Title

A Simple Lagrange Multiplier F-Test for Multivariate Regression Models

Authors

Beatty, Timothy K.M.
LaFrance, Jeffrey T.
Yang, Muzhe

Publication Date

2005-02-01

A Simple Lagrange Multiplier F-Test for Multivariate Regression Models

Timothy K. Beatty
University of British Columbia

Jeffrey T. LaFrance
University of California, Berkeley

Muzhe Yang
University of California, Berkeley

A Simple Lagrange Multiplier F-Test for Multivariate Regression Models

Abstract

This paper proposes a straightforward, easy to implement approximate F-test which is useful for testing restrictions in multivariate regression models. We derive the asymptotics for our test statistic and investigate its finite sample properties through a series of Monte Carlo experiments. Both theory suggests and simulations confirm that our approach will result in strictly better inference than the leading alternative

DEPARTMENT OF AGRICULTURAL AND RESOURCE ECONOMICS AND POLICY
DIVISION OF AGRICULTURAL AND NATURAL RESOURCES
UNIVERSITY OF CALIFORNIA AT BERKELEY

Working Paper No. 996

A SIMPLE LAGRANGE MULTIPLIER *F*-TEST FOR MULTIVARIATE REGRESSION MODELS

by

Timothy K. M. Beatty, Jeffrey T. LaFrance, and Muzhe Yang

California Agricultural Experiment Station
Giannini Foundation of Agricultural Economics
February, 2005

A Simple Lagrange Multiplier F-test for Multivariate Regression Models

Timothy K.M. Beatty, Jeffrey T. LaFrance, Muzhe Yang*
University of British Columbia, Vancouver, British Columbia
University of California, Berkeley, California
University of California, Berkeley, California

January 28, 2005

Abstract

This paper proposes a straightforward, easy to implement approximate F-test which is useful for testing restrictions in multivariate regression models. We derive the asymptotics for our test statistic and investigate its finite sample properties through a series of Monte Carlo experiments. Both theory suggests and simulations confirm that our approach will result in strictly better inference than the leading alternative.

1 Introduction

The tendency of the three commonly used asymptotic tests, the Wald (W), likelihood ratio (LR) and Lagrange multiplier (LM) tests, to over-reject in multivariate regression models is well established; see Bera, Byron, and Jarque (1981), Italianer (1985), Theil, Shonkwiler, and Taylor (1985), and Taylor, Shonkwiler, and Theil (1986). This is particularly troublesome in the context of estimating demand systems where it increases the likelihood of

*Timothy Beatty would like to thank the Canada Research Chair Program and Social Sciences and Humanities Research Council of Canada.

errors in inference concerning the basic predictions of utility theory, namely homogeneity (Laitinen, 1978) and symmetry (Meisner, 1979).

In this paper, we construct an easy to implement approximate LM F-test to improve the finite sample empirical size of testing for multivariate regression models. This proposed test guarantees an F-statistic strictly larger than the degrees of freedom adjusted LM test with probability one. We focus on the LM test because a simple degree-adjustment is known to be insufficient to correct W and LR tests, and the modified test procedure developed in this paper will not improve the finite sample properties of either of these alternatives. This proposed F-test simplifies computation compared to previous efforts, and offers a good approximation to an exact test even in highly nonlinear multivariate regressions.

This is far from the first attempt to address these problems. Bartlett and Bartlett-type corrections have been widely investigated as possible means to address the size problem of the LR, LM and W tests. Typically these involve re-scaling the test statistic, such that the adjusted statistic is closer to the asymptotic distribution than the unmodified one. The most frequently encountered correction is a simple degrees of freedom re-scaling of the form $(NT - K) / (G \cdot NT)$, where N refers to the number of regression equations; T denotes the number of periods per equation; K is the number of parameters, and G represents the number of restrictions. We will refer to this as the Laitinen-Meisner correction. However, the Monte Carlo results of Bera, Byron and Jarque (1981) show that this approach, when combined with critical values taken from the $F(G, NT - K)$ tables, under-corrects W and LR, but over-corrects LM. While a number of other Bartlett-type corrections have been proposed, see Cribari-Neto and Cordeiro (1996) for a complete discussion, they typically involve functions of cummulants of derivatives of the log-likelihood function. This complexity, in addition to doubts about their effectiveness (see Rocke, 1989), seems to have impeded their widespread adoption.

Hashimoto and Ohtani (1990) propose an exact test for linear restrictions in seemingly unrelated regressions (SUR). However, the applicability of their approach may be somewhat limited due to the fact that it is only valid in the context of linear regressions where the same regressors are used in each equation. Additionally, under certain circumstances the power of this test may be quite low.

One promising area of research involves simulation tests. Dufour and Khalaf (2002) employ simulation methods to test hypothesis in multivariate and in SUR models. Their approach results in important gains over the Bartlett and Bartlett-type corrections. However, for a sufficiently complex model (such as the nonlinear demand models estimated in this paper), the relatively large computational cost could make this method somewhat impractical.

We now turn our attention to the proposed approximate F-test. We begin with an informal overview of the construction and properties of this test statistic. We then formally prove its characteristics and that, with probability one, it must improve upon the Laitinen-Meisner corrected LM test. We expect this implement to increase very quickly with T, as a result of the argument below. To illustrate the usefulness of our approach, we apply the proposed F-test to a series of progressively more non-linear demand systems. For each model, we compute the empirical cumulative distribution function (CDF) of a test for symmetry of the price coefficient matrix using our proposed test statistic and the Laitinen-Meisner corrected LM statistic. The results are compared to a true F distribution. We focus on the implications of nonlinearity in the estimation procedure, estimation of the error variance-covariance matrix, and convergence in nonlinear models for the finite sample behavior of our approach.

