Computers Math. Applic. Vol. 17, No. 8/9, pp. 1215–1245, 1989
Printed in Great Britain.

0097-4943/89 \$3.00 + 0.00

TIME-VARYING LINEAR REGRESSION VIA FLEXIBLE LEAST SQUARES†

R. KALABA[1] and L. TESFATSION[2]‡

[1]Department of Electrical Engineering and Department of Biomedical Engineering and [2]Department of Economics, Modelling Research Group MC-0152, University of Southern California, Los Angeles, CA 90089, U.S.A.

Abstract—Suppose noisy observations obtained on a process are assumed to have been generated by a linear regression model with coefficients which evolve only slowly over time, if at all. Do the estimated time-paths for the coefficients display any systematic time-variation, or is time-constancy a reasonably satisfactory approximation? A "flexible least squares" (FLS) solution is proposed for this problem, consisting of all coefficient sequence estimates which yield vector-minimal sums of squared residual measurement and dynamic errors conditional on the given observations. A procedure with FORTRAN implementation is developed for the exact sequential updating of the FLS estimates as the process length increases and new observations are obtained. Simulation experiments demonstrating the ability of FLS to track linear, quadratic, sinusoidal, and regime shift motions in the true coefficients, despite noisy observations, are reported. An empirical money demand application is also summarized.

1. INTRODUCTION

1.1. *Overview*

Suppose an investigator undertaking a time-series linear regression study suspects that the regression coefficients might have changed over the period of time during which observations were obtained. The present paper proposes a conceptually and computationally straightforward way to guard against such a possibility.

The dynamic equations governing the motion of the coefficients will often not be known. Nevertheless, for many linear regression applications in the natural and social sciences, an assumption that the coefficients evolve only *slowly* over time seems reasonable. In this case two kinds of model specification error can be associated with each choice of an estimate $b = (b_1, \ldots, b_N)$ for the sequence of coefficient vectors b_n: residual *measurement* error given by the discrepancy between the observed dependent variable y_n and the estimated linear regression model $x_n^T b_n$ at each time n; and residual *dynamic* error given by the discrepancy $[b_{n+1} - b_n]$ between coefficient vector estimates for each successive pair of times n and $n + 1$.

Suppose a vector of "incompatibility costs" is assigned to each possible coefficient sequence estimate b based on the specification errors which b would entail. For example, suppose the cost assigned to b for measurement error is given by the sum of squared residual measurement errors, and the cost assigned to b for dynamic error is given by the sum of squared residual dynamic errors.

The "flexible least squares" (FLS) solution is defined to be the collection of all coefficient sequence estimates b which yield vector-minimal sums of squared residual measurement and dynamic errors for the given observations—i.e. which attain the "residual efficiency frontier". The frontier characterizes the efficient attainable trade-offs between residual measurement error and residual dynamic error. In particular, the frontier reveals the cost in terms of residual measurement error that must be paid in order to attain the *zero* residual dynamic error (time-constant coefficients) required by ordinary least squares estimation.

Coefficient sequence estimates b which attain the residual efficiency frontier are referred to as "FLS estimates". Each FLS estimate has a basic efficiency property: no other coefficient sequence

†The present paper is a revised version of Ref. [1], presented at the *1987 Ninth Annual Conference of the Society for Economic Dynamics and Control*, in an April 1987 seminar at UC Berkeley, and in a January 1988 seminar at the University of Arizona. The authors are grateful to conference and seminar participants for numerous helpful suggestions.
‡Author for correspondence.

estimate yields both lower measurement error and lower dynamic error for the given observations. The time-paths traced out by the FLS estimates thus indicate how the regression coefficients could have evolved over time in a manner *minimally incompatible* with the prior measurement and dynamic specifications.

The time-varying linear regression problem treated in the present paper is formally set out in Section 2. The FLS approach to this problem, briefly outlined above, is more carefully developed in Section 3. A matrix representation for the FLS estimates is derived in Section 4. In Section 5 a procedure is developed for the exact sequential updating of the FLS estimates as the process length increases and additional observations are obtained. Section 6 develops various intrinsic geometric relationships between the FLS estimates and the ordinary least squares solution obtained by imposing constancy on the coefficient vectors prior to estimation.

In Section 7 it is established, analytically, that any unanticipated shift in the true coefficient vector will be reflected in the time-paths traced out by the FLS estimates. Section 8 describes some of the simulation studies undertaken with a FORTRAN program "FLS" which demonstrate the ability of the FLS estimates to track linear, quadratic, sinusoidal, and regime shift time-variations in the true coefficients, despite noisy observations. Section 8 also briefly summarizes the findings of an empirical money demand study [2] in which FLS is used to investigate coefficient stability for the well-known Goldfeld U.S. money demand model [3] over 1959:Q2–1985:Q3.

The final Section 9 discusses topics for future research. Proofs of theorems are provided in Appendix A. A list of FORTRAN statements for the computer program FLS is provided in Appendix B, together with a brief discussion of the program logic.

1.2. *Relationship to previous time-varying linear regression studies*

The time-varying linear regression problem has attracted considerable attention from econometricians and statisticians over the past several decades. Early studies of this problem include Quandt [4] on estimating the location of a shift from one regression scheme to another, and Quandt [5] and Chow [6] on testing the null hypothesis of a shift at a particular point in time. A synthesis of this work can be found in Fisher [7]. See also the later work of Guthery [8], Brown *et al.* [9], Ertel and Fowlkes [10], and Cooley and Prescott [11] on linear regression models with stochastically varying coefficients. Rosenberg [12] provides a general survey of this literature.

Other studies (e.g. [13–18]) have investigated the application of Kalman–Bucy filtering [19, 20] to linear regression models with various types of non-constancy assumed for the coefficients. Finally, the relationship between statistical smoothing spline models (e.g. Craven and Wahba [21]) and time-varying linear regression models is clarified in [18, pp. 12–14].

All of these statistical time-varying linear regression studies require the specification of probabilistic properties for residual error terms and, ultimately, for test statistics. These requirements pose three potential difficulties.

First, most time-series data used in empirical economics is not generated within the framework of controlled experiments. The inability to replicate the same experiment a large number of times means that objective information concerning probabilistic properties for residual error terms may be difficult to obtain.† In addition, the complexity of many economic processes suggests that model specification errors are inevitable. However, the specification of probabilistic properties for residual error terms implies that these terms are to be interpreted as random shocks disturbing an otherwise correctly specified model rather than as potential discrepancies resulting from model misspecification. Finally, obtaining distributional properties for the test statistics relied on by conventional methods can require theoretically significant simplifications (e.g. linearizations) for computational reasons. If the test statistics then result in a rejection of the model, it may be difficult to pinpoint which maintained restrictions—theoretical, probabilistic, or computational—are responsible for the rejection.

---

†See, for example, the complex approximations undertaken by Doan *et al.* [22, pp. 6–26] and Miller and Roberds [23, pp. 5–10] in order to specify the initial mean values and second moment matrices required by the Kalman–Bucy filter.

In contrast to conventional statistical techniques, FLS is an exploratory data analysis tool for testing the basic compatibility of theory and observations. As clarified in previous studies [24–27], the theory may consist of nonlinear measurement, dynamic, and stochastic specifications. The form these specifications take is not restricted. In particular, investigators are not required to use an *ad hoc* stochastic framework when they have little knowledge of, or belief in, probabilistic properties for residual error terms. FLS determines the degree to which the theoretical specifications can be simultaneously satisfied, conditional on the given observations. Once a theoretical model is found which is basically compatible with the data, a more structured statistical approach can be used to refine the estimates.

Time-varying linear regression techniques are commonly applied when a process is undergoing some type of structural variation which is not yet well understood. The theoretical, simulation, and empirical results reported in the present study suggest that FLS provides a useful complement to existing statistical techniques for this class of problems.

1.3. Relationship to previous work in systems science and engineering

The idea of forming an incompatibility cost-of-estimation function as a suitably weighted sum of squared residual dynamic and measurement modelling errors was stressed by R. Sridhar, R. Bellman, and other associates in a series of studies [28–30] focusing on a class of continuous-time nonlinear filtering problems arising in rigid body dynamics. Invariant imbedding techniques [31, 32] were used to convert the first-order necessary conditions for minimization of the cost-of-estimation function (a two-point boundary value problem) into an initial value problem amenable to sequential solution techniques.

Building on this work, exact sequential filtering and smoothing equations were developed in [24, 25] for a discrete-time analog of the continuous-time Sridhar nonlinear filtering problem. As in previous studies, the exact sequential equations were obtained by converting the first-order necessary conditions for cost minimization into an initial value problem.

In [26] it is shown that sequential solution techniques can be devised for discrete-time processes modelled in terms of general nonlinear dynamic and measurement specifications *without* making direct use of the first-order necessary conditions for cost minimization. Specifically, two exact procedures are developed for the direct sequential minimization of the cost-of-estimation function as the duration of the process increases and new observation vectors are obtained. The first algorithm proceeds by an imbedding on the process length and the final *state* vector. The second algorithm proceeds by an imbedding on the process length and the final *observation* vector. Each algorithm generates optimal (least cost) filtered and smoothed state estimates, together with optimal one-step-ahead state predictions.

The basic conceptual idea of minimizing a weighted sum of squared residual dynamic and measurement modelling errors to obtain state estimates for nonlinear processes is extended in three directions in [27] to obtain a "flexible least cost" state estimation technique for a broader range of problems.

First, instead of focusing on the state estimates which minimize a cost-of-estimation function specified for one given set of weights, the solution to the state estimation problem is instead taken to be the collection of all state estimates which attain the "cost-efficiency frontier"—i.e. which yield vector-minimal sums of squared residual dynamic and measurement errors, conditional on the given observations. A cost-of-estimation function with *varying* weights is used to generate the cost-efficiency frontier. Second, it is shown that exact sequential updating equations can be obtained for more generally specified cost-of-estimation functions; e.g. cost-of-estimation functions for which the dynamic and measurement costs are specified to be arbitrary increasing functions of the absolute residual dynamic and measurement modelling errors. Third, it is shown that prior stochastic specifications can be incorporated into the cost-of-estimation function in addition to prior dynamic and measurement specificatons. The basic cost-efficiency frontier is then a surface in E^3 giving the locus of minimal attainable dynamic, measurement, and stochastic costs-of-estimation for a given set of observations.

The present paper undertakes a detailed theoretical and experimental study of the flexible least cost approach for processes characterized by linear state (coefficient) measurements, unknown state dynamics proxied by a smoothness prior, and squared residual error cost specifications.

2. TIME-VARYING LINEAR REGRESSION PROBLEM

Suppose noisy scalar observations $y_1, \ldots, y_N$ obtained on a process over a time-span $1, \ldots, N$ are assumed to have been generated by a linear regression model with coefficients which evolve only slowly over time, if at all. More precisely, suppose these prior theoretical beliefs take the following form:

Prior measurement specification [linear measurement]:

$$y_n - x_n^{\mathrm{T}} b_n \approx 0, \quad n = 1, \ldots, N. \tag{2.1a}$$

Prior dynamic specification [coefficient stability]:

$$b_{n+1} - b_n \approx \mathbf{0}, \quad n = 1, \ldots, N-1, \tag{2.1b}$$

where

$x_n^{\mathrm{T}} = (x_{n1}, \ldots, x_{nK}) = 1 \times K$ row vector of known exogenous regressors;
$b_n = (b_{n1}, \ldots, b_{nK})^{\mathrm{T}} = K \times 1$ column vector of unknown coefficients.

The measurement and dynamic specificatons (2.1) reflect the prior beliefs of linear measurement and coefficient stability in a simple direct way, without augmentation by any stochastic restrictions. These prior beliefs seem relevant for a wide variety of processes in both the natural and the social sciences.

A basic problem is then to determine whether the theory is compatible with the observations. That is, does there exist *any* coefficient sequence estimate $(b_1, \ldots, b_N)$ which satisfies the prior theoretical specifications (2.1) in an acceptable approximate sense for the realized sequence of observations $(y_1, \ldots, y_N)$? How might such a coefficient sequence estimate be found?

3. FLEXIBLE LEAST SQUARES (FLS)

3.1. The basic FLS approach

Two kinds of model specification error can be associated with each possible coefficient sequence estimate $b = (b_1, \ldots, b_N)$ for model (2.1). First, b could fail to satisfy the prior measurement specification (2.1a). Second, b could fail to satisfy the prior dynamic specification (2.1b).†

Suppose the cost assigned to b for the first type of error is measured by the sum‡ of squared residual measurement errors

$$r_{\mathrm{M}}^2(b; N) = \sum_{n=1}^{N} [y_n - x_n^{\mathrm{T}} b_n]^2, \tag{3.1}$$

and the cost assigned to b for the second type of error is measured by the sum of squared residual dynamic errors

$$r_{\mathrm{D}}^2(b; N) = \sum_{n=1}^{N-1} [b_{n+1} - b_n]^{\mathrm{T}} [b_{n+1} - b_n]. \tag{3.2}$$

Define the (time N) *residual possibility set* to be the collection

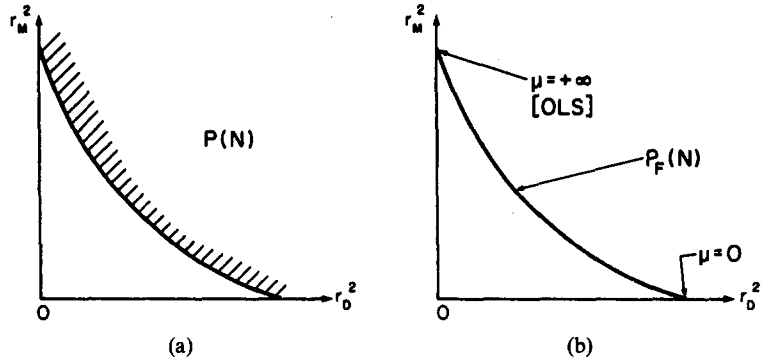
$$P(N) = \{r_{\mathrm{D}}^2(b; N), r_{\mathrm{M}}^2(b; N) \,|\, b \in E^{NK}\} \tag{3.3}$$

of all possible configurations of squared residual dynamic error and measurement error sums attainable at time N, conditional on the given observations $y_1, \ldots, y_N$. The residual possibility set is depicted in Fig. 1a.

If the prior theoretical specifications (2.1) are correct, the squared residual errors associated with the actual coefficient sequence will be approximately zero. In general, however, the lower

†This simple breakdown of costs into two categories, measurement and dynamic, can of course be generalized (see [24, 27, Section 4]).

‡It is assumed that preliminary scaling and transformations have been carried out as appropriate prior to forming the sums (3.1) and (3.2). In particular, the units in which the regressor variables are measured should be chosen so that the regressors are approximately of the same order of magnitude.



(a) (b)

Fig. 1. (a) Residual possibility set $P(N)$, (b) residual efficiency frontier $P_F(N)$.

envelope for the residual possibility set $P(N)$ will be bounded away from the origin in E^2. This lower envelope gives the locus of vector-minimal sums of squared residual dynamic and measurement errors attainable at time N, conditional on the given observations. In particular, the lower envelope reveals the cost in terms of residual measurement error that must be paid in order to achieve the *zero* residual dynamic error (time-constant coefficients) required by OLS estimation. Hereafter this lower envelope, denoted by $P_F(N)$, will be referred to as the (time N) *residual efficiency frontier*; and coefficient sequence estimates b which attain this frontier will be referred to as *FLS estimates.*

The FLS estimates along the residual efficiency frontier constitute a "population" of estimates characterized by a basic efficiency property: for the given observations, these are the coefficient sequence estimates which are *minimally incompatible* with the linear measurement and coefficient stability specifications (2.1). Three different levels of analysis can be used to compare the FLS estimates along the frontier with the time-constant OLS solution obtained at the frontier extreme point characterized by zero residual dynamic error.

At the most general level, the qualitative shape of the frontier indicates whether or not the OLS solution provides a good description of the observations. If the true model generating the observations has time-constant coefficients, then, starting from the OLS extreme point, the frontier should indicate that only small decreases in measurement error are possible even for large increases in dynamic error. The frontier should thus be rather flat (moderately sloped) in a neighborhood of the OLS extreme point in the r_D^2–r_M^2 plane. If the true model generating the observations has time-varying coefficients, then large decreases in measurement error should be attainable with only small increases in dynamic error. The frontier should thus be fairly steeply sloped in a neighborhood of the OLS extreme point. In this case the OLS solution is unlikely to reflect the properties exhibited by the typical FLS estimates along the frontier.

