Sankhyā: The Indian Journal of Statistics
2000, Volume 62, Series B, Pt. 3, pp. 433–447

ESTIMATING TIME-VARYING PARAMETERS IN LINEAR REGRESSION MODELS USING A TWO-PART DECOMPOSITION OF THE OPTIMAL CONTROL FORMULATION

By M.J. MANOHAR RAO
University of Bombay, Mumbai

SUMMARY. This paper discusses an econometric technique based on optimal control theory which, by employing a variation of the near-neighbourhood search problem, is seen to be suitable for the type of research that requires estimating time-varying parameters for linear regression models. The methodology is based on the characterization of the time-varying parameter (TVP) problem as an optimal control problem, with an explicit allowance for welfare loss considerations, which leads to an algorithm capable of updating the values of the time-varying parameters as well as their covariance matrices. The technique adopts an instruments-targets approach, with the initial condition and the emphasis on parameter flexibility being the instruments; and the total welfare loss and the norm of the error vector being the targets. The methodology is a blend of the flexible least squares and Kalman filter techniques. By determining all the required priors endogenously, it is seen to overcome some of the drawbacks associated with these two earlier approaches to the TVP problem. The method works on the premise that the dynamics of the system are determined by the system itself without being specified by the user in an arbitrary fashion.

1. Introduction

It is always tempting to argue that the parameters in econometric models cannot, in general, be expected to remain constant and hence we have to consider a time varying parameter (TVP) model in almost all circumstances. However specious this argument may be, in some situations it becomes necessary to resort to TVP models. The difficulty in estimating such models is however often exacerbated by the fact that the econometrician would have only some idea regarding the most likely value that a parameter may assume (as indicated by, say, the OLS and maximum likelihood estimators); with a range of uncertainty surrounding this 'nominal' value and consequently, misleading policy prescriptions are likely to arise from a straightforward optimization exercise based on such a set of 'nominal' values especially in the presence of structural breaks in the underlying economic, technological, behavioural

Paper received June 1995; revised November 1999.
AMS (1991) *subject classification.* 62F03.
Key words and phrases. TVP models, near-neighbourhood search problem, parameter flexibility, tracking error, smoothing vector, welfare loss, norm function, updating algorithm, instruments-targets approach.

and institutional patterns. As such discontinuities are a crucial feature of modern economic systems, several estimation methods have been proposed in the literature to deal with the problem of estimating TVP models.

Such TVP models can be classified into three types (see Judge *et al* 1985). First, the parameters can vary across subsets of observations within the sample but be non-stochastic. Examples of such models include a general systematically varying parameter model (see Belsley 1973) and a variety of "switching regression" models with either known joint points (see Quandt 1958, McGee and Carlton 1970, Hinkley 1971, Goldfeld and Quandt 1973) or unknown joint points (see Gallant and Fuller 1973, Poirier 1976).

A second class of models is where the parameters are stochastic, and are assumed to be generated by a stationary stochastic process. Examples of such models include the pure random coefficient model of Rao (1965), which was extended by Swami (1970, 1971) to combine time-series and cross-sectional data, the Hildreth-Houck model (see Hildreth and Houck 1968) and its generalization, the return to normality model (see Harvey and Phillips 1982) which includes the adaptive and varying-parameter regression models of Cooley and Prescott (1973a, b) and the stochastically convergent parameter model of Rosenberg (1973).

Finally, the third class of models consists of those where the stochastic parameters are generated by a process that is not stationary. These include the mixed estimation model of Cooper (1973), the Kalman filter model of Athans (1974), the stochastic variations model of Cooley and Prescott (1976), the random and changing coefficients model of Chow (1984), the systematically varying parameter model of Kalaba and Tesfatsion (1980) which was then extended to the flexible least squares (FLS) approach (see Kalaba and Tesfatsion 1988), the recursive model of Rao (1991) and the optimal control model of Rao and Nachane (1988).

Since the current paper attempts to estimate time-varying parameters using an optimal control formulation on lines broadly similar to the FLS method, it is essential to compare these two methods.

At the outset, it needs to be noted that the FLS method is a particular case of the larger class of problems envisaged by the trajectory sensitivity minimization method of Byrne and Burke (1976) which permits the direct incorporation of robustness considerations in the objective function (see Nachane and Rao 1994). The FLS method proposed for the time-varying linear regression problem consists of tracing out the residual possibility set and then determining the optimal coefficient sequence estimates which attain the so-called 'residual efficiency frontier' i.e., which yield minimal pairs of squared residual dynamic error and measurement error sums. The use of a quadratic loss function implies that the resulting problem can be solved within the framework of optimal control.

While the FLS method is computationally straightforward and permits the exact sequential updating of the FLS estimates as additional observations are obtained, it does have several limitations (see Tucci 1990). The most important one is that it assigns a single weight (called the incompatibility cost) to the squared residual dynamic errors which measure parameter flexibility. It then attempts to attain the residual efficiency frontier by varying this incompatibility cost. By doing so, it fails to distinguish between large and small coefficient values in a regression equation and as a squared loss function mechanically penalizes higher absolute deviations without considering their relative magnitudes, it automatically implies that larger coefficients would be relatively more stable vis-a-vis smaller ones in any application of the FLS method.

