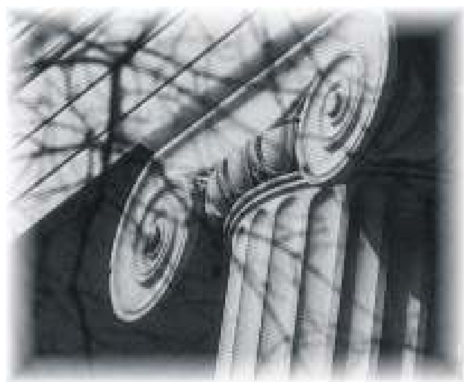
THE UNIVERSITY OF ALABAMA

ECONOMICS, FINANCE AND LEGAL STUDIES

WORKING PAPER SERIES



Working Paper No. 08-02-01

Valid Inference for a Class of Models Where Standard Inference Performs Poorly; Including Nonlinear Regression, ARMA, GARCH, and Unobserved Components

Jun Ma, and Charles R. Nelson

Culverhouse College of Commerce and Business Administration
The University of Alabama

*Valid Inference for a Class of Models Where Standard Inference Performs Poorly; Including Nonlinear Regression, ARMA, GARCH, and Unobserved Components***

Jun Ma*
Department of Economics, Finance and Legal Studies
University of Alabama
and

Charles R. Nelson*
Department of Economics
University of Washington

First draft June 18, 2006

This draft January 19, 2008

JEL *classification*: C120, C220, C330

Keywords: ARMA, Unobserved Components, State Space, GARCH, Zero-Information-Limit-Condition.

Abstract
Nelson and Startz (2007) showed that standard inference performs poorly in models where information about the parameter of interest depends on an unknown nuisance parameter. Standard errors are generally too small and the size of the *t*-test far from its nominal level in finite samples. This paper explores a testing strategy for models of the form $y = \gamma \bullet g(\beta, x) + \varepsilon$, including non-linear regression, ARMA, GARCH, and Unobserved Components. Monte Carlo experiments suggest that the size of the proposed test is close to correct in situations where standard inference is spurious while power is at least as good.

* Corresponding author: Box 353330, University of Washington, Seattle, WA 98195, USA. Tel: +1 206 543 5955; fax +1 206 685 7477.
E-mail address: cnelson@u.washington.edu.

** The authors gratefully acknowledge support from the Ford and Louisa Van Voorhis Endowment at the University of Washington. We thank the following for helpful comments: Michael Dueker, Walter Enders, Evan Koenig, Junsoo Lee, Christian Murray, David Papell, Ruxandra Prodan, Bent Sørensen, Richard Startz, and Eric Zivot, but responsibility for all errors is ours.

1. Introduction

This paper is concerned with inference in the class of econometric models that have a representation of the form

$$y_i = \gamma \bullet g(\beta, \mathbf{x}_i) + \varepsilon_i; \gamma \neq 0; i = 1,.., N. \tag{1.1}$$

We assume that errors ε_i are i.i.d. $\mathrm{N}(0,\sigma^2)$ so that Maximum Likelihood estimates of parameters γ and β are obtained by least squares given data y and $\mathbf{x}$. Examples include non-linear regression models such as production functions and the Phillips curve model of Staiger, Stock and Watson (1997) and, perhaps less obviously, time series models including ARMA and GARCH, and State Space models such as Unobserved Components used to decompose GDP into trend and cycle. The parameter of interest in this paper is β and we assume that the condition for identification, namely $\gamma \neq 0$, holds. (We note that testing for identification poses a special set of issues that are dealt with in a literature that includes Davies(1977, 1987) and Hansen (1996) but not here.) Additional regressors may be present but are suppressed here for simplicity.

The amount of information that the data contain about β is controlled by γ. The reciprocal of the asymptotic variance of $\hat{\beta}$, a natural measure of information, is given by

$$V_{\hat{\beta}}(\gamma,\beta,\sigma)^{-1} = \frac{\gamma^2}{\sigma^2} \bullet \frac{\sum g^2 \bullet \sum (g')^2 - [\sum g \bullet g']^2}{\sum g^2} \equiv I_{\hat{\beta}}(\gamma,\beta,\sigma); \gamma \neq 0 \tag{1.2}$$

where g and g' denote respectively $g(\beta, x_i)$ and its first derivative with respect to β, both evaluated at the true value of β. The notation $V_{\hat{\beta}}(\gamma,\beta,\sigma)$ is intended to highlight the fact that the asymptotic variance is a function of the parameters.

In practice a test of the null hypothesis $\beta = \beta_0$ is based on the computable statistic $t = (\hat{\beta} - \beta_0) \bullet \sqrt{I_{\hat{\beta}}(\hat{\gamma}, \hat{\beta}, \hat{\sigma})}$ where parameter estimates replace unknown true values. Nelson and Startz (2007) (hereafter NS) study the class of models for which information is at least approximately proportional to γ^2 and they argue that estimated information tends to be upward biased because if $\hat{\gamma}$ is unbiased with variance $V_{\hat{\gamma}}$

$$E(\hat{\gamma}^2) = \gamma^2 + V_{\hat{\gamma}} \,. \tag{1.3}$$

This upward bias suggests that standard errors will tend to be too small, and NS find this to be the case. One might be tempted to conjecture that the resulting *t*-statistic would then tend to be too large in absolute value, producing too many rejections of the null hypothesis when it is true, but NS show that test size can go in either direction. However, asymptotic theory does hold in model (1.1), so as sample size grows the size of the *t*-test approaches its nominal level.

The objective of this paper is to explore a strategy for obtaining a valid test statistic for β - one that has correct size in finite samples – based on linearization of $g(\beta, x_i)$. Expanding $g(\beta, x_i)$ around the null hypothesis $\beta = \beta_0$ gives

$$y_i = \gamma \bullet [g(\beta_0, x_i) + (\beta - \beta_0) \bullet g'(\beta_0, x_i)] + e_i \tag{1.4}$$

where e_i may incorporate a remainder term. The reduced form of the linear approximation is:

$$y_i = \gamma \bullet g(\beta_0, x_i) + \lambda \bullet g'(\beta_0, x_i) + e_i \tag{1.5}$$

The implication of the null hypothesis is that the reduced form regression coefficient λ is zero. Intuitively, the first term captures the contribution of β to the model if the null hypothesis is correct, the second term being required only if the null is wrong. In some models of practical importance $g(.)$ is actually linear and $g'(.)$ is simply data. In those cases the reduced form *t*-test is exact and has correct size. More generally, linearization of $g(.)$ is only an approximation; how well the reduced form test performs in size and power relative to the standard *t*-test is the subject of this paper.

The paper is organized as follows. Section 2 undertakes an investigation of the relative performance of the standard *t*-test and the proposed reduced form test when $g(\beta, x_i)$ is linear since useful analytical results are available in that case. Section 3 looks at how close the proposed test comes to achieving correct size in models of importance in practice where the reduced form regression is not exact but only a linear approximation. These include nonlinear regression, ARMA (1,1) and GARCH(1,1) models, and a simple Unobserved Components model. Section 4 concludes.

2. Size and power of the asymptotic and reduced form *t*-tests when *g(.)* is linear.

In the case where the function *g(.)* in (1.1) is linear the model takes the form:

$$y_i = \gamma \bullet (x_i + \beta \bullet z_i) + \varepsilon_i \tag{2.1}$$

where x_i and z_i denote independent variables. This case is both an archetype of the general class and important in its own right in applied econometrics. For example, the Phillips curve model of Staiger, Stock and Watson (1997) takes the form:

$$y_i = \gamma \bullet (x_i + \beta) + \varepsilon_i$$

where y is the change in inflation, and $g = (x_i + \beta)$, where x is the unemployment rate and β is the unknown natural rate, g' being simply one. The reduced form of (2.1) is

$$y_i = \gamma \bullet x_i + \lambda \bullet z_i + \varepsilon_i \tag{2.2}$$

where $\lambda = \gamma \bullet \beta$. Since β is exactly identified, the least squares estimate from (2.1) is equal to the indirect least squares estimate from the reduced form, that is $\hat{\beta} = \hat{\lambda} / \hat{\gamma}$. When the independent variables are exogenous and $\varepsilon_i \sim$ i.i.d. $N(0, \sigma^2)$, the asymptotic variance of $\hat{\beta}$ derived either from the information matrix for (2.1) under maximum likelihood, or using the 'delta method' for indirect least squares, is given by:

$$V_{\hat{\beta}} = \frac{1}{\gamma^2} \bullet \frac{\sigma^2}{N} \bullet \frac{m_{xx} + 2\beta \bullet m_{xz} + \beta^2 \bullet m_{zz}}{m_{xx} \bullet m_{zz} - m_{xz}^2} \tag{2.3}$$

where '*m*' denotes the raw sample second moment of the subscripted variables.