The next logical step is to construct summary statistics for the time-paths traced out by the FLS estimates along the frontier. For example, at any point along the frontier the average value attained by the FLS estimates for the kth coefficient can be compared with the OLS estimate for the kth coefficient, $k = 1, \ldots, K$. The standard deviation of the FLS kth coefficient estimates about their average value provides a summary measure of the extent to which these estimates deviate from constancy. These average value and standard deviation statistics can be used to assess the extent to which the OLS solution is representative of the typical FLS estimates along the frontier.

Finally, the time-paths traced out by the FLS estimates along the frontier can be directly examined for evidence of systematic movements in individual coefficients—e.g. unanticipated shifts at dispersed points in time. Such movements might be difficult to discern from the summary average value and standard deviation characterizations for the estimated time-paths.

This three-level analysis proved to be useful for interpreting and reporting the findings of the empirical money demand study [2].

3.2. Parametric representation for the residual efficiency frontier

How might the residual efficiency frontier be found? In analogy to the usual procedure for tracing out Pareto-efficiency frontiers, a parameterized family of minimization problems is considered.

Thus, let $\mu \geqslant 0$ be given, and suppose the $K \times N$ matrix of regressor vectors $[x_1, \ldots, x_N]$ has full rank K. Let each possible coefficient sequence estimate $b = (b_1, \ldots, b_N)$ be assigned an *incompatibility cost*

$$C(b; \mu, N) = \mu r_D^2(b; N) + r_M^2(b; N), \tag{3.4}$$

consisting of the μ-weighted average of the associated dynamic error and measurement error sums (3.1) and (3.2).† Expressing these sums in terms of their components, the incompatibility cost $C(b; \mu, N)$ takes the form

$$C(b; \mu, N) = \mu \sum_{n=1}^{N-1} [b_{n+1} - b_n]^T[b_{n+1} - b_n] + \sum_{n=1}^{N} [y_n - x_n^T b_n]^2. \tag{3.5}$$

As (3.5) indicates, the incompatibility cost function $C(b; \mu, N)$ generalizes the goodness-of-fit criterion function for ordinary least squares estimation by permitting the coefficient vectors b_n to vary over time.

If $\mu > 0$, let the coefficient sequence estimate which uniquely minimizes the incompatibility cost (3.4) be denoted by

$$b^{FLS}(\mu, N) = (b_1^{FLS}(\mu, N), \ldots, b_N^{FLS}(\mu, N)) \tag{3.6}$$

(uniqueness of the minimizing sequence for $\mu > 0$ is established below in Section 4). If $\mu = 0$, let (3.6) denote any coefficient sequence estimate b which minimizes the sum of squared residual dynamic errors $r_D^2(b; N)$ subject to $r_M^2(b; N) = 0$. Hereafter, (3.6) will be referred to as the *flexible least squares (FLS) solution at time N, conditional on μ.*

Finally, let the sums of squared residual measurement errors and dynamic errors corresponding to the FLS solution (3.6) be denoted by

$$r_M^2(\mu, N) = r_M^2(b^{FLS}(\mu, N); N); \tag{3.7a}$$

$$r_D^2(\mu, N) = r_D^2(b^{FLS}(\mu, N); N). \tag{3.7b}$$

By construction, a point (r_D^2, r_M^2) in E^2 lies on the residual efficiency frontier $P_F(N)$ if and only if there exists some $\mu \geqslant 0$ such that $(r_D^2, r_M^2) = (r_D^2(\mu, N), r_M^2(\mu, N))$. The residual efficiency frontier $P_F(N)$ thus takes the parameterized form‡

$$P_F(N) = \{r_D^2(\mu, N), r_M^2(\mu, N) \mid 0 \leqslant \mu < \infty\}. \tag{3.8}$$

The parameterized residual efficiency frontier (3.8) is qualitatively depicted in Fig. 1b. As μ approaches zero, the incompatibility cost function (3.4) ultimately places no weight on the prior dynamic specifications (2.1b); i.e. r_M^2 is minimized with no regard for r_D^2. Thus r_M^2 can generally be brought down close to zero and the corresponding value for r_D^2 will be relatively large. As μ becomes arbitrarily large, the incompatibility cost function (3.4) places absolute priority on the prior dynamic specifications (2.1b); i.e. r_M^2 is minimized subject to $r_D^2 = 0$. The latter case coincides with OLS estimation in which a single $K \times 1$ coefficient vector is used to minimize the sum of squared residual measurement errors r_M^2 (see Section 6, below).

The next two sections of the paper develop explicit procedures for generating the FLS solution (3.6).

†When a least-squares formulation such as (3.4) is used as the incompatibility cost function, a common reaction is that the analysis is implicitly relying on normality assumptions for residual error terms. To the contrary, (3.4) assesses the costs associated with various possible deviations between theory and observations; it bears no necessary relation to any intrinsic stochastic properties of the residual error terms. Specifically, (3.4) indicates that residual measurement errors of equal magnitude are specified to be equally costly, not that these errors are anticipated to be symmetrically distributed around zero; and similarly for residual dynamic errors. More general specifications for the incompatibility cost function can certainly be considered. See, for example, [27, Section 4].

‡In numerous simulation experiments the residual efficiency frontier (3.8) has been adequately traced out by evaluating the residual error sums (3.7) over a rough grid of μ-points increasing by powers of ten. In other words, the generation of the residual efficiency frontier is not a difficult matter. All of the numerically generated frontiers have displayed the convex shape qualitatively depicted in Fig. 1b (see Section 8, below, for a brief summary of these simulation experiments).

4. THE FLS SOLUTION: MATRIX REPRESENTATION

Matrix representations for the incompatibility cost function (3.4) and the FLS solution (3.6) will now be derived.

Let I denote the $K \times K$ identity matrix. Also, define

$$X(N)^{\mathrm{T}} = (x_1, \ldots, x_N) = K \times N \text{ matrix of regressors;} \tag{4.1a}$$

$$b(N) = (b_1^{\mathrm{T}}, \ldots, b_N^{\mathrm{T}})^{\mathrm{T}} = NK \times 1 \text{ column vector of coefficients;} \tag{4.1b}$$

$$y(N) = (y_1, \ldots, y_N)^{\mathrm{T}} = N \times 1 \text{ column vector of observations;} \tag{4.1c}$$

$$G(N) = \begin{bmatrix} x_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & x_N \end{bmatrix} = NK \times N \text{ matrix formed from the regressors;} \tag{4.1d}$$

$$A_n(\mu) = \begin{cases} x_1 x_1^{\mathrm{T}} + \mu I & \text{if } n = 1; \\ x_n x_n^{\mathrm{T}} + 2\mu I & \text{if } n \neq 1, N; \\ x_N x_N^{\mathrm{T}} + \mu I & \text{if } n = N; \end{cases} \tag{4.1e}$$

$$A(\mu, N) = \begin{bmatrix} A_1(\mu) & -\mu I & \mathbf{0} & \cdots & \mathbf{0} \\ -\mu I & A_2(\mu) & -\mu I & & \vdots \\ \mathbf{0} & -\mu I & \ddots & & \mathbf{0} \\ \vdots & & & \ddots & -\mu I \\ \mathbf{0} & \cdots & \mathbf{0} & -\mu I & A_N(\mu) \end{bmatrix}. \tag{4.1f}$$

The following results are established in Section A.1 of Appendix A. The incompatibility cost function (3.4) can be expressed in matrix form as

$$C(b(N); \mu, N) = b(N)^{\mathrm{T}} A(\mu, N) b(N) - 2b(N)^{\mathrm{T}} G(N) y(N) + y(N)^{\mathrm{T}} y(N). \tag{4.2}$$

The first-order necessary conditions for a vector $b(N)$ to minimize (4.2) thus take the form

$$A(\mu, N) b(N) = G(N) y(N). \tag{4.3}$$

The matrix $A(\mu, N)$ is positive semidefinite for every $\mu \geqslant 0$ and $N \geqslant 1$. Moreover, if $\mu > 0$ and the $N \times K$ regressor matrix $X(N)$ has rank K, then $A(\mu, N)$ is positive definite and the incompatibility cost function (4.2) is a strictly convex function of $b(N)$. In the latter case it follows from (4.3) that (4.2) is uniquely minimized by the $NK \times 1$ column vector

$$b^{\mathrm{FLS}}(\mu, N) = A(\mu, N)^{-1} G(N) y(N). \tag{4.4}$$

Thus, given $\mu > 0$ and rank $X(N) = K$, (4.4) yields an explicit matrix representation for the FLS solution (3.6).

To obtain the FLS solution (3.6) by means of equation (4.4), the $NK \times NK$ matrix $A(\mu, N)$ must be inverted. One could try to accomplish this inversion directly, taking advantage of the special form of the matrix $A(\mu, N)$. Alternatively, one could try to accomplish this inversion indirectly, by means of a lower-dimensional sequential procedure.

As will be clarified in the following sections, the latter approach yields a numerically stable algorithm for the exact sequential derivation of the FLS solution (3.6) which is conceptually informative in its own right. The sequential procedure gives directly the estimate $b_n^{\mathrm{FLS}}(\mu, n)$ for the time-n coefficient vector b_n as each successive observation y_n is obtained. This permits a simple direct check for coefficient constancy. Once the estimate for the time-n coefficient vector is obtained, it is a simple matter to obtain smoothed (back-updated) estimates for all intermediate coefficient vectors for times 1 through $n - 1$, as well as an explicit smoothed estimate for the actual dynamic relationship connecting each successive coefficient vector pair.

5. EXACT SEQUENTIAL DERIVATION OF THE FLS SOLUTION

In Section 5.1, below, a basic recurrence relation is derived for the exact sequential minimization of a "cost-of-estimation" function as the duration of the process increases and additional observations are obtained. In Section 5.2 it is shown how this basic recurrence relation can be more concretely represented in terms of recurrence relations for a $K \times K$ matrix, a $K \times 1$ vector, and a scalar.

In Sections 5.3 and 5.4 it is shown how the recurrence relations derived in Sections 5.1 and 5.2 can be used to develop exact sequential updating equations for the FLS solution (3.6). Specifically, these recurrence relations allow the original NK-dimensional problem of minimizing the incompatibility cost function (3.4) with respect to $b = (b_1, \ldots, b_N)$ to be decomposed into a sequence of N linear-quadratic cost-minimization problems, each of dimension K, a significant computational reduction.

5.1. The basic recurrence relation

Let $\mu > 0$ and $n \geqslant 2$ be given. Define the total cost of the estimation process at time $n-1$, conditional on the coefficient estimates $b_1, \ldots, b_n$ for times 1 through n, to be the μ-weighted sum of squared residual dynamic and measurement errors

$$W(b_1, \ldots, b_n; \mu, n-1) = \mu \sum_{s=1}^{n-1} [b_{s+1} - b_s]^{\mathrm{T}}[b_{s+1} - b_s] + \sum_{s=1}^{n-1} [y_s - x_s^{\mathrm{T}} b_s]^2. \tag{5.1}$$

Let $\phi(b_n; \mu, n-1)$ denote the smallest cost of the estimation process at time $n-1$, conditional on the coefficient estimate b_n for time n; i.e.

$$\phi(b_n; \mu, n-1) = \inf_{b_1, \ldots, b_{n-1}} W(b_1, \ldots, b_n; \mu, n-1). \tag{5.2}$$

By construction, the function $W(.; \mu, n-1)$ defined by (5.1) is bounded below over its domain E^{nK}. It follows by the principle of iterated infima that the cost-of-estimation function $\phi(.; \mu, n-1)$ defined by (5.2) satisfies the recurrence relation

$$\phi(b_{n+1}; \mu, n) = \inf_{b_n} [\mu[b_{n+1} - b_n]^{\mathrm{T}}[b_{n+1} - b_n] + [y_n - x_n^{\mathrm{T}} b_n]^2 + \phi(b_n; \mu, n-1)] \tag{5.3a}$$

for all b_{n+1} in E^K.

The recurrence relation (5.3a) is initialized by assigning a prior cost-of-estimation $\phi(b_1; \mu, 0)$ to each b_1 in E^K. Given the incompatibility cost function specification (3.4), this prior cost-of-estimation takes the form

$$\phi(b_1; \mu, 0) \equiv 0. \tag{5.3b}$$

In general, however, $\phi(b_1; \mu, 0)$ could reflect the cost of specifying an estimate b_1 for time 1 conditional on everything that is known about the process prior to obtaining an observation y_1 at time 1.

The recurrence relation (5.3) can be given a dynamic programming interpretation. At any current time n the choice of a coefficient estimate b_n incurs three types of cost conditional on an anticipated coefficient estimate b_{n+1} for time $n+1$. First, b_n could fail to satisfy the prior dynamic specification (2.1b). The cost incurred for this dynamic error is $\mu[b_{n+1} - b_n]^{\mathrm{T}}[b_{n+1} - b_n]$. Second, b_n could fail to satisfy the prior measurement specification (2.1a). The cost incurred for this measurement error is $[y_n - x_n^{\mathrm{T}} b_n]^2$. Third, a cost $\phi(b_n; \mu, n-1)$ is incurred for choosing b_n at time n based on everything that is known about the process through time $n-1$.

These three costs together comprise the total cost of choosing a coefficient estimate b_n at time n, conditional on an anticipated coefficient estimate b_{n+1} for time $n+1$. Minimization of this total cost with respect to b_n thus yields the cost $\phi(b_{n+1}; \mu, n)$ incurred for choosing the coefficient estimate b_{n+1} at time $n+1$ based on everything that is known about the process through time n.

As will be clarified in future studies, a recurrence relation such as (5.3) for the updating of incompatibility cost provides a generalization of the recurrence relation derived in Larson and Peschon [33, equation (14)] for the Bayesian updating of a probability density function.

5.2. *A more concrete representation for the basic recurrence relation*

It will now be shown how the basic recurrence relation (5.3) can be more concretely represented in terms of recurrence relations for a $K \times K$ matrix $Q_n(\mu)$, a $K \times 1$ vector $p_n(\mu)$, and a scalar $r_n(\mu)$.

The prior cost-of-estimation function (5.3b) can be expressed in the quadratic form

$$\phi(b_1; \mu, 0) = b_1^{\mathrm{T}} Q_0(\mu) b_1 - 2 b_1^{\mathrm{T}} p_0(\mu) + r_0(\mu), \tag{5.4a}$$

where

$$Q_0(\mu) = [0]_{K \times K}; \tag{5.4b}$$

$$p_0(\mu) = 0_{K \times 1}; \tag{5.4c}$$

$$r_0(\mu) = 0. \tag{5.4d}$$

Suppose it has been shown for some $n \geqslant 1$ that the cost-of-estimation function $\phi(\cdot\,; \mu, n-1)$ for time $n-1$ has the quadratic form

$$\phi(b_n; \mu, n-1) = b_n^{\mathrm{T}} Q_{n-1}(\mu) b_n - 2 b_n^{\mathrm{T}} p_{n-1}(\mu) + r_{n-1}(\mu) \tag{5.5}$$

for some $K \times K$ positive semidefinite matrix $Q_{n-1}(\mu)$, $K \times 1$ vector $p_{n-1}(\mu)$, and scalar $r_{n-1}(\mu)$.

The cost-of-estimation function $\phi(\cdot\,; \mu, n)$ for time n satisfies the recurrence relation (5.3a). Using the induction hypothesis (5.5), the first-order necessary (and sufficient) conditions for a vector b_n to minimize the bracketed term on the right-hand side of (5.3a), conditional on b_{n+1}, reduce to

$$\mathbf{0} = -2 y_n x_n^{\mathrm{T}} + 2[x_n^{\mathrm{T}} b_n] x_n^{\mathrm{T}} - 2\mu b_{n+1}^{\mathrm{T}} + 2\mu b_n^{\mathrm{T}} + 2 b_n^{\mathrm{T}} Q_{n-1}(\mu) - 2 p_{n-1}^{\mathrm{T}}(\mu). \tag{5.6}$$

The vector b_n which satisfies (5.6) is a linear function of b_{n+1} given by

$$b_n^*(\mu) = e_n(\mu) + M_n(\mu) b_{n+1}, \tag{5.7a}$$

where

$$M_n(\mu) = \mu[Q_{n-1}(\mu) + \mu I + x_n x_n^{\mathrm{T}}]^{-1}; \tag{5.7b}$$

$$e_n(\mu) = \mu^{-1} M_n(\mu)[p_{n-1}(\mu) + x_n y_n]. \tag{5.7c}$$

By the induction hypothesis (5.5), the $K \times K$ matrix $M_n(\mu)$ is positive definite.