Secondly, the solution of any optimal control problem presupposes the existence of an initial condition. In the FLS method, the initial (or starting) values of the coefficients can be selected a priori and therefore it is possible to run the control algorithm with alternative initial parameter estimates (for example, OLS or 2SLS or maximum likelihood estimates). Because alternative values of the initial conditions would yield completely different optimal trajectories (or coefficient sequence estimates) even for the same value of the incompatibility cost, it implies that the FLS method is incapable of providing a truly optimal or unique coefficient sequence.

Finally, the FLS method, being cast in a completely deterministic framework, does not have the capability to automatically update the covariance matrices of the system state.

In order to overcome these shortcomings of the FLS method, this paper characterizes the TVP problem as an optimal control problem, with a two-part decomposition into a deterministic and stochastic component, incorporating an explicit allowance for welfare loss considerations. This variation leads us to an algorithm which, by endogenously determining the relative emphasis on parameter flexibility and tracking error, optimally updates the values of the parameters in each time period based on past prediction errors. The methodology is then extended to specify a first-order difference equation for updating the covariance matrices for the estimated optimal control trajectory.

2. Optimal Control of the Time Varying Parameter Model

The classical linear multiple regression model is given by:

$$y_t = x_t'\beta + u_t, \quad t = 1, \ldots, N \tag{2.1}$$

where y_t is an observation on the dependent variable, x_t is a $(k \times 1)$ vector of observations on the k explanatory variables, β is the $(k \times 1)$ vector of parameters and u_t is the random disturbance term with $u_t \sim N(0, \sigma^2)$.

If we assume a TVP model, then the optimal control problem, employing the near-neighbourhood search concept, would be to find a sequence β_t^* $(t = 1, 2, \ldots, N)$ such that:

$$y_t^* = x_t'\beta_t^* \tag{2.2}$$

and the quadratic cost functional:

$$W = \frac{1}{2}\sum_{t=1}^{N} Q_t(y_t^* - y_t)^2 + \frac{1}{2}\sum_{t=1}^{N}(\beta_t^* - \beta_{t-1}^*)' R_t(\beta_t^* - \beta_{t-1}^*) \tag{2.3}$$

is minimized. The advantage of such a loss function is that unlike other schemes which expect β_t^* to track some given β (see Lee and Griffiths 1979), the present one just constrains β_t^* to lie within the neighbourhood of β_{t-1}^*.

To solve the control problem, we introduce the Lagrange multipliers ϕ_t and differentiate the Lagrangian expression:

$$J = \frac{1}{2}\sum_{t=1}^{N} Q_t(y_t^* - y_t)^2 + \frac{1}{2}\sum_{t=1}^{N}(\beta_t^* - \beta_{t-1}^*)' R_t(\beta_t^* - \beta_{t-1}^*) - \sum_{t=1}^{N}\phi_t(y_t^* - x_t'\beta_t^*) \quad (2.4)$$

to obtain the following vectors of derivatives:

$$\delta J/\delta y_t^* = Q_t(y_t^* - y_t) - \phi_t = 0 \quad (t = 1, \ldots, N) \quad (2.5)$$

$$\delta J/\delta \beta_t^* = R_t(\beta_t^* - \beta_{t-1}^*) + x_t\phi_t = 0 \quad (t = 1, \ldots, N) \quad (2.6)$$

$$\delta J/\delta \phi_t = -(y_t^* - x_t'\beta_t^*) = 0 \quad (t = 1, \ldots, N) \quad (2.7)$$

To solve for the unknowns, we use eq. (2.5) which yields:

$$\phi_t = Q_t y_t^* - Q_t y_t \quad (2.8)$$

Substituting the expression for y_t^*, derived from eq. (2.7), into eq. (2.8) above yields:

$$\phi_t = Q_t x_t'\beta_t^* - Q_t y_t \quad (2.9)$$

Substituting the above result into eq. (2.6) yields an expression for β_t^* is terms of β_{t-1}^* given by:

$$\beta_t^* = (R_t + x_t Q_t x_t')^{-1} R_t \beta_{t-1}^* + (R_t + x_t Q_t x_t')^{-1} x_t Q_t y_t \quad (2.10)$$

The following features of this control scheme are important:

(1) If we set $Q_t = 0,\ t = 1, \ldots, N$, then eq. (2.10) reduces to $\beta_t^* = \beta_{t-1}^*$, for all t, implying time-invariant parameters.

(2) If we set $R_t = 0,\ t = 1, \ldots, N$, then eq. (2.6) becomes:

$$x_t\phi_t = 0 \quad (2.11)$$

Substituting eq. (2.9) in eq. (2.11), we have:

$$x_t Q_t x_t'\beta_t^* = x_t Q_t y_t \quad (2.12)$$

which reduces to $x_t'\beta_t^* = y_t$ implying that the total flexibility accorded to the parameter values forces the regression equation to track all the observations of the dependent variable.