2 An Approximate LM F-test

We begin by writing down a nonlinear system of N equations with T periods each:

$$y_{jt} = f_j(x_{jt}, \boldsymbol{\beta}_{0j}) + \epsilon_{jt} \ \ (j = 1, \cdots, N; t = 1, \cdots, T) \tag{2.1}$$

Assume the following disturbance structure:

$$\boldsymbol{\epsilon}_t | \mathbf{x}_t \sim N\left(\mathbf{0}, \Sigma_{N \times N}\right) \tag{2.2}$$

where $\boldsymbol{\epsilon}_t$ is an N-vector, $\mathbf{x}_t$ is an $N \times K$ matrix, and $\boldsymbol{\beta}_{0j}$ is a K_j-vector.

A common statistical hypothesis of G restrictions imposed on K parameters (with $K = \sum_{j=1}^{N} K_j$, $G < K \ll NT$) can be written in the general form:

$$H_0 : \mathbf{g}\left(\boldsymbol{\beta}_K\right) = \mathbf{0}_G \tag{2.3}$$

To test these restrictions, we propose the following approximate LM F-statistic:

$$F = \frac{(\widetilde{S}(\widetilde{\Sigma}) - \widehat{S}(\widetilde{\Sigma}))/G}{\widehat{S}(\widehat{\Sigma})/(NT - K)} \tag{2.4}$$

where the error variance-covariance matrix for the first-round estimates of the *unrestricted* structural model is denoted $\widehat{\Sigma}$, and the second-round of *unrestricted* weighted residual sum of squares is given by $\widehat{S}(\widehat{\Sigma})$[1]; where the error variance-covariance matrix for the first-round estimates of the *restricted* structural model is denoted $\widetilde{\Sigma}$, and the second-round *restricted* weighted residual sum of squares is given by $\widetilde{S}(\widetilde{\Sigma})$; finally, where the *unrestricted* weighted residual sum of squares, based on the *restricted* error variance-covariance matrix from the first-round, is denoted $\widehat{S}(\widetilde{\Sigma})$.

The numerator, $(\widetilde{S}(\widetilde{\Sigma}) - \widehat{S}(\widetilde{\Sigma}))/G$, converges in distribution to a χ^2_G/G random variable. We use the first-round error variance-covariance matrix from the restricted model to remain consistent with the Lagrange multiplier principle. The LM test is well-known to have the smallest empirical size amongst the three classical tests for multivariate linear regression model.

The denominator, $\widehat{S}(\widehat{\Sigma})/(NT-K)$, calculated using the unrestricted model for both the first and second rounds, converges in distribution to a $\chi^2_{(NT-K)}/(NT-K)$ random variable under joint normality of the true residuals. However, note that even if this distributional assumption does not hold, the denominator will asymptotically converge to one, while $F(G, NT-K) \xrightarrow{d} \chi^2_G/G$ as $T \to \infty$. This implies that even if the residuals are not normally distributed the test proposed here will outperform the alternatives. In either case, with probability one, the empirical size of this F-test will be closer to the nominal size than a Laitinen-Meisner LM test that uses the F tables.

The numerator in our F-statistic is simply one form of the Lagrange multiplier test statistic, divided by its degrees of freedom, while the denominator is simply the unrestricted weighted residual sum of squares also divided by its degrees of freedom. The latter, and a single iteration on the estimated error variance-covariance matrix in the denominator, can be motivated as follows. First, the second-round unrestricted generalized least squares (GLS) criterion has an asymptotic $\chi^2_{(NT-K)}$ distribution. Second, at the beginning of the second-round generalized nonlinear least squares (GNLS) estimation step, we begin with the first-round nonlinear least squares (NLS) estimates for the structural parameters. This implies a starting value of the GNLS criterion of NT. Minimization implies that the denominator's GNLS criterion is strictly less than NT (i.e., $\widehat{S}(\widehat{\Sigma}) \leqslant NT$), with equality if and only if the first-round estimated error variance-covariance matrix just happens to coincide with the convergent iterative solution when one follows the standard

[1] Also denoted SSR_U in the figures.

Malinvaud (1980, Chapter 9) iteration on Σ. Note that this is a probability zero event. The approximate F-statistic, therefore, with probability one, is strictly larger than the Laitinen-Meisner adjusted LM.

$$\begin{aligned} F &= \frac{LM}{\widehat{S}(\widehat{\Sigma})} \cdot \frac{NT-K}{G} \\ &> \frac{LM}{NT} \cdot \frac{NT-K}{G} = LM_{\text{Laitinen-Meisner}} \ a.s.. \end{aligned} \tag{2.5}$$

Since a chi-squared random variable divided by its degrees of freedom converges almost surely to one, however, the approximate F-statistic converges (in probability and distribution) to the LM statistic, and therefore is an optimal test in large samples.

We formalize this argument as follows.