In practice the asymptotic variance is estimated at the point estimates of the parameters so the *t*-statistic for $\hat{\beta}$ is given by:

$$t_{\hat{\beta}}^2 = (\hat{\beta} - \beta_0)^2 \bullet \left[\hat{\gamma}^2 \bullet \frac{N}{\hat{\sigma}^2} \bullet \frac{m_{xx} \bullet m_{zz} - m_{xz}^2}{m_{xx} + 2\hat{\beta} \bullet m_{xz} + \hat{\beta}^2 \bullet m_{zz}} \right]. \tag{2.4}$$

where the null hypothesis is $\beta = \beta_0$. We confine our attention to the case $\beta_0 = 0$, noting that a non-zero value of β_0 corresponds to a transformed model with zero as the null hypothesis. For expository purposes we normalize the independent variables so that $m_{xx} = m_{zz} = 1$, and we denote their sample correlation as $\left(m_{xz} \big/ \sqrt{m_{xx} m_{zz}}\right) = \rho$. With this simplification the *t*-statistic for $\hat{\beta}$ is given by:

$$t_{\hat{\beta}}^2 = \frac{\hat{\lambda}^2}{\hat{\sigma}^2} \bullet N \bullet (1-\rho^2) \bullet \frac{1}{1+2\hat{\beta} \bullet \rho + \hat{\beta}^2} = t_{\hat{\lambda}}^2 \bullet \frac{1}{1+2\hat{\beta} \bullet \rho + \hat{\beta}^2} \qquad (2.5)$$

The second equality uses the fact that the first three factors are simply the *t*-statistic for $\hat{\lambda}$ in the reduced form linear regression of *y* on *x* and *z*. Since the reduced form is a classical linear regression this test has correct size, and it provides an alternative test of the null hypothesis $\beta = 0$. We note that if the standard error for $\hat{\beta}$ were computed under the null hypothesis instead of using point estimates, the two statistics would be numerically identical.

What can be said about the relative magnitude of the two *t*-statistics? It is clear that if the two explanatory variables are orthogonal, then in any given sample $t_{\hat{\beta}}^2 < t_{\hat{\lambda}}^2$ since the last term in (2.5) must be less than one. In that case the standard test will reject less often than its nominal size would indicate, a counter-intuitive result since the information measure $V_{\hat{\beta}}^{-1}$ for $\hat{\beta}$ is overestimated - as NS show. In the case that the explanatory variables are not orthogonal the outcome will depend on the distribution of $\hat{\beta}$ and how that is influenced by ρ. We now show that the distribution of $\hat{\beta}$ becomes concentrated when ρ is close to 1 or -1, but in either case the standard *t*-test based on $t_{\hat{\beta}}$ will reject too often.

Noting that $\hat{\lambda}$ and $\hat{\gamma}$ are jointly distributed Normal random variables, we may write:

$$\hat{\lambda} = \alpha + \kappa \bullet \hat{\gamma} + v \tag{2.6}$$

where α and κ are parameters implied by that joint distribution, and v is a Normal random variable that is uncorrelated with $\hat{\gamma}$. It is straightforward to show that in the standardized regressors case $\alpha = \rho \bullet \gamma$, $\kappa = -\rho$, and the variance of v is simply σ^2/N. Making these substitutions and dividing by $\hat{\gamma}$ one obtains:

$$\hat{\beta} = -\rho + \rho \bullet \left(\frac{\gamma}{\hat{\gamma}}\right) + \frac{v}{\hat{\gamma}} \tag{2.7}$$

If the standard deviation of $\hat{\gamma}$ is large relative to γ then the second term will typically be small, and if the standard deviation of $\hat{\gamma}$ is large relative to the standard deviation of v then the third term will also typically be small. That is indeed the situation when ρ is close to its bounds of 1 or -1 since the variance of $\hat{\gamma}$ is given by $\sigma^2 \bullet N^{-1}/(1-\rho^2)$ and can be made arbitrarily large. Thus, realizations of $\hat{\beta}$ will tend to concentrate around $-\rho$ when the regressors are strongly correlated.

(We note that a similar result holds for the instrumental variables estimator which in the simple just identified case is also a ratio of regression coefficients. A large literature discusses the IV case where concentration is associated with 'weak instruments' and the reader is referred to Phillips (1983) and Nelson and Startz (1990a, 1990b) for further discussion.)

Recalling from (2.5) that the multiplier applied to $t_{\hat{\lambda}}^2$ is the reciprocal of $(1+2\hat{\beta}\bullet\rho+\hat{\beta}^2)$, the effect of strong correlation between *x* and *z*, of either sign, working through the concentration of $\hat{\beta}$ around the value $-\rho$, is to drive this expression close to zero, making $t_{\hat{\beta}}^2$ very much larger than $t_{\hat{\lambda}}^2$. More generally, whether the test size is too large or too small depends on correlation between the regressors, regardless of sign.

To explore the degree of size distortion and its dependence on the information controlling parameter γ, sample size *N*, and correlation between regressors ρ, we have done a series of Monte Carlo experiments reported below. In each case the regressors have unit sample variance and are fixed in repeated samples, and the regression errors are i.i.d. *N*(0,1). The size of $t_{\hat{\lambda}}$ is of course exactly its nominal size – we focus on .05 – since the reduced form is a classical linear regression for this model. However, we will also be interested in its power against alternative hypotheses and how that compares to the power of the asymptotic test $t_{\hat{\beta}}$ as used in practice. Estimation is done in EViews™ using the non-linear regression routine. The number of replications is 10,000 in all the experiments in this paper, so the standard deviation of estimated size is .002 when the true size is .05.

Table 1 explores the effect of increasing γ when the regressors are orthogonal and *N*=100. The second line is the ratio $\gamma/\sqrt{V_{\hat{\gamma}}}$ which we expect to serve as a metric for spurious inference since from equation (1.3) the upward bias in estimating γ^2 is the variance of $\hat{\gamma}$ and it is the relative magnitude of these that plays a role both in estimating the standard error of $\hat{\beta}$ (the NS result) and also in the concentration of $\hat{\beta}$ away from zero. The third line presents the median of $\hat{\beta}$ rather than the mean since the moments of the ratio of normal random variables do not in general exist; see Fieller (1932) and Hinckley (1969). The estimated median and the inter-quartile range below it suggest that the sampling distribution of $\hat{\beta}$ is centered on zero and becomes more concentrated around zero with larger values of γ and more importantly with the metric $\gamma / \sqrt{V_{\hat{\gamma}}}$. The next two lines compare the true asymptotic information measure $\sqrt{1/V_{\hat{\beta}}} = \sqrt{I_{\hat{\beta}}}$ with the median of estimated values, expecting that it will be upward biased for small values of $\gamma / \sqrt{V_{\hat{\gamma}}}$, which is the case only at the smallest value reported in column one. Finally, the last row reports the empirical size of the *t*-test using $t_{\hat{\beta}}$, and it is evident that the test is undersized except at the largest value of the metric $\gamma / \sqrt{V_{\hat{\gamma}}}$. Thus, even when the standard error of $\hat{\beta}$ is well estimated, dispersion in $\hat{\beta}$ confounds the size of the conventional test. And the fact that the distribution of $t_{\hat{\beta}}$ depends on the unknown true value of γ demonstrates that it is non-pivotal in finite samples, its distribution depending instead on this nuisance parameter. We do not report the size of the test for $\lambda = 0$ because it is exactly .05 in this model.

Table 1: Monte Carlo Results for Orthogonal Regressors: *t*-test at nominal .05 level.
***N* = 100**

True γ	.01	.10	0.5	1.0
Asymptotic $\gamma / \sqrt{V_{\hat{\gamma}}}$	0.1	1	5	10
Median $\hat{\beta}$	0.10	0.03	-0.00	-0.00
Range (.25, .75)	(-0.95,1.17)	(-0.65,0.68)	(-0.14,0.13)	(-0.06,0.06)
Asymptotic $\sqrt{1/V_{\hat{\beta}}} = \sqrt{I_{\hat{\beta}}}$	0.1	1	5	10

Median $\sqrt{1/\hat{V}_{\hat{\beta}}} = \sqrt{\hat{I}_{\hat{\beta}}}$	0.43	0.85	5.00	10.13
Size of $t_{\hat{\beta}}$	0.0001	0.0002	0.0384	0.0505

Table 2 explores the effect of correlation between independent variables and its relation to the concentration of $\hat{\beta}$ and distortion of test size when the true value of γ is a 'moderate' .10. We note that the metric $\gamma/\sqrt{V_{\hat{\gamma}}}$ is smaller for larger values of the correlation, that the distribution of $\hat{\beta}$ does indeed become concentrated around $-\rho$, and test size becomes excessive as expected when correlation is strong. As observed above, when the metric $\gamma/\sqrt{V_{\hat{\gamma}}}$ is below unity the upward bias in estimated information for $\hat{\beta}$ is evident.

Table 2: Monte Carlo Results for Correlated Regressors: γ =.10; *t*-test at nominal .05 level, $N = 100$.

Correlation ρ	0	.50	.90	.99
Asymptotic $\gamma/\sqrt{V_{\hat{\gamma}}}$	1	0.87	0.44	0.14
Median $\hat{\beta}$	0.03	-0.21	-0.71	-0.96
Range (.25, .75)	(-0.65,0.68)	(-0.84,0.54)	(-1.27,-0.14)	(-1.17,-0.75)
Asymptotic $\sqrt{1/V_{\hat{\beta}}} = \sqrt{I_{\hat{\beta}}}$	1	0.87	0.44	0.14
Median $\sqrt{1/\hat{V}_{\hat{\beta}}} = \sqrt{\hat{I}_{\hat{\beta}}}$	0.85	0.81	0.99	2.67
Size of $t_{\hat{\beta}}$	0.0002	0.0194	0.2353	0.5653

Table 3 explores the response of rejection frequency to departures of the true value of β from the null value of zero, when the independent variables are uncorrelated and the true value of γ is again .10. If test size were correct, this would give us the power of the test against alternative hypotheses, but since test size is not generally correct it can at best suggest whether the test can convey some information about the null hypothesis. What we see is that although the inter-quartile range for $\hat{\beta}$ and the information measure is not distorted for this 'moderate' value of γ, the frequency of rejection increases very slowly as a function of the true β. In contract, the power of the correctly sized test $t_{\hat{\lambda}}$ rises steeply as the true β departs from zero.