Substituting (5.7a) into (5.3a), one obtains

$$\begin{aligned}\phi(b_{n+1}; \mu, n) &= [y_n - x_n^{\mathrm{T}} b_n^*(\mu)]^2 + \mu[b_{n+1} - b_n^*(\mu)]^{\mathrm{T}}[b_{n+1} - b_n^*(\mu)] + \phi(b_n^*(\mu); \mu, n-1) \\ &= b_{n+1}^{\mathrm{T}} Q_n(\mu) b_{n+1} - 2 b_{n+1}^{\mathrm{T}} p_n(\mu) + r_n(\mu),\end{aligned} \tag{5.8a}$$

where

$$Q_n(\mu) = [I - M_n(\mu)]; \tag{5.8b}$$

$$p_n(\mu) = \mu e_n(\mu); \tag{5.8c}$$

$$r_n(\mu) = r_{n-1}(\mu) + y_n^2 - [p_{n-1}(\mu) + x_n y_n]^{\mathrm{T}} e_n(\mu). \tag{5.8d}$$

Using (5.7b) and (5.7c) to eliminate $M_n(\mu)$ and $e_n(\mu)$ in (5.8), one obtains

$$Q_n(\mu) = \mu[Q_{n-1}(\mu) + \mu I + x_n x_n^{\mathrm{T}}]^{-1}[Q_{n-1}(\mu) + x_n x_n^{\mathrm{T}}]; \tag{5.9a}$$

$$p_n(\mu) = \mu[Q_{n-1}(\mu) + \mu I + x_n x_n^{\mathrm{T}}]^{-1}[p_{n-1}(\mu) + x_n y_n]; \tag{5.9b}$$

$$r_n(\mu) = r_{n-1}(\mu) + y_n^2 - [p_{n-1}(\mu) + x_n y_n]^{\mathrm{T}}[Q_{n-1}(\mu) + \mu I + x_n x_n^{\mathrm{T}}]^{-1}[p_{n-1}(\mu) + x_n y_n]. \tag{5.9c}$$

It is clear from (5.9a) that the $K \times K$ matrix $Q_n(\mu)$ is positive semidefinite (definite) if $Q_{n-1}(\mu)$ is positive semidefinite (definite). Equations (5.9) thus yield the sought-after recurrence relations for $Q_n(\mu)$, $p_n(\mu)$, and $r_n(\mu)$.

Note that the matrices $Q_n(\mu)$ are independent of the observations y_n. Their determination can thus be accomplished off-line, prior to the realization of any observations.

5.3. Filtered coefficient estimates

Let $\mu > 0$ be given, and suppose the $K \times n$ regressor matrix $[x_1, \ldots, x_n]$ has full rank K for each $n \geqslant K$. Using the recurrence relations derived in Sections 5.1 and 5.2, an exact sequential procedure will now be given for generating the unique FLS estimate $b_n^{\mathrm{FLS}}(\mu, n)$ for the time-n coefficient vector b_n, conditional on the observations $y_1, \ldots, y_n$, for each process length $n \geqslant K$.

At time $n = 1$, $Q_0(\mu)$, $p_0(\mu)$, and $r_0(\mu)$ are determined from (5.4b), (5.4c), and (5.4d) to be identically zero. A first observation y_1 is obtained. In preparation for time 2, the recurrence relations (5.9) are used to determine and store the matrix $Q_1(\mu)$, the vector $p_1(\mu)$, and the scalar $r_1(\mu)$. If $1 = K$, the unique FLS estimate for the time-1 coefficient vector b_1, conditional on the observation y_1, is given by

$$\begin{aligned} b_1^{\mathrm{FLS}}(\mu, 1) &= \arg\min_{b_1}([y_1 - x_1^{\mathrm{T}} b_1]^2 + \phi(b_1; \mu, 0)) \\ &= [Q_0(\mu) + x_1 x_1^{\mathrm{T}}]^{-1}[p_0(\mu) + x_1 y_1]. \end{aligned} \tag{5.10}$$

At time $n \geqslant 2$, $Q_{n-1}(\mu)$, $p_{n-1}(\mu)$, and $r_{n-1}(\mu)$ have previously been calculated and stored. An additional observation y_n is obtained. In preparation for time $n + 1$ the recurrence relations (5.9) are used to determine and store the matrix $Q_n(\mu)$, the vector $p_n(\mu)$, and the scalar $r_n(\mu)$. If $n \geqslant K$, the unique FLS estimate for the time-n coefficient vector b_n, conditional on the observations $y_1, \ldots, y_n$, is given by

$$\begin{aligned} b_n^{\mathrm{FLS}}(\mu, n) &= \arg\min_{b_n} ([y_n - x_n^{\mathrm{T}} b_n]^2 + \phi(b_n; \mu, n-1)) \\ &= [Q_{n-1}(\mu) + x_n x_n^{\mathrm{T}}]^{-1}[p_{n-1}(\mu) + x_n y_n]. \end{aligned} \tag{5.11}$$

If the $K \times K$ matrix $Q_{n-1}(\mu)$ has full rank K, the FLS estimate (5.11) satisfies the recurrence relation

$$b_n^{\mathrm{FLS}}(\mu, n) = b_{n-1}^{\mathrm{FLS}}(\mu, n-1) + F_n(\mu)[y_n - x_n^{\mathrm{T}} b_{n-1}^{\mathrm{FLS}}(\mu, n-1)], \tag{5.12a}$$

where the $K \times 1$ filter gain $F_n(\mu)$ is given by

$$F_n(\mu) = S_n(\mu) x_n / [1 + x_n^{\mathrm{T}} S_n(\mu) x_n]; \tag{5.12b}$$

$$S_n(\mu) = [Q_{n-1}(\mu)]^{-1}. \tag{5.12c}$$

It is easily established that (5.11) does yield the unique FLS estimate for the time-n coefficient vector b_n for each process length $n \geqslant K$. By assumption, the total incompatibility cost at time n, given the coefficient estimates $b_1, \ldots, b_n$, is

$$C(b_1, \ldots, b_n; \mu, n) = [y_n - x_n^{\mathrm{T}} b_n]^2 + W(b_1, \ldots, b_n; \mu, n-1), \tag{5.13}$$

where the function $W(.; \mu, n-1)$ is defined by (5.1). The simultaneous minimization of the incompatibility cost function (5.13) with respect to the coefficient vectors $b_1, \ldots, b_n$ can thus be equivalently expressed as

$$\begin{aligned} &\min_{b_1, \ldots, b_n} ([y_n - x_n^{\mathrm{T}} b_n]^2 + W(b_1, \ldots, b_n; \mu, n-1)) \\ &\quad = \min_{b_n}([y_n - x_n^{\mathrm{T}} b_n]^2 + \min_{b_1, \ldots, b_{n-1}} W(b_1, \ldots, b_n; \mu, n-1)) \\ &\quad = \min_{b_n}([y_n - x_n^{\mathrm{T}} b_n]^2 + \phi(b_n; \mu, n-1)). \end{aligned} \tag{5.14}$$

This establishes the first equality in (5.11). The second equality in (5.11) follows by direct calculation, using expression (5.5) for $\phi(b_n; \mu, n-1)$.

Relation (5.12) can be verified by tedious but straightforward calculations by use of (5.9), (5.10), and the well-known Woodbury matrix inversion lemma.

5.4. Smoothed coefficient estimates

Let $\mu > 0$ and $N \geqslant K$ be given. Suppose the procedure outlined in Section 5.3 has been used to generate the unique FLS estimate $b_N^{\mathrm{FLS}}(\mu, N)$ for the time-N coefficient vector b_N, conditional on the observations $y_1, \ldots, y_N$: i.e.

$$\begin{aligned} b_N^{\mathrm{FLS}}(\mu, N) &= \arg\min_{b_N}([y_N - x_N^{\mathrm{T}} b_N]^2 + \phi(b_N; \mu, N-1)) \\ &= [Q_{N-1}(\mu) + x_N x_N^{\mathrm{T}}]^{-1}[p_{N-1}(\mu) + x_N y_N]. \end{aligned} \tag{5.15}$$

The unique FLS estimates $(b_1^{\mathrm{FLS}}(\mu, N), \ldots, b_{N-1}^{\mathrm{FLS}}(\mu, N))$ for the coefficient vectors $b_1, \ldots, b_{N-1}$, conditional on the observations $y_1, \ldots, y_N$, can then be determined as follows.

In the course of deriving the FLS estimate (5.15), certain auxiliary vectors $e_n(\mu)$ and matrices $M_n(\mu)$, $1 \leqslant n \leqslant N-1$, were recursively generated in accordance with (5.7) and (5.8). Consider the sequence of relationships

$$\begin{aligned} b_1 &= e_1(\mu) + M_1(\mu) b_2 \quad ; \\ b_2 &= e_2(\mu) + M_2(\mu) b_3 \quad ; \\ &\ \ \vdots \\ b_{N-1} &= e_{N-1}(\mu) + M_{n-1}(\mu) b_N. \end{aligned} \tag{5.16}$$

By (5.6) and (5.7), each vector b_n appearing in the left column of (5.16) uniquely solves the minimization problem (5.3a) conditional on the particular vector b_{n+1} appearing in the corresponding right column of (5.16). Let equations (5.16) be solved for $b_1, \ldots, b_{N-1}$ in reverse order, starting with the initial condition $b_N = b_N^{\mathrm{FLS}}(\mu, N)$. These solution values yield† the desired FLS estimates for $b_1, \ldots, b_{N-1}$, conditional on the observations $y_1, \ldots, y_N$.

Consider any time-point n satisfying $K \leqslant n < N$. Using (5.7b), (5.7c), and (5.12), it follows by a straightforward calculation that the vector $e_n(\mu)$ takes the form

$$e_n(\mu) = [I - M_n(\mu)] b_n^{\mathrm{FLS}}(\mu, n). \tag{5.17}$$

Thus, for any given observations $y_1, \ldots, y_N$, the FLS smoothed estimate for b_n is a linear combination of the FLS filter estimate for b_n and the FLS smoothed estimate for b_{n+1}: i.e.

$$b_n^{\mathrm{FLS}}(\mu, N) = [I - M_n(\mu)] b_n^{\mathrm{FLS}}(\mu, n) + M_n(\mu) b_{n+1}^{\mathrm{FLS}}(\mu, N). \tag{5.18}$$

Note that the prior dynamic specifications (2.1b) constitute only a smoothness prior on the successive coefficient vectors $b_1, \ldots, b_N$. However complicated the actual dynamic relationships governing these vectors, their evolution as a function of n is only specified to be slow. Nevertheless, given the measurement prior (2.1a), the smoothness prior (2.1b), and the incompatibility cost specification (3.4), together with observations $\{y_1, \ldots, y_N\}$, the sequential FLS procedure generates explicit estimated dynamic relationships (5.16) for the entire sequence of unknown coefficient vectors $b_1, \ldots, b_N$ for each successive process length $N \geqslant K$.

6. FLS AND OLS: A GEOMETRIC COMPARISON

The FLS estimates for $b_1, \ldots, b_N$ can exhibit significant time-variation if warranted by the observations. Nevertheless, for every $\mu \geqslant 0$ and for every $N \geqslant K$, the FLS estimates are intrinsically related to the OLS solution which results if constancy is imposed on the coefficient vectors $b_1, \ldots, b_N$ prior to estimation.

Specifically, the following relationships are established in Section A.2 of Appendix A. First, as μ becomes arbitrarily large, the FLS estimate for each of the coefficient vectors $b_1, \ldots, b_N$ converges to the OLS solution $b^{\mathrm{OLS}}(N)$.

†To see this, express the minimized time-N incompatibility cost function $C(b; \mu, N)$ in terms of $\phi(b_N; \mu, N-1)$, analogous to (5.14), and then use the basic recurrence relation (5.3) to expand $\phi(b_N; \mu, N-1)$ into a recursive sequence of minimizations with respect to $b_1, \ldots, b_{N-1}$.

Theorem 6.1

Suppose the regressor matrix $X(N)$ has full rank K. Then

$$\lim_{\mu\to\infty} b_n^{\mathrm{FLS}}(\mu, N) = b^{\mathrm{OLS}}(N), \quad 1 \leqslant n \leqslant N. \tag{6.1}$$

Thus, OLS can be viewed as a limiting case of FLS in which absolute priority is given to the dynamic prior (2.1b) over the measurement prior (2.1a). As indicated in Fig. 1b, the squared residual error sums corresponding to the OLS solution do lie on the residual efficiency frontier $P_{\mathrm{F}}(N)$; but the investigator may have to pay a high price in terms of large residual measurement errors in order to achieve the zero residual dynamic errors required by OLS (see Section 8.4, below, for an empirical example).

Second, the OLS solution $b^{\mathrm{OLS}}(N)$ is a fixed matrix-weighted average of the FLS estimates for $b_1, \ldots, b_N$ for every $\mu \geqslant 0$.

Theorem 6.2

Suppose the regressor matrix $X(N)$ has full rank K. Then, for every $\mu \geqslant 0$,

$$\mathrm{b}^{\mathrm{OLS}}(N) = \left[\sum_{n=1}^{N} x_n x_n^{\mathrm{T}}\right]^{-1} \sum_{n=1}^{N} x_n x_n^{\mathrm{T}} b_n^{\mathrm{FLS}}(\mu, N). \tag{6.2}$$

The OLS solution can thus be viewed as a particular way of aggregating the information embodied in the FLS estimates for $b_1, \ldots, b_N$. A key difference between FLS and OLS is thus made strikingly apparent. The FLS approach seeks to understand which coefficient vector actually obtained at each time n; the OLS approach seeks to understand which coefficient vector obtained on average over time.

Finally, the FLS estimates for $b_1, \ldots, b_N$ are constant if and only if they coincide with the OLS solution and certain additional stringent conditions hold.

Theorem 6.3

Suppose $X(N)$ has full rank K. Then there exists a constant $K \times 1$ coefficient vector b such that

$$b_n^{\mathrm{FLS}}(\mu, N) = b, \quad 1 \leqslant n \leqslant N, \tag{6.3}$$

if and only if

$$b = b^{\mathrm{OLS}}(N) \quad \text{and} \quad [x_n^{\mathrm{T}} b^{\mathrm{OLS}}(N) - y_n] x_n = \mathbf{0}, \quad 1 \leqslant n \leqslant N. \tag{6.4}$$

7. REGIME SHIFT: A ROBUSTNESS STUDY FOR FLS

The FLS solution reflects the prior belief that the coefficient vectors b_n evolve only slowly over time, if at all. Suppose the true coefficient vectors actually undergo a time-variation which is contrary to this prior belief: namely, a single unanticipated shift at some time S.

More precisely, suppose the observations y_n for the linear regression model (2.1) are actually generated in the form

$$y_n = \begin{cases} x_n^{\mathrm{T}} z, & n = 1, \ldots, S; \\ x_n^{\mathrm{T}} w, & n = S+1, \ldots, N, \end{cases} \tag{7.1}$$

where N, S, and K are arbitrary integers satisfying $N > S \geqslant 1$ and $N > K \geqslant 1$, z and w are distinct constant $K \times 1$ coefficient vectors, and the $N \times K$ regressor matrix $X(N)$ has full rank K. Would an investigator using the FLS solution (3.6) be led to suspect, from the nature of the coefficient estimates he obtains, that the true coefficient vector shifted from z to w at time S? An affirmative answer is provided below in Theorems 7.1 and 7.2 (proofs are given in Section A.3 of Appendix A).

