Criterion-Based Inference for GMM in Linear Dynamic Panel Data Models

Criterion-Based Inference for GMM in Linear Dynamic Panel Data Models*

Stephen Bond
Nuffield College, Oxford and Institute for Fiscal Studies

Frank Windmeijer
Institute for Fiscal Studies
7 Ridgmount Street
London WC1E 7AE

January 31, 2000

Abstract

In this paper we examine the properties of some simple criterion-based, likelihood ratio type tests of parameter restrictions for standard GMM estimators in linear dynamic panel data models. A comparison is made with recent test proposals based on the continuously-updated GMM criterion (Hansen, Heaton and Yaron, 1996) or exponential tilting parameters (Imbens, Spady and Johnson, 1998). The preferred likelihood ratio type statistic is computed simply as the difference between the standard GMM tests of overidentifying restrictions in the restricted and unrestricted models. In Monte Carlo simulations we find this has similar properties to the two computationally more burdensome tests.

*__Acknowledgments:__ We thank Richard Spady and Clive Bowsher for alerting us to some of the possibilities that are explored in this paper. This work forms part of the research programme of the ESRC Centre for the Microeconomic Analysis of Fiscal Policy at the Institute for Fiscal Studies. The financial support of the ESRC is gratefully acknowledged.

1. Introduction

The problems with doing inference based on the efficient two-step GMM estimator for panel data are well known. Due to the fact that asymptotic standard errors are downward biased in small samples, standard Wald tests are oversized, see for example Arellano and Bond (1991) and Koenker and Machado (1999). Because of this, it has become standard practice to use the one-step estimation results for more reliable inference. However, as this estimator is not efficient, one would expect the power properties of tests based on it to be sub-optimal.

In this paper we compare the size and power properties of some alternative tests of parameter restrictions. Recent papers by Hansen, Heaton and Yaron (1996), (HHY), and Imbens, Spady and Johnson (1998), (ISJ), have proposed testing procedures and shown that the small sample properties of these are superior to those of the standard GMM Wald test, albeit not in the context of (dynamic) panel data models. HHY advocate use of a criterion-based test, using the continuously-updated GMM estimator, which is equivalent to robust LIML. This estimator requires numerical methods for optimisation, which is documented to have convergence problems and multi modality, see HHY, Arellano and Alonso-Borrego (1999) and ISJ (1998). ISJ use the empirical likelihood framework and advocate use of a weighted optimisation criterion, exponential tilting, and show that their criterion-based Hansen-Sargan test for overidentifying restrictions has better size properties than two-step, iterated and continuously-updated GMM. They also show superior size properties for the exponential tilting test of parameter restrictions compared to the Wald statistic in their model, but do not compare it to the HHY test.

We compare the properties of these test statistics with some simple tests based on the standard GMM criterion. These tests are of the "likelihood ratio" form, comparing the minimised GMM criterion function under the null to the criterion function under the alternative, as documented in for example Davidson and MacKinnon (1993). As the moment conditions we consider are linear in the parameters, different choices of weight matrices in the restricted and unrestricted models give rise to different LR statistics with LM and Wald test equivalences. The Monte Carlo investigation shows that the LR statistic that is computed simply as the difference between the standard GMM tests for overidentifying restrictions in the restricted and unrestricted models behaves well. This is found to have similar size and power properties as the computationally more burdensome tests based on the continuously-updated estimator or the exponential tilting parameters.

2. GMM and Test Statistics

Consider the moment conditions

$$E\left[g\left(X_i,\theta_0\right)\right]=E\left[g_i\left(\theta_0\right)\right]=0,$$

where $g\left(.\right)$ is vector of order q and θ_0 is a parameter vector of order k. The GMM estimator $\widehat{\theta}_N$ for θ_0 minimises[1]

$$\left[\frac{1}{N}\sum_{i=1}^{N}g_i\left(\theta\right)\right]'W_N\left[\frac{1}{N}\sum_{i=1}^{N}g_i\left(\theta\right)\right],$$

with respect to θ; where W_N is a positive semidefinite weight matrix which satisfies $\operatorname{plim}_{N\rightarrow\infty}W_N=W$, with W a positive definite matrix. Regularity conditions are assumed such that $\lim_{N\rightarrow\infty}\frac{1}{N}\sum_{i=1}^{N}g_i\left(\theta\right)=E\left[g_i\left(\theta\right)\right]$ and $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}g_i\left(\theta_0\right)\rightarrow N\left(0,\Psi\right)$. Let $\Gamma\left(\theta\right)=E\left[\partial g_i\left(\theta\right)/\partial\theta\right]$ and $\Gamma_0\equiv\Gamma\left(\theta_0\right)$, then $\sqrt{N}\left(\widehat{\theta}_N-\theta_0\right)$ has a limiting normal distribution, $\sqrt{N}\left(\widehat{\theta}_N-\theta_0\right)\rightarrow N\left(0,V_W\right)$, where

$$V_W=\left(\Gamma_0'W\Gamma_0\right)^{-1}\Gamma_0'W\Psi W\Gamma_0\left(\Gamma_0'W\Gamma_0\right)^{-1}. \tag{2.1}$$

[1] See Hansen (1982).

