Endogeneity in threshold nonlinearity tests

Alexander Dentler, *University of Wisconsin-Madison*

Gabriel Montes-Rojas,[*] *City University London*

Jose Olmo, *Centro Universitario de la Defensa de Zaragoza*

December 14, 2011

Abstract

This paper shows the conditions under which endogeneity of a regressor variable does not affect threshold nonlinearity tests. Inference on the values of the parameters derived from standard statistics is also appropriate. Simulation techniques are used to approximate the p-value of the test. Monte Carlo simulations confirm the validity of Wald tests in the presence of endogeneity in the regressors.

Keywords: Endogeneity, Threshold models, Wald tests.

JEL classification: C14, C22, C32, C50

---

[*]Corresponding Author: Department of Economics, City University London, Northampton Square, London EC1V 0HB, UK, email: `Gabriel.Montes-Rojas.1@city.ac.uk`, tel: +44 (0)20-7040-8919.

1 Introduction

Threshold models are simple yet efficient methods to capture nonlinearities in cross section and time series models. They split the sample into classes based on the value of observed variables according to threshold values. The theory of estimation and inference in threshold models with exogenous regressors has been extensively studied in the classical papers of Chan and Tong (1986), Chan (1993) and Hansen (1996,1997,2000). Consider a simple threshold nonlinear regression model

$$y_t = x_t\beta_1 + I(z_t > \delta)x_t\beta_2 + u_t, \tag{1}$$

where δ is the threshold defined on the variable z over a compact set $\Delta \subset \mathbb{R}$. Nonlinearity tests are based on the hypothesis that $H_O : \beta_2 = 0$. The most interesting case is when the threshold value δ is not known and must be estimated. There is an inherent statistical difficulty associated with this problem. For instance, conventional tests of the null of a linear model against the alternative have nonstandard distribution, since the threshold parameter is not identified under the null of linearity (see Hansen, 1996).

Caner and Hansen (2004) extend this test to make allowance for endogeneity of x but assume that the threshold variable is exogenous, i.e., $E[zu] = 0$, and extend Hansen's (1996,1997,2000) results to this case. However, as the authors note, "it may be desired to treat the threshold variable as endogenous (...), [and this] would be a substantially different model and would require a distinct estimator" (Caner and Hansen, 2004, p.814). Moreover, finding reliable instrumental variables may be a futile task in empirical settings. In fact, in many applications the threshold variable is $z_t = x_t$, such as in the self-exciting threshold autoregressive models where $x_t = y_{t-1}$.

Therefore, in this paper we consider the model,

$$y_t = x_t\beta_1 + I(x_t > \delta)x_t\beta_2 + u_t, \tag{2}$$

where $E[x_t u_t] \neq 0$, i.e., x is an endogenous variable. In a standard linear model with $\beta_2 = 0$, the endogeneity of x will make the OLS estimator of β_1 inconsistent. However, we show that this general endogeneity model, where the threshold variable is the endogenous one, produces no

distortions in the Hansen (1996,1997,2000) threshold nonlinearity tests for a large family of joint distributions describing the relationship between the explanatory variables and the error term. This is the Pearson family of distributions that includes as particular cases the multivariate Normal and t-Student distributions. The conclusion thus is that under these conditions there is no need to pursue two-stage least squares instrumental variables strategies to obtain correct size and power for these tests.

This paper is organized as follows. Section 2 presents the effect of endogeneity on threshold non-linearity tests and derives the main asymptotic results. Section 3 reports Monte Carlo experiments. Section 4 concludes.

2 Wald tests for threshold nonlinearity under endogeneity

Consider a sample of size n of $\{y_t, x_t, u_t\}_{t=1}^n$, satisfying equation (2) and $E[x_t u_t] \neq 0$. Define $x_{1t} = x_t$ and $x_{2t} = I(x_t > \delta)x_t$ and $x_t(\delta) = (x_{1t}, x_{2t})$.