Consider, first, the scalar coefficient case $K = 1$. Suppose $x_n \neq 0$, $1 \leqslant n \leqslant N$, and $z < w$. Then, as detailed in Theorem 7.1, below, the FLS estimates for $b_1, \ldots, b_N$ at any time $N > S$ exhibit the following four properties: (i) the FLS estimates monotonically increase between z and w; (ii) the FLS estimates increase at an increasing rate over the initial time points $1, \ldots, S$ and at a decreasing

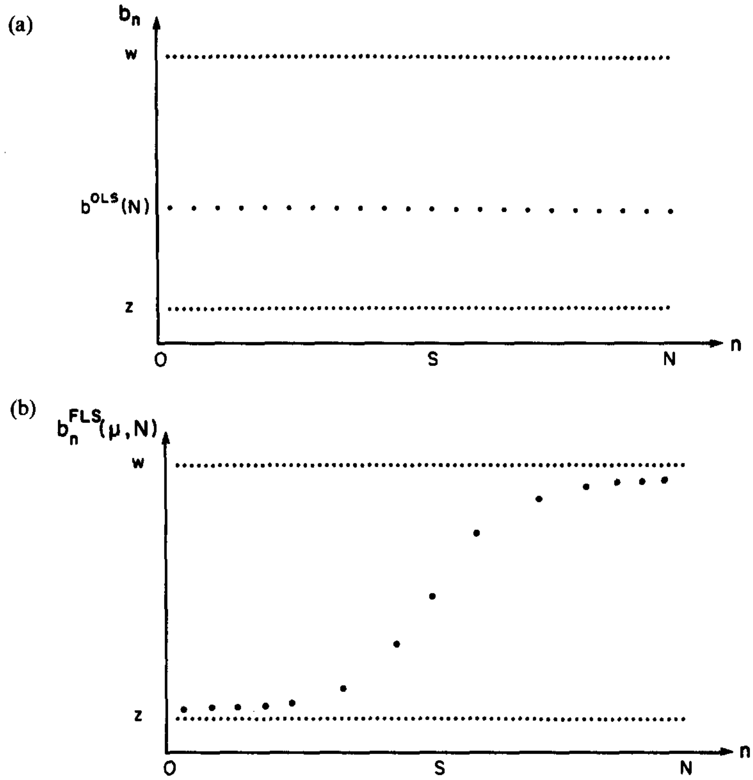


Fig. 2. (a) Qualitative properties of the OLS solution at time N with unanticipated shift from z to w at time S, (b) qualitative properties of the FLS solution for $b_1, \ldots, b_N$ at time N with an unanticipated shift from z to w at time S.

rate over the final time points $S+1, \ldots, N$; (iii) the initial S estimates cluster around z, with tighter clustering occurring for larger values of S and for smaller values of μ, and the final $N-(S+1)$ estimates cluster around w, with tighter clustering occurring for larger values of $N-(S+1)$ and for smaller values of μ; and (iv) if x_N remains bounded away from zero as N approaches infinity, the FLS estimate for b_N converges to w as N approaches infinity (see Fig. 2).

The statement of Theorem 7.1 makes use of certain auxiliary quantities $L_n(\mu)$, $1 \leqslant n \leqslant N$, defined as follows. Recall definition (4.1e) for the positive definite $K \times K$ matrices $A_n(\mu)$, $1 \leqslant n \leqslant N$. Let positive definite $K \times K$ matrices $L_n(\mu)$ be defined by

$$L_n(\mu) = \begin{cases} \mu A_n(\mu)^{-1} & \text{if} \quad n = 1 \text{ or } N; \\ 2\mu A_n(\mu)^{-1} & \text{if} \quad 1 < n < N. \end{cases} \tag{7.2}$$

It follows immediately from the well-known Woodbury matrix inversion lemma that

$$[I - L_n(\mu)] = A_n(\mu)^{-1} x_n x_n^{\mathrm{T}} = \begin{cases} x_n x_n^{\mathrm{T}}/[\mu + x_n^{\mathrm{T}} x_n] & \text{if} \quad n = 1 \text{ or } N; \\ x_n x_n^{\mathrm{T}}/[2\mu + x_n^{\mathrm{T}} x_n] & \text{if} \quad 1 < n < N. \end{cases} \tag{7.3}$$

In the special case $K = 1$, $L_n(\mu)$ is a scalar lying between zero and one, strictly so if $x_n \neq 0$. Moreover, $L_n(\mu) \to 1$ as $\mu \to \infty$ and $L_n(\mu) \to 0$ as $\mu \to 0$.

Theorem 7.1

Consider the linear regression model (2.1) with $K = 1$ and with $x_n \neq 0$ for $1 \leqslant n \leqslant N$. Suppose the observations y_n in (2.1) are actually generated in accordance with (7.1), where z and w are scalar coefficients satisfying $z < w$, and S is an arbitrary integer satisfying $1 \leqslant S < N$. Then the FLS solution (3.6) displays the following four properties for each $\mu > 0$:

(i) $z < b_1^{\mathrm{FLS}}(\mu, N) < \cdots < b_N^{\mathrm{FLS}}(\mu, N) < w$;

(ii) (a) $[b_{n+1}^{\mathrm{FLS}}(\mu, N) - b_n^{\mathrm{FLS}}(\mu, N)] > [b_n^{\mathrm{FLS}}(\mu, N) - b_{n-1}^{\mathrm{FLS}}(\mu, N)]$ for $1 \leqslant n \leqslant S$;

(b) $[b_{n+1}^{\mathrm{FLS}}(\mu, N) - b_n^{\mathrm{FLS}}(\mu, N)] < [b_n^{\mathrm{FLS}}(\mu, N) - b_{n-1}^{\mathrm{FLS}}(\mu, N)]$ for $S+1 \leqslant n < N$;

(iii) (a) $[b_n^{FLS}(\mu, N) - z] < \left[\prod_{k=n}^{S} L_k(\mu)\right][w - z]$ for $1 \leqslant n \leqslant S$;

(b) $[w - b_n^{FLS}(\mu, N)] < \left[\prod_{k=S+1}^{n} L_k(\mu)\right][w - z]$ for $S + 1 \leqslant n \leqslant N$;

(iv) $x_N^T[w - b_N^{FLS}(\mu, N)] \to 0$ as $N \to \infty$.

The next theorem establishes that, for the general linear regression model (2.1) with observations generated in accordance with (7.1), the FLS estimates for b_1 through b_S move successively away from z and the FLS estimates for b_{S+1} through b_N move successively toward w.

Theorem 7.2

Suppose the observations y_n for the linear regression model (2.1) are generated in accordance with (7.1), where N, S, and K are arbitrary integers satisfying $N > S \geqslant 1$ and $N > K \geqslant 1$, z and w are distinct constant $K \times 1$ coefficient vectors, and the $N \times K$ regressor matrix $X(N)$ has full rank K. Then the FLS solution (3.6) displays the following properties for each $\mu > 0$: (i) For $1 \leqslant n \leqslant S$,

$$[b_{n+1}^{FLS}(\mu, N) - z]^T[b_{n+1}^{FLS}(\mu, N) - z] \geqslant [b_n^{FLS}(\mu, N) - z]^T[b_n^{FLS}(\mu, N) - z],$$

with strict inequality holding for n if strict inequality holds for $n - 1$; and (ii) for $S + 1 \leqslant n < N$,

$$[b_{n+1}^{FLS}(\mu, N) - w]^T[b_{n+1}^{FLS}(\mu, N) - w] \leqslant [b_n^{FLS}(\mu, N) - w]^T[b_n^{FLS}(\mu, N) - w],$$

with strict inequality holding for n if strict inequality holds for $n + 1$.

8. SIMULATION AND EMPIRICAL STUDIES

A FORTRAN program "FLS" has been developed which implements the FLS sequential solution procedure for the time-varying linear regression problem (see Appendix B). As part of the program validation, various simulation experiments have been performed. In addition, the program has been used in [2] to conduct an empirical study of U.S. money demand instability over 1959:Q2–1985:Q3. A brief summary of these simulation and empirical studies will now be given.

8.1. Simulation experiment specifications

The dimension K of the regressor vectors x_n was fixed at 2. The first regressor vector, x_1, was specified to be $(1,1)^T$. For $n \geqslant 2$, the components of the regressor vector x_n were specified as follows:

$$x_{n1} = \sin(10 + n) + 0.01; \tag{8.1a}$$

$$x_{n2} = \cos(10 + n). \tag{8.1b}$$

The components of the two-dimensional coefficient vectors b_n were simulated to exhibit linear, quadratic, sinusoidal, and regime shift time-variations, in various combinations. The true residual dynamic errors $[b_{n+1} - b_n]$ were thus complex nonlinear functions of time.

The number of observations N was varied over $\{15, 30, 90\}$. Each observation y_n was generated in accordance with the linear regression model $y_n = x_n^T b_n + v_n$, where the discrepancy term v_n was generated from a pseudo-random number generator for a normal distribution $N(0, \sigma)$. The standard deviation σ was varied over $\{0, 0.5, 0.10, 0.20, 0.30\}$, where $\sigma = 0.x$ roughly corresponded to an x% error in the observations.

8.2. Simulation experiment results: general summary

The residual efficiency frontier $P_F(N)$ for each experiment was adequately traced out by evaluating the FLS estimates (3.6) and the corresponding residual error sums (3.7) over a rough grid of penalty weights μ increasing by powers of ten: namely, $\{0.01, 0.10, 1, 10, 100, 1000, 10000\}$.

No instability or other difficult numerical behavior was encountered. Each of the residual efficiency frontiers displayed the general qualitative properties depicted in Fig. 1b.

In each experiment the FLS estimates depicted the qualitative time-variations displayed by the true coefficient vectors, despite noisy observations. The accuracy of the depictions were extremely good for noise levels $\sigma \leqslant 0.20$ and for balanced penalty weightings $\mu \approx 1.0$. The accuracy of the depictions ultimately deteriorated with increases in the noise level σ, and for extreme values of μ. However, the overall tracking power displayed by the FLS estimates was similar for all three sample sizes, $N = 15$, 30, and 90. Presumably this experimentally observed invariance to sample size is a consequence of the fact that FLS provides a separate estimate for each coefficient vector at each time n rather than an estimate for the "typical" coefficient vector across time.

8.3. Illustrative experimental results for sinusoidal time-variations

Experiments were carried out with $N = 30$ and $\sigma = 0.05$ for which the components of the true time-n coefficient vector $b_n = (b_{n1}, b_{n2})$ were simulated to be sinusoidal functions of n. The first component, b_{n1}, moved through two complete periods of a sine wave over $\{1, \ldots, N\}$, and the second component, b_{n2}, moved through one complete period of a sine wave over $\{1, \ldots, N\}$. For the penalty weight $\mu = 1.0$, the FLS estimates b_{n1}^{FLS} and b_{n2}^{FLS} closely tracked the true coefficients b_{n1} and b_{n2}. As μ was increased from 1.0 to 1000 by powers of ten, the FLS estimates b_{n1}^{FLS} and b_{n2}^{FLS} were pulled steadily inward toward the OLS solution (0.03, 0.04); but the two-period and one-period sinusoidal motions of the true coefficients b_{n1} and b_{n2} were still reflected (see Fig. 3).

Another series of experiments was conducted with $N = 30$ and σ varying over $\{0, 0.05, 0.10, 0.20\}$ for which the true coefficient vectors traced out an ellipse over the observation interval. The

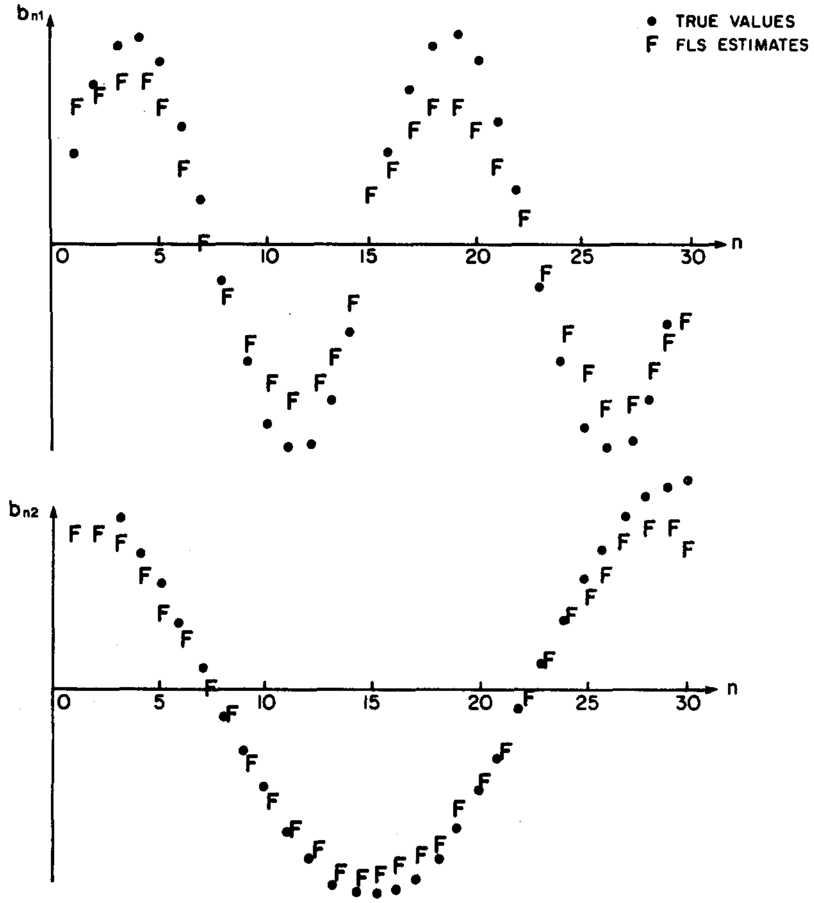


Fig. 3. Sine wave experiment with parameter values $\sigma = 0.05$ $\mu = 1$ and $N = 30$.

OLS solution for each of these experiments was approximately at the center (0, 0) of the ellipse. For $\mu = 1.0$, the FLS estimates closely tracked the true coefficient vectors. As μ was increased from 1.0 to 1000 by powers of ten, the FLS estimates were pulled steadily inward toward the OLS solution; but for each μ the FLS estimates still traced out an approximately elliptical trajectory around the OLS solution. The residual efficiency frontier and corresponding FLS estimates were surprisingly insensitive to the magnitude of σ over the range [0, 0.20]. The elliptical shape traced out by the FLS estimates began to exhibit jagged portions at a noise level $\sigma = 0.30$. Figure 4 plots the experimental outcomes for $\mu = 1.0$ and for $\mu = 100$ with noise level $\sigma = 0.05$.

A similar series of elliptical experiments was then carried out for the smaller sample size $N = 15$. The true coefficient vector traversed the same ellipse as before, but over fifteen successive observation periods rather than over thirty. Thus the true coefficient vector was in faster motion, implying larger residual dynamic errors $[b_{n+1} - b_n]$ would have to be sustained to achieve good coefficient tracking. For each given μ the FLS estimates still traced out an elliptical trajectory around the OLS solution, with good tracking achieved for $\sigma \leqslant 0.20$ and $\mu \approx 1.0$. However, in comparison with the corresponding thirty observation experiments, the elliptical trajectory was pulled further inward toward the OLS solution for each given μ.

The number of observations was then increased to ninety. The true coefficient vectors traced out the same ellipse three successive times over this observation interval. The noise level σ was set at 0.05 and the penalty weight μ was set at 1.0. The FLS estimates corresponding to $\mu = 1.0$ closely tracked the true coefficient vectors three times around the ellipse, with no indication of any tracking deterioration over the observation interval.

Finally, the latter experiment was modified so that the true coefficient vectors traced out the same ellipse six times over the ninety successive observation points. Also, the noise level σ was increased to 0.10. The FLS estimates corresponding to $\mu = 1.0$ then closely tracked the true coefficient vectors six times around the ellipse, with no indication of any tracking deterioration over the observation interval.

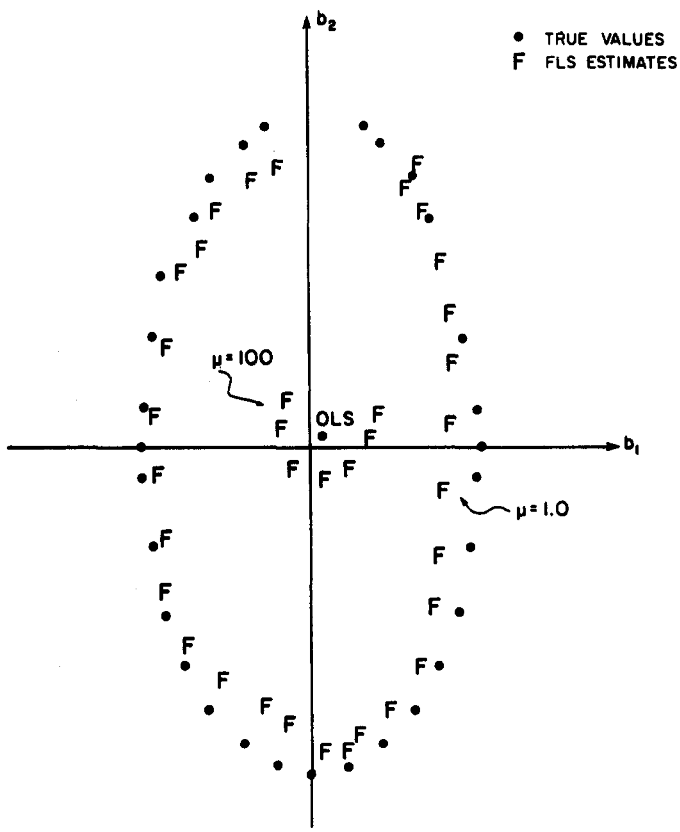


Fig. 4. Ellipse experiment with parameter values $\sigma = 0.05$, $N = 30$ and $\mu = 1$ and 100.