Table 3: Monte Carlo Results for Orthogonal Regressors: *t*-test of $H_0 : \beta = 0$ when the alternative is true, at nominal .05 level, γ=.10, *N* = 100.

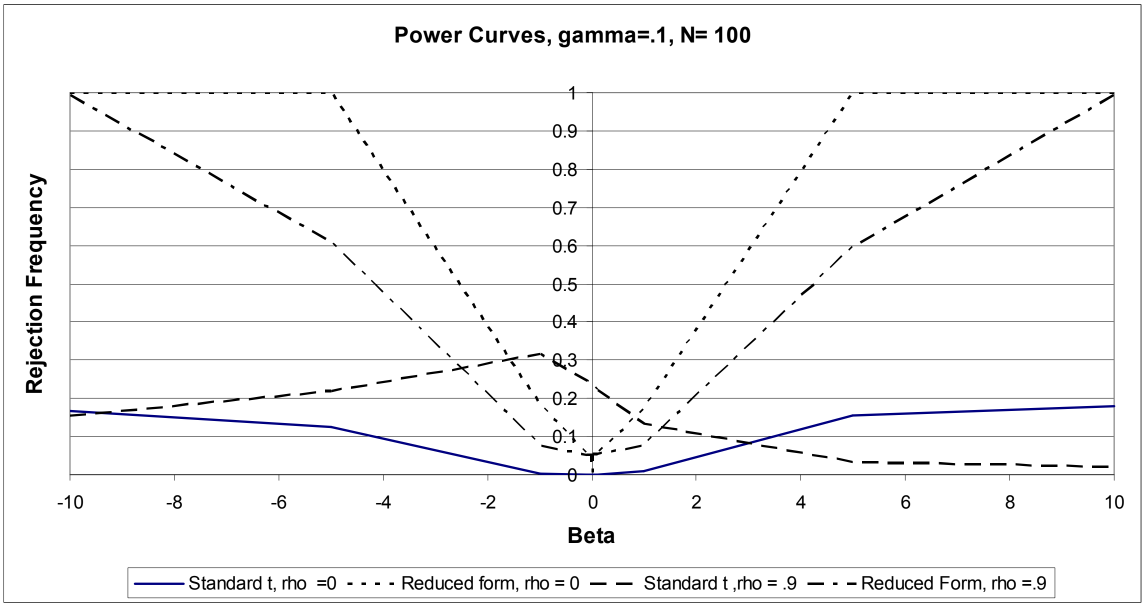
True β	.01	.10	1.0	5
Asymptotic $\gamma / \sqrt{V_{\hat{\gamma}}}$	1	1	1	1
Median $\hat{\beta}$	0.03	0.09	0.67	3.70
Range (.25, .75)	(-0.64,0.69)	(-0.57,0.75)	(0.01,1.51)	(2.23,6.64)
Asymptotic $\sqrt{1/V_{\hat{\beta}}} = \sqrt{I_{\hat{\beta}}}$	1	1.00	0.71	0.20
Median $\sqrt{1/\hat{V}_{\hat{\beta}}} = \sqrt{\hat{I}_{\hat{\beta}}}$	0.85	0.85	0.76	0.23
Frequency of rejection by $t_{\hat{\beta}}$	0.0010	0.0012	0.0086	0.1549
Frequency of rejection by $t_{\hat{\lambda}}$	0.0537	0.0556	0.1715	0.9993

Table 4 presents a similar comparison of rejection of the null under the alternative when the independent variables are strongly correlated. Strangely, for $t_{\hat{\beta}}$ rejections become less frequent as the true value of β departs farther from the null of zero, but the power of reduced form test $t_{\hat{\lambda}}$ does increase as expected. Figure 1 summarizes the results for 'power' of the two tests and their dependence on correlation between regressors.

Table 4: Monte Carlo Results for Correlated Regressors, ρ=.9, *t*-test of $H_0 : \beta = 0$ when the alternative is true, at nominal .05 level, γ=.10, *N* = 100.

True β	.01	.10	1.0	5
Asymptotic $\gamma / \sqrt{V_{\hat{\gamma}}}$	0.44	0.44	0.44	0.44
Median $\hat{\beta}$	-0.71	-0.68	-0.36	1.38
Range (.25, .75)	(-1.27,-0.14)	(-1.29,-0.09)	(-1.52,0.54)	(-0.93,4.52)
Asymptotic $\sqrt{1/V_{\hat{\beta}}} = \sqrt{I_{\hat{\beta}}}$	0.43	0.40	0.22	0.07
Median $\sqrt{1/\hat{V}_{\hat{\beta}}} = \sqrt{\hat{I}_{\hat{\beta}}}$	0.99	0.96	0.64	0.20
Frequency of rejection by $t_{\hat{\beta}}$	0.2344	0.2227	0.1344	0.0323
Frequency of rejection by $t_{\hat{\lambda}}$	0.0532	0.0543	0.0750	0.5966

Figure 1: Power Curves for tests of $H_0: \beta = 0$, N = 100, True γ = .1.



Asymptotic theory does take hold as sample size becomes large, but it does so very slowly as is evident in Table 5 below. We note that the size of $t_{\hat{\beta}}$ just approaches to its correct level only as the quantity $\gamma / \sqrt{V_{\hat{\gamma}}}$ increases to 10, requiring a sample size as large as 10,000 for $\gamma = .1$, and 1,000,000 for $\gamma = .01$!

Table 5: Monte Carlo Results for Orthogonal Regressors as sample size becomes large: *t*-test at nominal .05 level.

Sample Size N	100	10,000	1,000,000	10,000
True γ	.01	.01	.01	.1
Asymptotic $\gamma / \sqrt{V_{\hat{\gamma}}}$	.1	1	10	10
Median $\hat{\beta}$	0.10	0.01	-0.00	0.00
Range (.25, .75)	(-0.95, 1.17)	(-0.64, 0.63)	(-0.07, 0.07)	(-0.07, 0.07)
Asymptotic $\sqrt{1/V_{\hat{\beta}}} = \sqrt{I_{\hat{\beta}}}$	.1	1	10	10
Median $\sqrt{1/\hat{V}_{\hat{\beta}}} = \sqrt{\hat{I}_{\hat{\beta}}}$	0.43	0.84	9.96	9.89
Size of $t_{\hat{\beta}}$	0.0001	0.0006	0.0426	0.0447

To conclude this section we note that this strategy for obtaining an exact test for a ratio of regression coefficients has been known at least since Fieller (1954) who pointed out that under the null hypothesis $\lambda / \gamma = R$ the quantity $\hat{\gamma} \bullet R - \hat{\lambda}$ is a linear function of Normal random variables with mean zero with well defined variance. Our objective in this paper is to extend the idea to the case where $g(\beta, x_i)$ is not linear in β and the reduced form test based on linear approximation will not be exact. In the next section we explore how close this approach comes to achieving valid inference in non-linear regression, ARMA, GARCH, and unobserved components models.

3. Evaluation of the linearization test strategy in four nonlinear models.

3.1. A Production Function.

One example for regression form (1.1) with nonlinear *g(.)* is the Hicks-neutral Cobb-Douglas production function:

$$y_i = \gamma \bullet x_i^{\beta} + \varepsilon_i ; \gamma \neq 0 \qquad (3.1.1)$$

Here y_i and x_i are per capita output and capital input respectively and γ is Total Factor Productivity. Note that γ controls the amount of information data contains about β, an estimate of capital share.

To illustrate the potential spurious inference for β from this model we draw a sample of 100 observations of x_i from the log-normal distribution and pair it with 10,000 paths of simulated standard normal ε_i, each of sample size 100. Data *y* is then computed from (3.1.1) with various true parameter values. Estimation is done in EViews™ using the nonlinear regression routine. Table 6 reports estimation results for fixed value of γ at .01 and values of β in the economically relevant range from zero to .9. The negative bias in $\hat{\beta}$ is particularly notable here; for true values of .5 or larger the inter-quartile range does not include the true value. Recall from section 2 that the bias in $\hat{\beta}$ depends on the un-centered correlation ρ between the regressors. In this case the 'regressors' $g(\beta, x_i)$ and $g'(\beta, x_i)$ in the linear reduced-form and thus ρ depend on the value of β used to compute them. For positive β the value of ρ is positive and increasing with β, implying that $\hat{\beta}$ is downward biased as we see here. Note that the standard *t*-test rejects the null less often when the true β is zero but rejects too often as true β gets large. Also, note the size distortion is not as dramatic as in section 2 even as true β becomes as large. This attenuation is attributable to the fact that the non-linear estimation routine computes the regressors in the linear regression using the estimated $\hat{\beta}$ and the downward bias in that estimate attenuates the correlation between calculated regressors.

In Table 6 we also report the size of the reduced form test of $H_0 : \lambda = 0$, where $\lambda = \gamma \bullet (\beta - \beta_0)$, based on the linear approximation:

$$y_t = \gamma \bullet x_i^{\beta_0} + \lambda \bullet x_i^{\beta_0} \log(x_i) + e_i \qquad (3.1.2)$$

The proposed reduced-from test generally gives size close to correct in all cases.

