misspecification. This often proves hard to find. For a valid driver we need variables which are correlated with the misspecification. We would argue that this is much easier to achieve.

To formalize the idea of the coefficient drivers, we assume that each of the TVCs in (1) is generated in the following way.

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

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

where the π s *are fixed parameters, the* z_{dt} *are what we call the coefficient drivers; different coefficients of (2) can be functions of different sets of coefficient drivers.*

The regressors and the coefficients of (2) are conditionally independent of each other given the coefficient drivers.[3] These coefficient drivers are merely a set of variables that, to a reasonable extent, jointly explain the movement in γ_{jt}.

Under our method, the coefficient drivers included in equation (2) have two uses. Insertion of equation (2) into equation (1) parameterizes the latter equation. This is the first

[3] The distributional assumptions about the errors in (12) are given in Swamy, Tavlas, Hall and Hondroyiannis (2010).

use of the coefficient drivers. Here, the issue of identification of the parameterized model (1) is important.[4] The other important use of the drivers allows us to separate the bias and bias-free components of the coefficients.

Assumption 2 *The set of coefficient drivers and the constant term in (2) divides into three different subsets* A_{1j}, A_{2j}. and A_{3j} such *that the first set is correlated with any variation in the true parameter that is due to the underlying relationship being non-linear, the second set is correlated with bias in the parameter coming from any omitted variables, and the final set is correlated with bias coming from measurement error*

This assumption allows us to identify separately the bias-free, omitted-variables and measurement-error bias components of the coefficients of (1).

Assumption 2 is the key to making our procedure operational; it is the assumption that we can associate the various forms of specification biases with sets A_{2j}, and A_{3j}, which means that set A_{1j} simply explains the time-variation in the coefficients caused by the nonlinearity in the true functional form. If the true model is linear, then all that would be required for set A_{1j} would be to contain a constant. If the true model is nonlinear, then the bias-free components should be time-varying and the set of drivers belonging to A_{1j} will explain the time variation in these components.

3. A Suggestion for the Choice of Coefficient Drivers

Clearly, Assumptions 1 and 2 above are crucial for the successful implementation of the TVC approach. As noted above, the split of coefficient drivers stemming from these assumptions has always been a problematic part of the TVC-estimation procedure. There are, however, certain requirements that can help in selecting both the variables that make a good driver set and the split into the two subsets inherent under Assumption 2.

[4] To handle this issue, we use Lehmann and Casella's (1998, pp. 24 and 57) concept of identification.

Consider, first, the broad requirements that a complete set of drivers should fulfill; these relate to predictive power and relevance. Consider, again, the driver equations in equation (2).

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

Where $z_{0t} \equiv 1$. For this set of drivers to be a good set, the drivers must explain most of the variation in γ_{jt} Hence, we can define an analog of the conventional R^2 for the estimated counterparts to these equations as follows:

$$R^2 = 1 - \frac{SS\varepsilon_{jt}}{SS\gamma_{jt}} \tag{3}$$

(Where $SS\varepsilon_{jt}$ and $SS\gamma_{jt}$ are the sum of squared residuals and the total variation of the dependent variable, respectively). Here, we require the R^2 coefficient to be as close to 1 as possible, so that the drivers explain a large amount of the variation in the TVC. This result could, of course, be achieved simply by having a very large number of drivers. Therefore, we also require the drivers to be relevant in the sense that the π_{jd} are significantly different from zero. Estimation of the full TVC model produces a covariance matrix for the π_{jd} so conventional 't' statistics and probability levels may be produced in the standard way.

These two conditions are closely analogous to the idea of relevance in instrumental estimation, where the instruments must have explanatory power in explaining the variables being instrumented. If the R^2 in (3) is low, then we could infer that we had a weak set of coefficient drivers. There is, however, no requirement for the drivers to be independent of the coefficient γ_{jt}, as there is for instruments to be independent of the error term under an IV estimation procedure.

