Generalized empirical likelihood Pearson-type specification tests

Joaquim J. S. Ramalho
Departamento de Economia
Universidade de Evora

Richard J. Smith
Department of Economics
University of Bristol

This draft: February 2002

Abstract

This paper considers new ways of constructing generalized empirical likelihood (GEL) specification tests for moment condition models. In doing so, we explore an attractive feature of GEL estimation: the possibility of estimating a set of weights, the so-called GEL implied probabilities, which impose numerically all moment conditions on the data. The resultant GEL distribution function, as exploits the information contained in the moment conditions defining a particular model, is a more efficient estimator of the distribution of the data than the empirical distribution function. Contrasting certain functions of these two consistent estimators of the data generating process gives rise, as we show, to Pearson-type tests of overidentifying moment conditions. We develop two classes of Pearson-type tests. The first includes tests that are very similar in form to the classical Pearson χ^2 statistics. The other requires the partition of the sample space into several sets, the contrast between the empirical and the GEL implied probabilities estimated for each set forming the basis for the test. We show also that a similar approach can be used to construct Pearson-type tests for parametric restrictions, in which case two GEL distributions estimated under different assumptions are contrasted. A Monte Carlo simulation analysis concerning tests of overidentifying moment conditions shows that the size behaviour in finite samples of one of the Pearson-type statistics proposed is clearly superior to that of alternative tests.

Keywords: Pearson-type tests, generalized empirical likelihood, overidentifying moment conditions, parametric restrictions.

JEL classification: C12

1 Introduction

One of the popular model formalizations in econometrics requires only the specification of a set of moment conditions which the model to be estimated should satisfy. During the last two decades or so, the generalized method of moments (GMM) due to Hansen (1982) has been the conventional approach to estimation and inference in such models. However, there is increasing Monte Carlo evidence indicating that in finite samples GMM estimators may be badly biased and the associated tests may have actual sizes substantially different from the nominal ones. See, for example, the Special Issue of the *Journal of Business & Economic Statistics* (July 1996). Hence, a number of alternative estimators, which are asymptotically first-order equivalent to efficient GMM estimation, have been suggested recently. These include the empirical likelihood (EL) estimator of Qin and Lawless (1994) and Imbens (1997), and the exponential tilting (ET) estimator of Kitamura and Stutzer (1997) and Imbens, Spady and Johnson (1998). As shown by Smith (1997), these two estimators share a common structure, being members of a class of generalized empirical likelihood (GEL) estimators.

GEL estimation seems to possess many attractive theoretical features relative to GMM. Indeed, large sample analysis by Newey and Smith (2001) indicates that GEL estimators may be less prone to bias than GMM estimators. Furthermore, as likelihood-like methods, they allow the utilization of classical-type tests to evaluate various hypotheses concerning the specification of a particular model, including overidentifying moment conditions, whereas in the GMM framework Hansen's (1982) J test is the only way of assessing such issue. Also in finite samples GEL estimation seems to have diverse advantages over GMM: Imbens (1997) and Newey, Ramalho and Smith (2001) reported promising Monte Carlo results concerning the small sample bias of GEL estimators, while Imbens, Spady and Johnson (1998) found that GEL tests of overidentifying moment conditions, although also oversized in finite samples, possess actual sizes closer to the asymptotic ones than the J test.

This paper considers new ways of constructing GEL specification tests. In particular, we explore another attractive feature of GEL estimation: the possibility of estimating a set of weights, the so-called GEL implied probabilities, which impose numerically all moment conditions on the data, rather than only some linear combinations of them as in the GMM case. The resultant GEL distribution function, as exploits the information contained in the moment conditions defining a particular model, is a more efficient estimator of the distribution of the data than the empirical distribution function (EDF) implicitly used by GMM. Contrasting certain functions of these two consistent estimators of the data generating process gives rise, as we show, to Pearson-type tests of overidentifying moment conditions. We develop two classes of Pearson-type tests. The first includes tests that are very similar in form to the classical Pearson χ^2 statistics. The other requires the partition of the sample space into several sets, the contrast between the empirical and the GEL implied probabilities estimated for each set forming the basis for the test. We show also that a similar approach can be used to construct Pearson-type tests for parametric restrictions, in which case two GEL distributions estimated under different assumptions are contrasted.

In the second part of this paper we investigate, through a Monte Carlo simulation analysis based on two of the settings considered by Imbens, Spady and Johnson (1998), how Pearson-type statistics for overidentifying moment conditions perform in finite samples. We examine their size behaviour and compare it with some of the existing alternatives: Hansen's (1982) J test and Smith's (1997) distance metric and Wald statistics.

This paper is organized as follows. Section 2 briefly reviews the principal characteristics of GEL estimation and the concept of GEL implied probabilities. The Pearson-type tests for overidentifying moment conditions are derived in section 3 while the case of parametric restrictions is considered in section 4. The Monte Carlo simulation studies are discussed in section 5. Section 6 concludes. Some proofs of the results contained in the paper were relegated to the Appendix.

2 Generalized empirical likelihood

Let y_i $(i = 1 \quad n)$ be independent and identically distributed observations on the random vector y. Consider the s-vector of moment indicators $g(y \ \theta)$, which is known up to the k-element parameter vector θ, where $s \geq k$. It is assumed that there exists a unique value $\theta_0 \quad \Theta$, with the parameter space Θ compact, such that

$$E_F [g(y \ \theta_0)] = 0, \tag{1}$$

where $E_F [\cdot]$ denotes expectation taken with respect to the true but unknown distribution $F \equiv F(y)$ of the random vector y.