Proposition 1 *Let a "$\tilde{}$" denote restricted estimates, and a "$\hat{}$" denote unrestricted estimates. Let $\widehat{\Sigma}$ and $\widetilde{\Sigma}$ be the first-round nonlinear least squares estimates for Σ, under unrestricted and restricted models respectively. The generalized nonlinear least squares criteria for these estimates are:*

$$\begin{aligned} \widehat{S}(\widehat{\Sigma}) &\equiv \sum_{t=1}^{T} \widehat{\boldsymbol{\epsilon}}_t^T \widehat{\Sigma}^{-1} \widehat{\boldsymbol{\epsilon}}_t \\ &= \sum_{t=1}^{T} (\mathbf{y}_t - \mathbf{f}(\mathbf{x}_t, \widehat{\boldsymbol{\beta}}(\widehat{\Sigma}))^T \widehat{\Sigma}^{-1} (\mathbf{y}_t - \mathbf{f}(\mathbf{x}_t, \widehat{\boldsymbol{\beta}}(\widehat{\Sigma})) \\ \widehat{S}(\widetilde{\Sigma}) &\equiv \sum_{t=1}^{T} \widehat{\boldsymbol{\epsilon}}_t^T \widetilde{\Sigma}^{-1} \widehat{\boldsymbol{\epsilon}}_t \\ &= \sum_{t=1}^{T} (\mathbf{y}_t - \mathbf{f}(\mathbf{x}_t, \widehat{\boldsymbol{\beta}}(\widetilde{\Sigma}))^T \widetilde{\Sigma}^{-1} (\mathbf{y}_t - \mathbf{f}(\mathbf{x}_t, \widehat{\boldsymbol{\beta}}(\widetilde{\Sigma})) \\ \widetilde{S}(\widetilde{\Sigma}) &\equiv \sum_{t=1}^{T} \widetilde{\boldsymbol{\epsilon}}_t^T \widetilde{\Sigma}^{-1} \widetilde{\boldsymbol{\epsilon}}_t \\ &= \sum_{t=1}^{T} (\mathbf{y}_t - \mathbf{f}(\mathbf{x}_t, \widetilde{\boldsymbol{\beta}}(\widetilde{\Sigma}))^T \widetilde{\Sigma}^{-1} (\mathbf{y}_t - \mathbf{f}(\mathbf{x}_t, \widetilde{\boldsymbol{\beta}}(\widetilde{\Sigma})) \end{aligned}$$

The approximate LM F-statistic is:

$$F = \frac{(\widetilde{S}(\widetilde{\Sigma}) - \widehat{S}(\widetilde{\Sigma}))/G}{\widehat{S}(\widehat{\Sigma})/(NT-K)} \overset{a}{\sim} F\left(G, NT-K\right). \tag{2.6}$$

Proof. See appendix. ■

3 A Simulation Study

In order to characterize the finite sample properties of this test statistic, we conduct a series of Monte Carlo experiments. We investigate the performance of our proposed test statistic by applying it to three different models drawn from the demand literature: a simple linear model, a linear-quadratic model (AIDS like) and finally a quadratic utility demand system. The hypothesis of interest is Slutsky symmetry, following Meisner and Bera, Byron and Jarque. These models differ in a number of respects that make comparing results across models informative. In particular the models considered become progressively more nonlinear, which makes them progressively harder to estimate. This results in an increased probability that the estimation routine will not successfully converge.

In all cases, iterating to convergence of the GNLS criterion is essential to ensure comparisons between optima, especially in models that are highly nonlinear in the structural parameters. In addition, to help ensure that the GNLS criterion in the unrestricted model ($\widehat{S}(\widetilde{\Sigma})$), conditional on the restricted error variance-covariance matrix estimates ($\widetilde{\Sigma}$), is a global minimum and strictly less than the restricted GNLS criterion ($\widetilde{S}(\widetilde{\Sigma})$), it is useful to begin the unrestricted estimation step for the numerator at the second-round parameter estimates for the restricted model.

Given the central role played by Σ in our test statistic, we investigate the effects of estimating Σ on each model by running two Monte Carlo experiments, one in which Σ is assumed known and the other in which Σ is estimated. In each case we compare our results to a Laitinen-Meisner LM correction.

(1) Linear model:

$$\mathbf{y}_t = P_t \left(A\mathbf{s}_t + B\mathbf{p}_t + \boldsymbol{\gamma} m_t\right) + \boldsymbol{\epsilon}_t. \tag{3.1}$$

(2) Linear-quadratic model:

$$\mathbf{y}_t = P_t \left[A\mathbf{s}_t + B\mathbf{p}_t + \boldsymbol{\gamma}(m_t - \mathbf{p}_t^T A\mathbf{s}_t - \frac{1}{2}\mathbf{p}_t^T B\mathbf{p}_t)\right] + \boldsymbol{\epsilon}_t. \tag{3.2}$$

(3) Quadratic model:

$$\mathbf{y}_t = P_t \left[A\mathbf{s}_t + B\mathbf{p}_t + \frac{\left(m_t - \alpha^T\mathbf{s}_t - \mathbf{p}_t^T A\mathbf{s}_t\right) B\mathbf{p}_t}{\mathbf{p}_t^T B\mathbf{p}_t + 1}\right] + \boldsymbol{\epsilon}_t. \tag{3.3}$$

where our data consist of: $\mathbf{y}_t$ is an n-vector of food expenditures, $\mathbf{p}_t$ is an n-vector of food prices, m_t is income and $\mathbf{s}_t$ is an s-vector of demographic characteristics. We estimate, B an $n \times n$ matrix of parameters on prices, A an $n \times s$ matrix coefficients on demographic variables, and $\boldsymbol{\gamma}$ an n-vector of parameters on income, supernumerary income, or demographics, respectively. We assume that $\boldsymbol{\epsilon}_t$ is i.i.d $N(\mathbf{0}, \Sigma)$.