The more difficult issue, then, is how to perform the split in the coefficient drivers into the two sets outlined under Assumption 2 -- that is S_1, which contains the variables correlated with the unbiased coefficient, and S_2, which is the set of drivers correlated with the misspecification.

Let's assume that the true, unknown model is given by

$$y^*{}_t = f(x_{1t}{}^* ....x_{mt}{}^*) \qquad (4)$$

We are, therefore, interested in estimating

$$\frac{\delta y_t{}^*}{\delta x_{it}{}^*} \ for\ i = 1,..., K-1 \qquad (5)$$

where the values of all the arguments of $f(x_{1t}{}^* ....x_{mt}{}^*)$ other than x_{it} are held constant.

To understand how the split of drivers may be achieved, consider the following examples.

Example 1

If (4) is linear, then the S_1 set consists of just a constant, as the true parameter is a constant, and all other drivers explain the biases that stem from missing variables and measurement error.

Example 2

Suppose that (4) is a polynomial, such as a quadratic form. Consider, for simplicity, the case of only 2 explanatory variables. Then,

$$y^*{}_t = \beta_0 + \beta_1 x_{1t}{}^* + \beta_2 x_{1t}{}^{*2} + \beta_3 x_{2t}{}^* + \beta_4 x_{2t}{}^{*2} \qquad (6)$$

Then, we are interested in estimating

$$\frac{\delta y_t{}^*}{\delta x_{i1}{}^*} = \beta_1 + 2\beta_2 x_{1t}{}^* \qquad (7)$$

Thus, we estimate the TVC model

$$y_t = \beta_{0t} + \beta_{1t} x_{1t} \qquad (8)$$

And our driver equations would be

$$\beta_{0t} = \pi_{00} + \sum_{i=1}^{p-1} \pi_{0i} Z_{it} + \varepsilon_{0t}$$

$$\beta_{1t} = \pi_{10} + \pi_{1p+1}x_{1t} + \sum_{i=1}^{p-1} \pi_{1i}Z_{it} + \varepsilon_{1t} \quad ^{5} \tag{9}$$

where there are p-1 coefficient drivers, Z, which are correlated with the measurement errors and the missing variable x_2. Now we can see why this will give a consistent estimate of the true coefficients as when we remove the effect of the Z variables from (9) we get

$$\beta_{0t} - \sum_{i=1}^{p-1} \pi_{0i}Z_{it} - \varepsilon_{0t} = \pi_{0t} \tag{10}$$

$$\beta_{1t} - \sum_{i=1}^{p-1} \pi_{i}Z_{it} - \varepsilon_{1t} = \pi_{10} + \pi_{1p+1}x_{1t} \tag{11}$$

Thus, by equating the right hand side of (11) with (7) we can see that they are identical; in other words, the bias free component of the TVC will then be a consistent estimate of the true partial derivative.

$$E(\pi_{10} + \pi_{11}x_{1t}) \rightarrow \beta_1 + 2\beta_2 x_{1t} \tag{12}$$

It is also easy to see in this simple example exactly how the coefficient drivers correct for the omitted variable bias. What is required for a good coefficient driver is something that is well-correlated with the omitted variables. Consider an extreme example in which the omitted variables themselves, in this case x_2 and x_2^2, are used as the drivers. Then the equation for the time-varying constant becomes

$$\beta_{0t} = \pi_{00} + \pi_{01}x_{2t} + \pi_{02}x_2^{\ 2} + \varepsilon_{0t} \tag{13}$$

The time-varying constant, then, represents exactly the omitted variables and the final equation is correctly specified and will, therefore, give consistent parameter estimates. Of course, in general the coefficient drivers will not be perfectly correlated with the omitted variables, but as long as the correlation is relatively high the estimate of the unbiased effect of x_1 will be consistent.

Since we would not know whether the true model is quadratic, we could include higher order polynomial terms and test their significance in the usual way to see how many

[5] It is important to note that none of the variables included in the TVC equation such as (8) can be included as a coefficient driver in the time-varying constant equation since the model is then unidentified.

polynomial terms would be needed. If the non-linearity was not, in fact, a polynomial, then there are two possible courses of action.