Consider the following assumptions:

Assumptions A1-A4:

A1: (y_t, x_t) is strictly stationary, ergodic and ρ-mixing, with ρ-mixing coefficients satisfying $\sum_{m=1}^{\infty} \rho_m^{1/2} < \infty$; u_t is a martingale difference sequence with respect to the sigma-algebra determined by the set of all available information up to time $t-1$.

A2: δ lies in a compact set $\Delta \subset \mathbb{R}$, and $\beta \equiv (\beta_1, \beta_2) \in int\, B$, with B compact and convex;

A3: $E[|x_t|^{2+\epsilon}] < \infty$ with $|\cdot|$ the absolute value function;

A4: $\frac{1}{n}\sum_{t=1}^n x_t(\delta_1)x_t(\delta_2)'$ and $\frac{1}{n}\sum_{t=1}^n (x_t(\delta_1)u_t)(x_t(\delta_2)u_t)'$ converge almost surely to $E[x_t(\delta_1)x_t(\delta_2)']$ and $E[(x_t(\delta_1)u_t)(x_t(\delta_2)u_t)']$, respectively, uniformly over $\delta_1, \delta_2 \in \Delta$. Further, assume that $E[x_t(\delta)x_t(\delta)'] > 0$ for all $\delta \in \Delta$.

These assumptions are common in the regime switching literature. A1 and A3 guarantees that the process is stationary and that the series satisfies Hansen's (2000) Assumption 1.1. A2 imposes

3

that δ lies on a compact set. This assumption was used by Hansen (1996). A4 is equivalent to Assumption 3 in Hansen (1996) and it guarantees that the probability limit of these expressions exists for every δ. This assumption will be relevant for obtaining the asymptotic distribution of the nonlinearity Wald test discussed later in the section. The assumption shows that the empirical covariance function of the Wald test statistic converges uniformly almost surely to the covariance function of a zero-mean Gaussian process.

Let the capital letters (Y, X_1, X_2, U) denote the sample vectors containing the n observations for each variable. Simple OLS orthogonal projections algebra shows that

$$\hat{\beta}_2 = (X_2' M_1 X_2)^{-1}(X_2' M_1 Y), \tag{3}$$

where $M_1 = I_n - X_1(X_1' X_1)^{-1} X_1'$. Using the definition of $Y = X_1 \beta_1 + X_2 \beta_2 + U$ gets

$$\hat{\beta}_2 = \beta_2 + (X_2' M_1 X_2)^{-1}(X_2' M_1 U). \tag{4}$$

Consider now the last factor,

$$\frac{1}{n} X_2' M_1 U = \frac{1}{n} X_2' U - \frac{1}{n} X_2' X_1 (X_1' X_1)^{-1} X_1' U = \frac{1}{n}\sum_{t=1}^{n} x_{2t} u_t - \left(\frac{1}{n}\sum_{t=1}^{n} x_{2t} x_{1t}\right)\left(\frac{1}{n}\sum_{t=1}^{n} x_{1t}^2\right)^{-1}\left(\frac{1}{n}\sum_{t=1}^{n} x_{1t} u_t\right)$$

$$= \frac{1}{n}\sum_{t=1}^{n} I(x_t > \delta) x_t u_t - \left(\frac{1}{n}\sum_{t=1}^{n} I(x_t > \delta) x_t^2\right)\left(\frac{1}{n}\sum_{t=1}^{n} x_t^2\right)^{-1}\left(\frac{1}{n}\sum_{t=1}^{n} x_t u_t\right). \tag{5}$$

Then, under assumptions A1-A4, the following convergence in probability holds:

$$\frac{1}{n} X_2' M_1 U \xrightarrow{p} E[xu|x > \delta]P[x > \delta] - E[x^2|x > \delta]P[x > \delta](E[x^2])^{-1}E[xu]$$

$$= P[x > \delta]\left\{E[xu|x > \delta] - E[x^2|x > \delta](E[x^2])^{-1}E[xu]\right\}. \tag{6}$$