Table 6: Monte Carlo Results for Production Function: *t*-test at nominal .05 level. True γ = .01, *N* = 100

True β	0	.1	.5	.9
Asymptotic $\gamma / \sqrt{V_{\hat{\gamma}}}$	0.10	0.10	0.09	0.11
Median $\hat{\beta}$	-0.04	-0.09	-0.05	0.12
Range (.25, .75)	(-0.53, 0.50)	(-0.59, 0.42)	(-0.56, 0.48)	(-0.41, 0.71)
Asymptotic $\sqrt{1/V_{\hat{\beta}}} = \sqrt{I_{\hat{\beta}}}$	0.11	0.11	0.16	0.25
Median $\sqrt{1/\hat{V}_{\hat{\beta}}} = \sqrt{\hat{I}_{\hat{\beta}}}$	1.10	1.09	1.06	1.03
Size of $t_{\hat{\beta}}$	0.0276	0.0367	0.1142	0.1786
Size of $t_{\hat{\lambda}}$	0.0530	0.0543	0.0543	0.0545

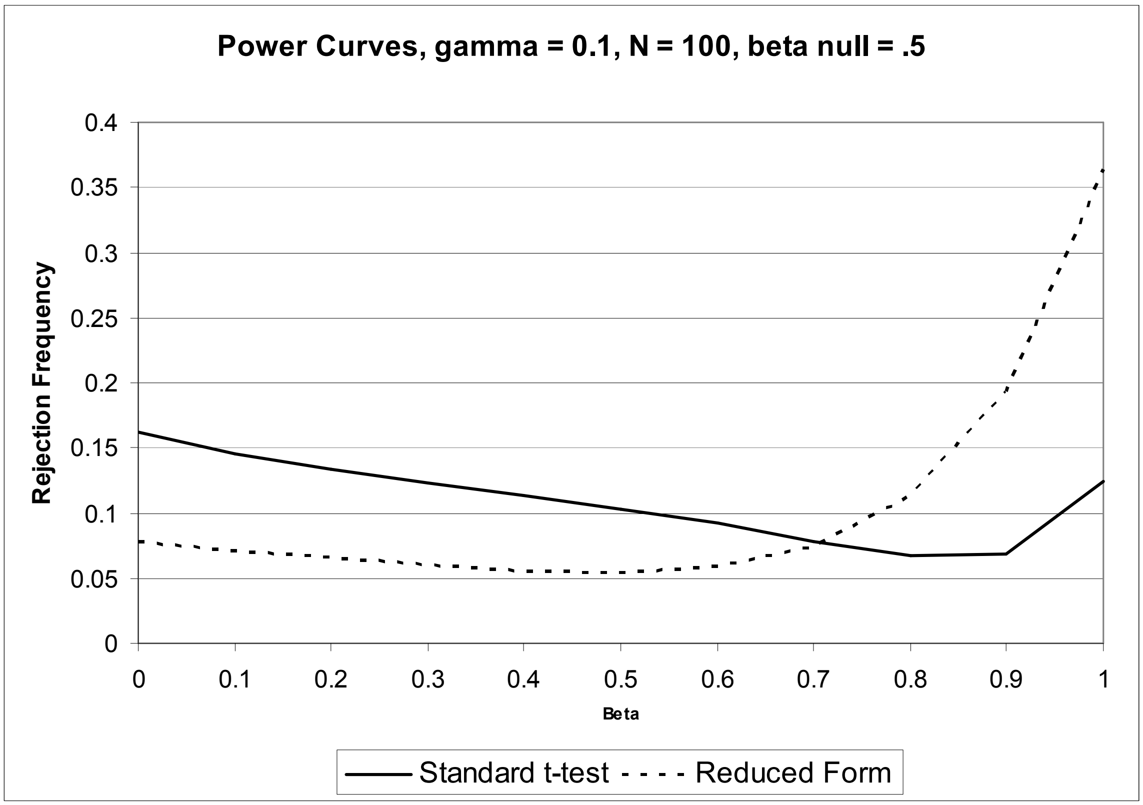
We report in Table 7 estimation results as γ increases from .01 to 1 for a fixed value of true β at .5. As the key metric $\gamma / \sqrt{V_{\hat{\gamma}}}$ approaches 10, the asymptotic distribution gradually takes hold and the actual size of the standard test is correct. We find that a value of 10 for this metric seems to be a rough rule of thumb for correct size of the standard test. The reduced-form test, however, maintains about the correct size across the range of parameter values considered here.

Table 7: Monte Carlo Results for Production Function: *t*-test at nominal .05 level, *N* = 100, true β = .5

True γ	.01	.1	1
Asymptotic $\gamma / \sqrt{V_{\hat{\gamma}}}$	0.09	0.91	9.10
Median $\hat{\beta}$	-0.05	0.27	0.50
Range (.25, .75)	(-0.56, 0.48)	(-0.26, 0.64)	(0.46, 0.54)
Asymptotic $\sqrt{1/V_{\hat{\beta}}} = \sqrt{I_{\hat{\beta}}}$	0.16	1.55	15.45
Median $\sqrt{1/\hat{V}_{\hat{\beta}}} = \sqrt{\hat{I}_{\hat{\beta}}}$	1.06	1.73	15.50
Size of $t_{\hat{\beta}}$	0.1142	0.1027	0.0523
Size of $t_{\hat{\lambda}}$	0.0543	0.0543	0.0543

Figure 2 below presents a power curve comparison between the reduced-form test and the standard *t*-test for testing the null $H_0 : \beta = .5$ as true β walks away from the null to both ends. The standard $t_{\hat{\beta}}$-test starts with a higher level of rejection frequency when the true β is .5, reflecting its size distortion and starts to decline as true β becomes higher than .5, only beginning climbing up as true β reaches as high as .9. The reduced form test using $t_{\hat{\lambda}}$, however, starts with correct size and has a higher level of rejection frequency *monotonically* as true β deviates further away from the null. As true β heads toward the left of the null, rejections by the standard *t*-test rise but not significantly more rapidly than the reduced form test. Neither test is very sensitive to departure from the null in the direction of zero. We surmise that the non-linearity of the model accounts for this asymmetry.

Figure 2: Power Curve for the test $H_0 : \beta = .5$, N = 100, True γ = .1.



3.2. The ARMA (1,1) Model.

The workhorse ARMA model in time series indirectly takes the form of (1.1). Consider a simple ARMA(1,1) model:

$$\begin{aligned} &(1-\phi L)y_t = (1-\theta L)\varepsilon_t ; t = 1,...,T \\ &\varepsilon_t \sim i.i.d.N(0,\sigma_\varepsilon^2), |\phi| < 1, |\theta| < 1 \end{aligned} \quad (3.2.1)$$

Given the invertibility in the moving average term, we may multiply both sides by $(1-\theta L)^{-1}$ to obtain:

$$y_t = \gamma \bullet g(\theta, \vec{y}_{t-1}) + \varepsilon_t . \quad (3.2.2)$$

Where $\gamma = \phi - \theta$, $g(\theta, \vec{y}_{t-1}) = \sum_{i=1}^{\infty} \theta^{i-1} y_{t-i}$ and $\vec{y}_{t-1} = (y_{t-1}, y_{t-2}, ...)$. NS show that when γ is small relative to the sample variance estimated information for either $\hat{\phi}$ or $\hat{\theta}$ is upward biased and the standard *t*-test rejects the null too often.

To achieve a valid test for θ, we may linearize *g(.)* around the null:

$$y_t = \gamma \bullet g(\theta_0, \vec{y}_{t-1}) + \lambda \bullet g_\theta(\theta_0, \vec{y}_{t-1}) + e_t . \quad (3.2.3)$$

Where $g_\theta(\theta, \vec{y}_{t-1}) = \frac{\partial g(\theta, \vec{y}_{t-1})}{\partial \theta} = \sum_{i=2}^{\infty} (i-1) \bullet \theta^{i-2} y_{t-i}$, $\lambda = \gamma \bullet (\theta - \theta_0)$ and e_i incorporates a remainder term. If the null $\theta = \theta_0$ is correct, the test statistic for the null $\lambda = 0$ in reduced form regression (3.2.3) should not be significant. In practice, to evaluate the regressors [$g(\theta_0, \vec{y}_{t-1})$, $g_\theta(\theta_0, \vec{y}_{t-1})$], we set y_t at its unconditional mean for all $t \leq 0$. To evaluate the improvement of this reduced form test relative to the conventional *t*-test, we implement a series of Monte Carlo experiments in EViews™ and report the results below.

Table 8: Monte Carlo Results for ARMA(1,1): *t*-test at nominal .05 level.

True θ = 0, *T* = 1,000.

True $\gamma(\phi)$	.01	.1	.2	.3
Asymptotic $\gamma / \sqrt{V_{\hat{\gamma}}}$	0.32	3.16	6.32	9.49
Median $\hat{\theta}$	-0.02	-0.01	-0.00	-0.00
Range (.25, .75)	(-0.65, 0.64)	(-0.26, 0.24)	(-0.11, 0.11)	(-0.07, 0.07)
Asymptotic $\sqrt{1/V_{\hat{\theta}}} = \sqrt{I_{\hat{\theta}}}$	0.32	3.16	6.32	9.49
Median $\sqrt{1/\hat{V}_{\hat{\theta}}} = \sqrt{\hat{I}_{\hat{\theta}}}$	2.75	3.89	6.59	9.62
Size of $t_{\hat{\theta}}$	0.4585	0.2237	0.1051	0.0734
Size of $t_{\hat{\lambda}}$	0.0506	0.0518	0.0526	0.0522

Table 8 explores the effect of increasing γ when true θ = 0 with a fixed sample size T = 1000. The first column where γ is .01 and the metric $\gamma / \sqrt{V_{\hat{\gamma}}}$ is only .32 confirms the results reported by NS: estimated information for $\hat{\theta}$ is too large and the conventional *t*-test rejects the null too often. In the last column, where γ is .3 corresponding to the key metric $\gamma / \sqrt{V_{\hat{\gamma}}}$ = 9.49, upward bias in estimated information is not evident and the size of conventional *t*-test is much closer to the nominal level 0.05. The fact that the sampling distribution of the conventional *t*-test statistic depends on the nuisance parameter γ implies again that the test is not pivotal. Note that the proposed $t_{\hat{\lambda}}$-test for the null θ = 0 is equivalent to testing the second lag in an AR(2) regression, which, is approximately the Box-Ljung *Q*-test with one lag for the residuals from an

AR(1) regression. The estimated size of the reduced form test is correct within sampling error.

