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

Jun Ma
University of Alabama

and

Charles R. Nelson
University of Washington

Presentation October 13, 2008

Copyright © 2006-2008 by Jun Ma and Charles R. Nelson

Models of the form

$$y_i = \gamma \bullet g(\beta, x_i) + \varepsilon_i; i = 1, ..., N \,. \qquad (1.1)$$

The parameter of interest is β, identified only if $\gamma \neq 0$.

Additional regressors and parameters would often be present in practice.

We assume ε_i are *i.i.d.* $N(0, \sigma^2)$ so ML estimates $\hat{\gamma}$ and $\hat{\beta}$ are obtained by non-linear least squares.

Class includes ARMA, GARCH and Unobserved Components as well as other State Space models.

Standard inference based on asymptotic theory often works poorly in finite samples.

Distribution of $\hat{\beta}$ will generally be displaced away from the true value.

Nelson and Startz (2007) show that estimated standard error for $\hat{\beta}$ generally too small.

But, size distortion may go either way!

This paper demonstrates that linear approximation to $g(\beta, \mathbf{x}_i)$ provides *t*-test that works well.

Bias and test size when *g(.)* is linear.

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

For example, the Phillips curve of Staiger, Stock and Watson (1997) where *y* is the change in inflation, *x* is actual unemployment and β is the natural rate.

Reduced form is

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

where $\lambda = \gamma \bullet \beta$, thus $\hat{\beta} = \hat{\lambda} / \hat{\gamma}$.

Moments of ratio of Normals do not exist, see Fieller (1932) and Hinckley (1969).

But $\hat{\lambda}$ and $\hat{\gamma}$ are jointly Normal across samples so:

$$\hat{\lambda} = \alpha + \kappa \bullet \hat{\gamma} + v \tag{2.3}$$

where v is Normal and uncorrelated with $\hat{\gamma}$ by construction.

To simplify, $\beta = 0$ and standardized regressors with sample correlation ρ .

Noting $\alpha = \rho \bullet \gamma$, $\kappa = -\rho$, and the variance of v is σ^2 / N .

$$\hat{\beta} = -\rho + \rho \bullet \left(\frac{\gamma}{\hat{\gamma}} \right) + \frac{v}{\hat{\gamma}} \tag{2.4}$$

Consider how the distribution of $\hat{\beta}$ is affected by γ and correlation ρ .

Large γ means that the ratio $\gamma / \hat{\gamma}$ tends to be closer to 1, so first two terms tend to cancel,

third term relatively small, so $\hat{\beta}$ will be located more tightly around its true value, zero.

Small value of γ means $\gamma/\hat{\gamma}$ will be small, so $\hat{\beta}$ around $-\rho$,

But with greater dispersion since the third term will tend to be large.

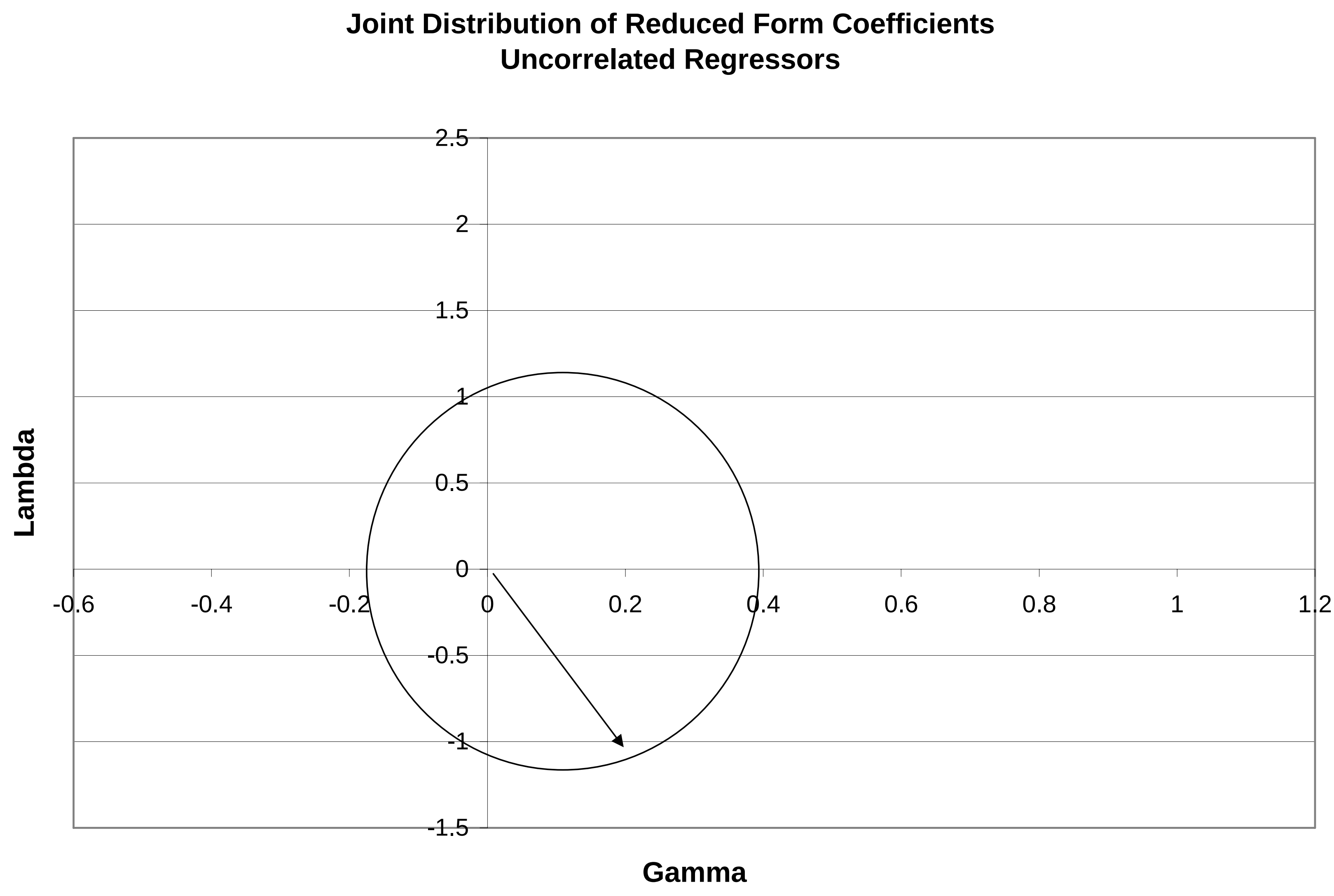
Now the effect of ρ :