(3) The 'adaptive' nature of the control scheme is apparent if we transform eq. (2.10). Initially, we assume that:

$$K_t = (R_t + x_t Q_t x_t')^{-1} x_t Q_t \quad (2.13)$$

Thus, the coefficient of β^*_{t-1} in eq. (2.10) becomes:

$$(R_t + x_t Q_t x'_t)^{-1}[(R_t + x_t Q_t x'_t) - x_t Q_t x'_t] = I - K_t x'_t \tag{2.14}$$

Substituting eqs. (2.13) and (2.14) into eq. (2.10) yields:

$$\beta^*_t = \beta^*_{t-1} + K_t(y_t - x'_t \beta^*_{t-1}) \tag{2.15}$$

We now have to find a set of restrictions on the covariance matrices of β^*_t when the above control rule is operative so as to estimate time-varying standard errors for the TVP model.

Subtracting eq. (2.2) from eq. (2.1) yields:

$$e_t = x'_t \overline{\beta}_t + u_t \tag{2.16}$$

where we have used the definitions:

$$e_t = y_t - y^*_t \tag{2.17}$$

$$\overline{\beta}_t = \beta - \beta^*_t \tag{2.18}$$

The existence of e_t makes the actual y_t deviate from the predicted y^*_t using the path β^*_t for the deterministic control problem. It is therefore necessary to modify the optimal setting β^*_t by $\overline{\beta}_t$ each period in order to control the deviations e_t.

Substituting eq. (2.2) into eq. (2.17) and the resulting expression for y_t into eq. (2.15) yields after simplification:

$$\beta^*_t = \beta^*_{t-1} + \Sigma_t e_t \tag{2.19}$$

where:

$$\Sigma_t = (I - K_t x'_t)^{-1} k_t \tag{2.20}$$

Lagging eq. (2.18) by one period, subtracting the result from eq. (2.18) and substituting the expression for $(\beta^*_t - \beta^*_{t-1})$ obtained from eq. (2.19) into it yields:

$$\overline{\beta}_t = \overline{\beta}_{t-1} - \Sigma_t e_t \tag{2.21}$$

Post-multiplying eq. (2.21) by $\overline{\beta}'_t$ and taking expectations:

$$E(\beta_t \overline{\beta}'_t) = E(\overline{\beta}_{t-1} \overline{\beta}'_t) - \Sigma_t E(e_t \overline{\beta}'_t) \tag{2.22}$$

Transposing eq. (2.21), pre-multiplying the result by $\overline{\beta}_{t-1}$ and taking expectations yields:

$$E(\overline{\beta}_{t-1} \overline{\beta}'_t) = E(\overline{\beta}_{t-1} \overline{\beta}'_{t-1}) - E(\overline{\beta}_{t-1} e_t)\Sigma'_t \tag{2.23}$$

Similarly, transposing eq. (2.21), pre-multiplying the result by e_t and taking expectations yields:

$$E(e_t \overline{\beta}'_t) \;=\; E(e_t \overline{\beta}'_{t-1}) - E(e^2_t)\Sigma'_t \;=\; E(e_t \overline{\beta}'_{t-1}) - V_t \Sigma'_t \tag{2.24}$$

where V_t is the variance of e_t.

Substituting eqs. (2.23) and (2.24) into eq. (2.22) yields:

$$E(\overline{\beta}_t\overline{\beta}_t') = E(\overline{\beta}_{t-1}\overline{\beta}_{t-1}') - E(\overline{\beta}_{t-1}e_t)\Sigma_t' - \Sigma_t E(e_t\overline{\beta}_{t-1}') + \Sigma_t V_t \Sigma_t' \tag{2.25}$$

We now obtain the expression for $E(e_t\overline{\beta}_{t-1}')$. To do so, we post-multiply eq. (2.16) by $\overline{\beta}_{t-1}'$ and take expectations to yield:

$$E(e_t\overline{\beta}_{t-1}') = x_t' E(\overline{\beta}_t\overline{\beta}_{t-1}') \tag{2.26}$$

because by definition:

$$E(u_t\overline{\beta}_{t-1}') = 0 \tag{2.27}$$

Transposing eq. (2.23) and substituting the result into eq. (2.26) above yields:

$$E(e_t\overline{\beta}_{t-1}') = x_t'[E(\overline{\beta}_{t-1}\overline{\beta}_{t-1}') - \Sigma_t E(e_t\overline{\beta}_{t-1}')] \tag{2.28}$$

which when solved uniquely in terms of $E(e_t\overline{\beta}_{t-1}')$ yields:

$$E(e_t\overline{\beta}_{t-1}') = \phi_t' E(\overline{\beta}_{t-1}\overline{\beta}_{t-1}') \tag{2.29}$$

where:

$$\phi_t = (1 + x_t'\Sigma_t)^{-1} x_t \tag{2.30}$$

Substituting eq. (2.29) into eq. (2.25) yields the desired difference equation for updating the covariance matrices, $E(\overline{\beta}_t\overline{\beta}_t') = \Gamma_t$, which is given by:

$$\Gamma_t = \Gamma_{t-1} - \Gamma_{t-1}\phi_t\Sigma_t' - \Sigma_t\phi_t'\Gamma_{t-1} + \Sigma_t V_t \Sigma_t' \tag{2.31}$$

Eq. (2.31) is solved forwards in time from $t = 1$ ato $t = N$, using the initial condition:

$$\Gamma_0 = E(\overline{\beta}_0\overline{\beta}_0') = 0 \tag{2.32}$$

which is due to the definition $\beta_0^* = \beta$ or $\overline{\beta}_0 = 0$.