The median and inter-quartile range of $\hat{\theta}$ suggest that the sampling distribution of $\hat{\theta}$ is centered on zero. However, the histogram of $\hat{\theta}$ in Figure 3 for the case γ = .01 suggests that the estimates tend to pile up at boundaries. The tendency for estimates in MA models to occur close to the boundary was noted previously by Kang (1975). Our preliminary analysis of this phenomenon in the ARMA case suggests that it is related to the bias effect demonstrated in section 2, but in this instance the correlation between regressors is itself a function of the parameter estimate. As the non-linear estimation routine iterates on θ and provisional values of $\hat{\theta}$ wander away from zero, the resulting correlation and bias pushes the estimate farther from zero. The relative success of the reduced form test comes from the fact that evaluates the test statistic under the null hypothesis θ_0 instead.

Table 9 explores the effect of increasing sample size mostly for the case when true γ = .01. Asymptotic theory does take hold, but the conventional *t*-test approaches correct size very slowly (requiring a sample size as large as 10,000 for γ = .1!). In contrast, the proposed test consistently has correct size within sampling error.

Figure 3: Histogram of $\hat{\theta}$ in the Monte Carlo. True γ = .01, θ = 0, *T* = 1,000.

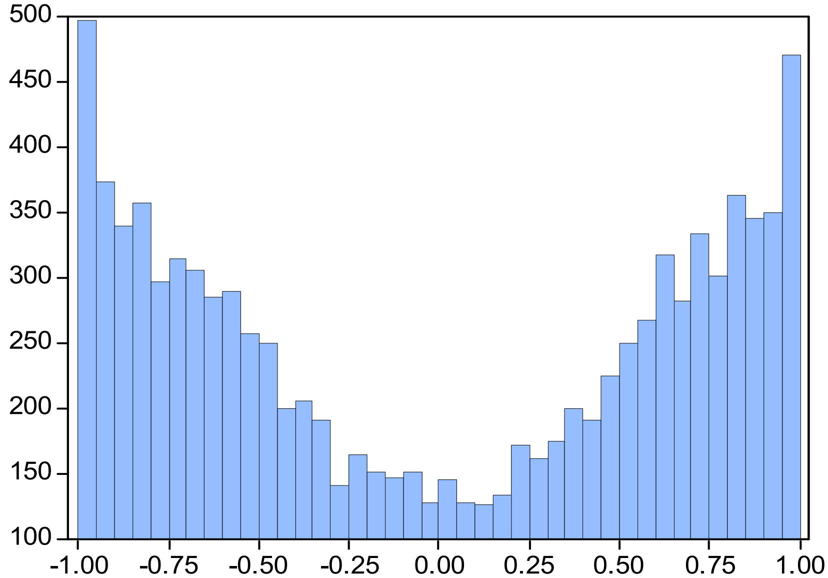


Table 9: Monte Carlo Results for ARMA(1,1): *t*-test at nominal .05 level for various sample sizes, True θ = 0.

Sample size	100	1000	10,000	10,000
True $\gamma(\phi)$	.01	.01	.01	.1
Asymptotic $\gamma/\sqrt{V_{\hat{\gamma}}}$	0.1	0.32	1	10
Median $\hat{\theta}$	-0.04	-0.02	-0.02	-0.00
Range (.25, .75)	(-0.69, 0.67)	(-0.65, 0.64)	(-0.58, 0.55)	(-0.07, 0.07)
Asymptotic $\sqrt{1/V_{\hat{\theta}}} = \sqrt{I_{\hat{\theta}}}$	0.1	0.32	1	10
Median $\sqrt{1/\hat{V}_{\hat{\theta}}} = \sqrt{\hat{I}_{\hat{\theta}}}$	2.84	2.75	2.74	10.18
Size of $t_{\hat{\theta}}$	0.4826	0.4585	0.3995	0.0660
Size of $t_{\hat{\lambda}}$	0.0506	0.0506	0.0487	0.0479

Often it is the AR root ϕ that is of a greater economic interest since it measures persistence. For example, if consumption growth g_t follows an ARMA (1,1) process, then one can show that the economic agent's conditional expectation of consumption growth is governed by the following AR(1) process:

$$E[g_{t+1} \mid I_t] = \mu + \phi \bullet E[g_t \mid I_{t-1}] + \gamma \bullet \varepsilon_t . \tag{3.2.4}$$

A large value of ϕ implies that any shock to the economy persists for a long period in agent's expectation. Recently, Bansal and his coauthors (see e.g., Bansal and Yaron (2000, 2004) and Bansal and Lundblad (2002)) show that a high level of persistence, interpreted as long-run risk, may explain the equity premium puzzle of Mehra and Prescott (1985) and the volatility puzzle of Shiller (1981) and LeRoy and Porter (1981). Ma (2007) finds that the estimated ARMA(1,1) implies a small estimated γ relative to its sampling variance and explores the implications of possible test size distortion in the conventional test as well as valid inference following the strategy suggested in this paper.

To obtain the reduced form test statistic for ϕ, rewrite the ARMA(1,1) model by multiplying both sides of (3.2.1) by $(1-\phi L)^{-1}$ given it is stationary:

$$y_t = \gamma \bullet g(\phi, \vec{\varepsilon}_{t-1}) + \varepsilon_t \tag{3.2.5}$$

Where $g(\phi, \vec{\varepsilon}_{t-1}) = \sum_{i=1}^{\infty} \phi^{i-1} \varepsilon_{t-i}$ and $\vec{\varepsilon}_{t-1} = (\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots)$. Taking a linear approximation of *g(.)* around the null, we propose to test $\lambda = 0$ in the following regression:

$$y_t = \gamma \bullet g(\phi_0, \vec{\varepsilon}_{t-1}) + \lambda \bullet g_\phi(\phi_0, \vec{\varepsilon}_{t-1}) + e_t \tag{3.2.6}$$

Where $g_\phi(\phi, \vec{\varepsilon}_{t-1}) = \dfrac{\partial g(\phi, \vec{\varepsilon}_{t-1})}{\partial \phi} = \sum_{i=2}^{\infty} (i-1) \bullet \phi^{i-2} \varepsilon_{t-i}$, and $\lambda = \gamma \bullet (\phi - \phi_0)$.

Since ε is not observable, we need a consistent estimate for ε to evaluate the regressors in (3.2.6) so as to make this test feasible. One way of doing this is to implement a restricted estimation by imposing the null and then compute $\tilde{\varepsilon}$ with restricted estimates. We experiment with this idea by generating data with true parameter values $\gamma = 0.1, \phi = 0, \sigma_{\varepsilon} = 1$ and sample size T = 100. Estimation is done in EViews™. The rejection frequency of the proposed test is 0.0461, at the nominal level 0.05, in contrast to 0.4233, that of standard *t*-test.

This reduced-form test may be generalized to address an ARMA model of any arbitrarily higher order, for example, an ARMA(*p*,*q*) model as follows:

$$[1-\phi_p(L)]y_t = [1-\theta_q(L)]\varepsilon_t; t = 1,...,T, \varepsilon_t \sim i.i.d.N(0,\sigma_{\varepsilon}^2) \qquad (3.2.7)$$

Where $\phi_p(L) = \sum_{i=1}^{p} \phi_i L^i$, $\theta_q(L) = \sum_{i=1}^{q} \theta_i L^i$, and the roots for $1-\phi(z) = 0$ and $1-\theta(z) = 0$ are all outside unit circle. A general representation similar to (3.2.2) may be obtained as:

$$y_t = \gamma_1 \bullet [(1-\theta_m(L))^{-1} \bullet y_{t-1}] + \cdots + \gamma_m \bullet [(1-\theta_m(L))^{-1} \bullet y_{t-m}] + \varepsilon_t \qquad (3.2.8)$$

Here $\gamma_k = \phi_k - \theta_k, 1 \le k \le m$, $m = \max(p,q)$, and $\phi_k = 0$ for $p < k \le m$ or $\theta_k = 0$ for $q < k \le m$. To test the null $\theta_k = \theta_{k,0}, 1 \le k \le q$, we suggest to linearize the last term associated with y_{t-m} to obtain the following regression with *q* augmented terms:

$$\begin{aligned} y_t &= \gamma_1 \bullet [(1-\theta_{m,0}(L))^{-1} \bullet y_{t-1}] + \cdots + \gamma_m \bullet [(1-\theta_{m,0}(L))^{-1} \bullet y_{t-m}] \\ &+ \lambda_1 \bullet [(1-\theta_{m,0}(L))^{-2} \bullet y_{t-(m+1)}] + \cdots + \lambda_q \bullet [(1-\theta_{m,0}(L))^{-2} \bullet y_{t-(m+q)}] + e_t \end{aligned} \qquad (3.2.9)$$

Where $\lambda_k = \gamma_k \bullet (\theta_k - \theta_{k,0}), 1 \le k \le q$. Again, if the null is correct the first *m* terms on the right hand side of (3.2.9) should be enough to capture the serial correlation in the data and the remaining *q* terms should not be significant. Notice here to compute the regressors for nonzero $\theta_{k,0}$'s, the coefficients $\varphi_{l,j}$'s in $\sum_{j=0}^{\infty} \varphi_{l,j} L^j = (1-\theta_{m,0}(L))^{-l}, l = 1,2$ may be obtained as the (1,1) element of matrix $(F_l)^j$, where F_l is the $(l \times m)$ by $(l \times m)$ transition matrix $(1-\theta_{m,0}(L))^l, l = 1,2$ in the state-space representation.