Define $g_i(\theta) \equiv g(y_i \ \theta)$ $(i = 1 \quad n)$ and $g_n(\theta) \equiv n^{-1} \ {}_{i=1}^{n} g_i(\theta)$. It is further assumed that the normalized sample counterparts of the moment conditions (1), $g_n(\theta)$ and $\ \overline{n} g_n(\theta_0)$, obey, respectively, a uniform (in θ) weak law of large numbers, $g_n(\theta) \ ^p \ E_F[g(y \ \theta)]$, and a central limit theorem, $\ \overline{n} g_n(\theta_0) \ -^d \ N(0 \ V)$, where the asymptotic variance matrix $V \equiv E_F\left[g_i(\theta_0)\, g_i(\theta_0)'\right]$ is positive definite and ^{p} and $-^d$ denote convergence in probability and convergence in distribution, respectively.

2.1 GEL estimation

Let $\rho(v)$ be a function of a scalar $v \equiv \phi' g_i(\theta)$ $(i = 1, \ldots, n)$, where ϕ is an s-vector of auxiliary parameters. The GEL estimators $\hat{\theta}$ and $\hat{\phi}$ result from the optimization of the criterion function

$$Q(\theta, \phi) = \sum_{i=1}^{n} \rho[\phi' g_i(\theta)], \tag{2}$$

which is a saddle point function with first-order conditions

$$\sum_{i=1}^{n} \hat{\rho}_1 \begin{pmatrix} g_i\left(\hat{\theta}\right) \\ G_i\left(\hat{\theta}\right)' \hat{\phi} \end{pmatrix} = 0, \tag{3}$$

where $\hat{\rho}_1 \equiv \rho_1\left(\hat{\phi}' g_i\left(\hat{\theta}\right)\right)$, $\rho_j(v) \equiv \partial^j \rho(v) / \partial v^j$ $(j = 1, 2, \ldots)$ and $G_i(\theta) \equiv \partial g_i(\theta) / \partial \theta$. There are several different ways of specifying $\rho(\cdot)$, each of which gives rise to a distinct GEL estimator. In this paper we focus on the two most well known special cases of GEL estimators, the EL and ET estimators, which result from considering $\rho(v) = -\ln(1 + v)$ and $\rho(v) = \exp(v)$, respectively.

Under suitable regularity conditions, it can be proved that the GEL estimator $\hat{\theta}$ satisfying (3) is a consistent estimator of θ_0 in (1). Moreover, GEL estimators are asymptotically normal distributed:

$$\sqrt{n} \begin{pmatrix} \hat{\phi} \\ \hat{\theta} - \theta_0 \end{pmatrix} \xrightarrow{d} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \left(\frac{\rho_1(0)}{\rho_2(0)}\right)^2 M & 0 \\ 0 & \left(G' V^{-1} G\right)^{-1} \end{pmatrix} \right), \tag{4}$$

where $G \equiv E_F[\partial g(y, \theta) / \partial \theta]$, $M = I - G\left(G' V G\right)^{-1} G' V^{-1}$ and $rk(M) = s - k$. An intermediate result that will be useful later is

$$\sqrt{n} \begin{pmatrix} \hat{\phi} \\ \hat{\theta} - \theta_0 \end{pmatrix} = - \begin{pmatrix} \frac{\rho_1(0)}{\rho_2(0)} V^{-1} M \\ \left(G' V^{-1} G\right)^{-1} G' V^{-1} \end{pmatrix} \sqrt{n} g_n(\theta_0) + O_p\left(n^{-\frac{1}{2}}\right). \tag{5}$$

See Newey and Smith (2001) for a rigorous analysis of the asymptotic properties of GEL estimators.

2.2 GEL implied probabilities

After GEL estimation of θ and ϕ, it is possible to calculate the ratios

$$\hat{p}_i \equiv p_i\left(\hat{\theta}, \hat{\phi}\right) = \frac{\rho_1\left(\hat{\phi}' g_i\left(\hat{\theta}\right)\right)}{\sum_{i=1}^{n} \rho_1\left(\hat{\phi}' g_i\left(\hat{\theta}\right)\right)} \quad (i = 1, \ldots, n), \tag{6}$$

which can be interpreted as implied probability measures, cf. Back and Brown (1993), being positive and summing to unity. These probabilities give rise to a consistent estimator for the distribution $F(y)$ in (1),

$$F_{gel}(y) = \sum_{i=1}^{n} p_i\left(\theta, \phi\right) 1\left(y_i \leq y\right), \tag{7}$$

where $1(\cdot)$ is an indicator function, which is a more efficient estimator than the EDF

$$F_n(y) = \frac{1}{n}\sum_{i=1}^{n} 1\left(y_i \leq y\right), \tag{8}$$

since it takes into account the information contained in (1). Indeed, in the GEL context each observation is reweighted in such a way that the empirical moment conditions $\sum_{i=1}^{n} p_i g_i\left(\theta\right)$ are numerically imposed, see (3) and (6), mirroring the population moment condition (1). For this reason, the GEL implied probabilities (6) can be used to form an efficient estimator $\sum_{i=1}^{n} p_i a\left(y_i, \theta\right)$ of $E_F\left[a\left(y, \theta_0\right)\right]$, for any function $a(y, \theta)$.

