Volume Title: Annals of Economic and Social Measurement, Volume 2, number 4

Volume Author/Editor: Sanford V. Berg, editor

Volume Publisher:

Volume URL: http://www.nber.org/books/aesm73-4

Publication Date: October 1973

Chapter Title: A Test for Systematic Variation in Regression Coefficients

Chapter Author: David A. Belsley

Chapter URL: http://www.nber.org/chapters/c9940

Chapter pages in book: (p. 492 - 500)

Annals of Economic and Social Measurement, 2/4, 1973

A TEST FOR SYSTEMATIC VARIATION IN REGRESSION COEFFICIENTS

BY DAVID A. BELSLEY*

This paper offers a statistical test of the constancy of the parameters of a linear regression. The F test is based on transformed residuals which result from OLS applied to the given equation under the null hypothesis of constancy.

SOME NOTATION

We consider the model

(1) $$y(t) = x'(t)\beta(t) + \varepsilon(t)$$

$$\beta(t) = \Gamma z(t) + u(t)$$

where

$x(t)$, $z(t)$ K and R vectors, respectively,
$\varepsilon(t)$ spherically distributed with $E\varepsilon\varepsilon' = \sigma^2 I$,
$u(t)$ independent over time with $Euu' = \sigma_u^2\Omega$.

(See preceding article for motivation.

In what follows we consider the special case $\sigma_u^2 = 0$, i.e., variation in $\beta(t)$ is systematic and non random. Hence, we may write

(2) $$y(t) = x'(t)\Gamma z(t) + \varepsilon(t) \qquad \Gamma = [\gamma_1 \dots \gamma_R]$$

$$= [x'(t) \otimes z'(t)]\Lambda + \varepsilon(t)$$

where

$$\Lambda = \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_R \end{bmatrix}.$$

Let

$$Y = [y(t)], \qquad X = \underset{T \times K}{\begin{bmatrix} x'(1) \\ \vdots \\ x'(T) \end{bmatrix}}, \qquad Z = \underset{T \times R}{\begin{bmatrix} z'(1) \\ \vdots \\ z'(T) \end{bmatrix}} \qquad D = \underset{T \times KR}{\begin{bmatrix} x'(1) \otimes z'(1) \\ \vdots \\ x'(T) \otimes z'(T) \end{bmatrix}}.$$

Then (2) becomes

(3) $$Y = D\Lambda + \varepsilon$$

* Research supported by National Science Foundation Grant GJ-1154x to the National Bureau of Economic Research, Inc. Research Report W0006. This report has not undergone the full critical review accorded the National Bureau's studies, including review by the Board of Directors.

and we note that we may write

$$D = [\mathcal{Z}_1 \mathcal{Z}_2 \ldots \mathcal{Z}_R][X \otimes I]. \tag{4}$$

where $\mathcal{Z}_r = \mathrm{diag}\, Z_r$ and Z_r is the rth column of Z.

Thus, (3) becomes

$$Y = \sum_{r=1}^{R} \mathcal{Z}_r X \gamma_r + \varepsilon. \tag{5}$$

Remarks

Our purpose here is to determine a test of the null hypothesis that $\beta(t) = \beta$, i.e., is constant, for all t. Clearly a regression could be run on (3) directly if the z's were known, but alternative modeling tests would be cumbersome given the size of $(D'D)^{-1}$ even for moderate K and R.

In what follows a two-step test is determined that looks to be efficient and does not require inversion of $D'D$. Alternative Z matrices may be compared with a minimum of computation. The first step is OLS of Y on X without regard to Z. The second step consists of regressing a transformed set of residuals from step one on the similarly transformed z's. H_0 may be tested with the results of the second regression.

Step One: OLS Y on X

First regress Y on X to get

$$\begin{aligned} b &= (X'X)^{-1}X'Y \\ &= (X'X)^{-1}X'D\Lambda + (X'X)^{-1}X'\varepsilon \\ &= (X'X)^{-1}X'\sum_r \mathcal{Z}_r X\gamma_r + (X'X)^{-1}X'\varepsilon \end{aligned} \tag{6}$$

and

$$\begin{aligned} e &\equiv Y - Xb = HY \qquad (H = I - X(X'X)^{-1}X') \\ &= H(D\Lambda + \varepsilon) \\ &= [H\mathcal{Z}_1 X \ldots H\mathcal{Z}_R X]\Lambda + H\varepsilon \\ &\equiv [V_1 \ldots V_R]\Lambda + H\varepsilon \\ &= \sum_{r=1}^{R} V_r\gamma_r + H\varepsilon \end{aligned}$$

where the V_r are the residual matrices from an auxiliary regression of $\mathcal{Z}_r X$ on X.