We apply this idea to an ARMA(2,2) model. With true parameter values $\phi_1 = 0.01, \phi_2 = 0.01, \theta_1 = 0, \theta_2 = 0, \sigma_\varepsilon = 1$ and sample size T = 100 we find that the standard *t*-test for $\hat{\theta}_1$ and $\hat{\theta}_2$ has empirical sizes of 0.5712 and 0.6981 at a nominal level 0.05. In contrast the proposed test for $\lambda_1 = 0$ and $\lambda_2 = 0$ based on regression (3.2.9) gives rejection frequencies of 0.0491 and 0.0487 respectively. Notice here since the null is $\theta_1 = 0$ and $\theta_2 = 0$, our proposed test is equivalent to testing the third and fourth lag in an AR(4) regression.

3.3. The Unobserved Component Model for Decomposing Trend and Cycle

The Unobserved Component model (hereafter UC) is proposed by Harvey (1985) and Clark (1987) to decompose GDP into a stochastic trend and cycle. In contrast to the decomposition of Beveridge and Nelson (1981) which typically attributes most output variation to trend, the UC decomposition tends to imply output variation is mostly due to a persistent cycle component. Morley, Nelson and Zivot (2003) find that this seeming contradiction is due to a parameter restriction typically imposed in UC models. Here we will show that small sample inference for persistence in the cycle component is also problematic.

In the UC model, the log of real GDP y_t is assumed to be a sum of unobserved stochastic trend τ_t and cycle c_t:

$$y_t = \tau_t + c_t \,. \tag{3.3.1}$$

Trend is a random walk with drift:

$$\tau_t = \tau_{t-1} + \mu + \eta_t, \eta_t \sim i.i.d N(0, \sigma_\eta^2) \,. \tag{3.3.2}$$

Cycle has a stationary ARMA representation; here we take the AR(1) as an example:

$$(1 - \phi L)c_t = \varepsilon_t, \varepsilon_t \sim i.i.d N(0, \sigma_\varepsilon^2) \,. \tag{3.3.3}$$

The UC model is estimated by maximizing the likelihood computed using the Kalman filter. (In practice one typically obtains a large value of $\hat{\phi}$ with a tight confidence interval. However, the information of $\hat{\phi}$ may be overestimated and standard inference of it may be spurious when the true cycle innovation variance (σ_ε^2) is small relative to the true trend innovation variance (σ_η^2).) Following Morley, Nelson and Zivot (2003), we note that the univariate representation of this particular UC model is ARMA(1,1) with parameters implied by the equality:

$$(1-\phi L)\Delta y_t = \mu(1-\phi)+(1-\phi L)\eta_t + \varepsilon_t - \varepsilon_{t-1} = \mu(1-\phi)+u_t - \theta u_{t-1} \qquad (3.3.4)$$

Where $u_t \sim i.i.d.N(0,\sigma_u^2)$. Thus the AR coefficient of the ARMA(1,1) is simply ϕ, while the MA parameter θ is identified (under the restriction $\sigma_{\eta,\varepsilon}=0$) by matching the zero and first-order autocovariances of the equivalent MA parts:

$$\psi_0 = (1+\phi^2)\sigma_\eta^2 + 2\sigma_\varepsilon^2 + 2(1+\phi)\sigma_{\eta\varepsilon}^2 = (1+\theta^2)\sigma_u^2 \qquad (3.3.5)$$

$$\psi_1 = -\phi\sigma_\eta^2 - \sigma_\varepsilon^2 - (1+\phi)\sigma_{\eta\varepsilon}^2 = -\theta\sigma_u^2 \qquad (3.3.6)$$

We may then solve for a unique θ by imposing invertibility, obtaining:

$$\theta = \frac{(1+\phi^2)+2(\frac{\sigma_\varepsilon^2}{\sigma_\eta^2})+2(1+\phi)\rho_{\eta\varepsilon}(\frac{\sigma_\varepsilon}{\sigma_\eta}) - \sqrt{[(1+\phi)^2+4(\frac{\sigma_\varepsilon^2}{\sigma_\eta^2})+4(1+\phi)\rho_{\eta\varepsilon}(\frac{\sigma_\varepsilon}{\sigma_\eta})]\bullet[(1-\phi)^2]}}{2[\phi+\frac{\sigma_\varepsilon^2}{\sigma_\eta^2}+(1+\phi)\rho_{\eta\varepsilon}(\frac{\sigma_\varepsilon}{\sigma_\eta})]} \qquad (3.3.7)$$

It is straightforward to show that θ comes arbitrarily close to ϕ as ${\sigma_\varepsilon}/{\sigma_\eta}$ approaches zero. By an analogy to ARMA model, the information for $\hat{\phi}$ may be overestimated when ${\sigma_\varepsilon}/{\sigma_\eta}$ is small relative to sample variation, and the uncertainty of cycle estimate $\hat{c}_t$ may appear too small.

To visualize spurious inference in this case, we have implemented a Monte Carlo experiment. Data is generated from the UC model as laid out by (3.3.1) – (3.3.3) with true parameter values $\mu=0.8, \phi=0, \sigma_\eta^2=0.95, \sigma_\varepsilon^2=0.05$ and the estimation is done in MATLAB 6.1. The sample size T is chosen to be 200 since it is approximately the one encountered in practice for postwar quarterly GDP dataset. To avoid local maxima, various starting values are used.

The standard *t*-test for $\hat{\phi}$ indeed rejects the null much too often; size is 0.4810 at the nominal level 0.05. This is partly because the standard error for $\hat{\phi}$ is underestimated; the median is 0.2852 compared with its true value 1.4815. Furthermore, $\hat{\phi}$ is upward biased as illustrated in Figure 4, its median being 0.58. Many $\hat{\phi}'s$ occur close to the positive boundary. This is consistent with Nelson's (1988) finding that a UC model with persistent cycle variation fits better than the true model even when all variation is due to stochastic trend (i.e., $\sigma_{\varepsilon}^{2}=0$).

At the same time, the cycle innovation variance estimate $\hat{\sigma}_{\varepsilon}^{2}$ is upward biased, having a median of .20, while the trend innovation variance estimate $\hat{\sigma}_{\eta}^{2}$ is instead downward biased, with a median of .73. What is the underlying driving force for the upward bias of $\hat{\phi}$ and $\hat{\sigma}_{\varepsilon}^{2}$ and the downward bias of $\hat{\sigma}_{\eta}^{2}$? The scatter plot in Figure 5 shows that there is a positive co-movement between $\hat{\phi}$ and $\hat{\sigma}_{\varepsilon}^{2}$, thus persistence in the estimated cycle tends to occur in estimations that also show large variance in the cycle. This is driven by the necessity that the model must account for the small amount of serial correlation in our data generating process for Δy_t. Setting its autocovariance at lag one equal to the true value for the sake of illustration, one would have combinations of the parameters that satisfy $-\frac{1-\phi}{1+\phi}\bullet\sigma_{\varepsilon}^{2}=-.05$. One solution would be the true combination $\phi=0;\sigma_{\varepsilon}^{2}=.05$, and another is $\phi=.9;\sigma_{\varepsilon}^{2}=.95$. Thus $\hat{\sigma}_{\varepsilon}^{2}$ will be far greater than its true value when $\hat{\phi}$ is close to its positive boundary, implying a dominating persistent cycle that tends to mimic the true underlying stochastic trend. Finally, Figures 4 and 5 show that large negative values of $\hat{\phi}$ are possible, but occur only infrequently because positive variances place restrictions on the parameter space.

In light of the connection between UC model and ARMA model we suggest implementing the reduced-form test in the following steps: first impose the null $\phi = \phi_0$ and estimate all other parameters in the UC model; secondly, take advantage of (3.3.7) to compute the implied restricted estimate $\tilde{\theta}$ and $\tilde{u}_t$ in the reduced-form ARMA(1,1) model; lastly, resort to regression (3.2.6) to compute the reduce-form test statistic. Using the same set of simulated data in the above MC experiment, the reduced-form test statistic for ϕ has the rejection frequency of 0.054, close to the nominal level 0.05. As illustrated in the ARMA section, the above strategy can be easily generalized to address a UC model with higher AR orders in the cycle component.