The efficient GMM estimator is based on a weight matrix that satisfies $\text{plim}_{N\to\infty} W_N = \Psi^{-1}$, with $V_W = \left(\Gamma_0'\Psi^{-1}\Gamma_0\right)^{-1}$. A weight matrix that satisfies this property is given by

$$W_N = \widehat{\Psi}_N^{-1} = \left(\frac{1}{N}\sum_{i=1}^{N} g_i\left(\overline{\theta}_N\right) g_i\left(\overline{\theta}_N\right)'\right)^{-1}, \tag{2.2}$$

where $\overline{\theta}_N$ is a consistent estimator for θ_0.

Denote $g(\theta) = \frac{1}{N}\sum_{i=1}^{N} g_i(\theta)$. The standard test for overidentifying restrictions is N times the minimised GMM criterion, given by

$$J\left(\widehat{\theta}_N\right) = Ng\left(\widehat{\theta}_N\right)' \widehat{\Psi}_N^{-1} g\left(\widehat{\theta}_N\right),$$

and has an asymptotic chi-squared distribution with $q-k$ degrees of freedom when the moment conditions are valid.

For testing r restrictions of the form

$$r(\theta_0) = 0,$$

the criterion-based test statistics we consider are given by

$$D_{AB} = N\left(g\left(\widetilde{\theta}_A\right)' A_N^{-1} g\left(\widetilde{\theta}_A\right) - g\left(\widehat{\theta}_B\right)' B_N^{-1} g\left(\widehat{\theta}_B\right)\right),$$

where A_N and B_N are consistent estimates of Ψ under the null, $\widetilde{\theta}_A$ is the GMM estimator imposing the restrictions under the null and using the weight matrix A_N^{-1}, and $\widehat{\theta}_B$ is the unrestricted GMM estimator using the weight matrix B_N^{-1}. Under the null, D_{AB} has an asymptotic chi-squared distribution with r degrees of freedom.

We consider the following three choices of weight matrices A_N and B_N:

$$\begin{aligned} D_{UU} &: \; A_N = B_N = \widehat{\Psi}_N; \\ D_{RR} &: \; A_N = B_N = \widetilde{\Psi}_N; \\ D_{RU} &: \; A_N = \widetilde{\Psi}_N,\; B_N = \widehat{\Psi}_N, \end{aligned}$$

where $\widehat{\Psi}_N$ is the estimator for Ψ in the unrestricted model based on an initial consistent estimator $\overline{\theta}_N$, as in (2.2), and $\widetilde{\Psi}_N$ is the estimator for Ψ in the restricted model based on an initial consistent estimator of the restricted parameter vector, $\overline{\overline{\theta}}_N$.

As shown by Newey and West (1987), in the case of moment conditions that are linear in the parameters, D_{UU} is equivalent to the standard Wald test, D_{RR} is equivalent to a Wald test in the unrestricted model using the efficient weight matrix under the null,[2] and is further equivalent to the LM test. D_{RU} is the "likelihood ratio" equivalent for GMM, see also Davidson and MacKinnon (1993, pp. 614-620).

HHY proposed the use of a statistic similar to D_{RU} for the continuously-updated GMM estimator. This estimator is equivalent to robust LIML and is defined as the value of θ, denoted $\widehat{\theta}_N^{CU}$, that minimizes

$$J^{CU}(\theta) = g(\theta)' \left(\frac{1}{N} \sum_{i=1}^{N} g_i(\theta) g_i(\theta)' \right)^{-1} g(\theta).$$

The test statistic D_{RU}^{CU} is then defined as

$$D_{RU}^{CU} = N \left(J^{CU}\left(\widetilde{\theta}_N^{CU}\right) - J^{CU}\left(\widehat{\theta}_N^{CU}\right) \right),$$

where $\widehat{\theta}_N^{CU}$ is the continuously updated GMM estimator for the unrestricted model and $\widetilde{\theta}_N^{CU}$ is the continuously updated GMM estimator for the restricted model.

The ISJ test statistic is based on the empirical likelihood method. Their "exponential tilting" estimator for θ_0 minimises the Kullback-Leibler information criterion

$$\min_{\pi,\theta} \sum_{i=1}^{N} \pi_i \ln \pi_i \quad \text{subject to} \quad \sum_{i=1}^{N} g_i(\theta) \pi_i = 0 \ \text{ and } \ \sum_{i=1}^{N} \pi_i = 1.$$

[2] For example, if the null hypothesis is $H_0 : \theta_0 = \delta$, we form the weight matrix in (2.2) using $g_i(\delta)$ in place of $g_i\left(\overline{\theta}_N\right)$.