This regression need not be run in practice. The relevance of V_r is seen from

$$H\mathcal{Z}_r X_r = \mathcal{Z}_r X - X(X'X)^{-1}X'\mathcal{Z}_r X = \mathcal{Z}_r X - XB_r \equiv V_r,$$

where B_r is the set of regression coefficients from $\mathcal{Z}_r X = XB_r + V_r$.[1]

Thus we have

$$e = \Sigma V_r \gamma_r + H\varepsilon. \tag{8}$$

We recall that H is idempotent, has rank $T - K$, and hence there exists an orthogonal C such that $C'HC = \begin{bmatrix} I_{T-K} & 0 \\ 0 & 0 \end{bmatrix} \equiv G$. Further we note $HV_r = V_r$, $r = 1 \ldots R$ and $He = e$. Hence, we may write

$$C'HCC'e = C'HCC'\Sigma V_r \gamma_r + C'HCC'\varepsilon \tag{9}$$

or

$$GC'e = GC'\Sigma V_r \gamma_r + GC'\varepsilon$$

and, partitioning $C = [C_1 C_2]$ so that the first $T - K$ rows of (9) become

$$\begin{aligned} f \equiv C_1'e &= C_1' \sum_{r=1}^{R} V_r \gamma_r + C_1'\varepsilon \\ &= C_1' \sum_{r=2}^{R} \mathcal{Z}_r X \gamma_r + \eta.^2 \end{aligned} \tag{10}$$

This last inequality comes from noting that $V_r = H\mathcal{Z}_r X$, and hence $C'V_r = C'H\mathcal{Z}_r X = C'HCC'\mathcal{Z}_r X = GC'\mathcal{Z}_r X$, which implies $C_1'V_r = C_1'\mathcal{Z}_r X$. We have also let $C_1\varepsilon \equiv \eta$.

We also note that η is spherically distributed, since $E\eta = 0$, $V\eta = E\eta\eta' = EC_1'\varepsilon\varepsilon'C_1 = \sigma_\varepsilon^2 C_1'C_1 = \sigma_\varepsilon^2 I_{T-K}$, due to the orthogonality of C.

It is the transformed residuals $f = C_1'e$ that we make use of in step two. The transformation C_1' comes from finding an orthogonal set of eigenvectors of $H = I - X(X'X)^{-1}X'$, and hence f depends only on knowledge of X and Y and does not require knowledge of Z.

Step Two

It is clear from (10) that the residuals from step one depend in a very involved way on the interrelation of X and Z through the terms $\mathcal{Z}_r X$. However, under the null hypothesis $H_0: \beta(t) = \beta$, these terms disappear, and a simpler test is available.

Consider a mechanical regression of f on Z transformed by C_1' (which depends only on X):

$$f = C_1'Z\delta + \psi. \tag{11}$$

[1] In passing we note from (6) that

$$\begin{aligned} b &= \Sigma(X'X)^{-1}X'\mathcal{Z}_r X\gamma_r + (X'X)^{-1}X'\varepsilon \\ &= \Sigma B_r \gamma_r + (X'X)^{-1}X'\varepsilon. \end{aligned}$$

Hence, $Eb = \Sigma B_r \gamma_r$, a weighted sum of the γ_r, and $V(b) = \sigma^2(X'X)^{-1}$.

[2] This latter sum goes from $r = 2$ to R since, if Z_1 (the first col. of Z) is a column vector of all ones, then $\mathcal{Z}_1 = I$ and hence $V_1 \equiv \mathcal{Z}_1 X - XB_1 = X - XB_1$, the least squares residuals of the auxiliary equation $X = XB_1 + V_1$. These residuals must necessarily be zero, since $B_1 = I$ does the trick of minimizing the sum of squares. Hence, $C_1'V_1 = 0 = C_1'\mathcal{Z}_1 X = C_1'X$.