Under the null hypothesis that the moment conditions (1) hold in the population of interest, the probability limit of ϕ is 0, so $\rho_1 = \rho_1(0)$, and, hence, the GEL probabilities $p_i\left(\theta, 0\right)$ $(i = 1, \ldots, n)$ are equal to the empirical measures $dF_n(y) = n^{-1}$; see (6). More rigorously, consider a first-order Taylor series expansion of $\sqrt{n}p_i$ about $\left(\theta, 0\right)$. As $p_i\left(\theta, 0\right) = n^{-1}$ and

$$\frac{\partial p_i(\theta, \phi)}{\partial \phi'} = \frac{\rho_2\left[\phi' g_i(\theta)\right] g_i(\theta) \sum_{i=1}^{n} \rho_1\left[\phi' g_i(\theta)\right] - \rho_1\left[\phi' g_i(\theta)\right] \sum_{i=1}^{n} \rho_2\left[\phi' g_i(\theta)\right] g_i(\theta)}{\left\{\sum_{i=1}^{n} \rho_1\left[\phi' g_i(\theta)\right]\right\}^2}, \tag{9}$$

it follows that:

$$\begin{aligned} \sqrt{n}\left(p_i - \frac{1}{n}\right) &= \frac{\rho_2(0)}{\rho_1(0)}\frac{1}{n}\left[g_i\left(\theta\right)' - \frac{1}{n}\sum_{i=1}^{n} g_i\left(\theta\right)'\right]\sqrt{n}\phi + O_p\left(n^{-\frac{3}{2}}\right) \\ &= \frac{\rho_2(0)}{\rho_1(0)}\frac{1}{n} g_i\left(\theta\right)' \sqrt{n}\phi + O_p\left(n^{-\frac{3}{2}}\right), \end{aligned} \tag{10}$$

since $n^{-1}\sum_{i=1}^{n} g_i\left(\theta\right) = O_p\left(n^{-\frac{1}{2}}\right)$. Equation (10), by expressing the asymptotic relationship between the empirical and GEL probability density functions, forms the basis for the construction of the Pearson-type test statistics derived in the next sections.

3 Tests of overidentifying moment conditions

In the GEL framework there are several ways of assessing the validity of the moment conditions (1). Indeed, as a sample version of each moment condition is associated with an element of the vector of auxiliary parameters ϕ, see (2), which can be interpreted as Lagrange multipliers in the

EL and ET context, the validity of (1) can be checked by testing the hypothesis $H_0 : \phi = 0$. Hence, the three classical tests can be used, as suggested by Smith (1997), who proposed testing H_0 employing the distance metric statistic

$$DM_n = 2\frac{\rho_2(0)}{[\rho_1(0)]^2} \quad n\rho(0) - Q\left(\theta \ \phi\right) \quad , \tag{11}$$

the Wald statistic

$$W_n = n\left[\frac{\rho_2(0)}{\rho_1(0)}\right]^2 \phi_n' V_n \phi_n, \tag{12}$$

and the Score statistic

$$J_n = n g_n\left(\theta\right)' V_n^{-1} g_n\left(\theta\right), \tag{13}$$

where $V_n \equiv n^{-1} \quad {}_{i=1}^{n} g_i\left(\theta\right) g_i\left(\theta\right)'$ is a consistent estimator for V. Note that the expression of the last test is identical to Hansen's (1982) J test. Under H_0, these three statistics have an asymptotic chi-square distribution with $s - k$ degrees of freedom.

In this section we consider other alternatives appropriate for testing the moment conditions (1). Namely, we develop two classes of Pearson-type test statistics. First, we show that a very simple adaptation of the standard Pearson statistics utilized in the parametric context allows their employment in the GEL framework as tests of overidentifying moment conditions. Then, we derive an alternative Pearson-type statistic which is based on the contrast between empirical and GEL probabilities estimated for each set into which the sample space is divided.

3.1 Classical Pearson statistics

Suppose that we have a data set containing some ties, where the distinct value y_i arises $n_i \geq 1$ times. Let u be the number of ties. In a parametric context, we may wish to test whether a given distribution function $\bar{F}(y)$ correctly describes the data. To this end, there are two versions of the Pearson statistic that are usually applied:

$$P_1^* = \sum_{i=1}^{u} \frac{(e_i - n_i)^2}{n_i} \tag{14}$$

and

$$P_2^* = \sum_{i=1}^{u} \frac{(e_i - n_i)^2}{e_i}, \tag{15}$$

where n_i and $e_i \equiv n \cdot d\bar{F}(y_i)$ denote, respectively, the actual and the expected number of observations of the distinct value y_i ($i = 1 \quad u$) under $\bar{F}(y)$. In (15) it is assumed that $e_i > 0$ for all $i = 1 \quad u$. If the parametric model is correctly specified, then the differences between the observed and expected number of outcomes are due solely to random fluctuations and both the statistics (14) and (15) have a limiting chi-square distribution.

In the GEL framework, we can ignore the ties in the data and deal with the probability associated with an observation, not a value; see *inter alia* Owen (2001). In other words, we can act as if a single data point was observed in each cell of a n-cell contingency table, that is, a GEL version of the above statistics may be directly obtained by setting $n_i = 1$, $u = n$ and $e_i = np_i$ $(i = 1, \ldots, n)$. In fact, as we show in the Appendix, the corresponding versions of (14) and (15) that allow the hypothesis (1) to be tested in models estimated by GEL methods are given by

$$P_1 = \sum_{i=1}^{n} (np_i - 1)^2 \tag{16}$$

and

$$P_2 = \sum_{i=1}^{n} \frac{(np_i - 1)^2}{np_i}. \tag{17}$$

Note that, from (10), it follows that $np_i = 1 + O_p\left(n^{-\frac{1}{2}}\right)$, so (16) and (17) are asymptotically equivalent. In the Appendix we show their asymptotic equivalence to the Wald statistic (12), which proves the following theorem:

Theorem 1 *Under (1), the test statistics P_1 and P_2 of (16) and (17) respectively each have a limiting distribution described by*

$$P_1, P_2 \xrightarrow{d} \chi^2_{s-k}.$$

3.2 Alternative Pearson-type tests

In this sub-section we develop an alternative Pearson-type test of overidentifying moment conditions. As discussed in section 2.2, the distribution $F(y)$ in (1) can be consistently estimated, under the hypothesis that those moment conditions hold in the population of interest, by either $F_{gel}(y)$ of (7) or $F_n(y)$ of (8). Therefore, we can think of testing the validity of the overidentifying moment conditions (1) by testing for $H_0 : F_{gel}(y) - F_n(y) = 0$. Indeed, if the null model is correctly specified, the limiting distribution of a test statistic based on the contrast $F_{gel}(y) - F_n(y)$ should be centred at zero.