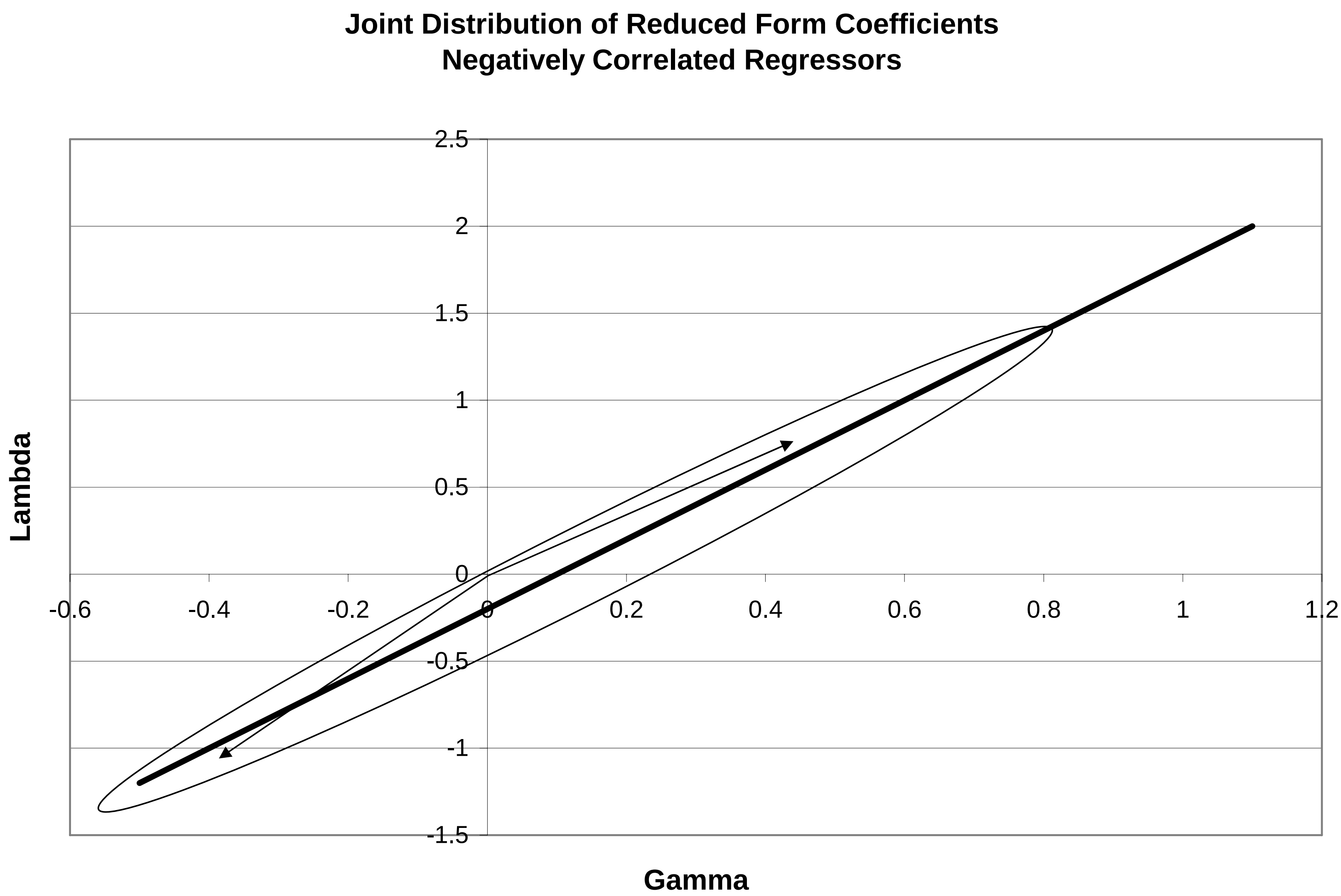
Stronger correlation increases sampling variation in $\hat{\gamma}$,

so second and third terms tend to be small,

concentrating distribution of $\hat{\beta}$ around $-\rho$.

Concentration effect can be seen graphically:


Joint Distribution of Reduced Form Coefficients
Uncorrelated Regressors
Lambda
Gamma
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1
1.2


Joint Distribution of Reduced Form Coefficients
Negatively Correlated Regressors
Lambda
Gamma
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1
1.2

Hypothesis testing

Asymptotic variance of $\hat{\beta}$ is given by:

$$V_{\hat{\beta}} = \frac{1}{\gamma^2} \bullet \frac{\sigma^2}{N} \bullet \frac{m_{xx} + 2\beta \bullet m_{xz} + \beta^2 \bullet m_{zz}}{m_{xx} \bullet m_{zz} - m_{xz}^2} \tag{2.5}$$

In practice unknown parameters replaced by the point estimates:

$$t_{\hat{\beta}}^2 = (\hat{\beta} - \beta_0)^2 \bullet \left[\hat{\gamma}^2 \bullet \frac{N}{\hat{\sigma}^2} \bullet \frac{m_{xx} \bullet m_{zz} - m_{xz}^2}{m_{xx} + 2\hat{\beta} \bullet m_{xz} + \hat{\beta}^2 \bullet m_{zz}} \right] \tag{2.6}$$

where null is $\beta = \beta_0$.