The symmetry of the Slutsky matrix is accommodated by a set of linear parameter restrictions on B,[2]

$$H_0 : B = B^T. \tag{3.4}$$

3.1 U.S. Food Demand Model and Data

In order to implement the demand systems described above, we combine three different time-series data sets. The first is data on per capita consumption of dairy products over the period 1919-2000. The second is a corresponding set of average retail prices for those products. The consumer price index for non-food items is used as the price of nonfood expenditures. The third data series are demographic factors that help to predict demand. These demographic factors include the first three central moments (the mean, variance and skewness) of the age distribution and the proportion of the U.S. population that is Black and the proportion that is neither Black or White.

3.2 Simulation Algorithm

We now turn to the algorithm by which the finite sample behavior of our test statistics will be investigated.

(1) For each of the three models (linear, linear-quadratic, quadratic utility) we estimate the model under the null (in this case $B = B^T$) hypothesis. We use the fitted values of $\widetilde{\mathbf{y}}_t$ and $\widetilde{\Sigma}$ under the null hypothesis as the truth for the purpose of the Monte Carlo experiment.

(2) For each round (s) we draw an N-dimensional multivariate standard normal random variable $\mathbf{z}_t^{(s)}$, where $\mathbf{z}_t^{(s)} \sim N(\mathbf{0}, I_N)$. These are rescaled by L, which is the upper triangular matrix resulting from a Cholesky decomposition of $\widetilde{\Sigma}$. This yields random variables $\boldsymbol{\epsilon}_t^{(s)} = L^T \mathbf{z}_t^{(s)}$, where $\boldsymbol{\epsilon}_t^{(s)} \sim N(\mathbf{0}, \widetilde{\Sigma})$. These

[2] Of course, Slutsky symmetry is feasible in (3.1) if and only if $\boldsymbol{\gamma} = \mathbf{0}$, but we include this model with $\boldsymbol{\gamma} \neq \mathbf{0}$ as a linear base point.

are added back, $\mathbf{y}_t^{(s)} = \widetilde{\mathbf{y}}_t^{(s)} + \boldsymbol{\epsilon}_t^{(s)}$, to generate a vector of random dependent variables for each good.

(3) The relevant model is then estimated. Given the importance of starting values in nonlinear estimation, the models are started at their true values. We then estimate the restricted and unrestricted models, and compute $\widetilde{S}(\widetilde{\Sigma})$, $\widehat{S}(\widehat{\Sigma})$ and $\widehat{S}(\widetilde{\Sigma})$.

(4) For each model and at each iteration we calculate the LM F-statistic, and Laitinen-Meisner corrected LM statistic.

(5) Empirical CDFs (ECDFs) are produced by plotting the results of 5,000 iterations. In each case the ECDFs of the F-test and the Laitinen-Meisner LM statistics are plotted against the CDF of the exact F distribution.

We discard estimates from iterations which fail to converge at any stage of the process and continue the simulation until 5,000 valid iterations are obtained.

3.3 Simulation Results

Figures 1-6 illustrate the results of the simulation exercise described above. Where informative, arrows make clear the horizontal distance between each statistic and the true value for the .90th , .95th and .99th percentiles respectively. In every instance, the proposed F-test first-order stochastically dominates the Laitinen-Meisner LM statistic. When Σ is known and the model is not very nonlinear, the proposed F-test is virtually indistinguishable from the CDF of an F distribution with the appropriate degrees of freedom. When Σ is estimated, the proposed F-test continues to do better than the Laitinen-Meisner LM statistic. In addition, for the most nonlinear of the models considered, the test becomes slightly conservative.

In every instance the ECDF of the Laitinen-Meisner LM statistic lies entirely above the true CDF. This confirms the earlier Bera, Byron and Jarque (1981) results, which appear to generalize to quite a high degree of nonlinearity. Using the Laitinen-Meisner LM statistic with the F tables results in an increased likelihood of committing a type I error. As the theory suggests and the simulations confirm, using the approach developed in this paper dominates the alternative Laitinen-Meisner LM statistic in every instance considered.

Figures 1 and 3 demonstrate that for linear and linear-quadratic models with Σ known, the proposed LM F-test is indistinguishable from a CDF of an F-statistic. Figures 2 and 4 make clear that, for these models, estimating

Σ results in a statistic undersized, although considerably less so than the Laitinen-Meisner LM statistic.

When applied to the most nonlinear model considered, the quadratic utility model in Figures 5 and 6, we see that the ECDF for our proposed LM corrected F-test lies everywhere somewhat below the true CDF. The result is a slightly conservative test where the empirical size is slightly too large. This results in a slight tendency to over-reject a true null. For the most nonlinear model, estimating Σ does not seem to have as large an impact as in the linear and linear-quadratic case. The reason for this result is likely to be the negative bias and low mean square error of the estimated Σ in nonlinear models (see LaFrance, 1993).

Under the three models (linear, linear-quadratic and quadratic utility models) with Σ estimated, comparisons between the Laitinen-Meisner LM (L-M) statistic, and the proposed LM F-statistic (L-B-W), in terms of critical values and sizes, are given in Table I and Table II.