Consider a first-order Taylor series expansion of $\sqrt{n}F_{gel}(y)$ around $\phi = 0$. From (10), it follows that

$$\sqrt{n}\left[F_{gel}(y) - F_n(y)\right] = \frac{\rho_2(0)}{\rho_1(0)} \frac{1}{n} \sum_{i=1}^{n} 1(y_i \le y)\, g_i\left(\hat{\theta}\right)' \sqrt{n}\hat{\phi} + O_p\left(n^{-\frac{1}{2}}\right). \tag{18}$$

Defining the s-vector $b \equiv E_F\left[1\left(y_i \leq y\right) g\left(y_i, \theta_0\right)\right]$, which is assumed to be nonzero, and noting that $\sqrt{n}\phi = -\frac{\rho_1(0)}{\rho_2(0)} V^{-1} M \sqrt{n} g_n\left(\theta_0\right) + O_p\left(n^{-\frac{1}{2}}\right)$, see (5), equation (18) can be written as:

$$\sqrt{n}\left(F_{gel}(y) - F_n(y)\right) = -b' V^{-1} M \sqrt{n} g_n\left(\theta_0\right) + O_p\left(n^{-\frac{1}{2}}\right). \tag{19}$$

Now, consider a partition of the sample space of y into the sets C_j $(j = 1, \ldots, L)$, where L is finite. Define

$$F_{gel}\left(C_j\right) = \sum_{i=1}^{n} p_i\left(\hat{\theta}, \hat{\phi}\right) 1\left(y_i \in C_j\right) \tag{20}$$

and

$$F_n\left(C_j\right) = \frac{1}{n} \sum_{i=1}^{n} 1\left(y_i \in C_j\right). \tag{21}$$

Using a similar argument to that above, we have, corresponding to (19),

$$\sqrt{n}\left(F_{gel}\left(C_j\right) - F_n\left(C_j\right)\right) = -b_j' V^{-1} M \sqrt{n} g_n\left(\theta_0\right) + O_p\left(n^{-\frac{1}{2}}\right) \qquad (j = 1, \ldots, L). \tag{22}$$

Stacking $B \equiv \left(b_1, \ldots, b_L\right)$, an $(s \times L)$ matrix, and defining

$$F_{gel} - F_n \equiv \begin{pmatrix} F_{gel}\left(C_1\right) - F_n\left(C_1\right) \\ \vdots \\ F_{gel}\left(C_L\right) - F_n\left(C_L\right) \end{pmatrix}, \tag{23}$$

an L-vector, it follows that

$$\sqrt{n}\left(F_{gel} - F_n\right) = -B' V^{-1} M \sqrt{n} g_n\left(\theta_0\right) + O_p\left(n^{-\frac{1}{2}}\right). \tag{24}$$

Hence, under the null hypothesis that the moment conditions (1) hold in the population, $\sqrt{n}\left(F_{gel} - F_n\right) \xrightarrow{d} N(0, \Psi)$, where $\Psi \equiv B' M' V^{-1} M B$ and, therefore, a Pearson-type test statistic suitable for testing (1) is given by

$$P_3 = n\left(F_{gel} - F_n\right)' \Psi_n^{-}\left(F_{gel} - F_n\right), \tag{25}$$

where $\Psi_n^{-} \equiv \left(B_n' M_n' V_n^{-1} M_n B_n\right)^{-}$ denotes a consistent estimator for a *g-inverse* of Ψ and B_n, V_n and M_n are consistent estimators for B, V and M, respectively.

Define $\upsilon = rk\left(B' M' V^{-1} M B\right)$. Then, we have the following result:

Theorem 2 *Under (1), the Pearson-type test statistic P_3 of (25) has a limiting distribution described by*

$$P_3 \xrightarrow{d} \chi_\upsilon^2.$$

Expression (25) may be simplified. Indeed, if $L \geq s$ and B is full row rank s, then a generalized inverse for Ψ is $B'(BB')^{-1} V (BB')^{-1} B$ and $\upsilon = s - k$. On the other hand, if $L = s$, the matrix B will be invertible, so $\Psi^{-} = B^{-1}(M'V^{-1}M)^{-} B^{-1} = B^{-1}VB^{-1}$. Hence:

Corollary 3 *If $L \geq s$ and B is full row rank s, then, under (1),*

$$P_3 = n\left(F_{gel} - F_n\right)' B_n' \left(B_n B_n'\right)^{-1} V_n \left(B_n B_n'\right)^{-1} B_n \left(F_{gel} - F_n\right) \; {}^{d} \; \chi^2_{s-k}. \tag{26}$$

Corollary 4 *If $L = s$ and B is full rank, then, under (1),*

$$P_3 = n\left(F_{gel} - F_n\right)' B_n^{-1} V_n B_n^{-1} \left(F_{gel} - F_n\right) \; {}^{d} \; \chi^2_{s-k}. \tag{27}$$

In the Appendix we show the asymptotic equivalence of the Pearson-type statistic P_3 (26) to some of the tests of overidentifying moment conditions presented before.