1. We could include a number of polynomial terms and think of this as a Taylor series approximation to the true unknown form.
2. We could try a range of specific non-linear forms, again testing one form against another.

By using the second option the standard TVC model is able to nest a number of popular non-linear models within a single framework, which also allows for measurement error and missing variables. This procedure is very different from other standard procedures. For example, a popular non-linear model is the smooth transition auto-regressive model (STAR). This allows a parameter to move smoothly between two values according to a function which responds to some threshold variable. If we were estimating a TVC such as (14) but believed the true non-linearity followed a STAR form, then we could specify the driver equations as

$$\beta_{0t} = \pi_{00} + \sum_{i=1}^{p} \pi_{01+i} Z_{it} + \varepsilon_{0t} \tag{16}$$

$$\beta_{1t} = \pi_{10} + \pi_{11} G(z_t, \zeta, c) + \pi_{12}(1 - G(z_t, \zeta, c)) + \sum_{i=1}^{p} \pi_{12+i} Z_{it} + \varepsilon_{1t} \tag{17}$$

Where $G(z_t, \zeta, c)$ is the transition function -- typically a logistic function for the LSTAR model, Z an exponential function for the ESTAR model or a second order logistic function, z is the transition variable and ζ and c are parameters. The model given by (17) is more general than the standard STAR model as it includes the drivers associated with measurement error and missing variables and will, hence, correct for these misspecifications. Model (17) also has a stochastic error term and thus becomes a stochastic STAR model. The split into the two sub sets is again obvious as the two terms capturing the STAR effect are clearly the set which is appropriate for S_1.

Another interesting non-linearity would be to use a set of combinations of simple non-linear functions such as the log or exponent function. In this case the TVC model would begin to encompass a neural net and the universal approximation theorems of Hal White would suggest that with sufficient complexity the model could then approximate any unknown functional form to any degree of accuracy.

Thus, to formalize the above approach for deriving the split in the coefficient drivers, we propose to choose the initial drivers in two ways. The first way involves specifically choosing nonlinear variables so as to capture any non-linearity that may be present. The second set captures omitted variables and measurement error and would normally not involve nonlinear transformations of the basic variables in the model. Once the drivers are chosen in this way, the split of the drivers so that they reveal the unbiased coefficient is straightforward -- it simply follows the two sets that were initially created.

Equation (1) and (2) are essentially a state space form where (1) is the measurement equation and (2) are the state equations. This means that all of these models may be estimated using the Kalman filter, which can estimate all of the parameters of the model by maximizing the likelihood function. The Kalman filter and the state space form must be linear in the state variables, but it can easily handle non-linearities in the other variables; all of the variants above may be estimated in standard software such as EVIEWS. An Appendix to this paper provides the EVIEWS code that was used to estimate the example below.

4 An Application:

In this section we investigate the effects of ratings agencies decisions on the sovereign bond spread between Greece and Germany. The underlying hypothesis is that this relationship is highly non-linear, a decline in ratings by one category from say AA to A will have a relatively small effect on spreads while a decline from say B to CCC will have a much larger effect. So as the rating goes down the effect on spreads get proportionally larger.

We believe there is a highly non-linear relationship between spreads and ratings. Of course, there are also many other things which might affect spreads, debt, deficits relative prices, politics etc. Therefore, if we examine a simple relationship between spreads and ratings there will be many omitted variables, which would cause bias for a standard OLS regression.

Our basic TVC model is then

$$sp_t = \alpha_{0t} + \alpha_{1t} rate$$

Where sp is the spread between 2 year Greek government bonds and rate is the agencies' ratings, turned into a numerical score where 1 is the best rating and 20 is the worse.