For $\beta_0 = 0$ and standardized regressors:

$$t_{\hat{\beta}}^2 = \frac{\hat{\lambda}^2}{\hat{\sigma}^2} \bullet N \bullet (1 - \rho^2) \bullet \frac{1}{1 + 2\hat{\beta} \bullet \rho + \hat{\beta}^2} = t_{\lambda}^2 \bullet \frac{1}{1 + 2\hat{\beta} \bullet \rho + \hat{\beta}^2} \tag{2.7}$$

Note $t_{\hat{\lambda}}$ has correct size and provides alternative test of null $\beta = 0$.

Exact test of Fieller (1954) for ratio of regression coefficients (incorrectly attributed in literature).

If regressors are orthogonal, then $t_{\hat{\beta}}^2 < t_{\hat{\lambda}}^2$ since the last term must be less than one.

But strong correlation concentrates $\hat{\beta}$ around $-\rho$, making $t_{\hat{\beta}}^2 > t_{\hat{\lambda}}^2$.

Whether test size is too large or too small depends on the correlation between the regressors.

NB: If identification condition $\gamma \neq 0$ fails, information matrix is singular.

Nevertheless, the reduced form test statistic $t_{\hat{\lambda}}$ still has an exact *t*-distribution since reduced form is still a properly specified classical regression.

Monte Carlo experiments with $\beta = 0$, standardized regressors, and errors are *i.i.d.* $N(0,1)$.

Estimation is done in EViews™ using the non-linear regression routine.

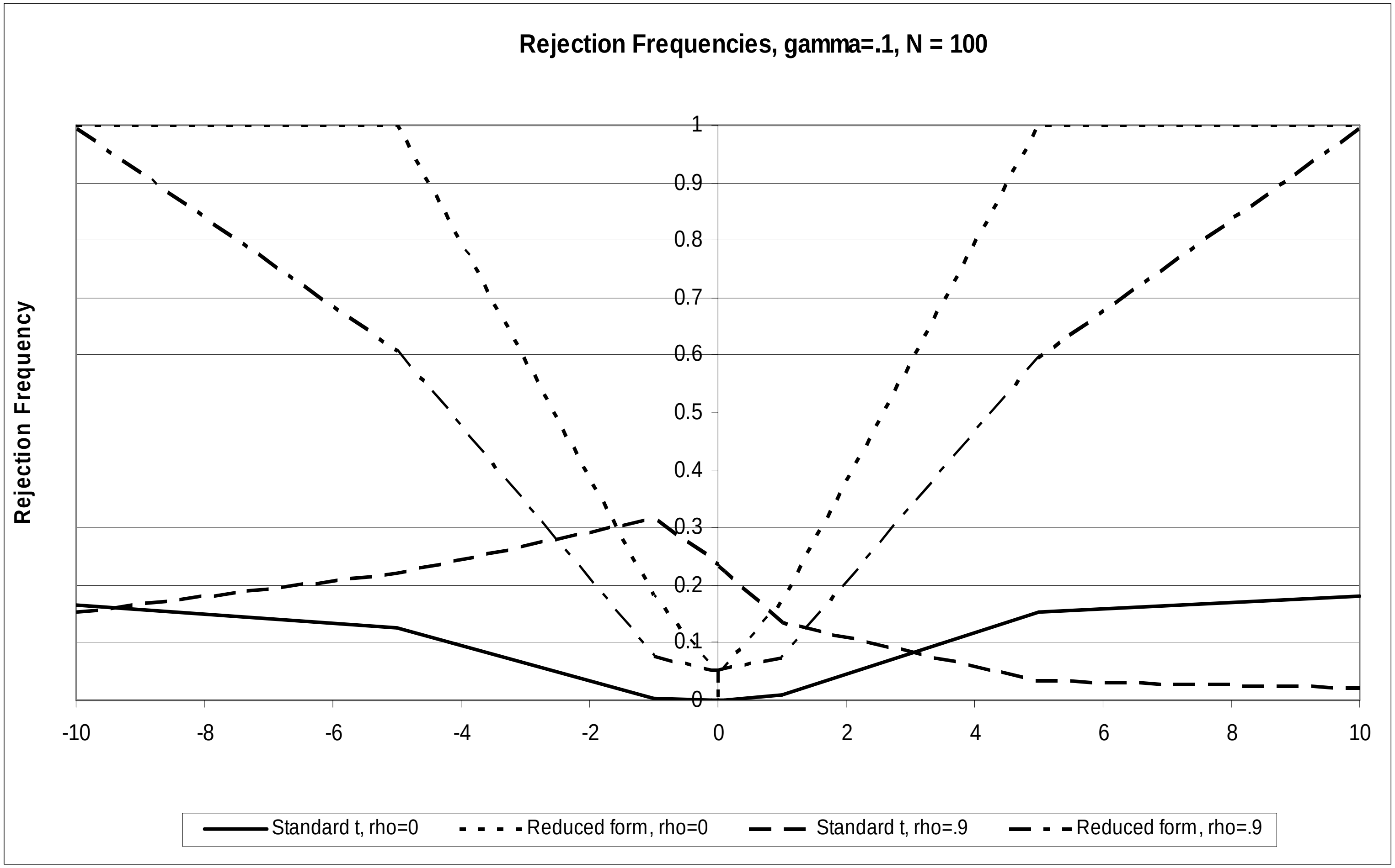
Table 1: The Effect of γ on the Distribution of $\hat{\beta}$ and Size of $t_{\hat{\beta}}$ with Orthogonal Regressors; N = 100

True γ	.01	.10	.5	1.0
Asymptotic $\gamma/\sqrt{V_{\hat{\gamma}}}$	.1	1	5	10
Median $\hat{\beta}$	.10	.03	-.00	-.00
Range (.25, .75)	(-.95, 1.17)	(-.65, .68)	(-.14, .13)	(-.06, .06)
Size of $t_{\hat{\beta}}$	.0001	.0002	.038	.050

Table 2: The Effect of Correlation ρ Between Regressors on Distribution of $\hat{\beta}$ and on Size of $t_{\hat{\beta}}$; $\gamma = .1$, N = 100.