3. Estimating Time Varying Parameters: A Numerical Illustration

To illustrate the applicability of this algorithm, we consider a numerical example providing the steps involved in the computation of the time-varying coefficients of a one-parameter regression equation, by means of which we can conceptualize its generalization towards a k-parameter regression model.

The following TVP model specifies that money supply (M) is a linear function of high-powered money (H):

$$M_t = m_t H_t \tag{3.1}$$

where the coefficient m_t is the money multiplier at time t. The financial year-end data on M and H for the Indian economy over the 30-year period 1960/61-1990/91 are provided in Table 1. Both of these variables are measured in crores of Indian rupees (Rs.) where Rs. 1 crore = U.S. \$0.21 million. The estimated OLS model based on this data is $M_t = 1.167455H_t$.

Table 1. DATA FOR THE ESTIMATION OF TIME-VARYING PARAMETERS

Year	Narrow Money (M)	Reserve Money (H)	Year	Narrow Money (M)	Reserve Money (H)
1960-61	2869	2161	1975-76	13325	7808
1961-62	3057	2275	1976-77	16024	9798
1962-63	3316	2462	1977-78	14388	10941
1963-64	3792	2710	1978-79	17292	14083
1964-65	4127	2882	1979-80	20000	16573
1965-66	4570	3146	1980-81	23424	19452
1966-67	4950	3328	1981-82	24937	20998
1967-68	5401	3557	1982-83	28535	23110
1968-69	5838	3901	1983-84	33398	28993
1969-70	6470	4251	1984-85	39915	35216
1970-71	7373	4823	1985-86	44095	38165
1971-72	8322	5382	1986-87	51516	44808
1972-73	9700	6033	1987-88	58555	53489
1973-74	11200	7273	1988-89	71107	62958
1974-75	11975	7604	1989-90	85921	73788
			1990-91	99081	82808

In the *FIRST* iteration, the computation proceeds as follows:

STEP 1. Initially we assign time-varying weights, Q_t and R_t, $t = 1, \ldots, 31$, so as to ensure *equal* emphasis on tracking ability and parameter flexibility. Consider the question of determining Q_1 and R_1. In this first period, M_1 has to track Rs. 2869 crores while m_1 has to track m_0 which we will assume for the time being to be the OLS estimate of m, i.e., $m_0 = 1.167455$.

As we are penalizing for percent deviations from the nominal in each case and because the variables are scaled differently, their assigned weights will have different scales. Thus, as the desired value of M_1 was 2869, a one percent deviation from this path would have a magnitude of 28.69 and when this is squared, we end up with a number of magnitude 823.12. Alternatively, a one percent deviation from the desired time path of m_1 will have a magnitude of 0.01167455 and when this is squared, we obtain a number of magnitude 1.36295×10^{-4}. Therefore, our quadratic cost functional, e.q. (2.3), for the first period would be given by:

$$W_1 = \tfrac{1}{2}Q_1(823.12) + \tfrac{1}{2}R_1(1.36295) \times 10^{-4}) \qquad (3.2)$$

Thus, if the effective weights are to be identical, then the coefficient R_1, which corresponds to the money multiplier (m_1), must be higher than the coefficient Q_1, which corresponds to the money supply (M_1). This implies the following normalized weights:

$$Q_1 : R_1 = 1 : 823.12/(1.36295 \times 10^{-4}) \qquad (3.3)$$

which generalizes to:

$$Q_t : R_t = 1 : (M_t)^2/(m_{t-1})^2, \quad t = 1, \ldots, 31 \qquad (3.4)$$

In the case of a k-parameter regression model, such a scheme would yield the following generalization given by:

$$Q_t : R_{j,t} = 1 : (y_t)^2/(\beta^*_{j,t-1})^2, \quad j = 1, \ldots, k; \; t = 1, \ldots, N \qquad (3.5)$$

STEP 2: We now have to estimate the smoothing coefficient for the first time period, K_1. From eq. (2.13), this is given by:

$$K_1 = (R_1 + x_1 Q_1 x_1')^{-1} x_1 Q_1 \tag{3.6}$$

From eq. (3.3), we have:

$$Q_1 = 1 \quad \text{and} \quad R_1 = 6,039,253 \tag{3.7}$$

Substituting the above values, along with the value of x_1 set out in Table 1, i.e., H_1= 2161, we obtain:

$$K_1 = 0.000201789606 \tag{3.8}$$

which is the correction factor that has to be applied to the first-period *á priori* prediction error in order to update m_0.