Our coefficient driver equations take the form:

$$\alpha_{0t} = \pi_{00} + \pi_{11}\text{pol} + \pi_{21}\text{dgdp} + \pi_{31}\text{cnewssq} + \pi_{41}\text{relp} + \pi_{51}\text{debtogdp} + \varepsilon_{1t}$$
$$\alpha_{1t} = \pi_{10} + \pi_{11}\text{rate} + \pi_{12}\text{pol} + \pi_{13}\text{dgdp} + \pi_{14}\text{cnewssq} + \pi_{15}\text{relp} + \pi_{16}\text{debtogdp} + \varepsilon_{2t}$$

In this driver set, rate gives a quadratic effect to the coefficient on ratings and allows for a strong non-linearity.

We begin by estimating this general model, to obtain the following results

$$\underset{(0.6)}{\alpha_{0t} = -4.2} \underset{(0.4)}{- 0.05\text{pol}} + \underset{(13.4)}{-0.9\text{dgdp}} + \underset{(2.9)}{0.03\text{cnewssq}} + \underset{(3.2)}{23.27\text{relp}} + \underset{(1.6)}{0.05\text{debtogdp}} + \varepsilon_{1t}$$

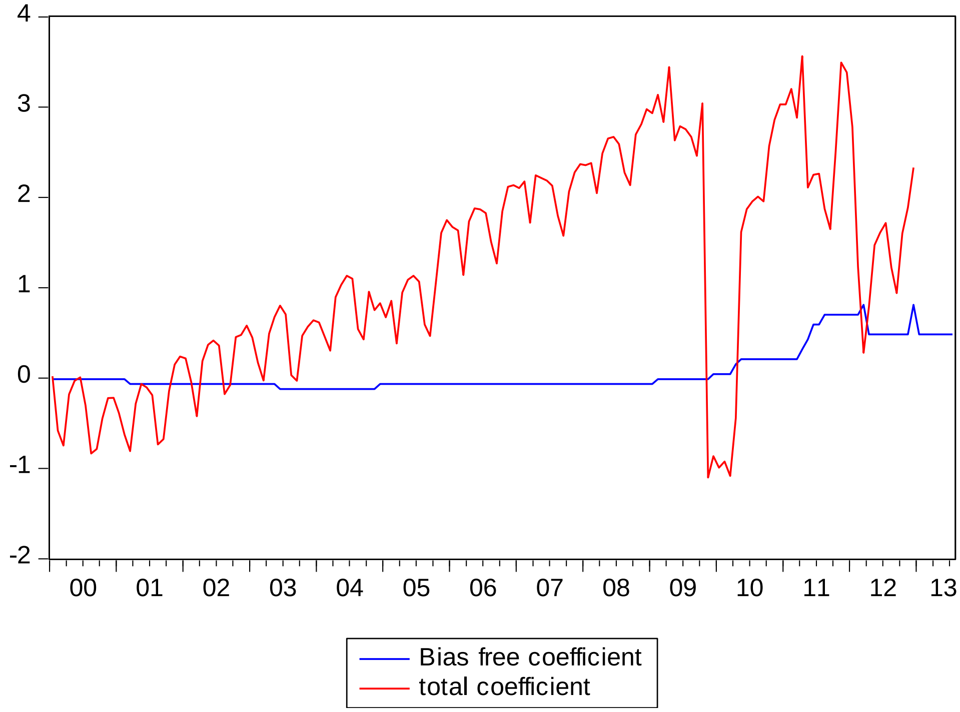
$$\underset{(1.2)}{\alpha_{1t} = -0.6} + \underset{(17.4)}{0.07\text{rate}} + \underset{(0.1)}{0.004\text{pol}} + \underset{(0.18)}{0.27\text{dgdp}} \underset{(3.2)}{- 0.005\text{cnewssq}} \underset{(3.35)}{- 3.5\text{relp}} + \underset{(0.12)}{0.0004\text{debtogdp}} + \varepsilon_{2t}$$