Figure 4: Plot of $\hat{\phi}$ in the Monte Carlo Experiment with true parameter

$$\mu = 0.8, \phi = 0, \sigma_{\eta}^{2} = 0.95, \sigma_{\varepsilon}^{2} = 0.05$$

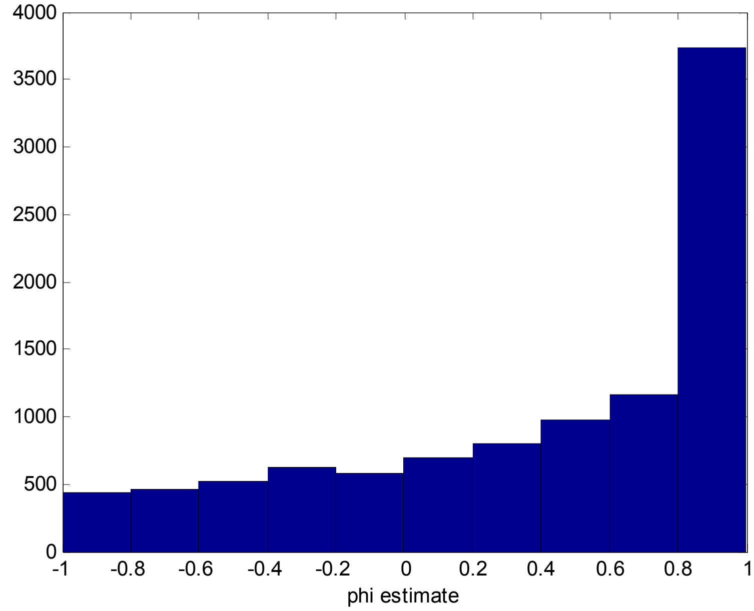
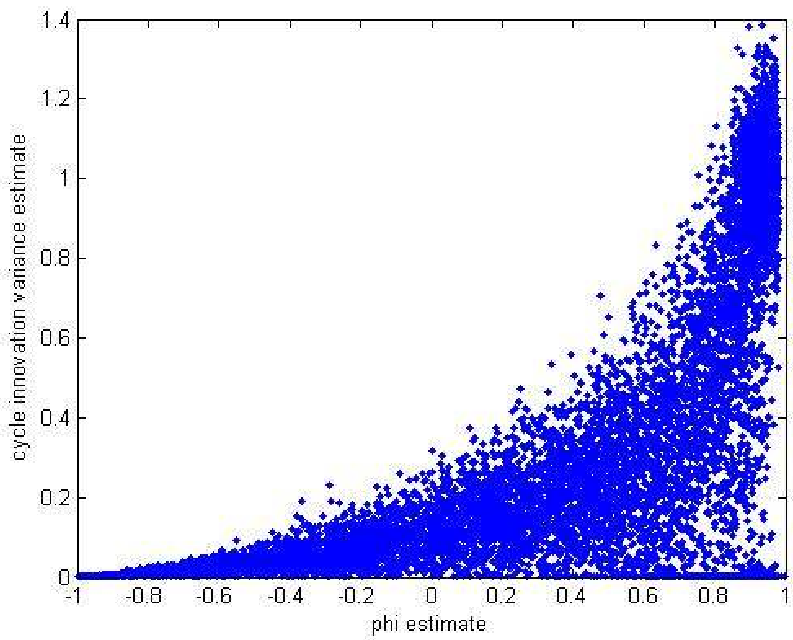


Figure 5: Scatter Plot of $\hat{\phi}$ and $\hat{\sigma}_{\varepsilon}^{2}$



3.4. The GARCH(1,1) Model.

The GARCH(1,1) developed by Bollerslev (1986) is perhaps one of the most popular approaches in capturing the time-varying volatility for time series data. Despite a great deal of rich extensions, the archetypal GARCH(1,1) model may be written:

$$\varepsilon_t = \sqrt{h_t} \cdot \xi_t, \xi_t \sim i.i.d. N(0,1) \qquad (3.4.1)$$

$$h_t = \omega + \alpha \cdot \varepsilon_{t-1}^{\ 2} + \beta \cdot h_{t-1} \qquad (3.4.2)$$

To see why the GARCH is among the models we are concerned with, write out the ARMA(1,1) representation of this GARCH model and make an analogy to ARMA model:

$$\varepsilon_t^{\ 2} = \omega + (\alpha + \beta) \cdot \varepsilon_{t-1}^{\ 2} + w_t - \beta \cdot w_{t-1} \qquad (3.4.3)$$

Where the innovation $w_t = \varepsilon_t^{\ 2} - h_t = h_t(\xi_t^2 - 1)$ is a Martingale Difference Sequence (MDS) with a time-varying volatility, and the $\alpha + \beta$ and β correspond to the AR and MA roots. Ma, Nelson and Startz (2007) show that in the GARCH(1,1) model as the identifying parameter α is small, relative to sample variation, the standard error for $\hat{\beta}$ is underestimated and the standard *t*-test rejects the null too often, implying a significant GARCH effect even when there is very little.

Our reduced-form test can be easily extended to the GARCH model to achieve a much better inference. If one wants to test the null $\beta = \beta_0$, write (3.4.2) into the following:

$$h_t = \frac{\omega}{1-\beta} + \alpha \cdot g(\beta, \vec{\varepsilon}_{t-1}^{\,2}) \qquad (3.4.4)$$

Where $g(\beta, \vec{\varepsilon}_{t-1}^{\,2}) = \sum_{i=1}^{\infty} \beta^{i-1} \varepsilon_{t-i}^2$ and $\vec{\varepsilon}_{t-1}^{\,2} = (\varepsilon_{t-1}^2, \varepsilon_{t-2}^2, ...)$. Take a linear expansion of $g(.)$ in the variance term around the null:

$$h_t = c + \alpha \cdot g(\beta_0, \vec{\varepsilon}^{\,2}_{t-1}) + \lambda \cdot g_\beta(\beta_0, \vec{\varepsilon}^{\,2}_{t-1}) \qquad (3.4.5)$$

Where $c = \dfrac{\omega}{1-\beta}$, $\lambda = \alpha \bullet (\beta - \beta_0)$, and $g_\beta(\beta, \vec{\varepsilon}^{\,2}_{t-1}) = \sum_{i=2}^{\infty} (i-1) \cdot \beta^{i-2} \varepsilon^2_{t-i}$. Our reduced-form test is the *t*-stat for $\lambda = 0$ in (3.4.5). To evaluate the performance of the test, we implement a Monte Carlo experiments for a few sets of relevant parameter values with a fixed sample size $T = 1000$. The estimation is done in MATLAB 6.1.

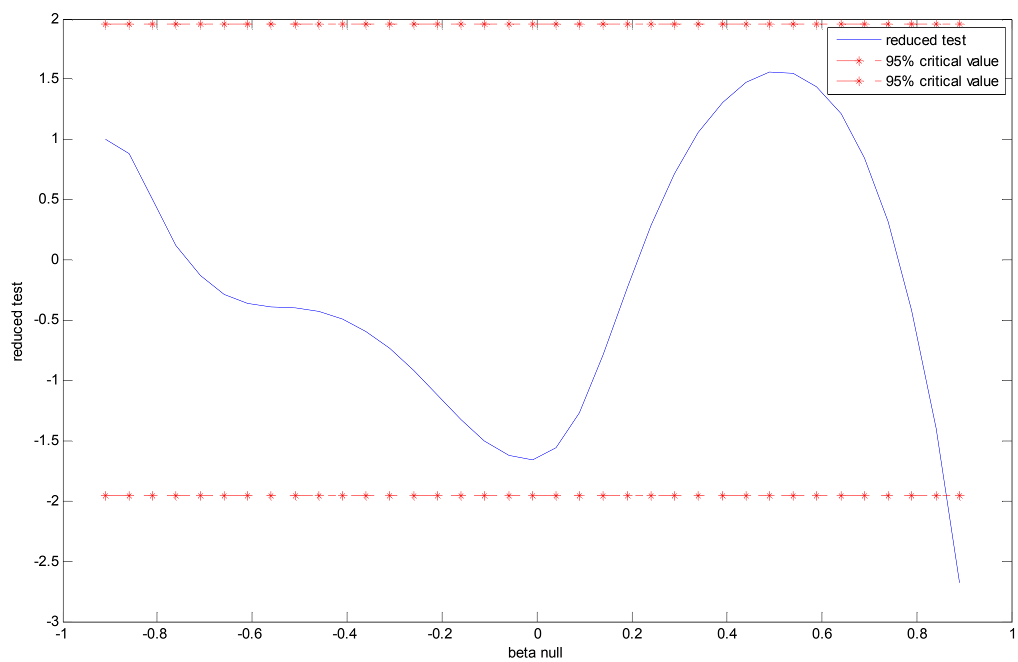
Table 10: Monte Carlo Results for GARCH(1,1): *t*-test at nominal .05 level. True $\beta = 0$, $T = 1{,}000$.

True $\gamma(\alpha)$	.01	.05	.1	.2
Asymptotic $\gamma / \sqrt{V_{\hat{\gamma}}}$	0.32	1.59	3.19	6.60
Median $\hat{\beta}$	0.33	0.08	-0.00	-0.01
Range (.25, .75)	(-0.30,0.74)	(-0.31,0.49)	(-0.22,0.22)	(-0.11,0.09)
Asymptotic $\sqrt{1/V_{\hat{\beta}}} = \sqrt{I_{\hat{\beta}}}$	0.32	1.59	3.19	6.60
Median $\sqrt{1/\hat{V}_{\hat{\beta}}} = \sqrt{\hat{I}_{\hat{\beta}}}$	3.13	3.42	4.13	6.93
Size of $t_{\hat{\beta}}$	0.4697	0.3436	0.1978	0.1063
Size of $t_{\hat{\lambda}}$	0.0778	0.0736	0.0761	0.0957
Size of $t_{\hat{\rho}}$	0.4690	0.3381	0.1877	0.0781
Size of $t_{\hat{\lambda}^*}$	0.0723	0.0685	0.0706	0.0698

Table 10 presents the estimation results along with comparisons of the reduced-form test and the standard *t*-test as α increases when the true $\beta = 0$ and sample size *T* = 1000. For the first column where the key metric $\gamma / \sqrt{V_{\hat{\gamma}}}$ is 0.32, indicating that α is very small relative to sample variation, the information of $\hat{\beta}$ is much overestimated. This upward bias declines as α becomes larger but is still visible for α as large as .2. As a result, standard *t*-test rejects the null too often. Contrary to the severe size distortion of standard *t*-test, our reduced-form test is consistently very close to the nominal level .05. We notice that for the case $\alpha = .01$, there is a tendency for $\hat{\beta}$ to occur close to boundary at both ends, but slightly more often on the right than left, leading to an upward bias of $\hat{\beta}$.