STEP 3: The first-period *á priori* prediction error is the one made by forecasting the value of M_1 on the basis of the money multiplier of the earlier period, i.e., m_0. From eq. (2.15), and given the estimates set out in Table 1, this is seen to be:

$$\text{First-period } \textit{á priori} \text{ prediction error} \;=\; M_1 - m_0 H_1 \;=\; 346.1297 \tag{3.9}$$

STEP 4: We now update the estimate of $m_0 = 1.167455$ on the basis of eq. (2.15) in order to derive a fresh estimate of the money multiplier for period 1, i.e., m_1. This yields the following:

$$m_1 \;=\; m_0 + K_1(M_1 - m_0 H_1) \;=\; 1.237300 \tag{3.10}$$

STEP 5: We now compute the welfare loss in period 1. With this value of m_1, the *á posteriori* predicted value of M_1 is :

$$M_1 = 1.237300(2161) = 2673.81 \tag{3.11}$$

as against the actual value of 2869 (see Table 1). The welfare loss in the first period, W_1, is therefore given by eq. (2.3) as:

$$W_1 \;=\; \tfrac{1}{2}(2673.81-2869)^2 + \tfrac{1}{2}(6,039,253)(1.237300-1.167455)^2 \;=\; 33,780 \tag{3.12}$$

We are now back in step 1 of the *next* round, but still the *first* iteration, where we have to compute afresh the second period time-varying weighting coefficient R_2. Using eq. (3.4), with $M_2 = 3057$ and the freshly estimated $m_1 = 1.237300$ we have:

$$Q_2 : R_2 \;=\; 1 : (M_2)^2/(m_1)^2 \;=\; 1 : 6,104,370 \tag{3.13}$$

The development from here on follows steps 2-5 as before with the estimates of m_t being recursively updated in each round.

At the end of the last round of the first iteration, we would have 31 time varying parameters and 31 corresponding welfare losses, one for each year of the sample period. What is important to note is that the combined welfare loss of the resulting trajectory, initialized by setting m_0 equal to its OLS estimate of $\hat{m} = 1.167455$, amounted to 17,130,066.

STEP 6: Therefore, as the initial condition, m_0, determines the combined welfare loss, $W(m_0)$, of the associated trajectory, it is possible to identify an optimal trajectory by choosing an initial condition, m_0^*, which minimizes the welfare loss from amongst the set of 'near-neighbourhood' nominal trajectories.

The results of the grid-search for m_0^* are shown in Table 2.

Table 2. GRID SEARCH FOR THE OPTIMAL INITIAL CONDITION

m_0	$W(m_0)$	m_0	$W(m_0)$	m_0	$W(m_0)$
1.000	17,383,976	1.355	17,056,006	1.360	17,055,987
1.100	17,204,154	1.356	17,055,996	1.400	17,058,607
1.200	17,104,929	1.357	17,055,989	1.600	17,128,451
1.300	17,061,911	1.358	17,055,986	1.800	17,256,153
1.350	17,056,107	1.359*	17,055,985	2.000	17,410,557

STEP 7: As the grid-search clearly identified a global minimum, we set $m_0^* = 1.359$ and estimated the entire sequence of money multipliers (m_t^*) and the corresponding predictions of the money supply $(\hat{M}_t)$ based on the TVP model set out in eq. (3.1).

In essence, therefore, the algorithm attempts to trade-off reduced (increased) prediction errors with increased (reduced) parameter instability. As it is the weighting pattern which determines this trade-off, we can use it to assess the overall performance of the TVP model in terms of a cost-benefit analysis.

STEP 8: To do so, we initially determine the extent of parameter flexibility and tracking capability inherent in the estimated control solution, in terms of RMS (root-mean-square) errors.

The RMS percent error for parameter flexibility is given by:

$$RMS\%(m_t^*) = \left[\frac{1}{30} \sum_{t=1}^{30} [100(m_t^* - m_{t-1}^*)/m_{t-1}^*]^2 \right]^{1/2} \tag{3.14}$$

The RMS percent error for tracking capability is given by:

$$RMS\%(\hat{M}_t) = \left[\frac{1}{30} \sum_{t=1}^{30} [100(\hat{M}_t - M_t)/M_t]^2 \right]^{1/2} \tag{3.15}$$

For this control run, the following RMS percent errors were obtained:
$RMS\%(m_t^*) = 3.2059$ and $RMS\%(\hat{M}_t) = 2.8415$.

STEP 9: We are now interested in the problem of minimizing a given loss functional defined on a specified subset of a suitable vector space. To do so, we define a 2-dimensional vector L whose components are the two RMS percent errors, i.e.,

$$L = (3.2059, 2.8415) \tag{3.16}$$

We can define the *distance between vectors* in this vector space in terms of a *norm*. We choose the Euclidean norm, denoted by $||\cdot||$, so that

$$||L|| = [(3.2059)^2 + (2.8415)^2]^{1/2} = 4.283941 \tag{3.17}$$

Such a norm of L represents the *combined* error due to excessive parameter flexibility and insufficient tracking capability. In any optimization set-up, it becomes essential to minimize $||L||$.

Step 10: Such a minimization can be carried out by varying the weighting coefficients Q and R appropriately. However, as Q has been normalized at unity, the relative emphasis between parameter flexibility and tracking capability can be specified merely by an adjustment in the choice of R. Table 3 provides the results of the grid-search for the optimal value of R which minimizes $||L||$.