't' statistics are in parenthesis. The R^2_1=0.80 and R^2_2=0.84, so both sets of drivers are producing a reasonably high degree of explanatory power for the two TVC's. A number of drivers in each equation are insignificant, however, so we now simplify the model to exclude insignificant drivers and then obtain the following more parsimonious model (note we do not exclude the constants even if insignificant)

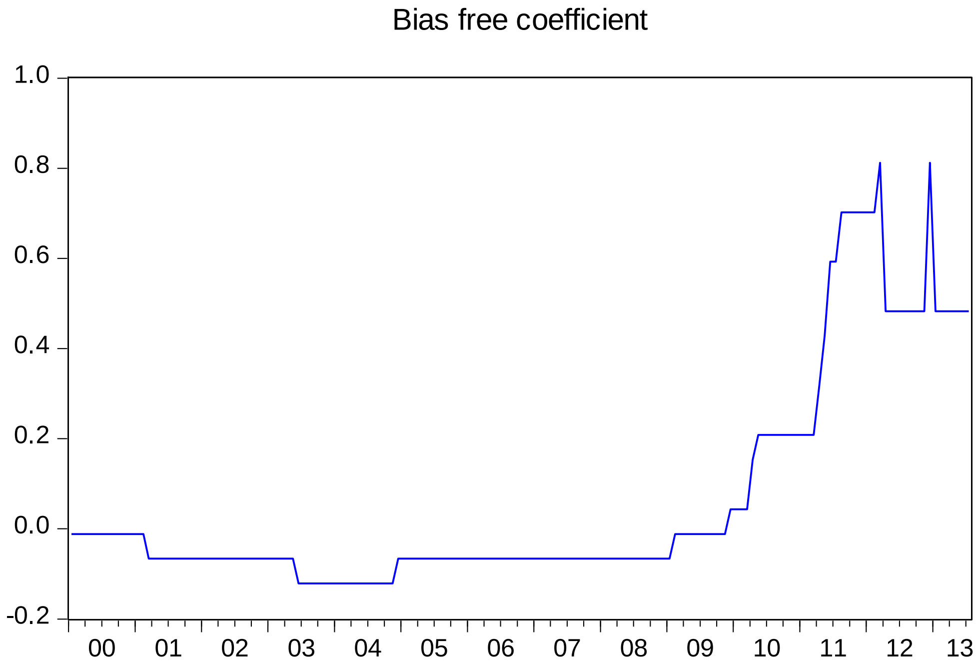
$$\underset{(2.0)}{\alpha_{0t} = -3.5} + \underset{(2.9)}{0.03\text{cnewssq}} + \underset{(3.9)}{15.6\text{relp}} + \underset{(2.7)}{0.04\text{debtogdp}} + \varepsilon_{1t}$$

$$\underset{(1.8)}{\alpha_{1t} = -0.64} + \underset{(20.9)}{0.05\text{rate}} \underset{(3.8)}{- 0.005\text{cnewssq}} \underset{(4.2)}{- 2.3\text{relp}} + \varepsilon_{2t}$$

The coefficient of interest is α_{1t}, The following figure shows the total value of this TVC along with the bias free effect which is given by subtracting the error term, and the effect from rate and cnewssq.



Because the scaling of the two coefficients is quite different the following figure shows just the bias free coefficient



This makes the strong non-linearity clear, as ratings started to rise (that is deteriorate). After 2008 the effect of ratings on spreads became increasingly powerful. We, therefore, have found a very strong quadratic link between ratings and spreads.

5 Conclusions

This paper has proposed a new way of selecting the split between coefficient drivers when deriving the bias free estimate of a coefficient within the TVC estimation framework. We have argued that if the true model is linear then only the constant in the coefficient driver set should be retained. If the true model is non-linear then an explicit set of drivers should be chosen which capture the nonlinearity, in the absence of any specific information this can best be done by using a set of polynomials in the explanatory variables. These drivers are then the only ones which should be retained when deriving the bias free component.