Correlation ρ	0	.50	.90	.99
Asymptotic $\gamma / \sqrt{V_{\hat{\gamma}}}$	1	.87	.44	.14
Median $\hat{\beta}$	.03	-.21	-.71	-.96
Range (.25, .75)	(-.65, .68)	(-.84, .54)	(-1.27, -.14)	(-1.17, -.75)
Size of $t_{\hat{\beta}}$	.0002	0.019	.235	.565

Figure 1: Rejection Frequencies for tests of $H_0: \beta = 0$, N = 100, γ = .10.



Asymptotic theory does take hold as sample size becomes large, albeit very slowly!

Table 3: The Effect of Sample Size N and γ on the Distribution of $\hat{\beta}$ and Size of $t_{\hat{\beta}}$; Orthogonal Regressors.

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 $\hat{\beta}$	.10	.01	-.00	.00
Range (.25, .75)	(-.95, 1.17)	(-.64, .63)	(-.07, .07)	(-.07, .07)
Size of $t_{\hat{\beta}}$	.0001	.0006	.043	.045

The Reduced Form Test in a Linear Approximation.

More generally $g(.)$ will not be linear and reduced form test can be based on the linear approximation:

$$y_i = \gamma \bullet [g(\beta_0, x_i) + (\beta - \beta_0) \bullet g_\beta(\beta_0, x_i)] + e_i \qquad (3.0.1)$$

where $g_\beta = dg(.)/d\beta$ and e_i includes a remainder. Thus

$$y_i = \gamma \bullet g(\beta_0, x_i) + \lambda \bullet g_\beta(\beta_0, x_i) + e_i; \text{ where } \lambda = \gamma \cdot \left(\beta - \beta_0\right). \qquad (3.0.2)$$

Least squares conditional on β_0 is the indirect least squares estimate $\hat{\beta} = \beta_0 + \left(\hat{\lambda}/\hat{\gamma}\right)$ and the implication of the null hypothesis $\beta = \beta_0$ is $\lambda = 0$.

Reduced form test will not have exact size when $g(.)$ is not linear, but how far off?

‘Bias’ in $\hat{\beta}$ and the size of standard *t*-test depend on correlation between $g(.)$ and $g_\beta(.)$.

Standard errors and *t*-statistic based on evaluation of $g(.)$ and $g_\beta(.)$ at $\beta = \hat{\beta}$.

Thus the correlation between ‘regressors’ is not fixed but ‘endogenous.’

This co-determination affects the distribution of the point estimate and the size of the standard *t*-test.,

but not the reduced form test since it evaluates $g(.)$ and $g_\beta(.)$ under the null hypothesis.

Non-linear Regression: A Production Function.

Consider Hicks-neutral Cobb-Douglas production function:

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

where linear reduced form approximation is

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

where $\lambda = \gamma \bullet (\beta - \beta_0)$.

We expect $\hat{\beta}$ and the size of the standard *t*-test to be biased in directions indicted by the correlation between x_i^{β} and $x_i^{\beta} \log(x_i)$, corresponding to $g(\beta, x_i)$ and $g_{\beta}(\beta, x_i)$.

Test based on the reduced form coefficient λ expected to have close to correct size.

If the value of β were .9, then the correlation is .92 and $\hat{\beta}$ would be biased downward; it is!

Size distortion is in the expected direction, but attenuated at downward biased point estimates.

Table 4: Small Sample Distribution of $\hat{\beta}$ and Test Size, True γ = .01, N = 100.

True β	0	.1	.5	.9
$\rho_{g(\beta),g_\beta(\beta)}$	.07	.29	.77	.92
Asymptotic $\gamma/\sqrt{V_{\hat{\gamma}}}$	.10	.10	.09	.11
Median $\hat{\beta}$	-0.04	-0.09	-0.05	0.12
Range (.25, .75)	(-.53, .50)	(-.59, .42)	(-.56, .48)	(-.41, .71)
Size of $t_{\hat{\beta}}$	0.027	0.037	0.114	0.179
Size of $t_{\hat{\lambda}}$	0.053	0.054	0.054	0.054

Table 5: Small Sample Distribution of $\hat{\beta}$ and Test Size, *N* = 100, true β = .5

True γ	.01	.1	1
Asymptotic $\gamma/\sqrt{V_{\hat{\gamma}}}$	.09	.91	9.10
Median $\hat{\beta}$	-.05	.27	.50
Range (.25, .75)	(-.56, .48)	(-.26, .64)	(.46, .54)
Size of $t_{\hat{\beta}}$	.114	.103	.052
Size of $t_{\hat{\lambda}}$	.054	.054	.054