4 Tests of parametric restrictions

The same principles used to construct Pearson-type tests of overidentifying moment conditions can be applied in other contexts. In this section we show how to develop GEL Pearson-type statistics appropriate for testing parametric restrictions.

4.1 Constrained GEL estimation

Consider the null hypothesis

$$H_0 : r(\theta_0) = 0, \tag{28}$$

where $r(\cdot)$ is a known continuously differentiable q-vector of parametric restrictions, where $q < k$. The $(q \times k)$ derivative matrix $R(\theta) \equiv \; _{\theta} r(\theta)$ is assumed full row rank q. Let $\left(\theta \;\; \phi\right)$ be the unconstrained estimators resulting from the optimization of the GEL criterion $Q(\theta \;\; \phi) = \rho[\phi' g_i(\theta)]$ and $\left(\theta \;\; \phi \;\; \psi\right)$ the estimators of the constrained model incorporating H_0, which are obtained by optimizing the modified GEL function $Q^*(\theta \;\; \phi \;\; \psi) = \rho^*[\phi' g_i(\theta) + \psi' r(\theta)]$ and can be written asymptotically as:

$$\sqrt{n} \begin{pmatrix} \phi \\ \psi \\ \theta - \theta_0 \end{pmatrix} = - \begin{pmatrix} \frac{\rho_1(0)}{\rho_2(0)}\left(V^{-1} - V^{-1}G\Omega P G' V^{-1}\right) \\ -\frac{\rho_1(0)}{\rho_2(0)}\left(R\Omega R'\right)^{-1} R\Omega G' V^{-1} \\ \Omega P G' V^{-1} \end{pmatrix} \sqrt{n} g_n(\theta_0) + O_p\left(n^{-\frac{1}{2}}\right), \tag{29}$$

where $\Omega \equiv (G'V^{-1}G)^{-1}$ and $P = I - R'(R\Omega R')^{-1} R\Omega$, with $rk(P) = q$; see Smith (1997).

Define the constrained GEL implied probabilities

$$p_i^* \equiv p_i^*\left(\theta \ \phi \ \psi\right) = \frac{\rho_1^* \ \phi' g_i\left(\theta\right) + \psi' r\left(\theta\right)}{\sum_{i=1}^{n} \rho_1^* \ \phi' g_i\left(\theta\right) + \psi' r\left(\theta\right)} \qquad (i = 1 \ \ldots \ n). \tag{30}$$

In this setting, assuming that the moment conditions (1) hold in the population, the EDF $F_n(y)$ and the unconstrained GEL distribution $F_{gel}(y)$ are still consistent estimators of the distribution F in (1), whether or not H_0 (28) holds. However, under this hypothesis, a more efficient estimator is given by

$$F_{gel}^*(y) = \sum_{i=1}^{n} p_i^*\left(\theta \ \phi \ \psi\right) 1\left(y_i \leq y\right). \tag{31}$$

The statistics suggested below for testing the parametric restrictions (28) are based on the contrasts $F_{gel}(y) - F_{gel}^*(y)$ or $p_i - p_i^*$ $(i = 1 \ \ldots \ n)$. Theorem 5 states the asymptotic relationship that occurs between the GEL implied probabilities p_i and p_i^* $(i = 1 \ \ldots \ n)$:

Theorem 5 *Under (1) and (28),*

$$\sqrt{n}\left(p_i - p_i^*\right) = \frac{\rho_2(0)}{\rho_1(0)} \frac{1}{n} g_i(\theta_0)' \sqrt{n}\left(\phi - \phi\right) + O_p\left(n^{-\frac{3}{2}}\right). \tag{32}$$

See the proof in the Appendix.

4.2 Classical Pearson tests

In the present framework, classical-type Pearson statistics for testing (28), similar to those derived in section 3.1, can be constructed. Corresponding to P_1 (16), we propose the statistic

$$P_1^{pr} = \sum_{i=1}^{n} \left(np_i - np_i^*\right)^2. \tag{33}$$

Corresponding to P_2, we suggest the following two alternatives:

$$P_{2a}^{pr} = \sum_{i=1}^{n} \frac{\left(np_i - np_i^*\right)^2}{np_i} \tag{34}$$

and

$$P_{2b}^{pr} = \sum_{i=1}^{n} \frac{\left(np_i - np_i^*\right)^2}{np_i}. \tag{35}$$

In the Appendix we show that these three statistics are asymptotically equivalent to the minimum chi-square statistic $MC_n = n \ \frac{\rho_2(0)}{\rho_1(0)}^2 \left(\phi - \phi\right)' V_n \left(\phi - \phi\right)$ proposed by Smith (2000) for testing the parametric restrictions (28), which allows us to give the following result:

Theorem 6 *Under (1) and (28), the GEL Pearson-type statistics P_1^{pr}, P_{2a}^{pr} and P_{2b}^{pr} of (33), (34) and (35) respectively each have limiting distributions described by*

$$P_1^{pr} \ P_{2a}^{pr} \ P_{2b}^{pr} \xrightarrow{d} \chi_q^2.$$

4.3 Alternative Pearson-type tests

By analogy with the overidentifying moment conditions case, a test statistic based on the normalized contrast $\sqrt{n}\ \ F_{gel}(y) - F^*_{gel}(y)$ constitutes an alternative way of assessing the null hypothesis H_0 (28). Expanding $F^*_{gel}(y)$ about $\left(\theta\ 0\ 0\right)$ yields:

$$\sqrt{n}\ \ F^*_{gel}(y) - F_n(y) \ = \frac{\rho_2(0)}{\rho_1(0)}\frac{1}{n}\sum_{i=1}^{n} 1(y_i \leq y)\, g_i\left(\theta\right)' \ \sqrt{n}\phi + O_p\left(n^{-\frac{1}{2}}\right); \tag{36}$$

see expression (54). Hence, using the same notation as in sub-section 3.2, we have:

$$\sqrt{n}\ \ F^*_{gel}(y) - F_n(y) \ = \frac{\rho_2(0)}{\rho_1(0)} b' \ \sqrt{n}\phi + O_p\left(n^{-\frac{1}{2}}\right). \tag{37}$$