Next we show how to apply our reduced-form test strategy to a real dataset. We take an example used by Ma, Nelson and Startz (2007), the monthly S&P 500 index return data, to illustrate how to obtain a confidence interval for $\hat{\beta}$ based on our reduced-form test. This dataset is from the Eviews 5.1 DRI Database. We focus on the sample period from 1947 January to 1984 September to make our result comparable to Bollerslev (1987). The GARCH estimates along with the Bollerslev and Wooldridge (1992) robust standard errors (in parenthesis), after accounting for the "Working" effect (see Working (1960)) in level, are reported from an estimation by MATLAB 6.1: $\hat{\omega} = 0.16 \cdot 10^{-3} (0.14 \cdot 10^{-3})$, $\hat{\alpha} = 0.077(0.048)$, $\hat{\beta} = 0.773(0.169)$. Standard *t*-test seems to imply a significantly large GARCH effect and the 95% confidence interval for β is [0.44, 1). We can instead numerically invert our reduced form test statistic, by creating a grid of β_0's, computing the corresponding $t_{\hat{\lambda}}$'s and plotting the latter against the former (see Figure 6), to obtain a 95% confidence interval for β: [-0.95, 0.87]. The confidence interval based on our reduced form test is fairly wide and does not indicate a significant GARCH effect.

Figure 6: 95% Confidence Interval for $\hat{\beta}$ based on $t_{\hat{\lambda}}$ for the monthly S&P 500 stock return data



Very often it is the sum $\alpha + \beta$ that is of a great interest in terms of its economic implications, since the volatility h_t is governed by a particular AR(1) process with $\alpha + \beta$ being the persistence measure:

$$h_t = \omega + (\alpha + \beta) \cdot h_{t-1} + \alpha \cdot w_{t-1} \tag{3.4.6}$$

Bansal and Yaron (2000, 2004) show that a large value of $\alpha + \beta$, interpreted as long run risk in uncertainty dynamics, may help to resolve a few asset pricing anomalies along with the non-separable utility specification. However, when α is small relative to sample variation, the standard inference for persistence estimate $\hat{\alpha} + \hat{\beta}$ may not be reliable and Ma (2007) discusses the economic implication of a corrected inference. To obtain a valid inference for $\hat{\alpha} + \hat{\beta}$, re-write the variance equation:

$$h_t = \frac{\omega}{1-\rho} + \alpha \bullet g(\rho, \vec{w}_{t-1}) \tag{3.4.7}$$

Where $\rho = \alpha + \beta$, $\vec{w}_{t-1} = (w_{t-1}, w_{t-2}, ...)$. Take a linear expansion of *g(.)* around the null ρ_0:

$$h_t = \frac{\omega}{1-\rho} + \alpha \cdot g(\rho_0, \vec{w}_{t-1}) + \lambda^* \cdot g_\rho(\rho_0, \vec{w}_{t-1}) \tag{3.4.8}$$

Where $\lambda^* = \alpha \bullet (\rho - \rho_0)$ and the reduced form test is the *t*-stat for λ^*. To make the reduced form test feasible one needs to have a consistent estimate for w_t which is readily obtained through a restricted GARCH estimation with imposed null.

Using the same set of simulated data we evaluate the performance of our reduced-form test for $\hat{\rho}$ and find it consistently has an empirical size close to the nominal level 0.05 while the standard *t*-test suffers from the size distortion of similar magnitudes to that of $\hat{\beta}$ (see Table 10).

4. Summary and Conclusions

This paper considers models that take the form $y = \gamma \bullet g(\beta, x) + \varepsilon$, where β is the parameter of interest and the identifying parameter γ is non-zero. Inference is problematic because the standard error for $\hat{\beta}$ depends on $\hat{\gamma}$. Nelson and Startz (2007) showed that although that standard error is downward biased in a broad class of models including this one that satisfy the Zero-Information-Limit-Condition, the *t*-statistic can be either too large or too small depending on the data generating process. In the special case $g(\beta, \mathbf{x}) = (x + \beta z)$, $\hat{\beta}$ is the ratio of regression coefficients and its finite sample distribution, as well as that of the *t*-ratio based on asymptotic theory, will depend on the correlation between regressors *x* and *z* as discussed in section 2. Fortunately, we may obtain an exact test in this case, following Fieller (1954). The remainder of this paper explores how well this approach works when $g(\beta, x)$ is not linear but we use a linear approximation instead. Applications to a production function, ARMA, GARCH, and Unobserved Components time series models uniformly suggest that the approximation strategy works well in sample sizes and parameter values encountered in econometric practice, while standard inference based on asymptotic theory suffers from severe size distortions.

References

Bansal, R., and A. Yaron (2000): "Risks for the long run: A potential resolution of asset pricing puzzles," NBER Working Paper 8059.

Bansal, R., and C. Lundblad (2002): "Market Efficiency, Asset Returns, and the Size of The Risk Premium in Global Equity Markets," *Journal of Econometrics*, 109, 195-237.

Bansal, R., and A. Yaron (2004): "Risks for the Long Run: a Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, LIX, 1481-1509.

Beveridge, S. and C.R.Nelson (1981): "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the Business Cycle," *Journal of Monetary Economics*, 7, 151-174.

Bollerslev, T. (1986): "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307-327.

Bollerslev, T. (1987): "A Conditional Heteroskedastic Time SeriesM for Speculative Prices and Rates of Return," *The Review of Economics and Statistics*, 69, 542-547.

Bollerslev, T., and J. Wooldridge (1992): "Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-varying Covariances," *Econometric Reviews*, 11, 143-172.

Clark, P.K. (1987): "The Cyclical Component of U.S. Economic Activity," *The Quarterly Journal of Economics*, 102, 797-814.

Davies, R. B. (1977): "Hypothesis Testing When a Nuisance Parameter is Present Only Under the Alternative," *Biometrika*, 64, 247-254.

Davies, R. B. (1987): "Hypothesis Testing When a Nuisance Parameter is Present Only Under the Alternative," *Biometrika*, 74, 33-43.

Fieller, E. C. (1932): "The Distribution of the Index in a Normal Bivariate Population," *Biometrika*, 24, 428-440.

Fieller, E. C. (1954): "Some Problems in Interval Estimation," *Journal of the Royal Statistical Society. Series B (Methodological)*, 16, 175-85.

Hansen, Bruce E. (1996): Inference When a Nuisance Parameter is Not Identified Under the Null Hypothesis," *Econometrica*, 64, 413-430.

Harvey, A. C. (1985): "Trends and Cycles in Macroeconomic Time Series," *Journal of Business and Economic Statistics*, 3, 216-27.

Hinkley, D. V. (1969): "On the Ratio of Two Correlated Normal Random Variables," *Biometrika*, 56 (3), 635-639.

Kang, K.M. (1975): "A Comparison of Estimators for Moving Average Processes," working paper, Australian Bureau of Statistics.

LeRoy, Stephen F., and Richard D. Porter, 1981, The present value relation: Tests based on implied variance bounds, *Econometrica* 49, 555-574.

Ma, J. (2007): "The Long-Run Risk in Consumption and Equity Premium Puzzle: New Evidence Based on Improved Inference," working paper, Department of Economics, Finance and Legal Studies, University of Alabama.

Ma, J., C. R. Nelson., and R. Startz (2007): "Spurious in the GARCH(1,1) Model When It Is Weakly Identified," *Studies in Nonlinear Dynamics & Econometrics*, Vol.11, No. 1, Article 1.

Mehra, Rajnish, and Edward C. Prescott, 1985, The equity premium: A puzzle, *Journal of Monetary Economics* 15, 145-161.

Morley, James, C.R.Nelson and E.Zivot (2003): "Why Are Beveridge-Nelson and Unobserved Component Decomposition of GDP So Different?" *Review of Economics and Statistics*, Vol. 85, No.2, 235-243.

Nelson, C. R. (1988): "Spurious Trend and Cycle in the State Space Decomposition of a Time Series with a Unit Root," *Journal of Economic Dynamics & Control*, 12, 475-488.

Nelson, C R., and R. Startz (1990a): "The Distribution of the Instrumental Variables Estimator and Its t-Ratio When the Instrument Is a Poor One," *The Journal of Business*, 63, 125-40.

Nelson, C. R., and R. Startz (1990b): "Some Further Results on the Exact Small Sample Properties of the Instrumental Variable Estimator," *Econometrica*, 58, 967-76.

Nelson, C. R., and R. Startz (2007): "The Zero-Information-Limit Condition and Spurious Inference in Weakly Identified Models," *Journal of Econometrics*, Vol. 138, 47-62.

Phillips, P. C. B. (1983): "Exact Small Sample Theory in the Simultaneous Equations Model," Handbook of Econometrics, Vol. 1, Chapter 8. Amsterdam: North Holland.

Shiller, R., (1981): "Do Stock Prices move too much to be justified by subsequent changes in dividends?" *American Economic Review* 71, 421-436.

Staiger, D., J. H. Stock., and M. W. Watson (1997): "The NAIRU, Unemployment and Monetary Policy," *The Journal of Economic Perspectives*, 11, 33-49.

Working, H. (1960): "Note on the correlation of first differences of a random chain," *Econometrica*, 28, 916-918.

END.