The estimated probabilities have the form

$$\pi_i = \frac{\exp\left(\gamma' g_i\left(\theta\right)\right)}{\sum_{j=1}^{N} \exp\left(\gamma' g_j\left(\theta\right)\right)}$$

where γ are the tilting parameters. Intuitively these measure how much the sample has to be re-weighted in order for the moment conditions to hold exactly. Tilting parameters can also be estimated conditional on the standard GMM estimator of the parameters θ. A test based on the restricted and unrestricted GMM estimators is the difference

$$D_{RU}^{ET} = N\left(\widetilde{\gamma}'_{N(GMM)} R_N\left(\widetilde{\theta}_N\right)\widetilde{\gamma}_{N(GMM)} - \widehat{\gamma}'_{N(GMM)} R_N\left(\widehat{\theta}_N\right)\widehat{\gamma}_{N(GMM)}\right)$$

where

$$\widehat{\gamma}_{N(GMM)} = \max_{\gamma} \frac{1}{N}\sum_{i=1}^{N} \exp\left(\gamma' g_i\left(\widehat{\theta}_N\right)\right),$$

and $\widehat{\theta}_N$ is the efficient two-step GMM estimator in the unrestricted model; $\widetilde{\gamma}_{N(GMM)}$ is the equivalent estimator of the tilting parameters based on the efficient two-step GMM estimator in the unrestricted model, $\widetilde{\theta}_N$; and

$$R_N\left(\theta\right) = \left[\frac{1}{N}\sum_{i=1}^{N} g_i\left(\theta\right) g_i\left(\theta\right)'\pi_i\right]\left[\frac{1}{N}\sum_{\iota=1}^{N} g_i\left(\theta\right) g_i\left(\theta\right)'\pi_i\pi_i\right]^{-1}\left[\frac{1}{N}\sum_{i=1}^{N} g_i\left(\theta\right) g_i\left(\theta\right)'\pi_i\right].$$

Both D_{RU}^{CU} and D_{RU}^{ET} have an asymptotic chi-squared distribution with r degrees of freedom, and have been shown by respectively HHY and ISJ to have better finite sample properties than the conventional Wald tests in particular contexts. Also, the corresponding tests of overidentifying restrictions have been shown to have better finite sample behaviour than the standard GMM test. Sofar as we are aware, the HHY $\left(D_{RU}^{CU}\right)$ and ISJ $\left(D_{RU}^{ET}\right)$ tests of linear restrictions have not been compared to either the LM (D_{RR}) test or the criterion-based test using the standard GMM criterion (D_{RU}). We consider this in the context of linear dynamic panel data models.

3. AR1 Process with Individual Effects

To evaluate the finite sample behaviour of the various test statistics described in the previous section, we consider the linear first order autoregressive panel data model with individual effects (η_i)

$$y_{it} = \alpha y_{it-1} + \eta_i + u_{it}$$

where $i = 1, ..., N$ and $t = 2, ..., T$, with T fixed. Under some basic assumptions (see Ahn and Schmidt, 1995) the following $(T-1)(T-2)/2$ linear moment conditions are valid

$$E\left[y_i^{t-2}\left(\Delta y_{it} - \alpha \Delta y_{it-1}\right)\right] = 0; \tag{3.1}$$

where $y_i^{t-2} = [y_{i1}, y_{i2}, ..., y_{it-2}]$. We call these moment conditions the DIFF moment conditions, see Arellano and Bond (1991). Under some further assumptions on initial conditions (see Blundell and Bond, 1998), the additional $(T-2)$ linear moment conditions

$$E\left[\Delta y_{it-1}\left(y_{it} - \alpha y_{it-1}\right)\right] = 0 \tag{3.2}$$

are valid. The joint moment conditions (3.1) and (3.2) are the so-called SYSTEM moment conditions, see Arellano and Bover (1995) and Blundell and Bond (1998).

Let Z_i be the matrix of instruments for observation i, then the moment conditions can be written as $E\left[Z_i' v(\alpha)\right] = 0$. The efficient one-step GMM weight matrix for the DIFF moment conditions when the u_{it} are homoscedastic and not serially correlated is given by $W_N = \left(\frac{1}{N}\sum_{i=1}^N Z_i' H Z_i\right)^{-1}$, where H is a $(T-2)$ square matrix which has 2's on the main diagonal, -1's on the first subdiagonals and zeros elsewhere. For the SYSTEM moment conditions there is no simple one-step efficient weight matrix, and often the one-step weight matrix is set to $W_N = \left(\frac{1}{N}\sum_{i=1}^N Z_i' Z_i\right)^{-1}$. The efficient weight matrix for both estimators under general conditions is given by $\widehat{\Psi}_N^{-1} = \left(\frac{1}{N}\sum_{i=1}^{N} Z_i' v_i(\overline{\alpha}) v_i(\overline{\alpha})' Z_i\right)^{-1}$, with $\overline{\alpha}$ the consistent one-step GMM estimator of α.

In Table 1 we present some Monte Carlo results for the AR1 panel data process. The data generating process is

$$\begin{aligned} y_{it} &= \alpha y_{it-1} + \eta_i + u_{it} \\ \eta_i &\sim N(0,1) \;\; ; \;\; u_{it} \sim N(0,1) \\ y_{i1} &= \frac{\eta_i}{1-\alpha} + v_i \;\; ; \;\; v_i \sim N\left(0, \frac{1}{1-\alpha^2}\right). \end{aligned}$$

The sample size is $N = 100$, $T = 6$, and we report the size properties of the test statistics for 10000 samples. We consider tests of the null hypothesis $H_0 : \alpha = \alpha_0$. Note that the test of overidentifying restrictions in the restricted model (i.e. imposing this null), using the efficient weight matrix under the null, is in this case the same as the Anderson-Rubin test statistic. Further, the D_{RR}, or LM test statistic is in this case similar in spirit to the T_{00} statistic of Wang and Zivot (1998),[3] see also Pagan and Robertson (1997).