8.4. *An empirical application: U.S. money demand instability*

In [2] two basic hypotheses are formulated for U.S. money demand: a measurement hypothesis that observations on real money demand have been generated in accordance with the well-known Goldfeld log–linear regression model [3]; and a dynamic hypothesis that the coefficients characterizing the regression model have evolved only slowly over time, if at all.

Time-paths are generated and plotted for all regression coefficients over 1959:Q2–1985:Q3 for a range of points along the residual efficiency frontier, including the extreme point corresponding to OLS estimation. At each point of the frontier other than the OLS extreme point, the estimated time-paths exhibit a clear-cut shift in 1974 with a partial reversal of this shift beginning in 1983. Since the only restriction imposed on the time-variation of the coefficients is a simple nonparametric smoothness prior, these results would seem to provide striking evidence that structural shifts in the money demand function indeed occurred in 1974 and 1983, as many OLS money demand studies have surmised. The shifts are small, however, in relationship to the pronounced and persistent downward movement exhibited by the estimated coefficient for the inflation rate over the entire sample period. Thus the shifts could be an artifact of model misspecification rather than structural breaks in the money demand relationship itself.

A second major finding of this study is the apparent fragility of inferences from OLS estimation, both for the whole sample period and for the pre-1974 and post-1974 subperiods. Specifically, the OLS estimates exhibit sign and magnitude properties which are not representative of the typical FLS coefficient estimates along the residual efficiency frontier. Moreover, the residual efficiency frontier is extremely attenuated in a neighborhood of the OLS solution, indicating that a high price must be paid in terms of residual measurement error in order to achieve the zero residual dynamic error (time-constant coefficients) required by OLS.

For example, at the extreme point corresponding to OLS estimation for the 1974:Q1–1985:Q3 subperiod, nominal money balances appear to be following a simple random walk $M_{t+1} \approx M_t$, indicating the presence of a severe "unit root" nonstationarity problem. These findings coincide with the findings of many other OLS money demand studies. In contrast, along more than 80% of the frontier for this same subperiod the FLS estimates for the coefficient on the log of lagged real money balances remain bounded in the interval [0.59, 0.81]; and the FLS coefficient estimates for other regressors (e.g. real GNP) are markedly larger than the corresponding OLS estimates. Thus the appearance of a unit root in money demand studies could be the spurious consequence of requiring absolute constancy of the coefficient vectors across time.

9. TOPICS FOR FUTURE RESEARCH

Starting from the rather weak prior specifications of locally linear measurement and slowly evolving coefficients, the sequential FLS solution procedure developed in Section 5 generates explicit estimated dynamic relationships (5.16) connecting the successive coefficient vectors $b_1, \ldots, b_N$ for each process length N. How reliably do these estimated dynamic relationships reflect the true dynamic relationships governing the successive coefficient vectors? The regime shift results analytically established in Section 7 and the simulation results summarized in Section 8 both appear promising in this regard.

More systematic procedures need to be developed for interpreting and reporting the time-variations exhibited by the FLS estimates along the residual efficiency frontier. As noted in Section 3, these estimates constitute a "population" characterized by a basic efficiency property: no other coefficient sequence estimate yields both lower measurement error and lower dynamic error for the given observations. Given this population, one can begin to explore systematically the extent to which any additional properties of interest are exhibited within the population. The frontier can be parameterized by a parameter $\delta \equiv \mu/[1+\mu]$ varying over the unit interval. For properties amenable to quantification, this permits the construction of an empirical distribution for the property. Such constructs were used in [2] to interpret and report findings for an empirical money demand study; other studies currently underway will further develop this approach.

Suppose y is actually a nonlinear function of x, and observations $y_1, \ldots, y_N$ have been obtained on y over a grid $x_1, \ldots, x_N$ of successive x-values. As the study by White [34] makes clear, the OLS estimate for a single (average) coefficient vector in a linear regression of $y_1, \ldots, y_N$ on $x_1, \ldots, x_N$ cannot be used in general to obtain information about the local properties of the nonlinear relation between y and x. Does the estimated relation $y_n = x_n^T b_n^{FLS}(\mu, N)$ between y_n and x_n generated via the FLS procedure for $n = 1, \ldots, N$ provide any useful information concerning the nonlinear relation between y and x? Encouraging results along these lines have been obtained in the statistical smoothing splines literature (e.g. [21]).

The geometric relationship between the FLS and OLS solutions established in Theorem 6.2 is suggestive of the "reflections in lines" construction for the OLS solution provided by D'Ocagne [35]. Can the D'Ocagne construction be used to provide a clearer geometric understanding of the FLS solution?

Finally, starting from the prior beliefs of locally linear measurement and slowly evolving coefficient vectors, the estimated dynamic relationships (5.16) connecting the successive coefficient vector estimates $b_n^{FLS}(\mu, N)$ represent the "posterior" dynamic equations generated by the FLS procedure, conditional on the given data set $\{y_1, \ldots, y_N\}$. An important question concerns the use of these posterior dynamic equations for prediction and adaptive model respecification.

These and other questions will be addressed in future studies.

REFERENCES

1. R. Kalaba and L. Tesfatsion, Time-varying regression via flexible least squares. Modelling Research Group Working Paper No. 8633, Department of Economics, University of Southern California, Los Angeles, Calif. (August 1986).
2. L. Tesfatsion and J. Veitch, Money demand instability: a flexible least squares approach. Modelling Research Group Working Paper No. 8809, Department of Economics, University of Southern California, Calif. (November 1988).
3. S. M. Goldfeld, The demand for money revisited. *Brookings Papers econ. Activity* **3**, 577–638 (1973).
4. R. Quandt, The estimation of the parameters of a linear regression system obeying two separate regimes. *J. Am. statist. Ass.* **53**, 873–880 (1958).
5. R. Quandt, Tests of the hypothesis that a linear regression system obeys two separate regimes. *J. Am. statist. Ass.* **55**, 324–330 (1960).
6. G. Chow, Tests of equality between sets of coefficients in two linear regressions. *Econometrica* **28**, 167–184 (1960).
7. F. Fisher, Tests of equality between sets of coefficients in two linear regressions: an expository note. *Econometrica* **38**, 361–366 (1970).
8. S. Guthery, Partition regression. *J. Am. statist. Ass.* **69**, 945–947 (1974).
9. R. Brown, J. Durbin and J. Evans, Techniques for testing the constancy of regression relations over time. *J. R. statist. Soc.* **37B**, 149–192 (1975).
10. J. Ertel and E. Fowlkes, Some algorithms for linear spline and piecewise multiple linear regression. *J. Am. statist. Ass.* **71**, 640–648 (1976).
11. T. Cooley and E. Prescott, Estimation in the presence of stochastic parameter variation. *Econometrica* **44**, 167–184 (1976).
12. B. Rosenberg, A survey of stochastic parameter regression. *Ann. econ. soc. Measur.* **2/4**, 381–397 (1973).
13. A. Sarris, A Bayesian approach to estimation of time-varying regression coefficients. *Ann. econ. soc. Measur.* **2/4**, 501–523 (1973).
14. P. Harrison and C. Stevens, Bayesian forecasting. *J. R. statist. Soc.* **B28**, 205–247 (1976).
15. K. Garbade, Two methods for examining the stability of regression coefficients. *J. Am. statist. Ass.* **72**, 54–63 (1977).
16. G. Chow, Random and changing coefficient models. In *Handbook of Econometrics* (Eds Z. Griliches and M. Intriligator), Vol. 2, Chap. 21 and pp. 1213–1245. Elsevier, Amsterdam (1984).
17. C. Carraro, Regression and Kalman filter methods for time-varying econometric models. Econometric Research Program Memorandum No. 320, Princeton University, Princeton, N.J. (1985).
18. R. Engle and M. W. Watson, Applications of Kalman filtering in econometrics. Discussion Paper 85-31, Department of Economics, University of California at San Diego, La Jolla, Calif. (1985).
19. R. E. Kalman, A new approach to linear filtering and prediction problems. *Trans. ASME, J. basic Engng* **82**, 35–45 (1960).
20. R. E. Kalman and R. S. Bucy, New results in linear filtering and prediction theory. *Trans. ASME, J. basic Engng* **83**, 95–108 (1961).
21. P. Craven and G. Wahba, Smoothing noisy data with spline functions. *Numer. Math.* **31**, 377–403 (1979).
22. T. Doan, R. Litterman and C. Sims, Forecasting and conditional projection using realistic prior distributions. *Econometric Rev.* **3**, 1–100 (1984).
23. P. Miller and W. Roberds, The quantitative significance of the Lucas critique. Staff Report 109, Federal Reserve Bank of Minneapolis, Research Department, 250 Marquette Avenue, Minneapolis, Minn. (April 1987).
24. R. Kalaba and L. Tesfatsion, A least-squares model specification test for a class of dynamic nonlinear economic models with systematically varying parameters. *J. Optim. Theory Applic.* **32**, 538–567 (1980).
25. R. Kalaba and L. Tesfatsion, An exact sequential solution procedure for a class of discrete-time nonlinear estimation problems. *IEEE Trans. autom. Control* **AC26**, 1144–1149 (1981).
26. R. Kalaba and L. Tesfatsion, Exact sequential filtering, smoothing, and prediction for nonlinear systems. *Nonlinear Analysis Theory Meth. Applic.* **12**, 599–615 (1988).
27. R. Kalaba and L. Tesfatsion, Sequential nonlinear estimation with nonaugmented priors. *J. Optim. Theory Applic.* **60** (1989).

28. R. Bellman, H. Kagiwada, R. Kalaba and R. Sridhar, Invariant imbedding and nonlinear filtering theory. *J. astronaut. Sci.* **13**, 110–115 (1966).
29. D. Detchmendy and R. Sridhar, Sequential estimation of states and parameters in noisy nonlinear dynamical systems. *Trans. ASME, J. basic Engng* 362–368 (1966).
30. H. Kagiwada, R. Kalaba, A. Schumitsky and R. Sridhar, Invariant imbedding and sequential interpolating filters for nonlinear processes. *Trans. ASME, J. basic Engng* **91**, 195–200 (1969).
31. R. Bellman and R. Kalaba, *Quasilinearization and Nonlinear Boundary-value Problems.* Elsevier, New York (1965).
32. R. Bellman, R. Kalaba and G. Milton Wing, Invariant imbedding and the reduction of two-point boundary value problems to initial value problems. *Proc. natn. Acad. Sci. U.S.A.* **46**, 1646–1649 (1960).
33. R. Larson and J. Peschon, A dynamic programming approach to trajectory estimation. *IEEE Trans. autom. Control* **AC11**, 537–540 (1966).
34. H. White, Using least squares to approximate unknown regressor functions. *Int. econ. Rev.* **21**, 149–170 (1980).
35. M. D'Ocagne, La determination geometrique du point le plus probable donne par un systeme de droites non convergentes. *J. Ecole Polytech.* **63**, 1–25 (1893).

APPENDIX A

Theorem Proofs

A.1. Proofs for Section 4

It will first be shown that the matrix representation (4.2) for the FLS incompatibility cost function $C(b(N); \mu, N)$ is correct. The proof will make use of the following preliminary lemma.

Lemma 4.1

Let N and K be arbitrary given integers satisfying $N \geqslant 2$ and $K \geqslant 1$, and let $A(\mu, N)$ be defined as in (4.1f). Then, for any $NK \times 1$ column vector $w = (w_1^T, \ldots, w_N^T)^T$ consisting of N arbitrary $K \times 1$ component vectors w_n, $1 \leqslant n \leqslant N$,

$$w^T A(\mu, N) w = \sum_{n=1}^{N} w_n^T x_n x_n^T w_n + \mu \sum_{n=1}^{N-1} [w_{n+1} - w_n]^T [w_{n+1} - w_n]. \tag{A1}$$

Proof. It is easily established, by straightforward calculation, that (A1) holds for all $2K \times 1$ column vectors w.

Suppose (A1) has been shown to hold for all $NK \times 1$ column vectors w for some $N \geqslant 2$. Note that $A(\mu, N+1)$ can be expressed in terms of $A(\mu, N)$ as follows:

$$A(\mu, N+1) = \left[\begin{array}{c|c} A(\mu, N) & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right] + \left[\begin{array}{cccc|c} \mathbf{0} \cdots \mathbf{0} & & \mathbf{0} & & \mathbf{0} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mu I & & -\mu I \\ \hline \mathbf{0} \cdots \mathbf{0} & & -\mu I & & A_{N+1}(\mu) \end{array}\right] \tag{A2}$$

where, as in Section 4, $A_{N+1}(\mu) = x_{N+1} x_{N+1}^T + \mu I$, and I denotes the $K \times K$ identity matrix. Let $(w_1, \ldots, w_N, w_{N+1})$ denote an arbitrary sequence of $N+1$ column vectors, each of dimension $K \times 1$, and let

$$w = (w_1^T, \ldots, w_N^T)^T; \tag{A3a}$$

$$v = (w^T, w_{N+1}^T)^T. \tag{A3b}$$

Then, using (A2),

$$v^T A(\mu, N+1) v = w^T A(\mu, N) w + w_{N+1}^T x_{N+1} x_{N+1}^T w_{N+1} + \mu [w_{N+1} - w_N]^T [w_{N+1} - w_N]. \tag{A4}$$

It follows by the induction step that (A1) holds for the arbitrary $(N+1)K \times 1$ column vector v. Q.E.D.

Theorem 4.1

The FLS incompatibility cost function $C(b(N); \mu, N)$ defined by (3.4) has the matrix representation

$$C(b(N); \mu, N) = b(N)^T A(\mu, N) b(N) - 2b(N)^T G(N) y(N) + y(N)^T y(N). \tag{A5}$$

Proof. Recall that $b(N) = (b_1^T, \ldots, b_N^T)^T$. It follows from Lemma 4.1 that

$$b(N)^T A(\mu, N) b(N) = \sum_{n=1}^{N} b_n^T x_n x_n^T b_n + \mu \sum_{n=1}^{N-1} [b_{n+1} - b_n]^T [b_{n+1} - b_n]. \tag{A6}$$

Thus,

$$[C(b(N); \mu, N) - b(N)^T A(\mu, N) b(N)] = -2 \sum_{n=1}^{N} [x_n^T b_n] y_n + \sum_{n=1}^{N} y_n^2 = -2b(N)^T G(N) y(N) + y(N)^T y(N). \quad \text{Q.E.D.} \tag{A7}$$

Theorem 4.2

Suppose $X(N)$ has full rank K. Then $A(\mu, N)$ is positive definite for every $\mu > 0$.

Proof. It must be shown that $w^T A(\mu, N) w > 0$ for every nonzero $NK \times 1$ column vector w. By Lemma 4.1, it is obvious that $w^T A(\mu, N) w \geqslant 0$. Suppose $w^T A(\mu, N) w = 0$ for some nonzero $NK \times 1$ column vector $w = (w_1^T, \ldots, w_N^T)^T$, where each component vector w_n is $K \times 1$. Then each of the nonnegative sums in (A1) must be zero; in particular, it must be true that $w_{n+1} = w_n$, $n = 1, \ldots, N-1$. Thus

$$0 = w^{\mathrm{T}}A(\mu, N)w = \sum_{n=1}^{N} w_n^{\mathrm{T}}x_nx_n^{\mathrm{T}}w_n = w_1^{\mathrm{T}}\left[\sum_{n=1}^{N} x_nx_n^{\mathrm{T}}\right]w_1 = w_1^{\mathrm{T}}X(N)^{\mathrm{T}}X(N)w_1, \tag{A8}$$

with $w_1 \neq \mathbf{0}$. However, (A8) contradicts the assumed nonsingularity of $X(N)^{\mathrm{T}}X(N)$. Q.E.D.