Table I: Comparison of Critical Values with Σ estimated

Size	Critical values	Linear		Linear-quadratic		Quadratic utility	
α	$F(10, 320)$	*L-M*	*L-B-W*	*L-M*	*L-B-W*	*L-M*	*L-B-W*
0.10	1.619	1.159	1.398	1.135	1.368	1.529	1.670
0.05	1.860	1.310	1.591	1.278	1.544	1.727	1.860
0.01	2.346	1.697	2.034	1.606	1.954	2.167	2.376

Table II: Comparison of Sizes with Σ estimated

Size	Critical values	Linear		Linear-quadratic		Quadratic utility	
α	$F(10, 320)$	*L-M*	*L-B-W*	*L-M*	*L-B-W*	*L-M*	*L-B-W*
0.10	1.619	0.014	0.046	0.009	0.037	0.078	0.113
0.05	1.860	0.004	0.020	0.002	0.015	0.035	0.055
0.01	2.346	0.0003	0.003	0.0000	0.002	0.006	0.010

4 Conclusion

This paper proposes a straightforward and easy to implement approximate F-test for a system of regression equations. In theory and practice, this approximate F-test partially overcomes the tendency of the Laitinen-Meisner degrees of freedom correction to over-reject. In every instance this simple alternative is more likely to lead to better inferences.

5 Appendix: Proof for Proposition 1

Consider a nonlinear system of N equations with T periods each[3]:

$$y_{jt} = f_j(x_{jt}, \boldsymbol{\beta}_{0j}) + \epsilon_{jt} \quad (j = 1, \cdots, N,\ t = 1, \cdots, T) \tag{1}$$

$$\text{Assume: } \boldsymbol{\epsilon}_t | \mathbf{x}_t \sim N(\mathbf{0}, \Sigma_{N\times N}) \tag{2}$$

A common statistical hypothesis of G restrictions imposed on K parameters (with $G < K \ll NT$) can be written in the general form:

$$H_0: \ \mathbf{g}(\boldsymbol{\beta}_K) = \mathbf{0}_G \tag{3}$$

Define:

$$\begin{aligned}
\underset{N\times 1}{\mathbf{u}_t} &\equiv \Sigma^{-\frac{1}{2}}\boldsymbol{\epsilon}_t \qquad \underset{NT\times 1}{\mathbf{u}} \equiv \left[\mathbf{u}_1^T, \mathbf{u}_2^T, \cdots, \mathbf{u}_T^T\right]^T \sim N(\mathbf{0}, I_{NT}) \\
\underset{N\times K}{F_t} &\equiv \frac{\partial \mathbf{f}(\mathbf{x}_t, \boldsymbol{\beta}_0)}{\partial \boldsymbol{\beta}^T} \underset{NT\times K}{F} \equiv \left[F_1^T, F_2^T, \cdots, F_T^T\right]^T \quad \underset{G\times K}{R} \equiv \frac{\partial \mathbf{g}(\boldsymbol{\beta}_0)}{\partial \boldsymbol{\beta}^T} \\
\underset{NT\times K}{F^*} &\equiv \left[F_1^T \Sigma^{-\frac{1}{2}}, F_2^T \Sigma^{-\frac{1}{2}}, \cdots, F_T^T \Sigma^{-\frac{1}{2}}\right]^T = \left[F_1^{*T}, F_2^{*T}, \cdots, F_T^{*T}\right]
\end{aligned}$$

Construct the following two symmetric and idempotent matrices:

$$\begin{aligned}
\underset{NT\times NT}{\mathbb{M}} &\equiv I_{NT} - F^*(F^{*T}F^*)^{-1}F^{*T} \\
\underset{NT\times NT}{\mathbb{A}} &\equiv F^*(F^{*T}F^*)^{-1}R^T[R(F^{*T}F^*)^{-1}R^T]^{-1}R(F^{*T}F^*)^{-1}F^{*T}
\end{aligned}$$

satisfying the mutual orthogonality condition $\mathbb{M}\mathbb{A} = \mathbf{0}$.

We take three steps to show :

$$F = \frac{(\widetilde{S}(\widetilde{\Sigma}) - \widehat{S}(\widetilde{\Sigma}))/G}{\widehat{S}(\widehat{\Sigma})/(NT-K)} \overset{a}{\sim} F(G, NT-K)$$

(1) Denominator: $\widehat{S}(\widehat{\Sigma}) \xrightarrow{p} \mathbf{u}^T \mathbb{M} \mathbf{u} \sim \chi^2(NT-K)$

Consider the F.O.C. of the *unrestricted* model $\widehat{\boldsymbol{\beta}} = \arg\min_{\widehat{\boldsymbol{\beta}}} \widehat{S}(\widehat{\Sigma})$, given $\widehat{\boldsymbol{\beta}} \xrightarrow{p} \boldsymbol{\beta}_0$ and $\widehat{\Sigma} \xrightarrow{p} \Sigma$, where $\widehat{\Sigma}$ is the first round estimated error covariance matrix from nonlinear least squares (NLS) on the unrestricted model:

$$\sum_{t=1}^{T} [\frac{\partial \mathbf{f}(\mathbf{x}_t, \widehat{\boldsymbol{\beta}})}{\partial \boldsymbol{\beta}^T}]^T \widehat{\Sigma}^{-1}(\mathbf{y}_t - \mathbf{f}(\mathbf{x}_t, \widehat{\boldsymbol{\beta}})) = 0 \tag{4}$$

[3] where $\boldsymbol{\epsilon}_t$ is an N-vector, $\mathbf{x}_t$ is an $N \times K$ matrix, $\boldsymbol{\beta}_{0j}$ is an K_j-vector, and $K = \sum_{j=1}^{N} K_j$.