Table 1 compares the size properties of the various test statistics for $\alpha = 0.3$. The statistic W_1 is the Wald test based on the one-step GMM estimator and its asymptotic standard error. D_{UU}/W_2 is the Wald test based on the efficient two-step GMM estimation results. W_{CU} is the Wald test based on the continuously-updated GMM estimation results. D_{RR}/LM is the LM test, and D_{RU}, D_{RU}^{CU} and D_{RU}^{ET} are the criterion-based tests described above.[4,5]

[3] The D_{RR} statistic is the same as the T_{00} statistic of Wang and Zivot (1998) for the DIFF moment conditions under the assumption of homoscedasticity and non-serial correlation of the u_{it}, using weight matrix $W_N = \left(\widehat{\sigma}_v^2 \frac{1}{N}\sum_{i=1}^{N} Z_i' H Z_i\right)^{-1}$ with $\widehat{\sigma}_v^2$ estimated using the value of α under the null.

[4] The test statistics D_{RU} and D_{RU}^{ET} can be negative in finite samples. When a statistic is negative, we interpret this as a non-rejection of the null hypothesis.

[5] For the calculation of the continuously-updated estimator and the exponential tilting parameters we used Maxlik 4.0 in Gauss with analytical derivatives.

Table 1. Size comparisons, AR1 model, $N = 100$, $T = 6$, $\alpha = 0.3$, 10000 replications

	DIFF					
size	W_1	D_{UU}/W_2	D_{RR}/LM	D_{RU}	D_{RU}^{CU}	D_{RU}^{ET}
0.20	0.2300	0.3071	0.2174	0.2119	0.2323	0.2174
0.10	0.1252	0.1917	0.1170	0.1086	0.1211	0.1176
0.05	0.0676	0.1245	0.0578	0.0517	0.0592	0.0626
0.01	0.0189	0.0453	0.0096	0.0088	0.0108	0.0124
	SYSTEM					
	W_1	D_{UU}/W_2	D_{RR}/LM	D_{RU}	D_{RU}^{CU}	D_{RU}^{ET}
0.20	0.2185	0.3553	0.2198	0.2357	0.2355	0.2128
0.10	0.1178	0.2364	0.1129	0.1186	0.1227	0.1175
0.05	0.0610	0.1583	0.0556	0.0636	0.0666	0.0609
0.01	0.0164	0.0637	0.0115	0.0135	0.0152	0.0168

For this value of α, the test statistics perform quite similarly, with the exception of the D_{RU}/W_2 statistic which is substantially oversized. For the DIFF moment conditions, D_{RU} has the best overall size properties, whereas for the SYSTEM moment conditions, D_{RR}/LM performs best in terms of size. Figures 1 to 6 show p-value plots (see Davidson and MacKinnon, 1996) for the sizes of the various test statistics for both the DIFF and SYSTEM moment conditions. We consider values of α of 0, 0.4, and 0.8 respectively, and the results are based on 10000 Monte Carlo replications. For high values of α the size properties of the statistics for the DIFF moment conditions diverge, with only D_{RU} and D_{RU}^{ET} having good size properties, and all other tests being oversized. This is due to the fact that the GMM estimator based on the DIFF moment conditions is downward biased in small samples for high values of α. This affects the standard Wald tests, and also the D_{RR}/LM test, as this is a Wald test on the unrestricted GMM estimator using the restricted weight matrix. The continuously-updated estimator has some convergence problems for the DIFF moment conditions when α is high.[6]

[6] For the DIFF moment conditions, when $\alpha = 0.8$, the continuously-updated estimation procedure did not converge in 0.3% of the samples. When this occurred, we discarded the

These problems are all due to the fact that instruments become weak for high α. We will discuss weak instruments further in section 5.

The GMM estimator based on the SYSTEM moment conditions has been shown by Blundell and Bond (1998) to have much smaller small sample bias, even for high values of α. As these moment conditions are much more informative, the D_{RR}/LM test statistic has superior size properties. The simple D_{RU} test behaves similarly to the tests based on the continuously-updated estimator and the tilting parameters. When $\alpha = 0.8$, the one-step Wald test is oversized, whereas D_{RU}^{ET} tends to be undersized. The continuously-updated estimator has no convergence problems in this case, but the D_{RU}^{CU} test statistic is oversized, more so than the D_{RU} test.

Figures 8 to 10 display the power of the tests at the 5% level of significance testing $H_0 : \alpha = 0$ and $H_0 : \alpha = 0.6$ respectively, again based on 10000 Monte Carlo replications. The power function is calculated for the values of $\alpha = 0, 0.1, ..., 0.8$, and corrected for size distortions. For the DIFF moment conditions, the D_{RR}/LM power properties for $H_0 : \alpha = 0$ are worst, due to the fact that the use of the weight matrix under the null biases the estimator towards zero, particularly at high values of α. The one-step Wald test has the best power for these moment conditions, as the efficient one-step weight matrix is used here.

The power properties of the criterion based test statistics are all very similar in the case of the SYSTEM moment conditions. The one-step Wald test now has least power when testing $H_0 : \alpha = 0$. It seems to have high power for the test $H_0 : \alpha = 0.6$, when α is large, but for these case its size properties are poor.

---

results and generated a new sample. However, the distribution of the converged estimation results shows some extremely outlying estimates.