Corollary 4.1

Suppose $X(N)$ has full rank K, and $\mu > 0$. Then the FLS incompatibility cost function $C(b(N); \mu, N)$ is a strictly convex function of $b(N)$ which attains its unique minimum at

$$b^{\mathrm{FLS}}(\mu, N) = A(\mu, N)^{-1}G(N)y(N). \tag{A9}$$

Proof. Strict convexity of $C(b(N); \mu, N)$ follows directly from Theorem 4.1 and Theorem 4.2. Thus, the first-order necessary conditions for minimization of $C(b(N); \mu, N)$ are also sufficient, and have at most one solution. By Theorem 4.1, these first-order conditions take the form

$$\mathbf{0} = A(\mu, N)b(N) - G(N)y(N), \tag{A10}$$

with unique solution (A9). Q.E.D.

A.2. Proofs for Section 6

Let N and K be arbitrary given integers satisfying $N \geqslant 2$ and $N \geqslant K \geqslant 1$. Suppose the $N \times K$ regressor matrix $X(N)$ for the linear regression model (2.1) has full rank K. Define $y(N)$ to be the $N \times 1$ column vector of observations $(y_1, \ldots, y_N)^{\mathrm{T}}$. and let a constant $K \times 1$ coefficient vector b replace b_n in (2.1a) for $n = 1, \ldots, N$. Then the ordinary least squares (OLS) problem is to estimate the constant coefficient vector b thought to underly the generation of the observation vector $y(N)$ by selecting b to minimize the sum of squared residual measurement errors

$$S(b, N) = \sum_{n=1}^{N} [y_n - x_n^{\mathrm{T}}b]^2 = [y(N) - X(N)b]^{\mathrm{T}}[y(N) - X(N)b]. \tag{A11}$$

The first-order necessary conditions (normal equations) for minimization of $S(b, N)$ take the form

$$\mathbf{0} = \sum_{n=1}^{N} [y_n - x_n^{\mathrm{T}}b]x_n = X(N)^{\mathrm{T}}y(\mathrm{N}) - X(N)^{\mathrm{T}}X(N)b. \tag{A12}$$

The OLS solution for b is thus uniquely given by

$$b^{\mathrm{OLS}}(N) = \left[\sum_{n=1}^{N} x_nx_n^{\mathrm{T}}\right]^{-1}\sum_{n=1}^{N} x_ny_n = [X(N)^{\mathrm{T}}X(N)]^{-1}X(N)^{\mathrm{T}}y(N). \tag{A13}$$

Proof of Theorem 6.1. In component form, the first-order necessary conditions (4.3) for minimization of the incompatibility cost function (4.2) take the following form: for $n = 1$:

$$\mathbf{0} = [x_1^{\mathrm{T}}b_1 - y_1]x_1 - \mu[b_2 - b_1]; \tag{A14a}$$

for $1 < n < N$:

$$\mathbf{0} = [x_n^{\mathrm{T}}b_n - y_n]x_n - \mu[b_{n+1} - b_n] + \mu[b_n - b_{n-1}]; \tag{A14b}$$

for $n = N$:

$$\mathbf{0} = [x_N^{\mathrm{T}}b_N - y_N]x_N + \mu[b_N - b_{N-1}]. \tag{A14c}$$

By a simple manipulation of terms, the first-order conditions (A14) can be given the alternative representation

$$\mu[b_{n+1} - b_n] = \sum_{s=1}^{n} [x_s^{\mathrm{T}}b_s - y_s]x_s, \quad 1 \leqslant n < N; \tag{A15a}$$

$$\mathbf{0} = \sum_{n=1}^{N} [x_n^{\mathrm{T}}b_n - y_n]x_n. \tag{A15b}$$

Introduce the transformation of variables

$$b_n = b^{\mathrm{OLS}}(N) + [b_n - b^{\mathrm{OLS}}(N)] \equiv b^{\mathrm{OLS}}(N) + u_n, \quad 1 \leqslant n \leqslant N. \tag{A16}$$

Then, letting u denote the vector $(u_1, \ldots, u_N)$, the FLS incompatibility cost function (4.2) can be expressed as

$$c(b(N); \mu, N) = \sum_{n=1}^{N} [x_n^{\mathrm{T}}b^{\mathrm{OLS}}(N) + x_n^{\mathrm{T}}u_n - y_n]^2 + \mu\sum_{n=1}^{N-1} [u_{n+1} - u_n]^{\mathrm{T}}[u_{n+1} - u_n] \equiv V(u; \mu, N). \tag{A17}$$

Using (A15), the first-order necessary conditons for a vector $u = (u_1, \ldots, u_N)$ to minimize $V(u; \mu, N)$ take the form

$$\mu[u_{n+1} - u_n] = \sum_{s=1}^{n} [x_s^{\mathrm{T}}b^{\mathrm{OLS}}(N) - y_s]x_s + \sum_{s=1}^{n} [x_sx_s^{\mathrm{T}}]u_s, \quad 1 \leqslant n < N; \tag{A18a}$$

$$\mathbf{0} = \sum_{n=1}^{N} [x_nx_n^{\mathrm{T}}]u_n, \tag{A18b}$$

where use has been made of the fact that $b^{\mathrm{OLS}}(N)$ satisfies the first-order necessary conditons (A12) for the minimization of $S(b, N)$.

The proof of Theorem 6.1 now proceeds by a series of lemmas.

Lemma 6.1

The FLS solution $u^*(\mu, N)$ to the minimization of $V(u; \mu, N)$ defined by (A17) satisfies

$$[u^*_{n+1}(\mu, N) - u^*_n(\mu, N)] \to \mathbf{0} \quad \text{as} \quad \mu \to \infty, \quad 1 \leq n \leq N-1. \tag{A19}$$

Proof. Suppose (A19) does not hold. Then for some n there exists $\epsilon > 0$ such that

$$[u^*_{n+1}(\mu, N) - u^*_n(\mu, N)]^{\mathrm{T}}[u^*_{n+1}(\mu, N) - u^*_n(\mu, N)] \geq \epsilon$$

for all sufficiently large μ. It follows from (A17) that $V(u^*(\mu, N); \mu, N) \geq \mu\epsilon$ for all sufficiently large μ, i.e. the minimum FLS incompatibility cost diverges to infinity as μ approaches infinity. However, it is also clear from (A17) that $V(\mathbf{0}; \mu, N) = S(b^{\mathrm{OLS}}(N)) < \infty$ for all $\mu > 0$, a contradiction. Thus (A19) must hold. Q.E.D.

Lemma 6.2

For each n, $1 \leq n \leq N$,

$$[u^*_n(\mu, N) - u^*_1(\mu, N)] \to \mathbf{0} \quad \text{as} \quad \mu \to \infty.$$

Proof. The proof is immediate from Lemma 6.1, since

$$[u^*_n(\mu, N) - u^*_1(\mu, N)] = [u^*_n(\mu, N) - u^*_{n-1}(\mu, N)] + [u^*_{n-1}(\mu, N) - u^*_{n-2}(\mu, N)] + \cdots + [u^*_2(\mu, N) - u^*_1(\mu, N)].\ \text{Q.E.D.}$$

Lemma 6.3

Suppose $X(N)$ has full rank K. Then, for each n, $1 \leq n \leq N$,

$$u^*_n(\mu, N) \to \mathbf{0} \quad \text{as} \quad \mu \to \infty. \tag{A20}$$

Proof. By Lemma 6.1, Lemma 6.2, and the first-order necessary condition (A18b),

$$\mathbf{0} = \sum_{n=1}^{N} [x_n x_n^{\mathrm{T}}] u^*_n(\mu, N) \to \left[\sum_{n=1}^{N} x_n x_n^{\mathrm{T}}\right] u^*_1(\mu, N) = X(N)^{\mathrm{T}} X(N) u^*_1(\mu, N) \tag{A21}$$

as $\mu \to \infty$; hence $u^*_1(\mu, N) \to [X(N)^{\mathrm{T}} X(N)]^{-1}\mathbf{0} = \mathbf{0}$ as $\mu \to \infty$. Claim (A20) then follows from Lemma 6.2. Q.E.D.

The proof of Theorem 6.1 now follows from Lemma 6.3. Specifically, by construction

$$u^*_n(\mu, N) = b^{\mathrm{FLS}}_n(\mu, N) - b^{\mathrm{OLS}}(N); \tag{A22}$$

hence, $u^*_n(\mu, N) \to \mathbf{0}$ as $\mu \to \infty$ if and only if (6.1) holds. Q.E.D.

Corollary 6.1

Suppose $X(N)$ has full rank K. Then, for each n, $1 \leq n \leq N$,

$$\lim_{\mu \to \infty} \mu[b^{\mathrm{FLS}}_{n+1}(\mu, N) - b^{\mathrm{FLS}}_n(\mu, N)] = \sum_{s=1}^{n} [x_s^{\mathrm{T}} b^{\mathrm{OLS}}(N) - y_s] x_s. \tag{A23}$$

Proof. Corollary 6.1 follows from Theorem 6.1, given the form (A15) for the first-order conditions satisfied by the FLS solution $b^{\mathrm{FLS}}(\mu, N)$. Q.E.D.

Proof of Theorem 6.2. As earlier shown, the FLS estimates satisfy the first-order conditions (A15b); i.e.

$$\mathbf{0} = \sum_{n=1}^{N} [x_n^{\mathrm{T}} b^{\mathrm{FLS}}_n(\mu, N) - y_n] x_n. \tag{A24}$$

Theorem 6.2 then follows immediately from (A24) and the analytical expression (A13) for the OLS solution $b^{\mathrm{OLS}}(\mathrm{N})$. Q.E.D.

Proof of Theorem 6.3. The proof that condition (6.3) implies condition (6.4) follows directly from the first-order conditions (A15) satisfied by the FLS solution and the first-order conditions (A12) satisfied by the OLS solution.

Conversely, condition (6.4) implies that the FLS cost function $V(u; \mu, N)$ defined in (A17) reduces to

$$V(u; \mu, N) = \sum_{n=1}^{N} [x_n^{\mathrm{T}} u_n]^2 + \mu \sum_{n=1}^{N-1} [u_{n+1} - u_n]^{\mathrm{T}}[u_{n+1} - u_n] + S(b^{\mathrm{OLS}}(N), N) \tag{A25}$$

for all $u = (u_1, \ldots, u_N)$. Thus, $V(u; \mu, N) \geq S(b^{\mathrm{OLS}}(N), N)$ for all u, with $V(\mathbf{0}; \mu, N) = S(b^{\mathrm{OLS}}(N), N)$. To establish that (6.4) implies (6.3), it thus suffices to show that $u^*(\mu, N) = \mathbf{0}$ is the unique minimizer of $V(u; \mu, N)$, given condition (6.4); for $u^*_n(\mu, N) = [b^{\mathrm{FLS}}_n(\mu, N) - b^{\mathrm{OLS}}(N)]$ by construction.

Suppose there exists a nonzero $\hat{u}$ such that $V(\hat{u}; \mu, N) = S(b^{\mathrm{OLS}}(N), N)$. Then, by (A25), it must hold that

$$[\hat{u}_{n+1} - \hat{u}_n] = \mathbf{0}, \quad 1 \leq n \leq N-1, \tag{A26}$$

hence

$$\hat{u}_n = \hat{u}_1 \neq \mathbf{0}, \quad 1 \leq n \leq N. \tag{A27}$$

Again using (A25), it follows that

$$V(\hat{u}; \mu, N) = \hat{u}_1^{\mathrm{T}}\left[\sum_{n=1}^{N} x_n x_n^{\mathrm{T}}\right]\hat{u}_1 + S(b^{\mathrm{OLS}}(N), N) = \hat{u}_1^{\mathrm{T}} X(N)^{\mathrm{T}} X(N) \hat{u}_1 + S(b^{\mathrm{OLS}}(N), N), \tag{A28}$$

with $\hat{u}_1 \neq \mathbf{0}$. Thus, in order to have $V(\hat{u}; \mu, N) = S(b^{\mathrm{OLS}}(N), N)$, it must hold that $X(N)^{\mathrm{T}} X(N)$ is singular. However, $X(N)$ has full rank K by assumption. It follows that no such nonzero $\hat{u}$ exists. Q.E.D.

A.3. Proofs for Section 7

The following preliminaries are needed for the proof of Theorem 7.1. Let $\mu > 0$ be given. For the particular observation sequence (7.1), the FLS incompatibility cost function (3.4) reduces to

$$C(b_1, \ldots, b_N; \mu, N) = \sum_{n=1}^{S} (x_n^{\mathrm{T}}[z - b_n])^2 + \sum_{n=S+1}^{N} (x_n^{\mathrm{T}}[w - b_n])^2 + \mu \sum_{n=1}^{N} [b_{n+1} - b_n]^{\mathrm{T}}[b_{n+1} - b_n]. \tag{A29}$$

The first-order conditions for a vector $(b_1, \ldots, b_N)$ to minimize $C(b_1, \ldots, b_N; \mu, N)$ in (A29) take the following form: for $n = 1$:

$$\mu[b_2 - b_1] = x_1 x_1^{\mathrm{T}}[b_1 - z]; \tag{A30a}$$

for $1 < n \leqslant S$:

$$\mu[b_{n+1} - b_n] = \mu[b_n - b_{n-1}] + x_n x_n^{\mathrm{T}}[b_n - z]; \tag{A30b}$$

for $S + 1 \leqslant n < N$:

$$\mu[b_{n+1} - b_n] = \mu[b_n - b_{n-1}] + x_n x_n^{\mathrm{T}}[b_n - w]; \tag{A30c}$$

for $n = N$:

$$\mu[b_N - b_{N-1}] = x_N x_M^{\mathrm{T}}[w - b_N]. \tag{A30d}$$

Combining terms, the first-order conditions (A30) take the form: for $1 \leqslant n \leqslant S$:

$$\mu[b_{n+1} - b_n] = \sum_{k=1}^{n} x_k x_k^{\mathrm{T}}[b_k - z]; \tag{A31a}$$

for $S + 1 \leqslant n < N$:

$$\mu[b_{n+1} - b_n] = \sum_{k=n+1}^{N} x_k x_k^{\mathrm{T}}[w - b_k]; \tag{A31b}$$

for $n = N$:

$$0 = \sum_{k=1}^{S} x_k x_k^{\mathrm{T}}[b_k - z] + \sum_{k=S+1}^{N} x_k x_k^{\mathrm{T}}[b_k - w]. \tag{A31c}$$

If $X(N)$ has full rank K, it follows from Section 4 that the solution to conditions (A30) [equivalently, (A31)] yields the unique FLS solution corresponding to the particular observation sequence (7.1).

Proof of Theorem 7.1

The four properties listed in Theorem 7.1 will be proved in order.

Proof of properties (i) and (ii). Suppose $z < b_1$ and $b_N < w$. Then properties (i) and (ii) follow directly from (A30) and (A31), respectively. Suppose $b_1 \leqslant z$. Then it follows directly from (A31a) that $b_n \leqslant z$ for $1 \leqslant n \leqslant S + 1$. In order for (A31c) to hold, it must then be true that $b_N \geqslant w$, implying $b_n \geqslant w$ for $S + 1 \leqslant n \leqslant N$ by (A31b). However, one then obtains $b_{S+1} \leqslant z < w \leqslant b_{S+1}$, a contradiction. A similar contradiction is obtained if one supposes that $b_N \geqslant w$.

Proof of property (iii). Recall definition (7.2) for the matrices $L_n(\mu)$. In the special case $K = 1$, with $x_n \neq 0$, $1 \leqslant n \leqslant N$, each $L_n(\mu)$ is a scalar lying strictly between zero and one. Moreover, $L_n(\mu) \to 1$ as $\mu \to \infty$ and $L_n(\mu) \to 0$ as $\mu \to 0$. The first-order necessary conditions (A30) can be expressed in terms of the matrices $L_n(\mu)$ as follows: for $n = 1$

$$[b_1 - z] = L_1(\mu)[b_2 - z]; \tag{A32a}$$

for $1 < n \leqslant S$:

$$[b_n - z] = L_n(\mu)([b_{n+1} - z] + [b_{n-1} - z])/2; \tag{A32b}$$

for $S + 1 \leqslant n < N$:

$$[b_n - w] = L_n(\mu)([b_{n+1} - w] + [b_{n-1} - w])/2; \tag{A32c}$$

for $n = N$:

$$[b_N - w] = L_N(\mu)[b_{N-1} - w]. \tag{A32d}$$

By (A32b) and property (i), for $n = S$ one has

$$b_S = [1 - L_S(\mu)]z + L_S(\mu)[b_{S+1} + b_{S-1}]/2 < [1 - L_S(\mu)]z + L_S(\mu)w. \tag{A33}$$

If $S = 1$, this completes the proof of part (a) of property (iii). Suppose $S > 1$ and, for some n satisfying $1 < n \leqslant S$, one has shown that

$$b_n < \left[1 - \prod_{k=n}^{S} L_k(\mu)\right]z + \left[\prod_{k=n}^{S} L_k(\mu)\right]w. \tag{A34}$$

Combining property (i), (A32) and the induction step (A34),

$$b_{n-1} < [1 - L_{n-1}(\mu)]z + L_{n-1}(\mu)b_n < \left[1 - \prod_{k=n-1}^{S} L_k(\mu)\right]z + \left[\prod_{k=n-1}^{S} L_k(\mu)\right]w. \tag{A35}$$

Thus, part (a) of property (iii) holds by induction for all n, $1 \leqslant n \leqslant S$.