The last expression determines whether the endogeneity in x has an effect on the estimation of β_2, and thus, whether it affects a test for threshold nonlinearities. Note that the last expression may indeed be close to zero for a broad range of (x, u) bivariate distributions. For instance, if we

assume that x and u follow a bivariate normal distribution with zero means, unit variances, and correlation ρ $(= E[xu])$. Then, $E[xu|x > \delta] = \rho E[x^2|x > \delta]$. In this case,

$$E[xu|x > \delta] = E[x^2|x > \delta](E[x^2])^{-1}E[xu]. \tag{7}$$

In fact, this condition can be also satisfied for all bivariate random variables in the Pearson family of distributions, such as the bivariate t-Student, where truncated moments can be expressed this way and (7) is satisfied (see Lee, 1984, p.847). In practice, of course, the distribution of (x, u) is not known. However, condition (7) is indeed satisfied for a large set of distributions.

We have explored the case of a single regressor x_t defining the threshold nonlinearity. The above results can be generalized under certain conditions to the following multivariate regression model characterized by an endogenous variable, x_t, that also defines the threshold nonlinearity, and w_t, a set of exogenous variables:

$$y_t = x_t\beta_1 + w_t'\gamma_1 + I(x_t > \delta)x_t\beta_2 + I(x_t > \delta)w_t'\gamma_2 + u_t.$$

For simplicity, we assume that both x_t and w_t have zero mean, although the result below can be also obtained if the model has a constant term. The design matrix X is now defined as $X_t = (x_t \ w_t \ I(x_t > \delta)x_t \ I(x_t > \delta)w_t)$ with partitions $X_{1t} = (x_t \ w_t)$, $X_{2t} = (I(x_t > \delta)x_t \ I(x_t > \delta)w_t)$ and define $X_1 = (x_1 \ w_1)$ and $X_2 = (x_2 \ w_2)$. The aim is to show the consistency of the OLS estimators of β_2 and γ_2 under the presence of an endogenous threshold variable. The difference with the previous case is the additional parameter γ_2; the OLS estimator is consistent if the following asymptotic condition is satisfied:

$$(X_2'M_1U) = [X_2'U] - [X_2'X_1(X_1'X_1)^{-1}X_1'U] \xrightarrow{p} \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

The first term is such that

$$X_2'U \xrightarrow{p} \begin{pmatrix} E[xu|x > \delta]P[x > \delta] \\ E[wu|x > \delta]P[x > \delta] \end{pmatrix}.$$

For the second term, our assumptions imply that

$$X_1'U \xrightarrow{p} \begin{pmatrix} E[xu] \\ 0 \end{pmatrix}.$$

Define the probability limit of the elements of the matrix $(X_2'X_1)$ as $(a_{xx}, a_{xw}, a_{wx}, a_{ww})$, that is,

$$\frac{1}{n}(X_2'X_1) = \begin{pmatrix} a_{xx} & a_{xw} \\ a_{wx} & a_{ww} \end{pmatrix}$$

and the probability limit of the elements of the matrix $(X_1'X_1)^{-1}$ as $(b_{xx}, b_{xw}, b_{wx}, b_{ww})$. Then,

$$X_2'X_1(X_1'X_1)^{-1}X_1'U \xrightarrow{p} \begin{pmatrix} (a_{xx}b_{xx} + a_{xw}b_{wx})E[xu] \\ (a_{wx}b_{xx} + a_{ww}b_{wx})E[xu] \end{pmatrix}$$

and the consistency condition becomes

$$\begin{pmatrix} E[xu|x > \delta]P[x > \delta] - (a_{xx}b_{xx} + a_{xw}b_{wx})E[xu] \\ E[wu|x > \delta]P[x > \delta] - (a_{wx}b_{xx} + a_{ww}b_{wx})E[xu] \end{pmatrix}$$

Two cases need to be considered here. First, assume that x_t and w_t are independent, which (together with the mean zero condition) implies that $E[wu|x > \delta] = 0$ and $a_{wx} = a_{xw} = b_{wx} = 0$. In this case, the second row in the asymptotic condition is zero and the study of the endogeneity is equal to the univariate regression model discussed above. Second, if x_t and w_t are not independent, then the endogeneity in x_t will have an effect in the estimation of γ_2. For this case, the effect of having endogeneity depends on the particular distribution of the triple (x_t, w_t, u_t).