Standard conditions and asymptotic results imply the following:

$$F^*(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) \doteq (I - \mathbb{M})\mathbf{u} \tag{5}$$

$$\widehat{S}(\widehat{\Sigma}) \doteq \mathbf{u}^T \mathbb{M} \mathbf{u} \tag{6}$$

$$\mathbf{u} \sim N(\mathbf{0}, I),\ \mathbb{M} \text{ is idempotent, } rank(\mathbb{M}) = tr(\mathbb{M}) = NT - K$$

$$\Rightarrow \widehat{S}(\widehat{\Sigma}) \xrightarrow{p} \mathbf{u}^T \mathbb{M} \mathbf{u} \sim \chi^2(NT - K) \tag{7}$$

(2) Numerator: $\widetilde{S}(\widetilde{\Sigma}) - \widehat{S}(\widetilde{\Sigma}) \xrightarrow{p} \mathbf{u}^T \mathbb{A} \mathbf{u} \sim \chi^2(G)$

Consider the F.O.C. of the *restricted* model, $\widetilde{\boldsymbol{\beta}} = \arg\min_{\widetilde{\boldsymbol{\beta}}} \left\{ \widetilde{S}(\widetilde{\Sigma}) | s.t. : \mathbf{g}(\widetilde{\boldsymbol{\beta}}) = \mathbf{0} \right\}$, given $\widetilde{\boldsymbol{\beta}} \xrightarrow{p} \boldsymbol{\beta}_0$ and $\widetilde{\Sigma} \xrightarrow{p} \Sigma$, where $\widetilde{\Sigma}$ is the first round estimated error covariance matrix from NLS on the restricted model:

$$\sum_{t=1}^{T} [\frac{\partial \mathbf{f}(\mathbf{x}_t, \widetilde{\boldsymbol{\beta}})}{\partial \boldsymbol{\beta}^T}]^T \widetilde{\Sigma}^{-1} (\mathbf{y}_t - \mathbf{f}(\mathbf{x}_t, \widetilde{\boldsymbol{\beta}})) = (\frac{\partial g(\widetilde{\boldsymbol{\beta}})}{\partial \boldsymbol{\beta}^T})^T \boldsymbol{\lambda} \tag{8}$$

$$\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta}_0 \doteq (F^{*T} F^*)^{-1} (F^{*T} \mathbf{u} - R^T \boldsymbol{\lambda}) \tag{9}$$

Since $\widetilde{\Sigma} \xrightarrow{p} \Sigma$ and $\widehat{\Sigma} \xrightarrow{p} \Sigma$ when H_0 is true, standard asymptotic arguments imply:

$$\widehat{\boldsymbol{\beta}}(\widetilde{\Sigma}) - \boldsymbol{\beta}_0 \doteq \left(F^{*T} F^*\right)^{-1} F^{*T} \mathbf{u} \doteq \widehat{\boldsymbol{\beta}}(\widehat{\Sigma}) - \boldsymbol{\beta}_0 \tag{10}$$

$$\widehat{S}(\widetilde{\Sigma}) \doteq \mathbf{u}^T \mathbb{M} \mathbf{u} \tag{11}$$

$$\mathbf{0} = g(\widetilde{\boldsymbol{\beta}}) \doteq R(\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) \tag{12}$$

Substituting (12) into (9), we obtain:

$$\boldsymbol{\lambda} = [R(F^{*T} F^*)^{-1} R^T]^{-1} (F^{*T} F^*)^{-1} F^{*T} \mathbf{u} \tag{13}$$

Substituting (13) into (9), we obtain:

$$\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta}_0 \doteq (F^{*T} F^*)^{-1} \left[F^{*T} \mathbf{u} - R^T [R(F^{*T} F^*)^{-1} R^T]^{-1} R(F^{*T} F^*)^{-1} F^{*T} \mathbf{u}\right]$$

$$F^*(\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) \doteq (I - \mathbb{M} - \mathbb{A})\mathbf{u} \tag{14}$$

$$\begin{aligned} \widetilde{S}(\widetilde{\Sigma}) &\equiv \sum_{t=1}^{T} (\mathbf{y}_t - \mathbf{f}(\mathbf{x}_t, \widetilde{\boldsymbol{\beta}}(\widetilde{\Sigma}))^T \widetilde{\Sigma}^{-1} (\mathbf{y}_t - \mathbf{f}(\mathbf{x}_t, \widetilde{\boldsymbol{\beta}}(\widetilde{\Sigma})) \\ &\doteq \sum_{t=1}^{T} (\mathbf{u}_t - F_t^*(\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta}_0))^T (\mathbf{u}_t - F_t^*(\widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)) \\ &= \mathbf{u}^T \mathbf{u} - 2\mathbf{u}^T (I - \mathbb{M} - \mathbb{A})\mathbf{u} + \mathbf{u}^T (I - \mathbb{M} - \mathbb{A})(I - \mathbb{M} - \mathbb{A})\mathbf{u} \\ &= \mathbf{u}^T (\mathbb{M} + \mathbb{A})\mathbf{u} \end{aligned} \tag{15}$$

$$\mathbf{u} \sim N(\mathbf{0}, I),\ \mathbb{M} \text{ and } \mathbb{A} \text{ are idempotent matrices}$$

$$\begin{aligned} rank(\mathbb{M}) &= tr(\mathbb{M}) = NT - K \\ rank(\mathbb{A}) &= tr(\mathbb{A}) = G \end{aligned}$$

$$\Rightarrow \widetilde{S}(\widetilde{\Sigma}) \xrightarrow{p} \mathbf{u}^T(\mathbb{M} + \mathbb{A})\mathbf{u} \sim \chi^2(NT - K + G) \tag{16}$$

Combining (11) and (16), it follows that:

$$\widetilde{S}(\widetilde{\Sigma}) - \widehat{S}(\widetilde{\Sigma}) \xrightarrow{p} \mathbf{u}^T(\mathbb{A})\mathbf{u} \sim \chi^2(G) \tag{17}$$