The proof of part (b) of property (iii) is entirely analogous.

Proof of property (iv). By property (i), the solution vectors $b_1, \ldots, b_N$ for the first-order conditions (A31) are bounded between z and w for all $N \geqslant 1$. It follows that the right-hand sum in (A31c) is bounded below by a finite negative number as $N \to \infty$. By property (i), the right-hand sum in (A31c) is a monotone decreasing function of N. A bounded monotone decreasing sequence must converge to a finite limit. A necessary condition for the right-hand sum in (A31c) to converge to a finite limit as $N \to \infty$ is $x_M^{\mathrm{T}}[b_N - w] \to 0$ as $N \to \infty$. Q.E.D.

Proof of Theorem 7.2

Proof of property (i). It follows immediately from the first-order necessary condition (A32a) and the definition (7.3) for $L_1(\mu)$ that the FLS solution satisfies

$$[b_2 - z]^{\mathrm{T}}[b_2 - z] = [b_1 - z]^{\mathrm{T}}[I + V_1(\mu)][I + V_1(\mu)][b_1 - z], \tag{A36}$$

where the $K \times K$ positive semidefinite matrix $V_1(\mu)$ is given by

$$V_1(\mu) = x_1 x_1^T/\mu. \tag{A37}$$

If $S = 1$, this completes the proof of property (i).

Suppose $S \geqslant 2$, and suppose for some $n - 1$ satisfying $1 \leqslant n - 1 < S$ it has been shown that the FLS solution satisfies

$$[b_n - z]^T[b_n - z] = [b_{n-1} - z]^T[I + V_{n-1}(\mu)][I + V_{n-1}(\mu)][b_{n-1} - z], \tag{A38}$$

where $V_{n-1}(\mu)$ is a $K \times K$ positive semidefinite matrix. Then there must exist a scalar $k_n(\mu) \geqslant 1$, and a symmetric orthogonal $K \times K$ matrix [reflection] of the form $P_n(\mu) = [I - 2u_n(\mu)u_n(\mu)^T]$, where $u_n(\mu)^T u_n(\mu) = 1$, such that

$$[b_n - z] = k_n(\mu)P_n(\mu)[b_{n-1} - z]. \tag{A39}$$

If strict inequality holds in (A38), then $k_n(\mu) > 1$.

Let $R_n(\mu)$ denote the inverse of the matrix $L_n(\mu)$ defined by (7.2). Thus, $R_n(\mu) = A_n(\mu)/2\mu = [x_n x_n^T + 2\mu I]/2\mu$. By (A32b),

$$2R_n(\mu)[b_n - z] = [b_{n+1} - z] + [b_{n-1} - z]. \tag{A40}$$

Combining (A39) and (A40), and noting that $P_n(\mu)^{-1} = P_n(\mu)$,

$$\begin{aligned}[b_{n+1} - z]^T[b_{n+1} - z] &= [b_n - z]^T 2R_n(\mu)2R_n(\mu)[b_n - z] - 2[b_{n-1} - z]^T 2R_n(\mu)[b_n - z] + [b_{n-1} - z]^T[b_{n-1} - z] \\ &= [b_n - z]^T[I + V_n(\mu)][I + V_n(\mu)][b_n - z],\end{aligned} \tag{A41}$$

where the positive semidefinite $K \times K$ matrix $V_n(\mu)$ satisfies

$$I + V_n(\mu) = I + [1 - k_n(\mu)^{-1}]I + 2k_n(\mu)^{-1}u_n(\mu)u_n(\mu)^T + x_n x_n^T/\mu = [2R_n(\mu) - k_n(\mu)^{-1}P_n(\mu)]. \tag{A42}$$

Note that $V_n(\mu)$ is positive definite if $k_n(\mu) > 1$. It follows that

$$[b_{n+1} - z]^T[b_{n+1} - z] \geqslant [b_n - z]^T[b_n - z], \tag{A43}$$

with strict inequality holding if $k_n(\mu) > 1$. Hence, by induction, property (i) holds for $1 \leqslant n \leqslant S$.

Proof of property (ii). The proof of property (ii) is entirely analogous. Q.E.D.

APPENDIX B

A FORTRAN Program for Finding the FLS Solution

A FORTRAN program "FLS" has been developed which implements the sequential FLS solution procedure developed in Section 5. FLS consists of a main program together with four subroutines: INPUT, WOOD, INV, and FOCTST. The main program and subroutines are currently dimensioned for regressor vectors with dimension $K \leqslant 10$, and for a number of observations $N \leqslant 110$.

The main program begins with a call to subroutine INPUT, which provides all the needed inputs to the program. Subroutine INPUT is the only part of the program requiring actions on the part of the user, aside from the write and format statements which appear in the main program and the dimension statements which appear at the beginning of each subroutine and the main program. (These write, format, and dimension statements should be tailored to conform to the specific dimensions of the user's problem.)

Specifically, subroutine INPUT initializes the penalty weight μ, the dimension K of the regressor vectors, and the number of observations N. It also fills a double precision array $X(10,110)$ with the $K \times N$ matrix of regressor values $[x_1, \ldots, x_N]$, and a double precision array $Y(110)$ with the $N \times 1$ vector of observations $(y_1, \ldots, y_N)^T$. For simulation studies, the observations $(y_1, \ldots, y_N)^T$ are generated in accordance with the linear regression model $y_n = x_n^T b_n + v_n$, $n = 1, \ldots, N$, for a $K \times K$ matrix of true coefficient values $[b_1, \ldots, b_N]$ and a specified sequence $(v_1, \ldots, v_N)$ of residual measurement errors. Subroutine INPUT stores the true coefficient values in a double precision array $TRUEB(10, 110)$ for later comparison with the numerically generated FLS coefficient estimates.

The main program next initializes a certain auxiliary array R. A DO loop for $n = 1, N$ then commences. The DO loop evaluates and stores the matrices M_n and vectors e_n in equations (5.7b) and (5.7c). The inversion required for the evaluation of M_n is accomplished in part by a call to subroutine WOOD, which implements the well-known Woodbury matrix inversion lemma. Subroutine WOOD in turn calls the matrix-inversion subroutine INV.

The main program next evaluates the FLS filter estimate $b_N^{FLS}(\mu, N)$ for the final coefficient vector b_N, using equation (5.15), and stores this $K \times 1$ vector in column N of a double precision array $B(10, 110)$. The FLS smoothed estimates for the $K \times 1$ coefficient vectors $b_1, \ldots, b_{N-1}$ are then determined in accordance with equations (5.16), and stored in columns 1 through $N - 1$ of the array $B(10, 110)$. The entire array B of FLS coefficient estimates for time 1 through N is printed out. For simulation studies, the array $TRUEB$ of true coefficient values for times 1 through N is also printed out for comparison with B.

Using the array B of FLS coefficient estimates, the main program then evaluates and prints out the sum of squared residual measurement errors (3.1), the sum of squared residual dynamic errors (3.2), and the total incompatibility cost (3.4).

The next portion of the main program consists of a validation check. The K-dimensional OLS solution $b^{OLS}(N)$ for the linear regression model (2.1a) is first evaluated as a matrix-weighted average (6.2) of the FLS estimates. This evaluation is stored in an array $BOLSE(10)$. The OLS solution $b^{OLS}(N)$ is then evaluated by means of the usual formula (A13). This evaluation is stored in an array $BOLS(10)$. Theoretically, the two expressions (6.2) and (A13) for $b^{OLS}(N)$ are equivalent. Thus, if the program is correct, the two evaluations should be in close agreement. Both of these evaluations are printed out.

The final portion of the main program consists of a second validation check. A call is made to subroutine FOCTST to determine how well the numerically generated FLS coefficient estimates satisfy the first-order necessary (and sufficient) conditions (A14) for minimization of the incompatibility cost function (3.4). Using the numerically generated FLS coefficient estimates stored in B, together with the inputs provided by subroutine INPUT, subroutine FOCTST evaluates the right-hand expression for each of the first-order conditions (A14) and prints out the resulting calculation.

A third validation check can be undertaken by letting μ increase over successive runs. As established in Section 6, Theorem 6.1, the FLS estimate $b_n^{FLS}(\mu, N)$ for the coefficient vector b_n converges to the OLS solution $b^{OLS}(N)$ as the penalty weight μ approaches infinity, for each $n = 1, \ldots, N$. Thus, the numerically generated FLS estimates should approach the numerically generated OLS solution $b^{OLS}(N)$ for large μ values.