Now, we turn to the analysis of the nonlinearity test. To do this we discuss the effect of the endogeneity of x_t in the variance of the parameter estimators. For simplicity, we concentrate on the univariate regression model and, in particular, on $\hat{\beta}_2$. The variance of the parameter estimator is

$$V[\sqrt{n}\hat{\beta}_2] = E\left[\left(\frac{X_2'M_1X_2}{n}\right)^{-1}\frac{(X_2'M_1U)(X_2'M_1U)'}{n}\left(\frac{X_2'M_1X_2}{n}\right)^{-1}\right]. \tag{8}$$

If we further assume homoscedasticity of the error term, this variance can be estimated by

$$\widehat{V}_n = \left(\frac{X_2' M_1 X_2}{n} \right)^{-1} \frac{1}{n} \sum_{t=1}^{n} \hat{u}_t^2, \tag{9}$$

with $\hat{u}_t$ the residuals of regression equation (2).

The linearity of model (2) is reflected in the null hypothesis $H_O : \beta_2 = 0$. In order to be able to implement this test under endogeneity of the regressor, condition (7) needs to be satisfied. The null hypothesis H_O can be tested using a Wald type test that reduces in the univariate case to a t-test. For the multivariate regression model the null hypothesis would be $H_O : \beta_2 = \gamma_2 = 0$ implying an F-test. Thus, for the univariate case the Wald test is

$$W = n \frac{(\hat{\beta}_2)^2}{\widehat{V}_n} = t_{\hat{\beta}_2}^2. \tag{10}$$

From the formulas above it follows that if δ is known and (7) holds, the t-test $t_{\hat{\beta}_2}$ is asymptotically standard normal under H_O.

The most interesting case, however, is when the threshold value δ is not known and must be estimated. There is an inherent statistical difficulty associated with this problem. For instance, conventional tests of the null of a linear model against the alternative have nonstandard distribution, since the threshold parameter is not identified under the null of linearity (see Hansen, 1996). In order to test the hypothesis we need to test the significance of the β_2 parameter for every $\delta \in \Delta$ (a compact set on $\mathbb{R}$). Consider $\hat{\beta}_2(\delta)$ and $\widehat{V}_n(\delta)$, where the dependence on δ is made explicit. Now define the Wald statistic as

$$W(\delta) = n \frac{(\hat{\beta}_2(\delta))^2}{\widehat{V}_n(\delta)} = t_{\hat{\beta}_2}^2(\delta). \tag{11}$$

Now, $W(.)$ is a process on δ and the test statistic is some functional of it, supremum and exponential average tests are usually considered. The former method also provides a candidate of the threshold parameter and the second defines an optimal test, as discussed by Andrews and Ploberger (1994).

As in Hansen (1996), we approximate the asymptotic distribution of these tests by simulation techniques. Under assumptions A1-A4 and H_O, the distribution of $W(\delta)$ can be approxi-

7

mated by the distribution of an auxiliary process $W^*(\delta) = \frac{((X_2'M_1X_2/n)^{-1}S_n^*(\delta))^2}{\widehat{V}_n(\delta)}$, with $S_n^*(\delta) = \frac{1}{\sqrt{n}}\sum_{t=1}^n I(x_t > \delta)x_t\hat{u}_t v_t$ the score function of (2), where $\hat{u}$ is the residual sequence of this regression equation and $(v_1, \ldots, v_n)'$ is an $i.i.d.$ vector of standard normal random variables. Then, under H_O, $\sqrt{n}\hat{\beta}_2(\delta) = (X_2'M_1X_2/n)^{-1}S_n^*(\delta) + o_p(1)$. The auxiliary process $W^*(\delta)$ can be simulated by obtaining independent vectors v to compute independent replicas of $W^*(\delta)$.