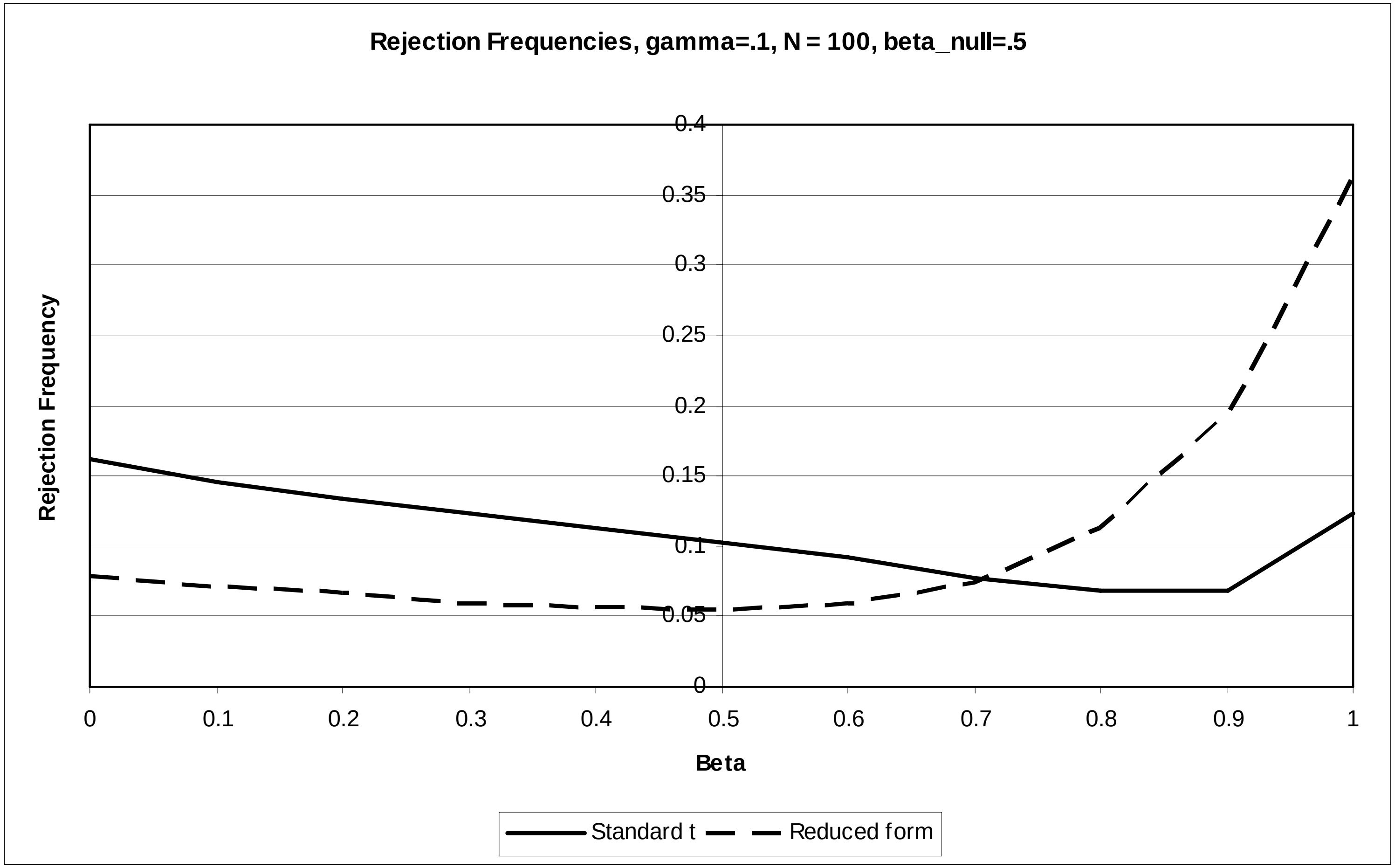
Case $\gamma = 0$ corresponds to failure of identification condition,

asymptotic theory underling the standard error and *t*-statistic not valid.

However, the reduced form test does not depend on that assumption.

Empirical size is 0.054, close to its nominal size and what we observed above.

Figure 2: Rejection Frequencies for the test $H_0 : \beta = .5$, *N* = 100, True γ = .1.



The ARMA (1,1) Model.

$$y_t = \phi \bullet y_{t-1} + \varepsilon_t - \theta \bullet \varepsilon_{t-1}; t = 1,\ldots,T$$
$$\varepsilon_t \sim i.i.d.N(0, \sigma_\varepsilon^2), |\phi| < 1, |\theta| < 1 \qquad (3.2.1)$$

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

where, $\gamma = (\phi - \theta)$, $g(\theta, \vec{y}_{t-1}) = \sum_{i=1}^{\infty} \theta^{i-1} y_{t-i}$ and $\vec{y}_{t-1} = (y_{t-1}, y_{t-2}, \ldots)$.

NS: when γ is small standard error too small and the standard *t*-test rejects the null too often.

Linearize the *g(.)* around the null to get the reduced-form test for θ:

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

where $g_\theta(\theta, \vec{y}_{t-1}) = \dfrac{\partial g(\theta, \vec{y}_{t-1})}{\partial \theta} = \sum_{i=2}^{\infty} (i-1) \bullet \theta^{i-2} y_{t-i}$, $\lambda = \gamma \bullet (\theta - \theta_0)$.

If the null $\theta = \theta_0$ is correct, the second term should not be significant.

Reduced-form test equivalent to testing second lag in an AR(2) regression,

approximately Box-Ljung *Q*-test with one lag from an AR(1) regression.

The estimated size of the reduced form test is correct within sampling error.

Case $\gamma = 0$ is failure of identification.

Reduced-form test still well defined in this case and the estimated size close to correct.

Table 6: Effect of γ on Inference for ARMA (1,1), True θ = 0, T = 1,000.

True $\gamma(=\phi)$	.01	.1	.2	.3
Asymptotic $\gamma/\sqrt{V_{\hat{\gamma}}}$	.32	3.16	6.32	9.49
Median $\hat{\theta}$	-.02	-.01	-.00	-.00
Range (.25, .75)	(-.65, .64)	(-.26, .24)	(-.11, .11)	(-.07, .07)
Size of $t_{\hat{\theta}}$	0.46	0.22	0.11	0.07
Size of $t_{\hat{\lambda}}$	0.051	0.052	0.053	0.052