(3) The symmetry, idempotency and orthogonality of the matrices, $\mathbb{A}$ and $\mathbb{M}$, imply that the quadratic form in (17), which is the numerator of our approximate LM F-statistic, and the quadratic form in (7), which is in the denominator, are asymptotically statistically independent, so we have:

$$F = \frac{(\widetilde{S}(\widetilde{\Sigma}) - \widehat{S}(\widetilde{\Sigma}))/G}{\widehat{S}(\widehat{\Sigma})/\left(NT - K\right)} \overset{a}{\sim} F\left(G, NT - K\right) \tag{18}$$

References

[1] Bera, A.K., Byron. R.P., Jarque, C.M., 1981. Further evidence on asymptotic tests for homogeneity and symmetry in large demand systems. Economics Letters 8, 101-105.

[2] Cribari-Neto, F., Cordeiro, G.M., 1996. On Bartlett and Bartlett-type corrections. Econometric Reviews 15, 339-367.

[3] Dufour, J.-M., Khalaf, L., 2002. Simulation based finite and large sample tests in multivariate regressions. Journal of Econometrics 111, 303-322.

[4] Hashimoto, N., Ohtani, K., 1990. An exact test for linear restrictions in seemingly unrelated regressions with the same regressors. Economics Letters 32, 243-246.

[5] Italiner, A., 1985. A small-sample correction for the likelihood ratio test. Economics Letters 19, 315-317.

[6] LaFrance, J.T., 1993. Weak separability in applied welfare analysis. American Journal of Agricultural Economics 75, 770-775.

[7] Laitinen, K., 1978. Why is demand homogeneity so often rejected? Economics Letters 1, 187-191.

[8] Malinvaud, E., 1980. Statistical Methods of Economics, Second edition. North-Holland, Amsterdam.

[9] Meisner, J.F., 1979. The sad fate of the asymptotic Slutsky symmetry test for large systems. Economics Letters 2, 231-233.

[10] Rocke, D.M., 1989. Bootstrap Bartlett adjustment in seemingly unrelated regressions. Journal of American Statistical Association 84, 598-601.

[11] Theil, H., Shonkwiler, J.S., Taylor, T.G., 1985. A Monte Carlo test of Slutsky symmetry. Economics Letters 19, 331-332.

[12] Taylor, T.G., Shonkwiler, J.S., Theil, H., 1986. Monte Carlo and bootstrap testing of demand homogeneity. Economics Letters 20, 55-57.

Figure 1.

Empirical and True F(10,320) CDF

Linear Model, Known Σ

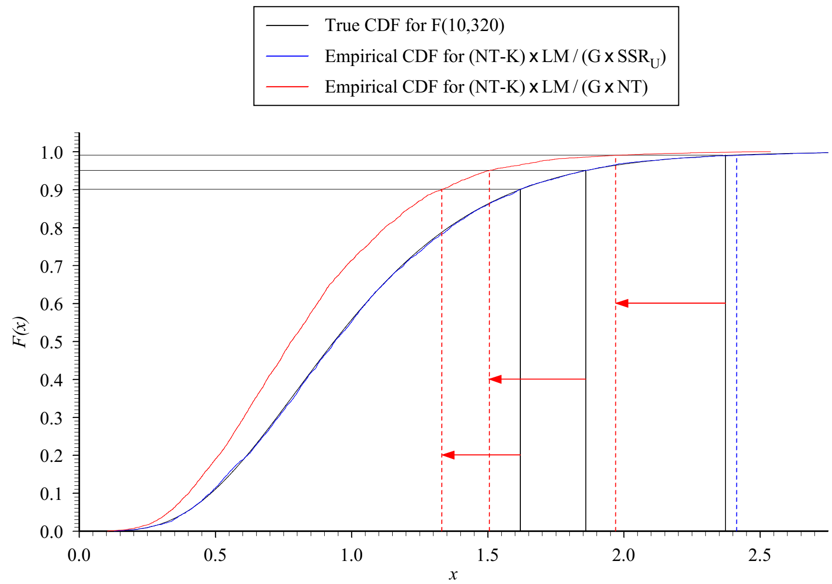

Figure 2.

Empirical and True F(10,320) CDF

Linear Model, Estimated Σ

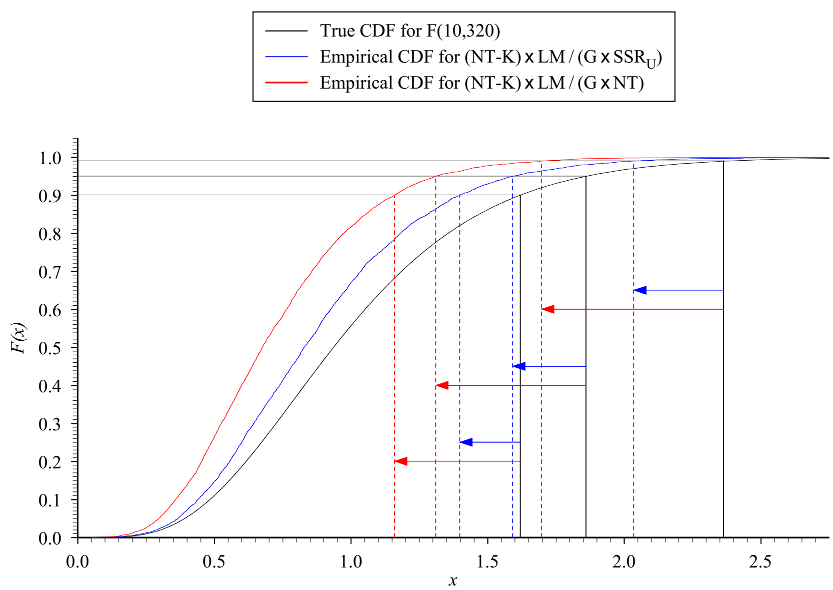

Figure 3.