3 Monte Carlo Simulation Studies

This section implements the p-value transformation of Hansen (1996) for a nonlinearity test under exogeneity and endogeneity. Consider the following data generating processes:

$$\text{DGP1: } y_t = x_t\beta_1 + u_{1t},$$

$$\text{DGP2: } y_t = x_t\beta_1 + u_{2t},$$

$$\text{DGP3: } y_t = x_t\beta_1 + I(x_t > \delta)x_t\beta_2 + u_{1t},$$

$$\text{DGP4: } y_t = x_t\beta_1 + I(x_t > \delta)x_t\beta_2 + u_{2t}.$$

The random errors u_{1t} and u_{2t} are i.i.d. and mutually independent. The regressor x_t is defined as $x_t = \rho u_{2t} + \sqrt{1-\rho^2}w_t$, with w_t an i.i.d. random variable mutually independent of u_{1t} and u_{2t}. This implies that DGP1 and DGP3 are defined by exogenous regressors and DGP2 and DGP4 by endogenous regressors. For DGP1 and DGP3 the error term u_1 is such that $E[x_t u_{1t}] = 0$ since $E[u_{1t}u_{2t}] = 0$, by construction. For DGP2 and DGP4, if $\rho \neq 0$ it follows that $E[x_t u_{2t}] = \rho\sigma_{u_2}^2 \neq 0$ with $\sigma_{u_2}^2$ the variance of u_{2t}, by construction of the regressor x_t.

We consider three distributions. First, a multivariate normal case where $u_1, u_2, w \sim i.i.d.\, N(0,1)$. Second, a multivariate t-Student case where $u_1, u_2, w \sim i.i.d.\, t_2$ and t_3. In these cases, except for the t_2, $\hat{\beta}_2 = \beta_2 + o_p(1)$ because (7) is satisfied. Finally, a mixture of distributions where $u_1, u_2 \sim i.i.d.\, (\chi_3^2 - 3)/\sqrt{6}$ and $w \sim i.i.d.\, N(0,1)$. In this case $\hat{\beta}_2 \neq \beta_2 + o_p(1)$ given that (7) is not satisfied. We set $\beta_1 = \beta_2 = 1$ and $\rho = 0.5$. The support of the threshold parameter δ (defined above

8

as Δ) is set to cover 90% of the endogenous regressor domain. We use the supremum and exponential average test statistics from the Wald-based p-value transformation method discussed above, using 200 replications. We repeat the Monte Carlo simulation 500 times in each case for sample sizes in $n = 100, 200, 500, 1000$. Finally we also consider three nominal sizes, $\alpha^* = 0.10, 0.05, 0.01$, and report the empirical size for each case.

[TABLE 1 ABOUT HERE]

Table 1 reports the empirical size for the multivariate normal case. Note that for both DGP1 and DGP2, the exponential average Wald tests achieve empirical size close to its nominal level, while the supremum Wald tests produces considerably small rejection rates. A more detailed look to the simulations show that the supremum test has a nonconservative size for all sample sizes and the exponential average appears the closest to their nominal rejection probabilities. In line with Andrews and Ploberger (1994) the latter method should be preferred to carry out this type of tests. Moreover, the exponential average provides better power in DGP3 and DGP4, although both tests are consistent. In all cases, we observe that the presence of endogeneity in x does not affect the tests for nonlinearities.

[TABLE 2 ABOUT HERE]

[TABLE 3 ABOUT HERE]

Tables 2 and 3 report the simulation results for the multivariate t_2 and t_3 distributions, respectively. In the first case, the variance of $\hat{\beta}_2(\delta)$ does not exist, provided that this distribution has infinite second moments. The size distortions produced by this distribution are of importance for small n, but they are considerably reduced for large n. Better size results are observed for the t_3 case. As in the normal case, the exponential average test appears the closest to their nominal rejection.

[TABLE 4 ABOUT HERE]

Table 4 reports the simulation results for the mixture of distributions where (7) is not satisfied. Note that even in this case, there are considerable differences between DGP1 and DGP2. As expected, DGP2 produces unacceptable size distortions, and the empirical rejection rates increse with n. This confirms the fact that condition (7) is necessary for the validity of the threshold nonlinearity tests based on (2).