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

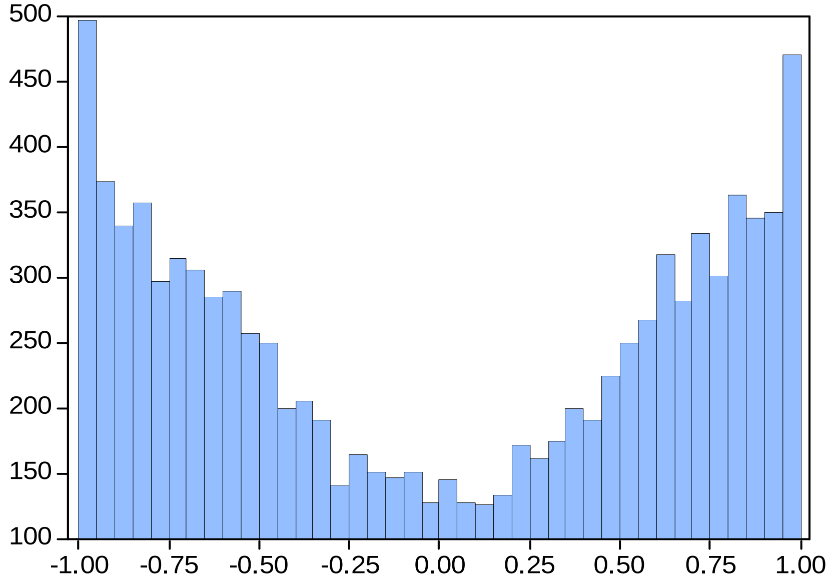


Figure 4: Computed un-centered correlation between $g(\theta, \vec{y}_{t-1})$ and $g_{\theta}(\theta, \vec{y}_{t-1})$ based on one sample draw. True γ = .01, θ = 0, T = 1,000.

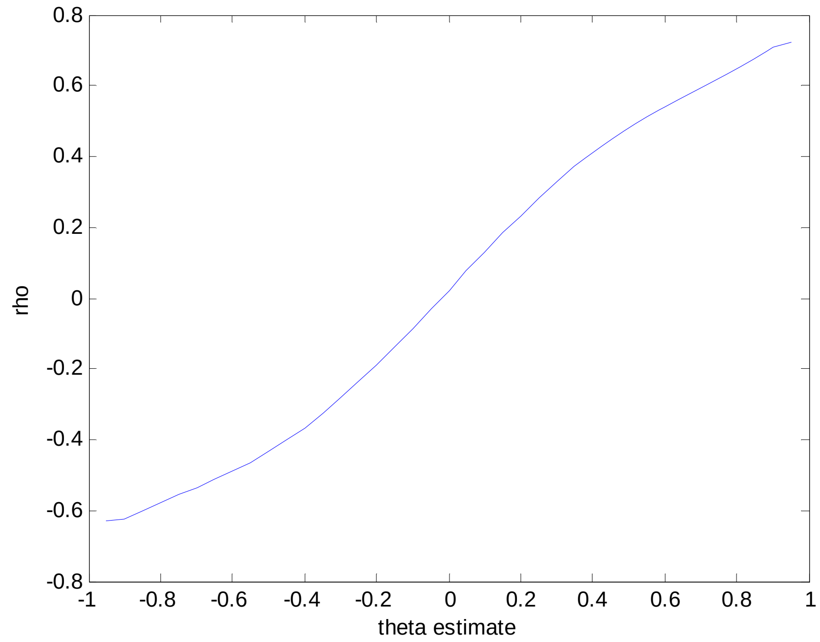


Table 7: Sample Size and Inference in the ARMA (1, 1), True θ = 0.

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 $\hat{\theta}$	-.04	-.02	-.02	-.00
Range (.25, .75)	(-.69, .67)	(-.65, .64)	(-.58, .55)	(-.07, .07)
Size of $t_{\hat{\theta}}$	.483	.458	.399	.066
Size of $t_{\hat{\lambda}}$	.051	.051	.049	.048

Often it is the AR root ϕ that is of a greater economic interest since it measures persistence.

For consumption growth a large value of ϕ implies that shock to the economy has a long-lasting impact on economic agent's conditional expectation of future consumption growth.

Bansal and Yaron (2000, 2004) show that high level of persistence, interpreted as long-run risk, may explain the equity premium puzzle of Mehra and Prescott (1985).

Ma (2007) finds that estimated ARMA(1,1) implies a small γ relative tosampling variance and explores the implications of possible test size distortion.

For the case $\gamma = 0.1, \phi = 0$ and T = 100 size of the reduced form test is 0.046

in contrast to 0.423 for standard *t*-test.

The reduced-form test can also be generalized to ARMA model of arbitrary order.

For the ARMA(2,2) model with parameter values $\phi_1 = 0.01, \phi_2 = 0.01, \theta_1 = 0, \theta_2 = 0$ and $T = 100$

we find that the standard *t*-test for $\hat{\theta}_1$ and $\hat{\theta}_2$ have empirical sizes of 0.571 and 0.698.

In contrast the reduced-form test gives rejection frequencies of 0.049 and 0.049 respectively.

The Unobserved Component Model for Decomposing Trend and Cycle

Unobserved Component model of Harvey (1985) and Clark (1987) :

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

trend assumed to be a random walk with drift:

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

and cycle has a stationary AR representation:

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

In practice, largest AR root is estimated close to unity, implying that the cycle is very persistent, and the trend variance is estimated to be very small, implying that the trend is very smooth.

We focus on the case that cycle is AR(1).

Following Morley, Nelson and Zivot (2003), univariate representation is ARMA(1,1):

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

MA parameter θ is identified (under the restriction $\sigma_{\eta,\varepsilon} = 0$) by matching autocovariances:

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

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

Solve for unique θ by imposing invertibility:

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

It is straightforward that θ becomes arbitrarily close to ϕ as $\sigma_\varepsilon / \sigma_\eta$ approaches zero.

Monte Carlo experiment: $\mu = 0.8, \phi = 0, \sigma_\eta^2 = 0.95, \sigma_\varepsilon^2 = 0.05$, T=200

Almost all the variation due to trend while cycle is small with no persistence at all.

The standard *t*-test for ϕ indeed rejects the null much too often; size is 0.481.

Standard error for $\hat{\phi}$ is underestimated; the median is 0.29 compared true value 1.48.

Further, $\hat{\phi}$ is upward biased, its median being 0.58, many close to the positive boundary.

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.

At the same time, the cycle innovation variance estimate is upward biased,

while the trend innovation variance estimate is instead downward biased.