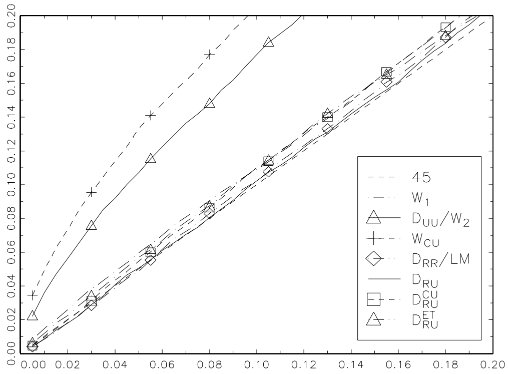

Figure 1. p-value plot, $\alpha = 0$, DIFF

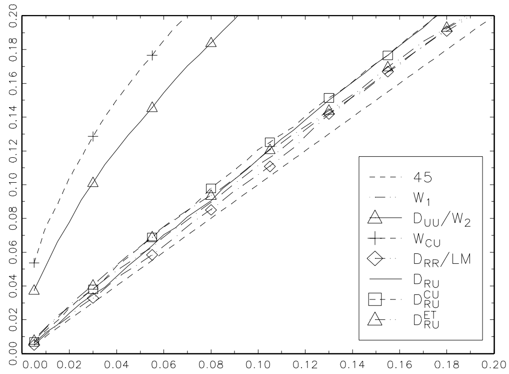

Figure 2. p-value plot, $\alpha = 0$, SYSTEM

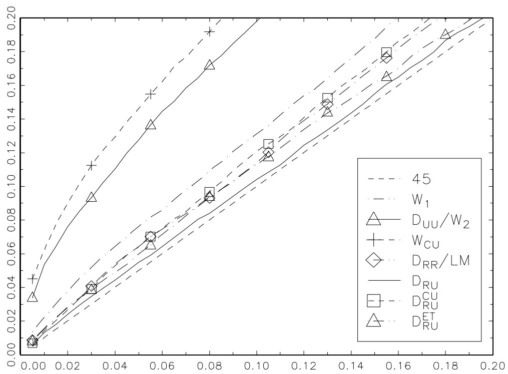

Figure 3, p-value plot, $\alpha = 0.4$, DIFF

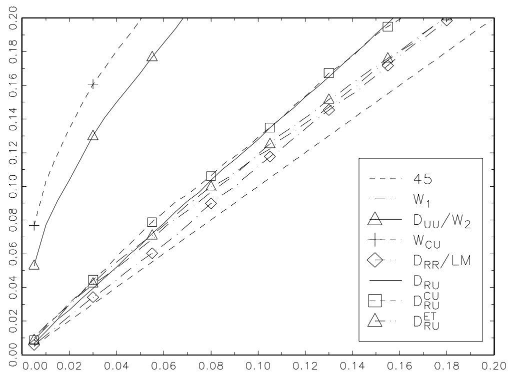

Figure 4. p-value plot, $\alpha = 0.4$, SYSTEM

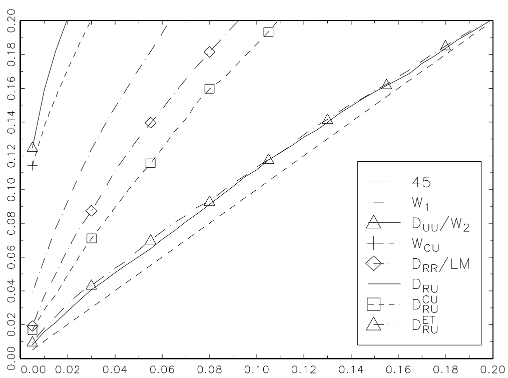

Figure 5. p-value plot, $\alpha = 0.8$, DIFF

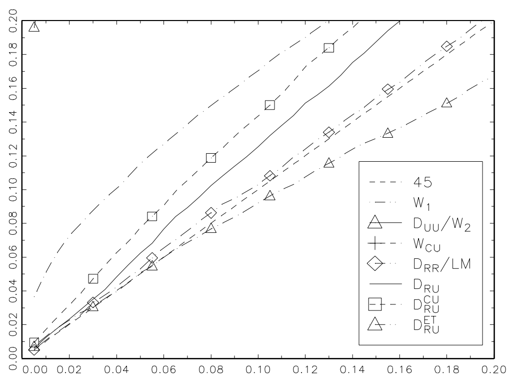

Figure 6, p-value plot, $\alpha = 0.8$, SYSTEM

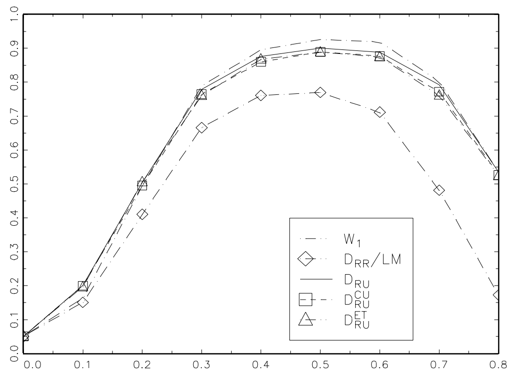

Figure 7. Power plot, $H_0 : \alpha = 0$, DIFF

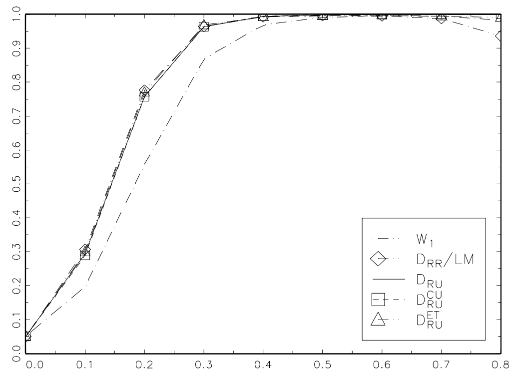

Figure 8. Power plot, $H_0 : \alpha = 0$, SYSTEM

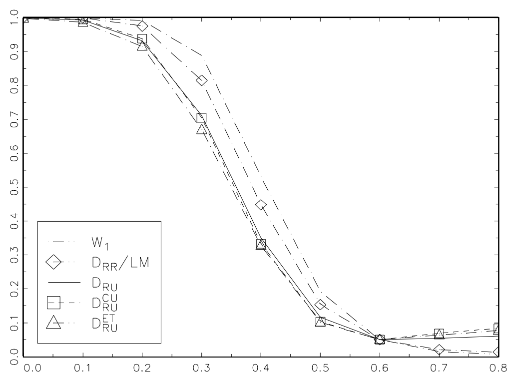

Figure 9. Power plot, $H_0 : \alpha = 0.6$, DIFF

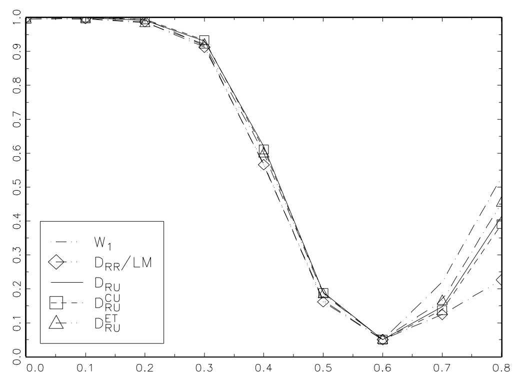

Figure 10. Power plot, $H_0 : \alpha = 0.6$, SYSTEM

In conclusion, the results of these Monte Carlo simulations show that the simple D_{RU} test, based on the standard GMM criterion, performs quite well, very similar and often better than the recently proposed computationally more burdensome test statistics D_{RU}^{CU} and D_{RU}^{ET}. The D_{RR}/LM statistic has very good size properties, and also reasonable power properties, in models where weak identification is not an issue.

4. AR2 Process with Individual Effects

For the AR1 process, there were no unknown parameters to be estimated in the restricted model. To evaluate the performance of the test statistics in the case where there are unknown parameters under the null, consider the AR2 process

$$y_{it} = \alpha_1 y_{it-1} + \alpha_2 y_{it-2} + \eta_i + u_{it}$$

and the simple test $H_0 : \alpha_2 = \delta$. There are various ways to construct the weight matrix under the null. We propose use of the weight matrix based on the one-step GMM estimate for α_1 in the restricted model