The FORTRAN statements for program FLS are listed below. The logical progression of the program statements, explained in the preceding paragraphs, is summarized in comment statements interspersed throughout the program. Print-out is given for one of the ellipse experiments discussed in Section 8.3.

```
 C
 C       FLS:  A FORTRAN PROGRAM FOR TIME-VARYING LINEAR REGRESSION
 C             VIA FLEXIBLE LEAST SQUARES
 C
 C       MAIN PROGRAM
 C
         IMPLICIT REAL*8(A-H,O-Z)
         REAL*8M
 C       THIS PROGRAM IS CURRENTLY DIMENSIONED FOR A MAXIMUM OF
 C       110 OBSERVATIONS, WITH REGRESSOR VECTORS OF A MAXIMUM
 C       DIMENSION 10.
         DIMENSION X(10,110),Y(110),R(10,10)
         DIMENSION XN(10),Z(10,10),M(10,10,110),V(10)
         DIMENSION U(10),E(10,110),W(10),Q(10,10),A(10,10)
         DIMENSION C(10,10),D(10),B(10,110),DIFVEC(10)
         DIMENSION VVV(10),ZZ(10,10),ZZINV(10,10),BOLSE(10)
         DIMENSION BOLS(10),XY(10),XXT(10,10),XXTINV(10,10)
         DIMENSION XTBOLS(110),OLSR(110)
         DIMENSION TRUEB(10,110)
 C        THE FOLLOWING SUBROUTINE INITIALIZES THE PENALTY WEIGHT
 C        AMU, THE DIMENSION K OF THE REGRESSOR VECTORS, AND
 C        THE NUMBER OF OBSERVATIONS NCAP.  IT ALSO GENERATES
 C        THE K BY NCAP MATRIX X OF REGRESSOR VALUES AND THE
 C        NCAP BY 1 VECTOR Y OF SCALAR OBSERVATIONS.
          CALL INPUT(AMU,K,NCAP,X,Y,TRUEB)
 C        INITIALIZATION FOR THE AUXILIARY MATRIX RN = QN-1 + AMU*I
 C        AT TIME N = 1
          DO 30 I=1,K
          DO 40 J=1,K
          R(I,J) = 0.0D+00
          IF(I.EQ.J) R(I,J) = AMU
      40  CONTINUE
      30  CONTINUE
          DO 50 N=1,NCAP
 C        FORM THE REGRESSOR VECTOR XN AND THE SCALAR
 C        OBSERVATION YN
          DO 60 I=1,K
          XN(I)=X(I,N)
      60  CONTINUE
          YN=Y(N)
 C        CALCULATE THE INVERSE Z OF THE MATRIX RN+XNXNT VIA
 C        THE WOODBURY MATRIX INVERSION LEMMA
          CALL WOOD(K,R,XN,Z)
 C        CALCULATE & STORE THE K BY K MATRICES MN AND THE K BY 1
 C        VECTORS EN APPEARING IN EQUATIONS (5.7B) AND (5.7C)
          DO 70 I=1,K
          DO 80 J=1,K
          M(I,J,N)=AMU*Z(I,J)
      80  CONTINUE
      70  CONTINUE
          DO 90 I=1,K
          V(I)=XN(I)*YN
      90  CONTINUE
          DO 100 I=1,K
          IF(N.EQ.1) U(I)=0.0D+00
          IF(N.GT.1) U(I)=AMU*E(I,N-1)
     100  CONTINUE
          DO 110 I=1,K
          W(I)=U(I)+V(I)
```

```
  110 CONTINUE
      DO 120 I=1,K
      E(I,N)=0.0D+00
      DO 130 J=1,K
      E(I,N)=E(I,N)+Z(I,J)*W(J)
  130 CONTINUE
  120 CONTINUE
      DO 140 I=1,K
      DO 150 J=1,K
      R(I,J)=-AMU*AMU*Z(I,J)
      IF(I.EQ.J)R(I,J)=(2.0D+00)*AMU+R(I,J)
  150 CONTINUE
  140 CONTINUE
   50 CONTINUE
C     CALCULATE THE (5.15) FLSV ESTIMATE BN FOR THE
C     FINAL TIME N = NCAP
      DO 160 I=1,K
      DO 170 J=1,K
      Q(I,J)=-AMU*M(I,J,NCAP-1)
      IF(I.EQ.J) Q(I,J)=Q(I,J)+AMU
  170 CONTINUE
  160 CONTINUE
C     OBTAIN THE INVERSE C OF THE MATRIX A=(QN-1 + XNXNT)
C     IN EQUATION (5.15)
      DO 180 I=1,K
      DO 190 J=1,K
      A(I,J)=Q(I,J)+X(I,NCAP)*X(J,NCAP)
  190 CONTINUE
  180 CONTINUE
      CALL INV(K,A,C)
C     FORM THE VECTOR D=(PN-1 + XNYN) IN EQUATION (5.15)
      DO 200 I=1,K
      D(I)=AMU*E(I,NCAP-1)+X(I,NCAP)*Y(NCAP)
  200 CONTINUE
C     POSTMULTIPLY C BY D TO OBTAIN BNCAP
      DO 210 I=1,K
      B(I,NCAP)=0.0D+00
      DO 220 J=1,K
      B(I,NCAP)=B(I,NCAP)+C(I,J)*D(J)
  220 CONTINUE
  210 CONTINUE
C     USE EQUATIONS (5.16) TO OBTAIN SMOOTHED FLS ESTIMATES
C     FOR B1,...,BNCAP-1
      NCAP1=NCAP-1
      DO 230 N=1,NCAP1
      L=NCAP-N
      DO 240 I=1,K
      B(I,L)=E(I,L)
      DO 250 J=1,K
      B(I,L)=B(I,L)+M(I,J,L)*B(J,L+1)
  250 CONTINUE
  240 CONTINUE
  230 CONTINUE
      WRITE(6,2020)
 2020 FORMAT(1X,'HERE ARE THE FLS ESTIMATES FOR B1 AND THE
     & TRUE B1')
      WRITE(6,37)(B(1,N),TRUEB(1,N),N=1,NCAP)
      WRITE(6,2030)
 2030 FORMAT(1X,'HERE ARE THE FLS ESTIMATES FOR B2 AND THE
     & TRUE B2')
      WRITE(6,37)(B(2,N),TRUEB(2,N),N=1,NCAP)
   37 FORMAT(1X,2D25.10)
C     CALCULATING RSUBM FROM EQUATION (3.1)
      SUM =0.0D+00
      DO 500 N=1,NCAP
      SUM1=0.0D+00
      DO 510 J=1,K
      SUM1=SUM1+X(J,N)*B(J,N)
  510 CONTINUE
      DIF=Y(N)-SUM1
      DIFSQ=DIF*DIF
      SUM=SUM+DIFSQ
  500 CONTINUE
      RSUBM=SUM
```

```
C         CALCULATING RSUBD FROM EQUATION (3.2)
          SUM=0.0D+00
          DO 520 N=1,NCAP1
          DO 530 J=1,K
          DIFVEC(J)=B(J,N+1)-B(J,N)
    530   CONTINUE
          SUM1=0.0D+00
          DO 540 J=1,K
          SUM1=SUM1+DIFVEC(J)*DIFVEC(J)
    540   CONTINUE
          SUM=SUM+SUM1
    520   CONTINUE
          RSUBD=SUM
C         CALCULATING THE INCOMPATIBILITY COST AMU*RSUBD + RSUBM
          COST=AMU*RSUBD+RSUBM
          WRITE(6,580) RSUBM,RSUBD,COST
    580   FORMAT(1HO,'HERE ARE RSUBM,RSUBD,COST'/1X,3D20.10)
C         FIRST VALIDATION CHECK:
C         CALCULATION OF THE ESTIMATE BOLSE FOR OLS FROM THE MATRIX
C         AVERAGE OF THE FLS ESTIMATES GIVEN BY EQUATION (6.2)
          DO 810 I=1,K
          VVV(I)=0.0D+00
          DO 820 N=1,NCAP
          SUM1=0.0D+00
          DO 830 J=1,K
    830   SUM1=SUM1+X(J,N)*B(J,N)
    820   VVV(I)=VVV(I)+X(I,N)*SUM1
    810   CONTINUE
          DO 840  J=1,K
          DO 850 I=1,K
    850   ZZ(I,J)=0.0D+00
    840   CONTINUE
          DO 860 N=1,NCAP
          DO 870 I=1,K
          DO 880 J=1,K
    880   ZZ(I,J)=X(I,N)*X(J,N)+ZZ(I,J)
    870   CONTINUE
    860   CONTINUE
          CALL INV(K,ZZ,ZZINV)
          DO 890 I=1,K
          BOLSE(I)=0.0D+00
          DO 900 J=1,K
    900   BOLSE(I)=BOLSE(I)+ZZINV(I,J)*VVV(J)
    890   CONTINUE

          WRITE(6,910)
    910   FORMAT(1HO,'COMPONENTS OF BOLSE')
          WRITE(6,920) (BOLSE(I),I=1,K)
    920   FORMAT(1X,D35.10)
C         CALCULATING THE OLS ESTIMATE BOLS FROM THE USUAL
C         FORMULA BOLS=(XXT)-1*XY
          DO 1600 I=1,K
          SUM=0.0D+00
          DO 1610 N=1,NCAP
   1610   SUM=SUM+X(I,N)*Y(N)
   1600   XY(I)=SUM
          DO 1620 I=1,K
          DO 1630 J=1,K
          XXT(I,J)=0.0D+00
          DO 1640 N=1,NCAP
   1640   XXT(I,J)=XXT(I,J)+X(I,N)*X(J,N)
   1630   CONTINUE
   1620   CONTINUE
          CALL INV(K,XXT,XXTINV)
          DO 1650 I=1,K
          BOLS(I)=0.0D+00
          DO 1660 J=1,K
   1660   BOLS(I)=BOLS(I)+XXTINV(I,J)*XY(J)
   1650   CONTINUE
          WRITE(6,1670)
   1670   FORMAT(1HO,'COMPONENTS OF BOLS')
          WRITE(6,1671) (BOLS(I),I=1,K)
   1671   FORMAT(1X,D35.10)
```

```
C           CALCULATING THE SUM OF SQUARED RESIDUAL MEASUREMENT
C           ERRORS OLSRM FOR OLS
            DO 1800 N=1,NCAP
            XTBOLS(N)=0.0D+00
            DO 1810 I=1,K
      1810  XTBOLS(N)=XTBOLS(N)+X(I,N)*BOLS(I)
      1800  CONTINUE
            DO 1820 N=1,NCAP
      1820  OLSR(N)=Y(N)-XTBOLS(N)
            OLSRM=0.0D+00
            DO 1830 N=1,NCAP
      1830  OLSRM=OLSRM+OLSR(N)*OLSR(N)
            WRITE(6,1840)
      1840  FORMAT(1H0,'SUM OF SQUARED RESIDUAL MEASUREMENT ERRORS
         &  FOR OLS')
            WRITE(6,1841) OLSRM
      1841  FORMAT(1X,D27.10)
C           SECOND VALIDATION CHECK: FIRST-ORDER CONDITION TEST
C           HOW WELL DO THE FLS ESTS SATISFY THE FOC CONDITIONS (A.14)
            CALL FOCTST(AMU,K,NCAP,X,Y,B)
            STOP
            END
```

```
C
            SUBROUTINE INPUT(AMU,K,NCAP,X,Y,TRUEB)
            IMPLICIT REAL*8(A-H,O-Z)
            DIMENSION X(10,110),Y(110),TRUEB(10,110)
C           RUN FOR ELLIPTICAL TRUE B
C           WITH NORMAL NOISE N(0,SIGMA) IN THE OBSERVATIONS
            K=2
            AMU=1.0D+00
            NCAP=30
            SIGMA=0.00D+00
            DO 3030 I=1,NCAP
            AI=DFLOAT(I)
            PI=(DATAN(1.0D+00))*4.0D+00
            TRUEB(1,I)=.5D+00*DSIN((2.0D+00*PI/30.0D+00)*AI)
            TRUEB(2,I)=DCOS((2.0D+00*PI/30.0D+00)*AI)
      3030  CONTINUE
            X(1,1)=1.0D+00
            X(2,1)=1.0D+00
            DO 3010 I=2,NCAP
            AI=DFLOAT(I)
            X(1,I)=DSIN(10.0D+00+(AI))+.01D+00
            X(2,I)=DCOS(10.0D+00+(AI))
      3010  CONTINUE
      4020  CONTINUE
            DO 3020 I=1,NCAP
            Y(I)=X(1,I)*TRUEB(1,I)+X(2,I)*TRUEB(2,I)+SIGMA*GNORM(0)
      3020  CONTINUE
            RETURN
            END
```

```
C
      SUBROUTINE WOOD(K,R,X,Z)
C     CALCULATES THE INVERSE Z OF A MATRIX OF THE FORM R+XXT
C     VIA THE WOODBURY MATRIX INVERSION LEMMA
      IMPLICIT REAL*8(A-H,O-Z)
      DIMENSION R(10,10),X(10),Z(10,10),S(10,10),V(10)
      DIMENSION XNUM(10,10),U(10,10)
C     CALCULATE THE INVERSE S OF THE K BY K MATRIX R
      CALL INV(K,R,S)
C     CALCULATE THE K BY 1 VECTOR V=SX=R-1*X
      DO 10 I=1,K
      V(I)=0.0D+00
      DO 20 J=1,K
      V(I)=V(I)+S(I,J)*X(J)
   20 CONTINUE
   10 CONTINUE
C     CALCULATE XNUM=VVT=R-1*XXT*R-1
      DO 30 I=1,K
      DO 40 J=1,K
      XNUM(I,J)=V(I)*V(J)
   40 CONTINUE
   30 CONTINUE
C     CALCULATE Y=(1+VTV)=(1+XT*R-1*X)
      Y=1.0D+00
      DO 50 I=1,K
      Y=Y+X(I)*V(I)
   50 CONTINUE
C     CALCULATE U=XNUM/Y
      DO 60 I=1,K
      DO 70 J=1,K
      U(I,J)=XNUM(I,J)/Y
   70 CONTINUE
   60 CONTINUE
C     CALCULATE Z=S-U
      DO 80 I=1,K
      DO 90 J=1,K
      Z(I,J)=S(I,J)-U(I,J)
   90 CONTINUE
   80 CONTINUE
      RETURN
      END
```

```
C
      SUBROUTINE INV(K,A,C)
C     CALCULATES THE INVERSE C OF A K BY K MATRIX A
      IMPLICIT REAL*8(A-H,O-Z)
      DIMENSION A(10,10),C(10,10),B(10,20)
      DO 5 J=1,K
      DO 6 I=1,K
    6 B(I,J)=A(I,J)
    5 CONTINUE
      K2=K*2
      DO 7 J=1,K
      DO 8 I=1,K
      B(I,K+J)=0.0D+00
      IF(I.EQ.J) B(I,K+J)=1.0D+00
    8 CONTINUE
    7 CONTINUE
C     THE PIVOT OPERATION STARTS HERE
      DO 9 L=1,K
      PIVOT=B(L,L)
      DO 13 J=L,K2
   13 B(L,J)=B(L,J)/PIVOT
C     TO IMPROVE THE ROWS
      DO 14 I=1,K
      IF(I.EQ.L) GO TO 14
      AIL=B(I,L)
      DO 15 J=L,K2
   15 B(I,J)=B(I,J)-AIL*B(L,J)
   14 CONTINUE
    9 CONTINUE
      DO 45 I=1,K
      DO 46 J=1,K
   46 C(I,J)=B(I,K+J)
   45 CONTINUE
      RETURN
      END
```

```
C
      SUBROUTINE FOCTST(AMU,K,NCAP,X,Y,B)
      IMPLICIT REAL*8(A-H,O-Z)
      DIMENSION X(10,110),Y(110),B(10,110),DIF(10)
      WRITE(6,100)
  100 FORMAT(1H0,'HERE ARE THE FOC TEST RESULTS FOR EQUATIONS (A.14)')
      DO 1 N=1,NCAP
      IF (N.NE.1) GO TO 9000
      SUM=0.0D+00
      DO 2 J1=1,K
      SUM=SUM+X(J1,N)*B(J1,N)
    2 CONTINUE
      SUM=SUM-Y(N)
      DO 3 J1=1,K
      DIF(J1)=SUM*X(J1,N)-AMU*(B(J1,N+1)-B(J1,N))
    3 CONTINUE
      WRITE (6,200) (N,(DIF(J1),J1=1,K))
  200 FORMAT(1X,'FOR N EQUAL TO',I5/1X,6D12.3)
      GO TO 1
 9000 IF (N.EQ.NCAP) GO TO 9001
      SUM=0.0D+00
      DO 4 J1=1,K
      SUM=SUM+X(J1,N)*B(J1,N)
    4 CONTINUE
      SUM=SUM-Y(N)
      DO 5 J1=1,K
      DIF(J1)=SUM*X(J1,N)-AMU*(B(J1,N+1)-B(J1,N))
     &+AMU*(B(J1,N)-B(J1,N-1))
    5 CONTINUE
      WRITE(6,200)(N,(DIF(J1),J1=1,K))
      GO TO 1
 9001 SUM=0.0D+00
      DO 6 J1=1,K
      SUM=SUM+X(J1,N)*B(J1,N)
    6 CONTINUE
      SUM=SUM-Y(N)
      DO 7 J1=1,K
      DIF(J1)=SUM*X(J1,N)+AMU*(B(J1,N)-B(J1,N-1))
    7 CONTINUE
      WRITE(6,200)(N,(DIF(J1),J1=1,K))
    1 CONTINUE
      RETURN
      END
```

```
HERE ARE THE FLS ESTIMATES FOR B1 AND THE                TRUE B1
               0.2664583662D 00          0.1039558454D 00
               0.2694731181D 00          0.2033683215D 00
               0.3316402261D 00          0.2938926261D 00
               0.3699068596D 00          0.3715724127D 00
               0.3953642205D 00          0.4330127019D 00
               0.4326236760D 00          0.4755282581D 00
               0.4605030736D 00          0.4972609477D 00
               0.4529753669D 00          0.4972609477D 00
               0.4207365504D 00          0.4755282581D 00
               0.3914008494D 00          0.4330127019D 00
               0.3464992703D 00          0.3715724127D 00
               0.2607197346D 00          0.2938926261D 00
               0.1728667415D 00          0.2033683215D 00
               0.1061502112D 00          0.1039558454D 00
               0.9113728284D-02          0.1743934249D-15
              -0.1095330453D 00         -0.1039558454D 00
              -0.1769458630D 00         -0.2033683215D 00
              -0.2538268719D 00         -0.2938926261D 00
              -0.3441702216D 00         -0.3715724127D 00
              -0.3958057314D 00         -0.4330127019D 00
              -0.4203932681D 00         -0.4755282581D 00
              -0.4479204450D 00         -0.4972609477D 00
              -0.4596577973D 00         -0.4972609477D 00
              -0.4328820104D 00         -0.4755282581D 00
              -0.3843627705D 00         -0.4330127019D 00
              -0.3460810812D 00         -0.3715724127D 00
              -0.2931581383D 00         -0.2938926261D 00
              -0.2024471117D 00         -0.2033683215D 00
              -0.1367832817D 00         -0.1039558454D 00
              -0.1366870612D 00         -0.3487868498D-15
HERE ARE THE FLS ESTIMATES FOR B2 AND THE                TRUE B2
               0.8186598318D 00          0.9781476007D 00
               0.8216745887D 00          0.9135454576D 00
               0.7298853419D 00          0.8090169944D 00
               0.5876975259D 00          0.6691306064D 00
               0.4437492687D 00          0.5000000000D 00
               0.2862222235D 00          0.3090169944D 00
               0.9637134191D-01          0.1045284633D 00
              -0.1037199882D 00         -0.1045284633D 00
              -0.2817904824D 00         -0.3090169944D 00
              -0.4419070669D 00         -0.5000000000D 00
              -0.6080218997D 00         -0.6691306064D 00
              -0.7451037872D 00         -0.8090169944D 00
              -0.8182393711D 00         -0.9135454576D 00
              -0.8779069682D 00         -0.9781476007D 00
              -0.9203720936D 00         -0.1000000000D 01
              -0.8849817721D 00         -0.9781476007D 00
              -0.8133885130D 00         -0.9135454576D 00
              -0.7401421998D 00         -0.8090169944D 00
              -0.6139096520D 00         -0.6691306064D 00
              -0.4433775477D 00         -0.5000000000D 00
              -0.2771113441D 00         -0.3090169944D 00
              -0.1040209829D 00         -0.1045284633D 00
               0.9253143197D-01          0.1045284633D 00
               0.2885775361D 00          0.3090169944D 00
               0.4503972622D 00          0.5000000000D 00
               0.5900936547D 00          0.6691306064D 00
               0.7316989686D 00          0.8090169944D 00
               0.8276481285D 00          0.9135454576D 00
               0.8455177477D 00          0.9781476007D 00
               0.8454327629D 00          0.1000000000D 01

HERE ARE RSUBM,RSUBD,COST
   0.6572307833D-01    0.6281822245D 00    0.6849053009D 00

COMPONENTS OF BOLSE
                         0.3846260631D-01
                         0.3743910190D-01

COMPONENTS OF BOLS
                         0.3846260631D-01
                         0.3743910190D-01

SUM OF SQUARED RESIDUAL MEASUREMENT ERRORS                FOR OLS
          0.9969611789D 01
```

```
HERE ARE THE FOC TEST RESULTS FOR EQUATIONS (A.14)
FOR N EQUAL TO    1
  -0.111D-15  -0.111D-15
FOR N EQUAL TO    2
   0.130D-16   0.153D-15
FOR N EQUAL TO    3
  -0.101D-15  -0.125D-15
FOR N EQUAL TO    4
  -0.200D-15   0.111D-15
FOR N EQUAL TO    5
   0.226D-16   0.278D-16
FOR N EQUAL TO    6
   0.954D-17  -0.139D-16
FOR N EQUAL TO    7
  -0.572D-16   0.0
FOR N EQUAL TO    8
  -0.633D-16   0.416D-16
FOR N EQUAL TO    9
  -0.954D-17  -0.278D-16
FOR N EQUAL TO   10
  -0.859D-16  -0.971D-16
FOR N EQUAL TO   11
  -0.416D-16  -0.139D-15
FOR N EQUAL TO   12
  -0.278D-16  -0.555D-16
FOR N EQUAL TO   13
  -0.555D-16  -0.180D-15
FOR N EQUAL TO   14
  -0.416D-16  -0.720D-16
FOR N EQUAL TO   15
   0.139D-16   0.153D-15
FOR N EQUAL TO   16
  -0.555D-16  -0.200D-15
FOR N EQUAL TO   17
   0.139D-16  -0.111D-15
FOR N EQUAL TO   18
   0.278D-16   0.278D-16
FOR N EQUAL TO   19
   0.139D-16  -0.167D-15
FOR N EQUAL TO   20
   0.231D-15   0.555D-16
FOR N EQUAL TO   21
  -0.156D-16   0.278D-16
FOR N EQUAL TO   22
  -0.269D-16   0.416D-16
FOR N EQUAL TO   23
   0.169D-15   0.416D-16
FOR N EQUAL TO   24
  -0.694D-17   0.0
FOR N EQUAL TO   25
  -0.226D-16   0.139D-16
FOR N EQUAL TO   26
   0.104D-15   0.167D-15
FOR N EQUAL TO   27
  -0.155D-15   0.208D-15
FOR N EQUAL TO   28
   0.139D-16  -0.694D-16
FOR N EQUAL TO   29
   0.555D-16   0.807D-16
FOR N EQUAL TO   30
   0.111D-15  -0.232D-15
```