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

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

Introduction

Most econometric relationships are subject to at least the following three problems

- Measurement Error
- The true functional form is unknown
- Omitted variables

TVC estimation is a technique which deals with all of these problems at once

However one weakness of the TVC approach is determining the split of the coefficient drivers into the set which determines the unbiased coefficient and the set which captures all the biases

This is a crucial weakness in TVC estimation.

This paper makes a suggestion for formalising this split

2. The Interpretation of Model Coefficients

Consider the relationship between y_t^* an endogenous variable and K-1 of its determinants

$x_{1t}^* \ldots\ldots x_{K-1,t}^*$

Where in particular K-1 may be only a subset of the full set of determinants so that we have omitted variables . In addition we have measurement error

$y_t = y_t^* + v_{0t}, \; x_{jt} = x_{jt}^* + v_{jt}$

And we may be estimating the wrong functional form

What is the econometrician trying to achieve?

Our answer to this is to derive an estimate of the partial derivative of y_t^* with respect to x_{jt}^*, and to test hypothesis about this.

If you want to estimate the true model then there really is no alternative to specifying it correctly.

However if you are only interested in partial effects then we offer another way forward.

Consider the following time varying parameter model

$$y_t = \gamma_{0t} + \gamma_{1t} x_{1t} + \ldots + \gamma_{K-1,t} x_{K-1,t}$$

All potential misspecification is captured in the time varying coefficients which offer a complete Explanation of y.

Now the key question is what are the stochastic assumptions about the TVCs

The correct stochastic assumptions about the TVC comes from an understanding of the misspecification which drives the time variation.

Notation and assumptions

Let m_t denote the total number of determinants of y, this can not generally be known, but in general m>k-1.

Now let α^*_{0t} and α^*_{jt} j=1…K-1 and α^*_{gt} g=k…m be the true coefficients on the underlying model, where the parameters vary because of either a non-linear functional form or truly changing parameters

Now for g=k…m_t let λ^*_{0gt} denote the intercept and λ^*_{jgt} j=1…K-1 denote the other coefficients of the regression of x^*_{gt} on x^*_{1t} …$x^*_{K-1,t}$

Then we can establish the following representation

Theorem 1.

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

And

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

The first term is the true variation, the second the measurement effect, the third the omitted variables.

And if we estimate a standard fixed coefficient regression model then the error term comprises all these effects.

This means that the error term can not be independent of the included Xs (as it contains them). It is also impossible for valid instruments to exist in this example as if a variable is correlated with the included X variables it must be correlated with the errors (as the error contains the included Xs)

Theorem 2:
For j=1…K-1 the component α_{jt}^{*} of γ_{jt} is the direct or 'biased free' effect of x_{jt}^{*} on y_t^{*} with all the other determinants of y_t^{*} held constant, and it is unique.

The direct effect will be constant if the relationship between y and all the Xs is linear and time invariant.

This is a useful interpretation of standard regression coefficients, which emphasises their potential biases.

To make this approach useful as an estimation strategy we must have some way of identifying the bias part of the TVC.

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

$$\gamma_{jt} = \pi_{j0} + \sum_{d=1}^{p-1} \pi_{jd} z_{dt} + \varepsilon_{jt}$$

Assumption 2: For j=1,…,K-1, the set of P-1 coefficient drivers and the constant term divides into two disjoint subsets S1 and S2 so that $\pi_{j0} + \sum_{d \in S_1} \pi_{jd} z_{dt}$ 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 RHS of (3) over the relevant estimation and forecasting periods.

So we assume the drivers identify the bias component. This is like the dual of IV

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

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

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

A vector formulation of the model

$$y_t = x_t' \gamma_t$$

Where

$$\gamma_t = \Pi z_t + \varepsilon_t$$

And

$$\Pi = \left[\pi_{jd} \right]_{0 \le j \le K-1, 0 \le d \le p-1}$$

Then $$y_t = (z_t' \otimes x_t') \pi^{Long} + x_t' \varepsilon_t$$

or $$y = X_z \pi^{Long} + D_x \varepsilon$$

Theorem 3

Under Assumptions 1-4

$$\mathrm{E}(y \mid X_z) = X_z \pi^{Long}$$

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

3. Identification and Consistent Estimation of TVC

The fixed coefficient vector π^{Long} is identified if $.X_z$ has full column rank. A necessary condition for this is that T>Kp.