4   Conclusion

This article shows that for the Pearson family of distributions endogeneity of the regressor variable does not produce distortions of associated nonlinearity tests based on self-exciting processes. Otherwise, the parameter estimator corresponding to the nonlinear component is biased and inconsistent. We show via a Monte-Carlo experiment this statistical phenomenon for different simulated processes.

The conclusions of this paper suggest that nonlinearity tests can be performed under the standard OLS with exogenous regressors paradigm under very general conditions, and complements the results of Caner and Hansen (2004).

References

- Andrews, D. W. K., and Ploberger, W. (1994). "Optimal Tests when a Nuisance Parameter is Present only Under the Alternative," Econometrica 62, 1383-1414.

- Caner, M. and Hansen, B.E. (2004). "Instrumental Variable Estimation of a Thereshold Model," Econometric Theory 20, 813-843.

- Chan, K. S. (1993). "Consistency and Limiting Distribution of the Least Squares Estimator of a Threshold Autoregressive Model," Annals of Statistics 21, 520-533.

- Chan, K. S., and Tong, H. (1986). "On estimating thresholds in autoregressive models," Journal of Time Series Analysis 7, 179-194.

- Hansen, B.E. (1996). "Inference when a Nuisance Parameter is not Identified Under the Null Hypothesis," Econometrica 64, 413-430.

- Hansen, B.E. (1997). "Inference in TAR models," Studies in Nonlinear Dynamics and Econometrics 2, 1-14.

- Hansen, B.E. (2000). "Sample Splitting and Threshold Estimation," Econometrica 68, 575-603.

- Lee, L.-F. (1984). "Tests for the Bivariate Normal Distribution in Econometric Models with Selectivity," Econometrica 52, 843-863.

- van der Vaart, A. W., and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes.* Springer, New York.

Table 1: Bivariate normal

		DGP1.000	DGP2	DGP3	DGP4
$\alpha^* = 0.10$					
n=100	sup	0.024	0.018	0.634	0.732
	expave	0.114	0.132	0.882	0.928
n=200	sup	0.012	0.014	0.910	0.966
	expave	0.118	0.122	0.988	1.000
n=500	sup	0.008	0.014	1.000	1.000
	expave	0.078	0.092	1.000	1.000
n=1000	sup	0.016	0.008	1.000	1.000
	expave	0.086	0.070	1.000	1.000
$\alpha^* = 0.05$					
n=100	sup	0.014	0.012	0.534	0.668
	expave	0.084	0.080	0.840	0.884
n=200	sup	0.010	0.006	0.874	0.952
	expave	0.074	0.076	0.976	0.998
n=500	sup	0.002	0.010	1.000	1.000
	expave	0.026	0.052	1.000	1.000
n=1000	sup	0.006	0.004	1.000	1.000
	expave	0.062	0.044	1.000	1.000
$\alpha^* = 0.01$					
n=100	sup	0.004	0.002	0.360	0.504
	expave	0.026	0.024	0.646	0.744
n=200	sup	0.004	0.000	0.756	0.886
	expave	0.018	0.022	0.894	0.960
n=500	sup	0.000	0.000	1.000	1.000
	expave	0.010	0.018	1.000	1.000
n=1000	sup	0.002	0.002	1.000	1.000
	expave	0.020	0.012	1.000	1.000