Table 3. Grid search for the optimal relative emphasis ($Q = 1.000$)

R	$\|\|L\|\|$	R	$\|\|L\|\|$	R	$\|\|L\|\|$
1.000	4.283941	0.594	4.125381	0.550	4.127545
0.800	4.172502	0.593	4.125375	0.500	4.136440
0.650	4.129520	0.592	4.125372	0.400	4.177830
0.600	4.125468	0.591*	4.125371	0.250	4.313265
0.595	4.125390	0.590	4.125372	0.000	4.827172

This brought us back to step 1 of the *SECOND* iteration where, instead of ensuring equal emphasis on parameter flexibility and tracking capability, we assigned, on the basis of the results spelt out above, $Q = 1.000$ and $R = 0.591$. We then worked through steps 1-5 as before, yielding a revised set of m_t^*.

We then executed Step 6 and carried out a grid-search for the optimal initial condition, m_0^*, which minimized the associated combined welfare loss, $W(m_0^*)$. Using this 'second-iteration' estimate, we re-worked Steps 7-10, this time carrying out a grid-search for the optimal relative emphasis to be placed on parameter flexibility, R^*. Successive iterations in this manner converged to yield the following set of estimates:

$$(m_0^*, R^*) = (1.342, 0.587) \tag{3.18}$$

after which there were no further changes in either m_0^* or R^*, indicating that the sequence of time-varying parameters had latched on to their optimal trajectory.

The results of such an iterative exercise involving a search around the 'near-neighbourhood' of the desired values of the parameters and targets yielded the following set of time-varying money multipliers of our model given by eq. (3.1). The results are displayed in Table 4 as well as in Figure 1.

Table 4. Optimal sequence of time-varying parameters

Year	m_t^*	Year	m_t^*
1960-61	1.332870	1975-76	1.649353
1961-62	1.339676	1976-77	1.640528
1962-63	1.344192	1977-78	1.404195
1963-64	1.377851	1978-79	1.282489
1964-65	1.410984	1979-80	1.232673
1965-66	1.436662	1980-81	1.214420
1966-67	1.467793	1981-82	1.197235
1967-68	1.498883	1982-83	1.220328
1968-69	1.497404	1983-84	1.175421
1969-70	1.512711	1984-85	1.148259
1970-71	1.522718	1985-86	1.152724
1971-72	1.537386	1986-87	1.150818
1972-73	1.580283	1987-88	1.114174
1973-74	1.554380	1988-89	1.123693
1974-75	1.567140	1989-90	1.148680
		1990-91	1.177902

In Table 5, we have provided the information necessary for the computation of the time-varying standard errors, $\sigma_t = (\Gamma_t)^{1/2}$, of money multipliers, m_t^*, provided in Table 4.

Table 5. Summary statistics for computing time-varying standard errors

Year	K_t	Σ_t	ϕ_t	V_t	Γ_t	σ_t
60-61	.000293	.000805	788.5	.0	.0	.0
61-62	.000275	.000736	850.1	105.8	.000057	.007577
62-63	.000254	.000684	916.8	83.5	.000024	.004947
63-64	.000225	.000580	1053.6	661.6	.000217	.014738
64-65	.000212	.000547	1118.3	852.0	.000206	.014371
65-66	.000195	.000510	1206.6	801.4	.000161	.012691
66-67	.000184	.000477	1285.2	848.2	.000156	.012522
67-68	.000172	.000447	1372.4	879.3	.000140	.011844
68-69	.000161	.000438	1440.1	954.7	.000146	.012103
69-70	.000146	.000387	1604.8	862.1	.000093	.009686
70-71	.000129	.000345	1807.7	786.2	.000070	.008399
71-72	.000115	.000306	2029.3	735.4	.000051	.007207
72-73	.000100	.000258	2358.9	1899.0	.000115	.010735
73-74	.000088	.000246	2603.1	3258.8	.000165	.012865
74-75	.000082	.000218	2859.0	3077.6	.000105	.010273
75-76	.000075	.000183	3204.5	12774.3	.000413	.020334
76-77	.000064	.000176	3585.5	12712.4	.000286	.016931
77-78	.000066	.000242	2996.5	67718.2	.003846	.062021
78-79	.000049	.000158	4362.8	93501.1	.000876	.029612
79-80	.000039	.000116	5667.9	95925.7	.001015	.031871
80-81	.000032	.000091	6984.3	92210.0	.000490	.022141
81-82	.000030	.000084	7549.4	88800.2	.000501	.022392
82-83	.000026	.000069	8882.9	91850.4	.000325	.018033
83-84	.000022	.000065	9956.8	103585.1	.000348	.018670
84-85	.000018	.000052	12434.4	106854.4	.000186	.013667
85-86	.000016	.000044	14226.6	104247.4	.000155	.012454
86-87	.000014	.000038	16518.8	100451.4	.000105	.010295
87-88	.000012	.000035	18555.3	127952.3	.000126	.011227
88-89	.000009	.000026	23687.6	131416.9	.000059	.007741
89-90	.000008	.000021	28528.5	179031.7	.000069	.008317
90-91	.000007	.000018	32219.9	253874.8	.000075	.008713