OLS gives

(12) $$d = (Z'C_1C_1'Z)^{-1}Z'C_1f \quad \text{and from (10)}$$
$$= (Z'C_1C_1'Z)^{-1}Z'C_1C_1'\Sigma \mathscr{Z}_r X\gamma_r + (Z'C_1C_1'Z)^{-1}Z'C_1C_1'\varepsilon$$
$$\equiv (Z'QZ)^{-1}Z'Q \sum_{r=2}^{R} \mathscr{Z}_r X\gamma_r + (Z'QZ)^{-1}Z'Q\varepsilon$$

where $Q \equiv C_1C_1'$.

Under the null hypothesis $H_0 : \beta(t) = \beta$, $\gamma_r = 0$ for $r = 2 \ldots R$, and hence the first term of (12) is 0. That is, under H_0:

(13) $$d = (Z'QZ)^{-1}Z'Q\varepsilon$$
$$= (Z'QZ)^{-1}Z'C_1f.$$

In addition, from (10) we have under H_0 that

(14) $$f = C_1'\varepsilon.$$

Further, we note for future reference that Q is idempotent—since $QQ = C_1C_1'C_1C' = C_1IC_1' = C_1C_1' = Q$—and of rank $T - K$.

Now consider the residuals g of this second step; using (13) and (14),

(15) $$g \equiv f - C_1'Zd$$
$$= C_1'\varepsilon - C_1'Z(Z'QZ)^{-1}Z'Q\varepsilon$$
$$= C_1'[I - Z(Z'QZ)^{-1}Z'Q]\varepsilon$$
$$\equiv N\varepsilon \quad \text{where we let } N = C_1'[I - Z(Z'QZ)^{-1}Z'Q].$$

Now

$$g'g = \varepsilon'N'N\varepsilon$$
$$= \varepsilon'[I - QZ(Z'QZ)^{-1}Z']C_1C_1'[I - Z(Z'QZ)^{-1}Z'Q]\varepsilon$$
$$= \varepsilon'[Q - QZ(Z'QZ)^{-1}Z'Q][Q - QZ(Z'QZ)^{-1}Z'Q]\varepsilon$$
(16) $$\equiv \varepsilon'MM\varepsilon \quad \text{where } M \equiv Q - QZ(Z'QZ)^{-1}Z'Q$$
$$\equiv \varepsilon M\varepsilon$$

since M is seen to be idempotent with $\rho(M) = \operatorname{tr} M = T - K - R$. And hence,

(17) $$g'g \leftrightarrow \sigma_\varepsilon^2 X^2_{T-K-R}.$$

From (13) we have

(18) $$\mathrm{d} = (Z'QZ)^{-1}Z'Q\varepsilon \equiv B\varepsilon$$

and

$$
\begin{aligned}
BM &= (Z'QZ)^{-1}Z'Q[Q - QZ(Z'QZ)^{-1}Z'Q] \\
&= (Z'QZ)^{-1}Z'Q - (Z'QZ)^{-1}Z'Q = 0.
\end{aligned}
$$

Hence, the linear form (18) is distributed independently of the quadratic form (17) and the usual tests of significance on d may take place. Under H_0: Ed = 0, and hence a t value for a specific d at $T - K - R$ degrees of freedom in excess of the test level rejects the null hypothesis.

Boston College, and
National Bureau of Economic Research

Annals of Economic and Social Measurement, 2/4, 1973

KALMAN FILTER MODELS
A BAYESIAN APPROACH TO ESTIMATION OF TIME-VARYING REGRESSION COEFFICIENTS

BY ALEXANDER H. SARRIS*

The origins of time-varying linear regression coefficients are discussed, and it is noted that time variation cannot be estimated unless some restrictions are placed on the infinite forms of possible time changes. For example, a Markov structure imposed a priori on the coefficients renders them estimable. The structure imposes an incompletely specified prior probability distribution on the coefficients. The prior becomes completely determined through fitting it to the data. Bayes' theorem is then used to derive an estimator of the parameters. Under the assumption of perfect prior fit, the Bayes estimator is unbiased, minimum variance, and orthogonal to the residuals. Under the assumption of incomplete prior fit, the optimality properties of the estimator hold asymptotically. Finally, the problem of identifying the best Markov structure that fits the parameters is examined, and a Bayesian solution is proposed. This last discussion indicates the limitations of any method that attempts to identify time-varying coefficients.