Subtracting (37) from (18) produces:

$$\sqrt{n}\ \ F_{gel}(y) - F^*_{gel}(y) \ = \frac{\rho_2(0)}{\rho_1(0)} b' \ \sqrt{n}\left(\phi - \phi\right) + O_p\left(n^{-\frac{1}{2}}\right). \tag{38}$$

As $\sqrt{n}\phi = -\frac{\rho_1(0)}{\rho_2(0)}(V^{-1} - V^{-1}G\Omega G'V^{-1})\ \sqrt{n}g_n(\theta_0) + O_p\left(n^{-\frac{1}{2}}\right)$, see (5), and $\sqrt{n}\phi = -\frac{\rho_1(0)}{\rho_2(0)}$ $(V^{-1} - V^{-1}G\Omega PG'V^{-1})\ \sqrt{n}g_n(\theta_0) + O_p\left(n^{-\frac{1}{2}}\right)$, see (29), it follows that

$$\sqrt{n}\left(\phi - \phi\right) = \frac{\rho_1(0)}{\rho_2(0)} V^{-1}G\Omega(I - P)G'V^{-1}\ \sqrt{n}g_n(\theta_0) + O_p\left(n^{-\frac{1}{2}}\right). \tag{39}$$

Substituting (39) into (38) yields:

$$\sqrt{n}\ \ F_{gel}(y) - F^*_{gel}(y) \ = b'V^{-1}G\Omega(I - P)G'V^{-1}\ \sqrt{n}g_n(\theta_0) + O_p\left(n^{-\frac{1}{2}}\right). \tag{40}$$

Now, consider a partition of the sample space of y into the sets C_j $(j = 1 \quad L)$ identical to that of sub-section 3.2. Define $F^*_{gel}(C_j) = \sum_{i=1}^{n} p_i 1(y_i \quad C_j)$ and

$$F_{gel} - F^*_{gel} \equiv \begin{matrix} F_{gel}(C_1) - F^*_{gel}(C_1) \\ \\ F_{gel}(C_L) - F^*_{gel}(C_L) \end{matrix}, \tag{41}$$

so

$$\sqrt{n}\left(F_{gel} - F^*_{gel}\right) = B'V^{-1}G\Omega(I - P)G'V^{-1}\ \sqrt{n}g_n(\theta_0) + O_p\left(n^{-\frac{1}{2}}\right). \tag{42}$$

Then, noting that $\Omega(I - P)\Omega^{-1}G(I - P)'\Omega = \Omega(I - P)$, it follows from (40) that, under (1) and (28), $\sqrt{n}\left(F_{gel} - F^*_{gel}\right) \overset{d}{\ } N(0\ \Psi)$, where $\Psi \equiv B'V^{-1}G\Omega(I - P)G'V^{-1}B$. Thus, a Pearson-type statistic for testing the parametric restrictions (28) is given by

$$P^{pr}_3 = n\left(F_{gel} - F^*_{gel}\right)' \Psi^-_n \left(F_{gel} - F^*_{gel}\right), \tag{43}$$

where Ψ_n^- denotes a consistent estimator for a *g-inverse* of Ψ. In the Appendix, we show the asymptotic equivalence of P_3^{pr} to Smith's (1997) MC_n test statistic for parametric restrictions. Hence:

Theorem 7 *Under (1) and (28), the Pearson-type test statistic P_3^{pr} of (43) has a limiting distribution described by*

$$P_3^{pr} \xrightarrow{d} \chi_q^2.$$

Assuming that B is full row rank s, a generalized inverse for Ψ is $B' (BB')^{-1} V (BB')^{-1} B$. In the case that B is a square matrix ($L = s$), a generalized inverse for Ψ is simply $B^{-1} V B^{-1}$.

5 Finite sample properties of tests of overidentifying moment conditions: Monte Carlo investigation

In this section we investigate the finite sample properties of some of the Pearson-type tests proposed in the previous sections. In particular, we examine the size behaviour of the P_1, P_2 and P_3 test statistics of overidentifying moment conditions suggested in section 3 and assess how they perform comparatively to the J, Wald (W) and distance metric (DM) tests.

5.1 Experimental designs

We follow closely the simulation study realized by Imbens, Spady and Johnson (1998) to compare the finite sample properties of the aforementioned tests, using their first two experimental designs as basis for our investigation. The first model simulated is a simplified version of an asset-pricing model, characterized by the moment indicators for unit i

$$g(X_i, Z_i, \theta) = \begin{pmatrix} \exp[-0.72 - \theta(X_i + Z_i) + 3Z_i] - 1 \\ Z_i \cdot \exp[-0.72 - \theta(X_i + Z_i) + 3Z_i] - 1 \end{pmatrix}, \tag{44}$$

where X and Z were generated independently from a $N(0, 0.16)$ distribution and the true value of θ is 3. The second Monte Carlo experiment is based on the moment vector

$$g(Z_i, \theta) = \begin{pmatrix} Z_i - \theta \\ Z_i^2 - \theta^2 - 2\theta \end{pmatrix}, \tag{45}$$

where Z has a chi-square distribution with one degree of freedom and $\theta_0 = 1$. We considered samples of 100, 200, 500 and 1000 observations, each one being replicated 10000 times.