We have also argued for two conditions to be applied to the drivers which we call predictive power and relevance, that is the drivers should explain a large proportion of the movement in the TVC and they should be statistically significant.

We illustrate this process by estimating a nonlinear relationship between country risk rating and sovereign bond spreads for Greece and show that there is a highly non-linear effect here. Finally all this may be done within standard software such as EVIEWS and the code for the estimated model is given in appendix A.

Appendix A: EVIEWS code for the reported model in section 4

```
@signal sp_gr = sv1 + sv2*rate_gr

@state sv1 = c(1) +c(6)*pol_gr+c(7)*dgdp_gr+c(8)*cnewssq_gr+c(9)*relp_gr+c(10)*debtogdp_gr+sv3(-
1)+c(16)*sv4(-1)
@state sv2 =
c(2)+c(3)*rate_gr+c(11)*pol_gr+c(12)*dgdp_gr+c(13)*cnewssq_gr+c(14)*relp_gr+c(15)*debtogdp_gr +sv5(-
1)+c(17)*sv6(-1)

@state sv3=[var = exp(-56.88)]
@state sv4=sv3(-1)
@state sv5= [var = exp(-4.748)]

@state sv6=sv5(-1)
```

In this code SP_gr is the spread between Greek and German 2 year government bond yields, Rate_gr is the ratings on country risk denoted as a variable scaled from 1 to 20 where 1 is equivalent to a AAA rating, pol_gr is an indicator of political unrest in Greece, cnewssq is an indicator of news about Greek GDP derived from ECB forecasts., relp_gr is relative process between Greece and Germany, debtogdp_gr is the debt to GDP ratio for Greece.

This model is also coded to have a first order moving average error process on the two time varying coefficients sv1 and sv2, which is usual in TVC models.

The c(?) parameters are the usual EVIEWS notation for parameters to be estimated by maximum likelihood.

References

Barndorff-Nielsen, O., and D. R. Cox (1979): "Edgeworth and saddlepoint approximations with statistical applications," Journal of the Royal Statistical Society B, 41, 279-312.

Basmann, R. L. (1988): "Causality tests and observationally equivalent representations of econometric models," Journal of Econometrics, Annals, 39, 69-104.

Brown, L. D. (1990): "An ancillarity paradox which appears in multiple linear regression (with discussion). Annals of Statistics, 18, 471-538.

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

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

Chang, I-L., P. A. V. B. Swamy, , C. Hallahan, and G. S. Tavlas (2000): "A computational approach to finding causal economic laws," Computational Econonomics, 16, 105-136.Cuthbertson K., S. G. Hall, and M. P. Taylor (1992): Applied Econometric Techniques. Ann Arbor: University of Michigan Press.

Cuthbertson K., and M. P. Taylor (1990): "The case of the missing money and the Lucas critique," Journal of Macreconomics, 12, 437-454. DeGroot, M. H. (1982): "Comment on 'Lindley's paradox' by G. Shafer," Journal of the American Statistical Association, **77**, 336-339.Durbin, J., and S. J. Koopman (2001): Time Series Analysis by State Space Methods. Oxford: Oxford University Press.
Granger, C. W. J. (2008): "Non-linear models: Where do we go next – time varying parameter models," Studies in Nonlinear Dynamics and Econometrics, 12, 1-10.

Granger, C. W. J., and P. Newbold (1974): "Spurious regression in econometrics," Journal of Econometrics, 2, 111-120.

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

Hall, S. G., G. Hondroyiannis, P. A. V. B. Swamy, and G. S. Tavlas (2009a): "The new Keynesian Phillips curve and lagged inflation: A case of spurious correlation?" Southern Economics Journal, 76, 467-481.

Hall, S. G., G. Hondroyiannis, P. A. V. B. Swamy, and G. S. Tavlas (2009b): "Where has all the money gone? Wealth and the demand for money in South Africa," The Journal of African Economics, 18, 84-112.

Hall, S. G., G. Hondroyiannis, P. A. V. B. Swamy, and G. S. Tavlas (2008): "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.