1. INTRODUCTION

Over the last two decades great effort has been spent by econometricians, statisticians and system theorists on the problem of system identification. This problem is concerned with construction of a model whose output is close in some sense to the observed data from the real system. The modeler is guided by experience, knowledge of the real thing he is trying to describe, and intuition in specifying some equations (dynamic or static) which he terms the "structure" of the model. The equations are usually specified to within a number of parameters or coefficients which must be estimated by fitting the equations to the available data. The unknown parameters are usually assumed *a priori* to be constant. Then the problem of system identification is reduced to one of constant parameter estimation. There is a wealth of methods for the solution of this problem. A good survey of the ones that have been developed by econometricians and statisticians can be found in Theil (1971), while Åström and Eykhoff (1971) have surveyed the methods that have been developed primarily in system theory.

There are several reasons for suspecting that the parameters of many models, constructed by both engineers and econometricians, are not constant but in fact time varying. In engineering the origins of parameter variation are usually not very hard to pinpoint. Component wear, metal fatigue or component failure are some very common reasons for parameter variations. The major objective of construction of engineering models is control and regulation of the real system

* Research supported by National Science Foundation Grant GJ-1154x to the National Bureau of Economic Research, Inc.; and by an M.I.T. endowed fellowship. This Research Report, W0013 has not undergone the full critical review accorded the National Bureau's studies, including review by the Board of Directors.

I wish to acknowledge most valuable discussions with, and feedback from, Dr. Paul W. Holland of the NBER Computer Research Center, and Dr. J. Phillip Cooper of the University of Chicago. I retain sole responsibility for errors and omissions.

modeled. Therefore, most of the research in that area has concentrated on devising ways to make the output of the model insensitive to parameter variations.

On the other hand the origins of time varying parameters in econometric models are not very easy to isolate. Suspicions that shocks in the economy lead to sometimes permanent changes in the parameters of econometric models, have been substantiated ever since it was noticed that models of the economy fitted with prewar data gave noticeably different parameters than when fitted with postwar data. However, if one examines the process of economic modeling he will see several other sources of parameter variation. I will mention four of the most common ones.

Many econometric equations are mis-specified in the sense that they exclude variables that could possibly be part of the equation. Consider an equation of the form

$$y_t = \sum_{i=1}^{k} \beta_i x_{it} + \sum_{j=1}^{s} \gamma_j z_{jt} + \varepsilon_t \tag{1}$$

where y is an endogenous variable and the x_i, z_j are the true explanatory variables. If the econometrician ignores the z_j and lumps them with the error term ε, then whenever the z_j's behave in a non-stationary fashion there will be time variations in the intercept of (1).

Nonlinearities also give rise to parameter variations. If, for instance, the true relation is:

$$y_t = \alpha_1 + \alpha_2 x_t + \alpha_3 x_t^2 + \varepsilon_t \tag{2}$$

and the analyst considers the linear relation

$$y_t = \beta_1 + \beta_{2t} x_t + \varepsilon_t \tag{3}$$

then

$$\frac{\partial y_t}{\partial x_t} = \beta_{2t} = \alpha_2 + 2\alpha_3 x_t \tag{4}$$

thus β_{2t} is not constant.

Finally proxy variables and aggregation are also sources of parameter variation. For a detailed exposition of the sources of parameter variation the reader is referred to Cooley (1971).

This paper is concerned with a Bayesian method of estimation of time varying parameters. In section 2 a survey of previous research is given. The problem posed here is described in section 3. In sections 4 through 6 the method proposed for parameter estimation is presented and the properties of the estimator analysed. Sections 7 and 8 consider some problems that arise in applying the estimation technique. In section 9 the question of identifiability of a particular Markov structure is taken up, and a Bayesian solution which is the only feasible one is proposed. The last section summarizes the results.

2. Previous Research on Estimation of Time Varying Parameters

The problem of estimation of time varying parameters has not received very much attention from econometricians. On the other hand system theorists have devoted many years of research to various aspects of it. The reasons for this apparent gap will become clearer later.

The model from this point on will be assumed to be linear in the parameters. The following three classes of non-constant parameters are distinguished

(a) Time varying but non-stochastic

(b) Random but stationary

(c) Random but not necessarily stationary.

The earliest time varying parameter in econometrics dealt with parameters that were piecewise constant (Quandt (1958. 1960)) namely in class (a). This work was continued later by McGee and Carleton (1970), Brown and Durbin (1971) and Belsley (1973) but is still far from solved.