$$y_{it} - \delta y_{it-2} = \alpha_1 y_{it-1} + \eta_i + u_{it},$$

which we denote by $\overline{\overline{\alpha}}_{1\delta}$, giving the two-step weight matrix

$$W_N\left(\overline{\overline{\alpha}}_{1\delta}, \delta\right) = \widetilde{\Psi}_N^{-1} = \left(\frac{1}{N}\sum_{i=1}^{N} Z_i' v_i\left(\overline{\overline{\alpha}}_{1\delta}, \delta\right) v_i\left(\overline{\overline{\alpha}}_{1\delta}, \delta\right)' Z_i\right)^{-1}.$$

The statistic D_{RU} is then simply the difference between the GMM tests of overidentifying restrictions in the restricted and unrestricted model, keeping the number of moment conditions constant.

Table 2 reports sizes and some size-corrected power properties of the various test statistics for the test $H_0 : \alpha_2 = 0$, for both the DIFF and SYSTEM moment conditions. The other parameters of the DGP are the same as in the previous section. Again, $N = 100$, but now $T = 7$. Results are largely the same as before. For the DIFF moment conditions, D_{RU} possesses the best size properties, and has power very similar to the power of D_{RU}^{CU} and D_{RU}^{ET}. D_{RU}^{CU} seems to have better power for high values of α_1, but this is accompanied by convergence problems for the continuously-updated estimator. The D_{RR}/LM test again shows problems similar to the one-step Wald test, being oversized and having very low power. For the SYSTEM moment conditions, the D_{RR}/LM test has superior size properties with similar power as the other criterion-based tests. For higher values of α_1 the one-step Wald test is the most powerful, but its size properties deteriorate with increasing α_1, as in the AR1 model. Size and power properties of D_{RU}, D_{RU}^{CU}, and D_{RU}^{ET} are quite similar, with D_{RU} having possibly the best overall size and power properties of the three.

Table 2. Size and size-corrected power comparisons, AR2 model
$N = 100$, $T = 7$, 10000 replications, $H_0 : \alpha_2 = 0$.