Persistence in estimated cycle tends to occur in samples that also show large variance in the cycle.

Why? Model must account for the small amount of serial correlation in our data generating process.

Roughly, $-\frac{1-\phi}{1+\phi} \bullet \sigma_{\varepsilon}^{2} = -.05$.

One solution is the combination of true values, $\phi = 0; \sigma_{\varepsilon}^{2} = .05$,

but another is $\phi = .9; \sigma_{\varepsilon}^{2} = .95$.

Large negative values of $\hat{\phi}$ are possible but infrequent because positive variances > 0.

Reduced-form test: impose the null $\phi = \phi_0$ and estimate all other parameters in the UC model;

secondly, impute from (3.3.7) the restricted estimate $\tilde{\theta}$ and $\tilde{u}_t$ in the reduced-form ARMA(1,1).

Size of reduced-form test for ϕ is 0.054.

When all variation is due to stochastic trend, i.e., $\sigma_{\varepsilon}^2 = 0$, identification for ϕ fails.

However, reduced-form test works well with size 0.058.

This reduced-form test can also be generalized to address a UC model with higher AR order.

Figure 5: Plot of $\hat{\phi}$ in the Monte Carlo Experiment with true parameter

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

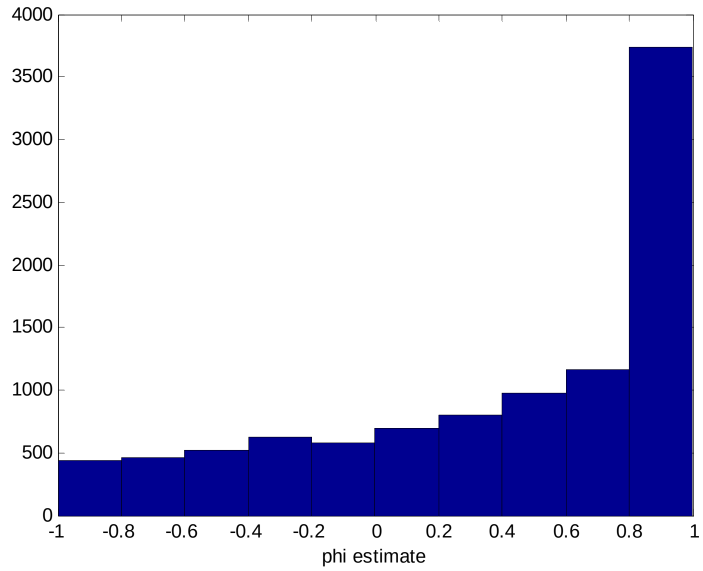
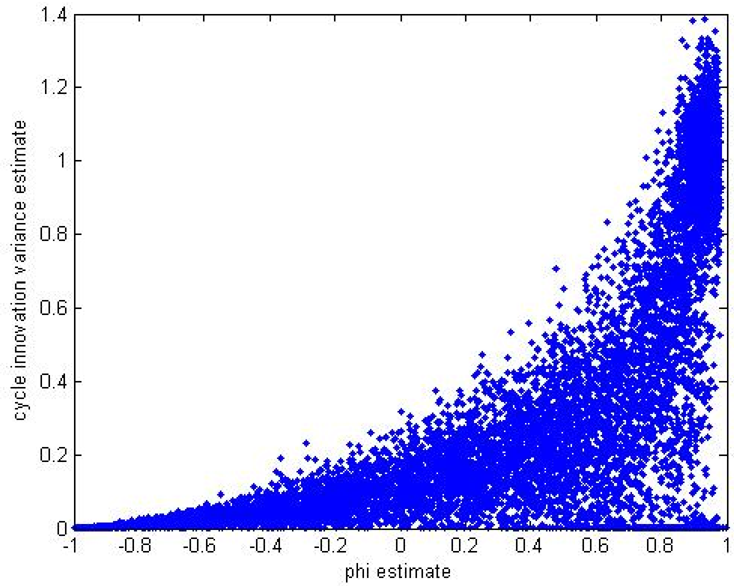


Figure 6: Scatter Plot of $\hat{\phi}$ and $\hat{\sigma}_{\varepsilon}^{2}$ in the Monte Carlo Experiment with true parameter

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



The GARCH(1,1) Model.

The GARCH model of Bollerslev (1986) in simple (1,1) case:

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

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

Analogy to the ARMA (1,1) model:

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

Ma, Nelson and Startz (2007) show that when α is small relative to its sampling 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 none.

Defining $g(\beta,\vec{\varepsilon}_{t-1}^{\,2}) = \sum_{i=1}^{\infty} \beta^{i-1}\varepsilon_{t-i}^2$ and $\vec{\varepsilon}_{t-1}^{\,2} = (\varepsilon_{t-1}^2, \varepsilon_{t-2}^2, \ldots)$ to rewrite (3.4.2) one obtains:

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

Taking a linear expansion of nonlinear *g*(.) around the null,

defining $c = \dfrac{\omega}{1-\beta}$, $\lambda = \alpha \bullet (\beta - \beta_0)$ and $g_\beta(\beta, \vec{\varepsilon}_{t-1}^{\,2}) = \sum_{i=2}^{\infty} (i-1) \bullet \beta^{i-2}\varepsilon_{t-i}^2$, we have:

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

The reduced-form test is the *t*-stat of the null $\lambda = 0$ in (3.4.5).

Table 8: Reduced form and standard *t*-tests for GARCH(1,1): True β = 0, *T* = 1,000.

True $\gamma(=\alpha)$	.01	.05	.1	.2
Asymptotic $\gamma\big/\sqrt{V_{\hat{\gamma}}}$	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)
Size of $t_{\hat{\beta}}$	0.470	0.344	0.198	0.106
Size of $t_{\hat{\lambda}}$	0.078	0.074	0.076	0.096

For the case $\alpha = 0$ identification fails and standard *t*-test does not have usual asymptotic distribution.

The reduced-form test, however, is still valid and has size of 0.076 for true β = 0 and T = 1,000.