For the tests requiring evaluation at GEL estimators (W, DM, P_1, P_2 and P_3), we considered both ET and EL estimation. In both cases, consistent estimators for the matrices needed to compute the W and P_3 tests were obtained in three different ways:

- $gel(n)$: uses sample means to estimate consistently V and G, for example:

$$V_n = \frac{1}{n} \sum_{i=1}^{n} g_i\left(\theta\right) g_i\left(\theta\right)'; \tag{46}$$

- $gel(s)$: uses the GEL implied probabilities p_i $(i = 1 \ldots n)$ for both V and G, for example:

$$V_n = \sum_{i=1}^{n} p_i g_i\left(\theta\right) g_i\left(\theta\right)'; \tag{47}$$

- $gel(r)$: the matrix G is estimated as in $gel\,(s)$ and V is estimated robustly as:

$$V_n = \sum_{i=1}^{n} p_i g_i\left(\theta\right) g_i\left(\theta\right)' \left[n \sum_{i=1}^{n} p_i^2 g_i\left(\theta\right) g_i\left(\theta\right)' \right]^{-1} \sum_{i=1}^{n} p_i g_i\left(\theta\right) g_i\left(\theta\right)'. \tag{48}$$

The same three procedures were followed to compute the J test but, in addition, we evaluate it also at two-step (J_{2s}), repeatedly-iterated (J_{ri}) and continuous-updating (J_{cu}) GMM estimators, in which cases we only use a consistent estimator for the matrix V based on, naturally, sample means; see Hansen, Heaton and Yaron (1996).

In their Monte Carlo simulation study, Imbens, Spady and Johnson (1998) analyzed the finite sample behaviour of the following tests: J_{2s}, J_{ri}, J_{cu}, $J_{et(s)}$, $W_{et(s)}$, $W_{et(r)}$, DM_{et} and DM_{el}. In this section we replicate their results for the two experimental designs described above and examine whether their conclusions remain valid when other estimators are employed to evaluate the J and Wald tests. In particular, we study the effects of using EL instead of ET estimation [$J_{el(s)}$, $W_{el(s)}$ and $W_{el(r)}$ tests], confirm the conjecture that robust estimation of the matrix V does not work well in the case of the J test [$J_{et(r)}$ and $J_{el(r)}$ tests], for reasons explained below, and investigate the consequences of using sample means to estimate that same matrix when GEL estimation is utilized [$J_{et(n)}$, $J_{el(n)}$, $W_{et(n)}$ and $W_{el(n)}$ tests].

The implementation of the P_3 test requires the previous partition of the sample space into L sets. In order to examine the sensitivity of this test to the number of classes into which the observations are divided, we considered two different values for L: 8 and 16. The definition of each set in each Monte Carlo sample was such that each class contains, approximately, $(100/L)\%$ of the observations.[1]

[1] For a discussion of alternative ways of partitioning the sample space see *inter alia* Andrews (1988).

5.2 Main results

Tables 1 and 2, for the asset-pricing model, and 3 and 4, for the chi-squared moments case, report the estimated size of each test at seven different levels of significance for the asset-pricing models. For each significance level, sample size and model considered, the actual size closest to the nominal size is underlined. For the tests analyzed in their paper, these results conform with those presented by Imbens, Spady and Johnson (1998).[2] As can be immediately seen from tables 1 and 3, all these tests are significantly oversized in almost all cases, even for $n = 1000$, particularly for the chi-squared moments model. Clearly, the $W_{et(r)}$ test registered the best behaviour in most experiments, the only exceptions being the largest nominal sizes, where the J_{cu} test, in the first model, and the $W_{el(r)}$ test, in both models, achieved superior performances. The J test evaluated at two-step GMM estimators, the most widely applied test to assess overidentifying moment condition models, had a disastrous behaviour in these experiments, even being the worst of all versions of the J test based on sample mean estimators for the matrix V [J_{2s}, J_{ri}, J_{cu}, $J_{et(n)}$ and $J_{el(n)}$] in the asset-pricing model. The DM tests also produced very modest results, with that based on the EL objective function performing substantially better than that using the ET criterion, particularly for the chi-squared moments model and for the smallest nominal sizes.

Table 1 about here

Table 2 about here

Table 3 about here

Table 4 about here

As noted by Imbens, Spady and Johnson (1998), robust estimation of the matrix V decisively influences the performance of the tests. However, the extraordinary benefits reported by them for the $W_{et(r)}$ statistic do not extend to all the other tests. They do not extend even to the $W_{el(r)}$ test for the smallest nominal sizes considered. The behaviour of the J test also deteriorates considerably. Although a theoretical analysis of the effects of using the $gel\,(r)$ method is not available, it is evident why the W and the J tests are affected in opposite ways: the matrix V appears in the expression of those tests in an inverse way.

The estimated sizes for the Pearson-type statistics are reported in tables 2 and 4. The P_1 and P_2 tests perform very modestly, being substantially oversized in all cases. Their size behaviour

[2] There is the following correspondence between the notation used here and that utilized by Imbens, Spady and Johnson (1998): $J_{2s} = T_{g1}^{AM}$, $J_{ri} = T_{g2}^{AM}$, $J_{cu} = T_{g3}^{AM}$, $J_{et(s)} = T_{et}^{AM}$, $W_{et(s)} = T_{et(s)}^{LM}$, $W_{et(r)} = T_{et(r)}^{LM}$, $DM_{et} = T_{klic(et)}^{CF}$ and $DM_{el} = T_{lr(el)}^{CF}$.

does not differ much from that described for the other tests.[3] However, the P_3 statistic shows a very promising performance. Whichever the number of classes considered, the general effects of evaluating this test at different estimators are similar in all cases. Analogously to the W test, the least number of rejections of the null hypothesis occurs when robust estimation of V is employed. This is not surprising since the matrix V appears in the expressions of both tests in a similar form. However, while this was always beneficial for the W test, the P_3 test becomes sometimes quite undersized, particularly for the smallest nominal sizes and sample sizes considered.