		DIFF					
		$\alpha_2 = 0$					
	size	W_1	D_{UU}/W_2	D_{RR}/LM	D_{RU}	D_{RU}^{CU}	D_{RU}^{ET}
	0.10	0.1435	0.2253	0.1291	0.1162	0.1285	0.1313
$\alpha_1 = 0.3$	0.05	0.0851	0.1506	0.0701	0.0585	0.0654	0.0762
	0.01	0.0257	0.0630	0.0166	0.0132	0.0143	0.0199
	0.10	0.1572	0.2439	0.1416	0.1199	0.1291	0.1280
$\alpha_1 = 0.5$	0.05	0.0945	0.1663	0.0794	0.0625	0.0692	0.0746
	0.01	0.0309	0.0744	0.0205	0.0143	0.0147	0.0204
	0.10	0.1845	0.2691	0.1608	0.1291	0.1409	0.1357
$\alpha_1 = 0.7$	0.05	0.1122	0.1939	0.0920	0.0710	0.0756	0.0778
	0.01	0.0404	0.0926	0.0250	0.0166	0.0156	0.0221
		$\alpha_2 = 0.1$					
		W_1	D_{UU}/W_2	D_{RR}/LM	D_{RU}	D_{RU}^{CU}	D_{RU}^{ET}
	0.10	0.1021	0.1003	0.1237	0.2322	0.2429	0.2327
$\alpha_1 = 0.3$	0.05	0.0456	0.0410	0.0607	0.1393	0.1487	0.1427
	0.01	0.0060	0.0051	0.0134	0.0397	0.0458	0.0447
	0.10	0.0480	0.0466	0.0726	0.1854	0.2129	0.1970
$\alpha_1 = 0.5$	0.05	0.0152	0.0146	0.0341	0.1070	0.1298	0.1172
	0.01	0.0014	0.0008	0.0056	0.0292	0.0351	0.0310
	0.10	0.0126	0.0140	0.0300	0.0979	0.1615	0.1129
$\alpha_1 = 0.7$	0.05	0.0037	0.0057	0.0092	0.0458	0.0899	0.0554
	0.01	0.0002	0.0001	0.0008	0.0046	0.0176	0.0106

Table 2 continued. Size and size-corrected power comparisons, AR2 model
$N = 100$, $T = 7$, 10000 replications, $H_0 : \alpha_2 = 0$.

		SYSTEM					
		$\alpha_2 = 0$					
	size	W_1	D_{UU}/W_2	D_{RR}/LM	D_{RU}	D_{RU}^{CU}	D_{RU}^{ET}
	0.10	0.1205	0.2723	0.1034	0.1301	0.1311	0.1432
$\alpha_1 = 0.3$	0.05	0.0652	0.1938	0.0539	0.0691	0.0702	0.0849
	0.01	0.0177	0.0935	0.0091	0.0132	0.0156	0.0224
	0.10	0.1211	0.2909	0.1000	0.1308	0.1326	0.1406
$\alpha_1 = 0.5$	0.05	0.0662	0.2114	0.0486	0.0677	0.0710	0.0837
	0.01	0.0186	0.1037	0.0084	0.0141	0.0153	0.0259
	0.10	0.1448	0.3204	0.0961	0.1271	0.1338	0.1356
$\alpha_1 = 0.7$	0.05	0.0837	0.2393	0.0468	0.0665	0.0704	0.0786
	0.01	0.0253	0.1306	0.0092	0.0161	0.0186	0.0256
		$\alpha_2 = 0.1$					
		W_1	D_{UU}/W_2	D_{RR}/LM	D_{RU}	D_{RU}^{CU}	D_{RU}^{ET}
	0.10	0.3539	0.3850	0.3683	0.3646	0.3465	0.3713
$\alpha_1 = 0.3$	0.05	0.2270	0.2531	0.2581	0.2571	0.2286	0.2567
	0.01	0.0739	0.0908	0.1050	0.1071	0.0787	0.0996
	0.10	0.4314	0.4168	0.3537	0.3560	0.3309	0.3639
$\alpha_1 = 0.5$	0.05	0.3116	0.2856	0.2373	0.2574	0.2254	0.2495
	0.01	0.1354	0.1032	0.0856	0.0982	0.0761	0.0841
	0.10	0.5807	0.5187	0.3727	0.3573	0.3123	0.3571
$\alpha_1 = 0.7$	0.05	0.4573	0.3875	0.2544	0.2521	0.1989	0.2462
	0.01	0.2234	0.1605	0.0999	0.0916	0.0656	0.0905

5. Weak Instruments

Under weak instruments, in the sense of Staiger and Stock (1997), the asymptotic distributions of the D_{RR}/LM and D_{RU} parameter tests are nonstandard. Following Dufour (1997), Wang and Zivot (1998) derive the asymptotic distributions of the boundedly pivotal TSLS LM and LIML LR statistics for IV regressions under weak instruments. They advocate the use of a pre-test method to determine the identification of the model, and the use of different critical values according to the outcome of the pre-test.