Havenner, A., and P. A. V. B. Swamy (1981): "A random coefficient approach to seasonal adjustment of economic time series," Journal of Econometrics, 15, 177-209.

Hondroyiannis, G., P. A. V. B. Swamy, and G. S. Tavlas (2009): "The new Keynesian Phillips curve in a time-varying coefficient environment: some European evidence," Macroeconomic Dynamics, 13, 149-166.

Judge, G. G., W. E. Griffiths, R. C. Hill, H. Lütkepohl, and T. Lee (1985): The Theory and Practice of Econometrics, 2nd edition. New York: John Wiley and Sons.

Kariya, T., and H. Kurata (2004): Generalized Least Squares. Hoboken, NJ: John Wley & Sons, Ltd.

Lehmann, E. L. (1999): Elements of Large-Sample Theory. Berlin Heidelberg New York: Springer. Lehmann, E. L., and G. Casella(1998): Theory of Point Estimation, 2nd edition. Berlin Heidelberg New York: Springer.

Lucas, R. E., Jr. (1976): "Econometric policy evaluations: a critique," in K. Brunner and A. Meltzer, eds.,The Phillips Curve and the Labor Market, Carnegie-Rochester Series in Public Policy. Amsterdam: North Holland, 19-46.

Pratt, J. W., and R. Schlaifer (1984): "On the nature and discovery of structure," Journal of American Statistical Association, 79, 9-21, 29-33.

Pratt, J. W., and R. Schlaifer(1988): "On the interpretation and observation of laws," Journal of Econometrics, Annals, 39, 23-52.

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

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

Swamy, P.A.V.B. (1974). Linear Models with Random Coefficients. In Zarembka (ed.), *Frontiers in Econometrics*, pp. 143-168. New York: Academic Press.

Swamy, P. A. V. B., and J. S. Mehta (1975): "Bayesian and non-Bayesian analysis of switching regressions and a random coefficient regression model," Journal of the American Statistical Association, 70, 593-602.

Swamy, P. A. V. B., J. S. Mehta, and R. N. Singamsetti (1996): "Circumstances in which different criteria of estimation can be applied to estimate policy effects," Journal of Statistical Planning and Inference, 50**,** 121-153.

Swamy, P. A. V. B., and G. S. Tavlas (2001): "Random coefficient models," in B. H. Baltagi, ed., A Companion to Theoretical Econometrics, Malden: Blackwell.

Swamy, P. A. V. B., and G. S. Tavlas (2006): "A note on Muth's rational expectations hypothesis: A time-varying coefficient interpretation," Macroeconomic Dynamics, 10, 415-425 (2006)

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

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

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

Swamy, P. A. V. B., J. S. Mehta, I-L Chang, and T. S. Zimmerman(2009): "An efficient method of estimating the true value of a population characteristic from its discrepant estimates," Computational Statistics & Data Analysis, 53, 2378-2389.

Swamy, P. A. V. B., and J. S. Mehta (2009): "Two-step versus simultaneous estimation of survey-non-sampling error and true value components of small area sample estimators," Unpublished manuscript.

Swamy, P.A.V.B., G.S. Tavlas, S.G. Hall, and G. Hondroyiannis. (2010). Estimation of Parameters in the Presence of Model Misspecification and Measurement Error. *Studies in Nonlinear Dynamics and Econometrics,* 14, 1, 1-33.

Yokum, J. T., A. R. Wildt, and P. A. V. B. Swamy (1998): "Forecasting disjoint data structures using appropriate constant and stochastic coefficient models," Journal of Applied Statistical Science, 8, 29-49.

Zellner, A. (1979): "Causality and econometrics," in K. Brunner and A. H. Meltzer, eds., Three Aspects of Policy and Policymaking, Amsterdam: North-Holland.

Zellner, A. (1988): "Causality and causal laws in economics," Journal of Econometrics, Annals, 39, 7-21.