The errors are not identified

Thus assumptions 1-4 make all the fixed parameters of the model identifiable.

This does not happen if we assume random walk TVCs

Practical Estimation

The TVC Model can be estimated by an iteratively rescaled generalized least squares (IRSGLS) method developed in Chang, Swamy, Hallahan and Tavlas (2000).

Or alternatively it can be specified in state space form and estimated by maximum likelihood.

Choice of coefficient drivers

The main contribution of this paper is to explore what are suitable drivers and in particular how to make the split between the two sets S1 and S2.

The essence of the proposal here is that the variables in S1 should only be there to reflect nonlinearity and hence time variation in the true unbiased coefficient. Hence the S1 variables should be chosen to reflect this.

What makes a good driver set?

The drivers should be

1. Relevant
2. With high explanatory power.

How to Judge this

Explanatory power

An analogue to the standard R^2

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

Relevance

The π_{jd} Should be individually significant

How to choose the split in the Driver set

The idea here is to choose special set of drivers to capture non linearity, everything else then goes into S2

If the true model is $y^*{}_t = f(x_{1t}{}^* \ldots . x_{mt}{}^*)$

Then we are interested in estimating

$$\frac{\delta y_t{}^*}{\delta x_{it}{}^*} \quad for\, i = 1, \ldots, K-1$$

Example 1

If the true model 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 the true model is a polynomial, such as a quadratic form. Consider, for simplicity, the case of only 2 explanatory variables. Then,

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

We wish to estimate

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

We estimate the TVC model

$$y_t = \beta_{0t} + \beta_{1t} x_{1t}$$

And the driver equations would be

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

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

Removing the Z drivers gives

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

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

And

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

We can also see how the Z drivers remove the omitted variable bias, the drivers should be correlated with the omitted variables so lets take an extreme case and make the drivers the two omitted variables, then

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

And the model is well specified as the missing variables are all in the time varying constant

The General case

Generally we do not know the form of the nonlinearity, options then are

- We could include a number of polynomial terms and think of this as a Taylor series approximation to the true unknown form.
- We could try a range of specific non-linear forms, again testing one form against another.
- We could include a number of simple non linear transformations such as a LOG of x, in which case the TVC model will work like a neural net.

For example the following pair of coefficient equations will allow us to capture a generalization of a STAR model

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

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

Where $G(z_t, \zeta, c)$ is the transition function and Z captures ommited variables and measurement error

The split into the two sub sets is again obvious

- **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,

- Our basic TVC model is then

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

- And the coefficient driver equations take the form

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

We estimate this general model to give

$$\alpha_{0t} = -4.2 - 0.05\text{pol} + -0.9\text{dgdp} + 0.03\text{cnewssq} + 23.27\text{relp} + 0.05\text{debtogdp} + \varepsilon_{1t}$$
$$(0.6) \quad (0.4) \quad (1.3) \quad (2.9) \quad (3.2) \quad (1.6)$$