The second class of varying coefficients mentioned above applies to many problems in econometrics and statistics. and especially to the analysis of cross-sectional data. The problem is usually posed in terms of a relation of the form

(5) $$y_t = \sum_{i=1}^{k} \beta_{it} x_{it} + \varepsilon_t$$

where at each period t the parameters β_{it} $(i = 1, \ldots, k)$ are a sample from a multivariate distribution with mean μ and covariance matrix Σ. The objective is usually to estimate μ and Σ. Work on this problem has been done by Rao (1965), Hildreth and Houck (1968), Burnett and Guthrie (1970). Swamy (1970). and Rosenberg (1972).

Under the third category mentioned above come the various sequential variation models of the form

(6) $$\beta_{t+1} = T\beta_t + u_t.$$

This model is very common in the engineering literature and can be utilized to represent a wide variety of sample paths. In the econometrics literature to my knowledge only Rosenberg (1967, 1968a, b) has dealt extensively with this kind of sequential variation. Cooley (1971) has also used it, mainly as a predictive tool.

On the other hand the engineering literature on estimation of models of the form (6) is huge. The earliest work was the one by Kalman and Bucy (1961). For extensive bibliographies and various aspects of the problem the reader can consult the textbooks of Sage and Melsa (1971). and Åström (1970) as well as the special issue of the IEEE (1971) Transactions on Automatic Control.

In most of the engineering literature the statistics of the uncertain quantities are assumed known. This is a severe restriction when one is transferred to the realm of statistics and econometrics and is one of the primary reasons for which there is a large gap between research in system theory and the quantitative social sciences. Interesting exceptions to the rule in the engineering literature are the papers by Mehra (1970, 1971, 1972). and Kashyap (1970). Furthermore. the engineers usually make strong *a priori* assumptions about the matrix T, which as will be seen in section 9 do not. in general. hold in an econometric framework.

3. Problem Description

Consider the following model

(7) $$y_t = x_t \beta_t + \varepsilon_t$$

where y_t is the response to the effects of the k explanatory variables $x_{1t}, x_{2t}, \ldots, x_{kt}$. x_t is a $1 \times k$ vector of the mentioned explanatory variables, β_t is a $k \times 1$ vector of time varying coefficients, and ε_t is a disturbance term that is assumed to be normally distributed with the following properties.

(8) $$E[\varepsilon_t] = 0$$

(9) $$E[\varepsilon_t^2] = \sigma_\varepsilon^2$$

(10) $$E[\varepsilon_t \varepsilon_l] = \sigma_\varepsilon^2 \delta_{kl}$$

(11) $$E[\varepsilon_t \beta_l] = 0$$

where δ_{kl} is the Kronecker delta, and σ_ε^2 is an unknown constant. The assumption is that there are N observations on the endogenous variable y and the k exogenous variables.

Define the following quantities

(12) $$y = [y_1, y_2, \ldots, y_N]'$$

where (′) denotes the transposition.

(13) $$\beta = [\beta_1', \beta_2', \ldots, \beta_N']'$$

(14) $$\varepsilon = [\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_N]'$$

(15) $$X = \begin{bmatrix} x_1 & 0 & \ldots & 0 \\ 0 & x_2 & \ldots & 0 \\ \vdots & & \ddots & \\ 0 & & & x_N \end{bmatrix}.$$

The available information now can be written as follows:

(16) $$y = X\beta + \varepsilon.$$

It can be readily seen now that it is impossible to estimate the vector β (a $Nk \times 1$ vector) from (16), via ordinary least squares (OLS) regression. To use the OLS formula the matrix $X'X$ must be invertible. It is easily seen, however, that this $Nk \times Nk$ matrix has rank at most equal to N. So there are not enough degrees of freedom to estimate β.

The conclusion from the above discussion is that there is no hope of estimating β unless some more information about the vector becomes available. I will assume that the β_t's can be generated by a Markovian structure of the form

(17) $$\beta_{t+1} = T\beta_t + u_{t+1} \qquad (t = 0, 1, \ldots, N-1)$$

where: T is a known $k \times k$ transition matrix and u_j is a $k \times 1$ vector of random shocks distributed as multivariate normal with zero mean and covariance matrix

(18) $$E[u_j u_k'] = \sigma_u^2 R \delta_{kj}$$

where R is a known $k \times k$ positive semidefinite matrix.

The vector β_0 will be assumed unknown.