Table 2: Bivariate t_2

		mixture of two t distr with 2 dof			
		DGP1.000	DGP2	DGP3	DGP4
$\alpha^* = 0.10$					
n=100	sup	0.046	0.126	0.998	0.496
	expave	0.206	0.184	1.000	0.578
n=200	sup	0.022	0.096	1.000	0.504
	expave	0.162	0.120	1.000	0.550
n=500	sup	0.014	0.098	1.000	0.608
	expave	0.184	0.098	1.000	0.622
n=1000	sup	0.016	0.076	1.000	0.682
	expave	0.162	0.076	1.000	0.688
$\alpha^* = 0.05$					
n=100	sup	0.028	0.112	0.992	0.466
	expave	0.182	0.150	1.000	0.534
n=200	sup	0.018	0.082	1.000	0.460
	expave	0.126	0.096	1.000	0.498
n=500	sup	0.006	0.084	1.000	0.576
	expave	0.164	0.092	1.000	0.594
n=1000	sup	0.008	0.070	1.000	0.652
	expave	0.138	0.074	1.000	0.658
$\alpha^* = 0.01$					
n=100	sup	0.012	0.076	0.986	0.374
	expave	0.094	0.094	0.996	0.434
n=200	sup	0.006	0.056	1.000	0.350
	expave	0.072	0.058	1.000	0.384
n=500	sup	0.004	0.062	1.000	0.478
	expave	0.082	0.068	1.000	0.496
n=1000	sup	0.004	0.042	1.000	0.574
	expave	0.074	0.042	1.000	0.586

Table 3: Bivariate t_3

		DGP1.000	DGP2	DGP3	DGP4
$\alpha^* = 0.10$					
n=100	sup	0.020	0.092	0.966	0.562
	expave	0.160	0.154	0.994	0.668
n=200	sup	0.018	0.046	1.000	0.678
	expave	0.156	0.082	1.000	0.750
n=500	sup	0.022	0.028	1.000	0.836
	expave	0.122	0.040	1.000	0.852
n=1000	sup	0.016	0.036	1.000	0.904
	expave	0.122	0.042	1.000	0.910
$\alpha^* = 0.05$					
n=100	sup	0.016	0.070	0.954	0.492
	expave	0.126	0.120	0.986	0.628
n=200	sup	0.014	0.034	1.000	0.626
	expave	0.126	0.058	1.000	0.716
n=500	sup	0.014	0.022	1.000	0.810
	expave	0.092	0.028	1.000	0.826
n=1000	sup	0.008	0.026	1.000	0.884
	expave	0.080	0.032	1.000	0.892
$\alpha^* = 0.01$					
n=100	sup	0.008	0.036	0.916	0.370
	expave	0.058	0.060	0.966	0.468
n=200	sup	0.002	0.014	1.000	0.508
	expave	0.040	0.030	1.000	0.558
n=500	sup	0.002	0.012	1.000	0.734
	expave	0.032	0.018	1.000	0.760
n=1000	sup	0.000	0.016	1.000	0.828
	expave	0.030	0.016	1.000	0.832

Table 4: Mixture of χ_3^2 and normal

		DGP1	DGP2	DGP3	DGP4
$\alpha^* = 0.10$					
n=100	sup	0.018	0.244	0.628	0.962
	expave	0.130	0.430	0.856	0.988
n=200	sup	0.020	0.430	0.902	1.000
	expave	0.110	0.630	0.988	1.000
n=500	sup	0.016	0.840	1.000	1.000
	expave	0.100	0.944	1.000	1.000
n=1000	sup	0.016	0.988	1.000	1.000
	expave	0.096	0.996	1.000	1.000
$\alpha^* = 0.05$					
n=100	sup	0.006	0.186	0.554	0.920
	expave	0.086	0.334	0.794	0.974
n=200	sup	0.010	0.352	0.854	1.000
	expave	0.076	0.526	0.968	1.000
n=500	sup	0.012	0.770	1.000	1.000
	expave	0.078	0.886	1.000	1.000
n=1000	sup	0.012	0.984	1.000	1.000
	expave	0.068	0.992	1.000	1.000
$\alpha^* = 0.01$					
n=100	sup	0.002	0.080	0.398	0.800
	expave	0.018	0.162	0.618	0.890
n=200	sup	0.002	0.190	0.718	0.992
	expave	0.028	0.292	0.898	0.996
n=500	sup	0.002	0.580	1.000	1.000
	expave	0.018	0.680	1.000	1.000
n=1000	sup	0.006	0.940	1.000	1.000
	expave	0.018	0.964	1.000	1.000