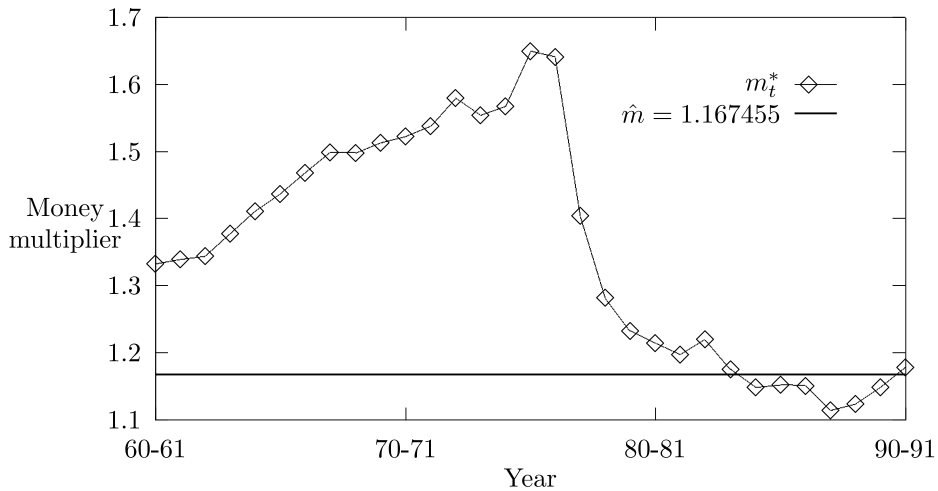

Figure 1. Time-varying money multipliers

The updating nature of the algorithm preempts any divergence in the estimated standard errors because: (i) the coefficient of Γ_{t-1} is invariably negative and (ii) increasing values of V_t are compensated by decreasing values of Σ_t. Based upon these standard errors, we have provided, in Figure 2, a confidence interval for the money multipliers ($m_t^* \pm 3\sigma_t$) over the sample period.

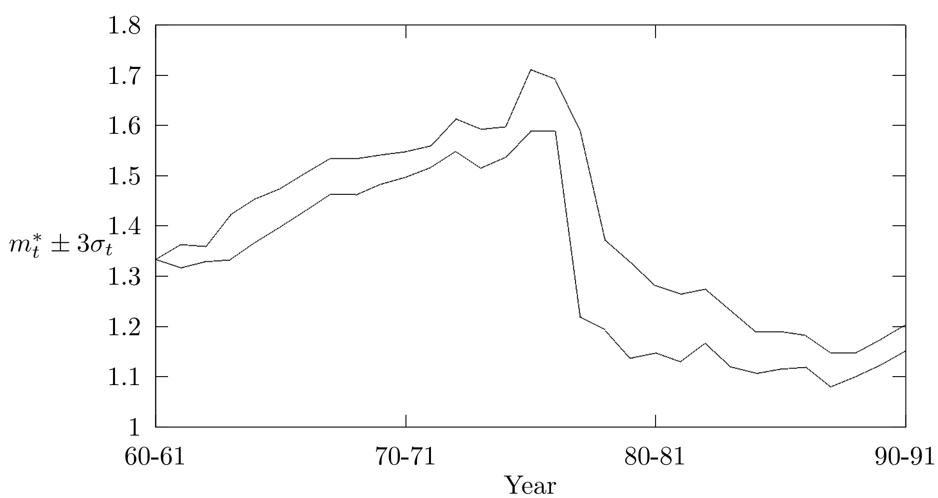

Figure 2. Confidence interval (money multipliers)

4. **Conclusions**

This paper discusses an econometric technique based on a two-part decomposition of the optimal control formulation which is seen to be suitable for the type of research that requires estimating time-varying parameters for linear regression models. The methodology is based on the characterization of the TVP problem as a near-neighbourhood search problem, with an explicit allowance for welfare loss considerations which leads to an updating algorithm capable of predicting the optimal values of the time-varying parameters as well as their covariance matrices.

The algorithm adopts an instruments-targets approach with m_0^*, the initial condition, and R^*, the emphasis on parameter flexibility, being the instruments; and $W(m_0^*)$, the total welfare loss, and $||L||$, the norm of the vector defined by the RMS percent errors, being the targets. The ensuing instrument-assignment problem becomes one of m_0^* tracking $W(m_0^*)$ and R^* tracking $||L||$.

The methodology overcomes the drawbacks associated with the FLS method as it endogenously: (i) derives the weighting pattern in the loss function, i.e., the relative costs of increasing parameter flexibility (R) *vis-a-vis* the benefits of a high level of tracking capability (Q), from the dynamic responses of the TVP system which take into consideration the comparative magnitudes of the coefficient values as well as the dependent variable; (ii) identifies the optimal initial conditions which minimize the welfare loss from amongst the set of 'near-neighbourhood' nominal trajectories; and (iii) determines the parameters of the updating equation for computation of the time-varying covariance matrices of the coefficients.

As our method updates the regression coefficients and their covariance matrices in each time period, it is a Bayesian method. As such, it would be natural to compare it with the Kalman filter estimation method which is also an updating method involving two parts: the state transition equation describing the evolution of the state; and the measurement equation describing how the data actually observed is generated by the state.