The sum $\alpha + \beta$ is of potentially greater economic interest.

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 the equity premium puzzle.

Similar results follow in this case; see Ma (2007) for further discussion.

Monthly S&P 500 index return data from the Eviews 5.1 DRI Database for the sample period

January 1947 to September 1984 corresponding to Bollerslev (1987).

GARCH estimates with Bollerslev and Wooldridge's (1992) robust standard errors:

$$\hat{\omega} = 0.16 \bullet 10^{-3}(0.14 \cdot 10^{-3}), \ \hat{\alpha} = 0.077(0.048), \ \hat{\beta} = 0.773(0.169)$$

Implies a significant and large GARCH effect,

95% confidence interval for β : [0.44, 1).

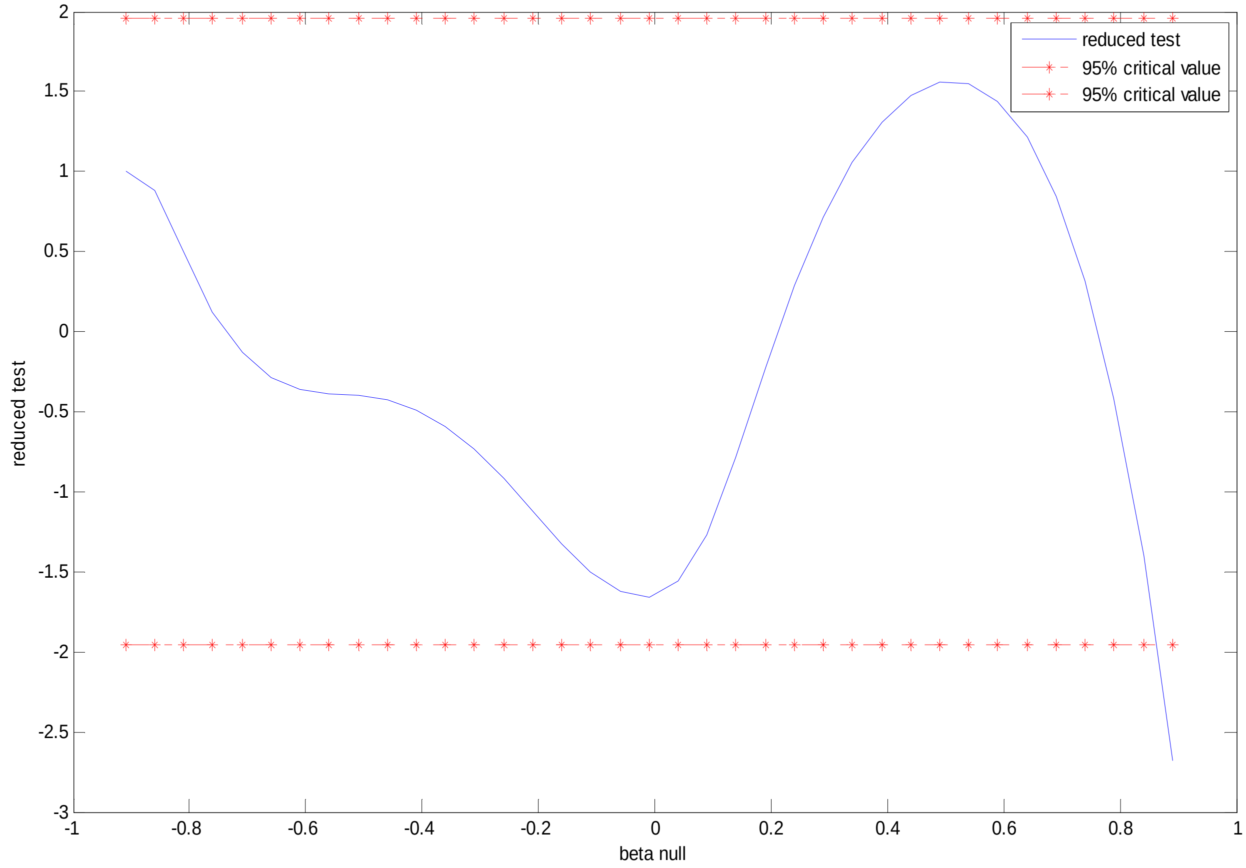
However, small value of $\hat{\alpha}$ relative to the sample size.

To obtain confidence interval we numerically invert the reduced-form test statistic,

create a grid of β_0's, compute the corresponding $t_{\hat{\lambda}}$'s and plot the latter against the former.

Resulting 95% confidence interval is [-0.95, 0.87], which covers almost the entire parameter space.

Figure 7: The 95% Confidence Interval for $\hat{\beta}$ based on the reduced-form test for monthly S&P 500 stock return data



Summary and Conclusions

This paper considers models of the form $y = \gamma \bullet g(\beta, x) + \varepsilon$, where β is of interest.

Small sample inference is usefully studied by working with the approximation

$$g(\beta, x) \approx g(\beta_0, x) + (\beta - \beta_0) \bullet g_\beta(\beta_0, x)$$

and the corresponding reduced form regression model

$$y_i = \gamma \bullet g(\beta_0, x_i) + \lambda \bullet g_\beta(\beta_0, x_i) + e_i; \text{ where } \lambda = \gamma \bullet (\beta - \beta_0).$$

Findings:

The distribution of $\hat{\beta}$ is biased in a direction determined by the correlation between the 'regressors' in the reduced form, and the distribution becomes concentrated when that correlation is strong.

The distribution of the standard *t*-statistic for $\hat{\beta}$ based on asymptotic theory is also dependent on that correlation, as is the size of the *t*-test.

Both of these distributions are also dependent on the true value of γ, so the conventional *t*-test is not pivotal in finite samples.

A reduced form test is exact when $g(.)$ is linear;

we show that is has nearly correct size when the reduced form model is only an approximation.

Further, its distribution does not depend on the identifying restriction $\gamma \neq 0$.

End.