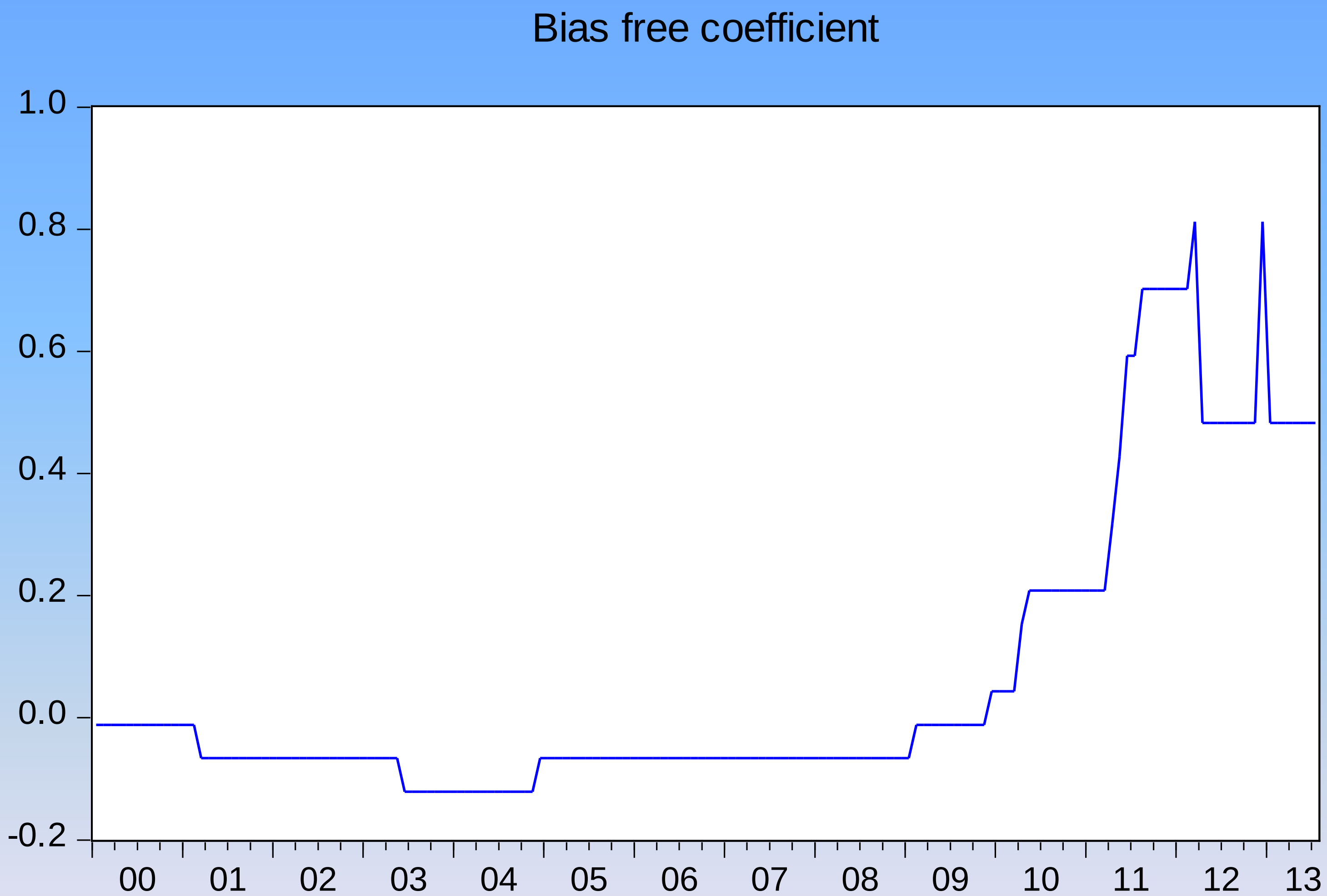
$$\alpha_{1t} = -0.6 + 0.07\text{rate} + 0.004\text{pol} + 0.27\text{dgdp} - 0.005\text{cnewssq} - 3.5\text{relp} + 0.0004\text{debtogdp} + \varepsilon_{2t}$$
$$(1.2) \quad (17.4) \quad (0.1) \quad (0.18) \quad (3.2) \quad (3.35) \quad (0.12)$$

The R^2_1=0.80 and R^2_2=0.84 which is reasonably high and we then eliminate insignificant drivers

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

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

This gives the following bias free coefficient on ratings



Kalman filter formulation in EVIEWS

```
@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)*cnewss
q_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)
```

Conclusion.

We have proposed a new way of selecting coefficient drivers in the TVC framework.

This allows us to make the split of the drivers much more easily and in an intuitive way.

It also allows us to generalise a number of standard non linear models to allow for both a stochastic term in the coefficient equation and to allow for biases from omitted variables and measurement error