However, the main problem with any application of the Kalman filter method is to develop appropriate *á priori* estimators for the five unknowns of the system that are required to generate the recursions, i.e., the initial coefficient estimates and their covariance matrix, the covariance matrices of the transition and measurement equations, and the process by means of which the coefficients evolve over time. Because of the difficulty in doing so, most computer programs which include the Kalman filter option (e.g. TSP) set these estimatiors at pre-specified default options which allow the time varying coefficients to evolve as random walks (with a signal-to-noise ratio of 1).

Because the correction factor (K) used in our scheme plays the same role as the Kalman gain, our method can be viewed as a blend of both the FLS and Kalman filter methods, In such a context, its most important characteristic is that it is able to determine all the required priors endogenously, implying that the dynamics of the system are determined by the system itself without being specified by the user in an arbitrary fashion.

In conclusion, we hope that the study, by providing certain hitherto unavailable answers, helps in obtaining even better solutions to the problem of estimating TVP models in the future.

References

Athans, M. (1974). The importance of Kalman filtering methods for economic systems, *Annals of Economic and Social Measurement* **3**, 49–64.

Byrne, P.C. and Burke, M. (1976). Optimization with trajectory sensitivity considerations, *IEEE Transactions on Automatic Control* **AC-21**, 282–283.

Belsley, D. (1973). A test for systematic variation in regression coefficients, *Annals of Economic and Social Measurement* **2**, 487–494.

Chow, G.C. (1984). Random and changing coefficient models, in *Handbook of Econometrics* **2**, Z. Griliches and M. Intriligator (Eds.), Elsevier Science Publishers, Amsterdam, 1213–1245.

Cooley, T.F. and Prescott, E.C. (1973a). An adaptive regression model, *International Economic Review* **14**, 364–371.

− − −− (1973b). Varying parameter regression: A theory and some applications, *Annals of Economic and Social Measurement* **2**, 463–473.

− − −− (1976). Estimation in the presence of stochastic parameter variation, *Econometrica* **44**, 167–184.

Cooper, J.P. (1973). Time varying regression coefficients: A mixed estimation approach and operational limitations of the general Markov structure, *Annals of Economic and Social Measurement* **2**, 525–530.

Gallant, A.R. and Fuller, W.A. (1973). Fitting segmented polynomials whose joint points have to be estimated, *Journal of the American Statistical Association* **68**, 144–147.

Goldfeld, S.M. and Quandt, R.E. (1973). The estimation of structural shifts by switching regressions, *Annals of Economic and Social Measurement* **2**, 475–485.

Harvey, A and Phillips, G. (1982). Estimation of regression models with time varying parameters, in *Games, Economic Dynamics and Time Series Analysis*, M. Deistler, E. Furst and G. Schwodiauer (Eds.), Physica-Verlag, Wien-Wurzburg, 306–321.

Hildreth, C. and Houck, J. (1968). Some estimators for a linear model with random coefficients, *Journal of the American Statistical Association* **63**, 584–595.

Hinkley, D.V. (1971). Inference in two-phase regression, *Journal of the American Statistical Association* **66**, 736–743.

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

Kalaba, R. and Tesfatsion, L. (1980). A least-squares model specification test for a class of dynamic nonlinear economic models with systematically varying parameters, *Journal of Optimization Theory and Applications* **32**, 538–567.

− − −− (1988). The flexible least squares approach to time-varying linear regression, *Journal of Economic Dynamics and Control* **12**, 43–48.

Lee, L.F. and Griffiths, W.E. (1979). *The Prior Likelihood and Best Linear Unbiased Prediction in Stochastic Coefficient Linear Models*, Working Paper **1**, University of New England.

McGee, V.E. and Carlton, W.T. (1970). Piecewise regression, *Journal of the American Statistical Association* **65**, 1109–124.

Nachane, D.M. and Rao, M.J.M. (1994). The time-varying parameter model: An optimal control approach using trajectory sensitivity methods, *Sankhyā Series B* **56**, 286-304.

Poirier, D. (1976). *The Economics of Structural Change*, North-Holland, Amsterdam.

Quandt, R.E. (1958). The estimation of the parameters of a linear regression system obeying two separate regimes, *Journal of the American Statistical Association* **53**, 873–880.

Rao, C.R. (1965). The theory of least squares when the parameters are stochastic, *Biometrika* **52**, 447–458.

Rao, M.J.M. and D.M. Nachane (1988). Varying-parameter models: An optimal control formulation, *Journal of Quantitative Economics* **4**, 59–79.

Rao, M.J.M. (1991). The recursive estimation of time-varying parameters: A variation on the near-neighbourhood search problem, *Journal of Quantitative Economics* **7**, 303–311.

Rosenberg, B. (1973). The Analysis of a cross-section of time series by stochastically convergent parameter regression, *Annals of Economic and Social Measurement* **2**, 399–450.

Swamy, P.A.V.B. (1970), Efficient inference in a random coefficient regression model, *Econometrica* **38**, 311–323.

– – –– (1971). *Statistical Inference in Random Coefficient Regression Models*, Springer-Verlag, New York.

Tucci, M. (1990), A note on flexible least squares, *Journal of Economic Dynamics and Control* **14**, 175–182.

M.J. Manohar Rao
Professor and Director
Department of Economics
University of Bombay
Vidyanagari Campus, C.S.T. Road
Mumbai 400 098, INDIA
E-mail: mjmrao@hotmail.com