Consider the AR1 model with the parameter test $H_0 : \alpha = \alpha_0$. Both D_{RR}/LM and D_{RU} are bounded by the Anderson-Rubin statistic $Ng(\alpha_0)' \widehat{\Psi}(\alpha_0) g(\alpha_0)$ which is asymptotically χ^2_q distributed under the null, regardless of identification issues. A pre-test of identification for GMM in panel data models is the Arellano, Hansen and Sentana (1999) test for underidentification. The Wang-Zivot strategy is to use the χ^2_1 critical values for the tests if the underidentification test does not indicate identification problems, and the use the χ^2_q critical values when it does.

For the AR1 model the test for underidentification is based on the moments

$$E\left[y_{it}^{t-1}\Delta y_{it}\right] = 0,$$

and the model is underidentified if the test statistic for overidentifying restrictions is smaller than the specified critical value of the χ^2_s distribution, with $s = T(T-1)/2$.

Table 3 shows the size properties for the two tests together with the results of the underidentification test. When the adjustment is made for underidentification, the size properties of the D_{RR}/LM test improve considerably. The adjusted D_{RU} test is undersized.

Table 3. AR1 model, $N = 100$, $T = 6$, size properties for test results unadjusted and adjusted (*) for underidentification, DIFF moment conditions, 10000 replications

	size	D_{RR}/LM	D_{RU}	SU	D^*_{RR}/LM^*	D^*_{RU}
	0.10	0.2164	0.1139	0.2731	0.1621	0.0873
$\alpha = 0.8$	0.05	0.1305	0.0628	0.4056	0.0825	0.0425
	0.01	0.0363	0.0164	0.6833	0.0141	0.0065
	0.10	0.3680	0.1741	0.7366	0.1092	0.0533
$\alpha = 0.9$	0.05	0.2438	0.1105	0.8447	0.0466	0.0212
	0.01	0.0807	0.0314	0.9632	0.0048	0.0014
	0.10	0.4537	0.2248	0.8533	0.0890	0.0499
$\alpha = 0.95$	0.05	0.3164	0.1438	0.9242	0.0388	0.0194
	0.01	0.1135	0.0478	0.9856	0.0046	0.0018

SU is test for underidentification. Results indicate frequency of underidentification.

6. Conclusions

In this paper we have considered the properties of a simple test of parameter restrictions based on standard two-step efficient GMM estimators. The test is computed simply as the difference between the minimised values of the GMM criterion function in the restricted and unrestricted models. We compared this to criterion-based tests of parameter restrictions based on the continuously-updated GMM estimator of Hansen, Heaton and Yaron (1996) and the exponential tilting proposal of Imbens, Spady and Johnson (1998), as well as to standard asymptotic Wald tests, and to the LM test statistic which is easily computed in the case of moment conditions that are linear in the parameters.

We investigated the properties of these tests using Monte Carlo experiments in the context of simple parameter restrictions in linear dynamic panel data models. Our main finding is that the test based on the standard GMM criterion function has very similar properties to the computationally more burdensome alternatives.

In future research we will investigate whether this finding holds in more general settings, for example in the context of non-linear restrictions and non-linear models.

References

[1] Ahn, S.C. and P. Schmidt (1995), Efficient Estimation of Models for Dynamic Panel Data, *Journal of Econometrics*, 68, 5-28.

[2] Alonso-Borrego, C. and M. Arellano (1999), Symmetrically Normalised Instrumental-Variable Estimation using Panel Data, *Journal of Business & Economic Statistics* 17, 36-49.

[3] Arellano, M. and S. Bond (1991), Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations, *Review of Economic Studies* 58, 277-98.

[4] Arellano, M. and O. Bover (1995), Another Look at the Instrumental-Variable Estimation of Error-Components Models, *Journal of Econometrics* 68, 29-51.

[5] Arellano, M., L.P. Hansen and E. Sentana (1999), Underidentification?, mimeo, CEMFI, Madrid.

[6] Davidson, R. and J.G. MacKinnon (1993), *Estimation and Inference in Econometrics*, Oxford University Press, New York.

[7] Davidson, R. and J.G. MacKinnon (1996), Graphical Methods for Investigating the Size and Power of Hypothesis Tests, *Manchester School* 66, 1-26.

[8] Dufour, J.-M., (1997), Some Impossibility Theorems in Econometrics with Applications to Structural and Dynamic Models, *Econometrica* 65, 1365-1387.

[9] Hansen, L.P., (1982), Large Sample Properties of Generalised Method of Moments Estimators, *Econometrica* 50, 1029-1054.

[10] Hansen, L.P., J. Heaton and A. Yaron (1996), Finite-Sample Properties of some Alternative GMM Estimators, *Journal of Business & Economic Statistics* 14, 262-280.

[11] Imbens, G.W., R.H. Spady and P. Johnson (1998), Information Theoretic Approaches to Inference in Moment Condition Models, *Econometrica* 66, 333-357.

[12] Koenker, R. and J.A.F. Machado (1999), GMM Inference when the Number of Moment Conditions is Large, *Journal of Econometrics* 93, 327-344.

[13] Newey, W.K. and K.D. West (1987), Hypothesis Testing with efficient Method of Moments Estimation, *International Economic Review* 28, 777-787.

[14] Pagan, A.R. and J. Robertson, GMM and its Problems, mimeo, Australian National University.

[15] Wang, J. and E. Zivot (1998), Inference on Structural Parameters in Instrumental Variables Regression with Weak Instruments, *Econometrica* 66, 1389-1404.