Empirical and True F(10,320) CDF

LinQuad Model, Known Σ

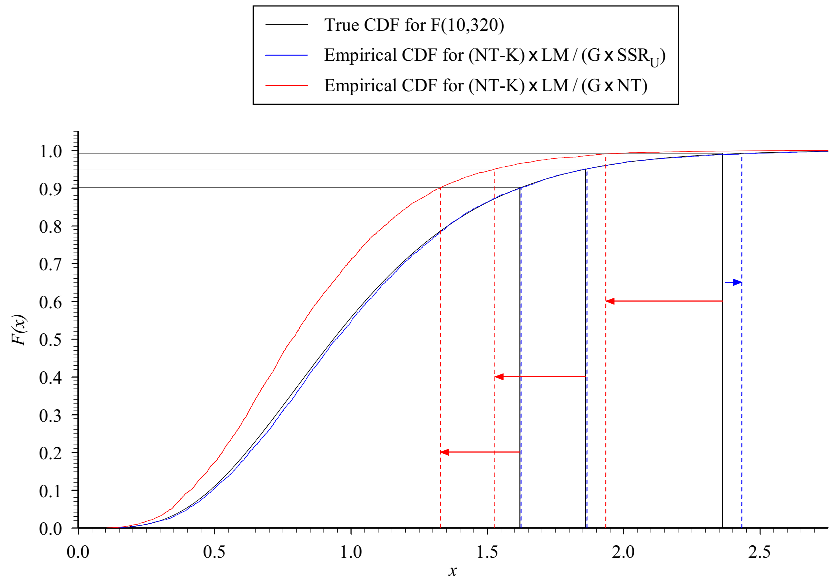

Figure 4.

Empirical and True F(10,320) CDF

LinQuad Model, Estimated Σ

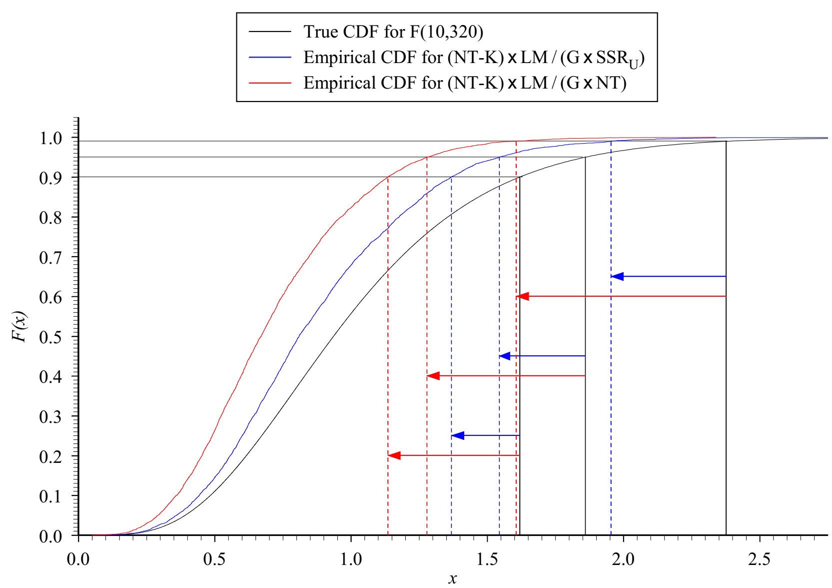

Figure 5.

Empirical and True F(10,320) CDF

Quadratic Utility, Known Σ

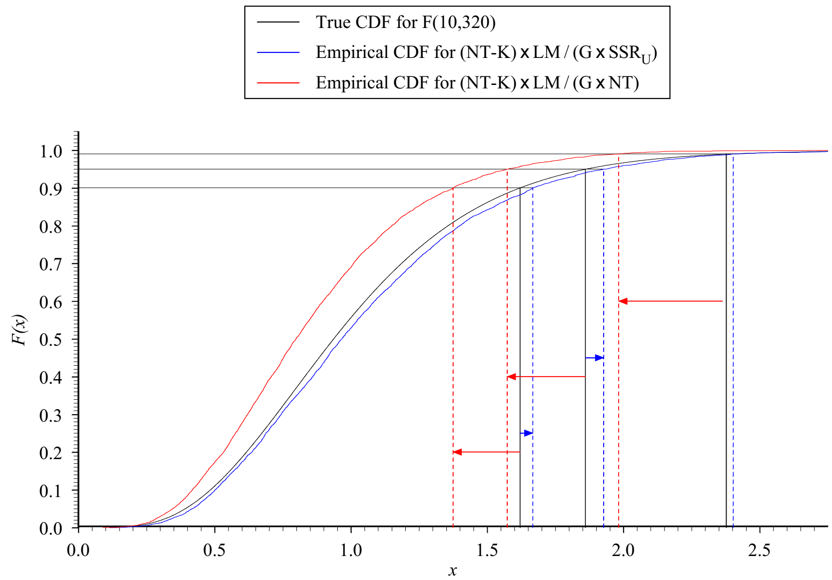

Figure 6.

Empirical and True F(10,320) CDF

Quadratic Utility, Estimated Σ