Figure 1 displays QQ-plots comparing the six versions of the P_3 test for the $L = 8$ case. Vertical coordinates are Monte Carlo estimates of quantiles of the finite sample distribution of those statistics and horizontal coordinates are quantiles of a chi-square variable with one degree of freedom. The vertical solid line marks the asymptotic critical value for a nominal size of 0.05. Clearly, the best performances are obtained by $P_3^{et(r)}$ and $P_3^{el(r)}$. Note how for $n \geq 500$ (first model) or $n = 1000$ (second model) the estimated quantiles of these tests are very close to the asymptotic ones while the other versions of P_3 are still significantly oversized. Notice also how, for small sample sizes, all three EL versions of the P_3 test tend to reject the null hypothesis significantly less than the corresponding ET variants.

Figure 1 about here

The performance of the P_3 test does not seem to depend significantly on L on finite samples. This is particularly evident for the asset-pricing model case. For the chi-squared moments model the differences between the $L = 8$ and $L = 16$ cases are more important being, however, attenuated as the sample size increases. Figure 2 illustrates this situation, displaying QQ-plots for the $P_3^{et(r)}$ test for the two distinct values of L simulated.

Figure 2 about here

Figure 3 compares the robust forms of the W and P_3 tests (for $L = 8$), both evaluated at ET and EL estimators. Recall that the $W_{et(r)}$ statistic registered the best behaviour of all tests analyzed in the previous sub-section. From figure 3 we see that the P_3 test clearly performs better for both models, its actual quantiles being in most cases closer to the asymptotic ones.

---

[3] The sizes estimated for the EL version of the P_2 test are numerically equal to those calculated for the $J_{el(s)}$ and $W_{el(s)}$ statistics which is due to the particular form assumed by the EL probabilities: $p_i = n^{-1} \left[1 + \phi' g\left(y_i, \theta\right)\right]^{-1}$ $(i = 1, \ldots, n)$. For example, as $\phi' g\left(y_i, \theta\right) = np_i - 1$ and V is estimated by $\sum_{i=1}^{n} p_i g\left(y_i, \theta\right) g\left(y_i, \theta\right)'$, we have $W_n = n\phi' V_n \phi = \sum_{i=1}^{n} \frac{\phi' g(y_i, \theta) g(y_i, \theta)' \phi}{1+\phi' g(y_i, \theta)} = P_2$.

Furthermore, while the P_3 test is relatively indifferent to the use of ET or EL estimation, at least for larger sample sizes, in the case of the Wald test EL estimation does not work well, even for $n = 1000$.

Figure 3 about here

6 Summary

In this chapter we developed new Pearson-type statistics suitable for testing overidentifying moment conditions and parametric restrictions. One of those statistics, the P_3 test, performed very well in two Monte Carlo simulation studies concerning tests of overidentifying moment conditions. Its size behaviour, when based on robust estimation of the matrix V, seems to be superior to that of alternative tests. Moreover, the P_3 statistic does not seem to be sensitive to the number of classes into which the sample space is divided.

7 Appendix

7.1 P_1 and P_2 tests

To demonstrate that the statistics (16) and (17) are appropriated for testing the moment conditions (1), we show the asymptotic equivalence of P_1 to the Wald test of overidentifying moment conditions presented in (12). The proof is very simple since, from (10), it follows that

$$
\begin{aligned}
n\left(p_i - \frac{1}{n}\right)^2 &= \left[\frac{\rho_2(0)}{\rho_1(0)}\right]^2 \frac{1}{n}\phi' g_i\left(\theta\right) g_i\left(\theta\right)' \phi + o_p(1) \\
\sum_{i=1}^{n} \frac{(np_i - 1)^2}{n} &= \left[\frac{\rho_2(0)}{\rho_1(0)}\right]^2 \phi' \frac{1}{n} \sum_{i=1}^{n} g_i\left(\theta\right) g_i\left(\theta\right)' \phi + o_p(1) \\
\sum_{i=1}^{n} (np_i - 1)^2 &= n\left[\frac{\rho_2(0)}{\rho_1(0)}\right]^2 \phi' V_n \phi + o_p(1) \\
P_1 &= W_n + o_p(1).
\end{aligned}
$$

7.2 P_3 test

To prove the asymptotic equivalence of the P_3 test to the other tests of overidentifying moment conditions, note that (24) can be rewritten both as

$$
\sqrt{n}\left(F_{gel} - F_n\right) = -B'V^{-1}\sqrt{n}g_n\left(\theta\right) + O_p\left(n^{-\frac{1}{2}}\right) \tag{49}
$$

and

$$\sqrt{n}\left(F_{gel} - F_n\right) = \frac{\rho_2(0)}{\rho_1(0)} B' \quad \sqrt{n}\phi + O_p\left(n^{-\frac{1}{2}}\right). \tag{50}$$

Expression (49) follows from a Taylor series expansion of $\sqrt{n} g_n\left(\theta\right)$ around $\sqrt{n} g_n(\theta_0)$, $\sqrt{n} g_n\left(\theta\right) = \sqrt{n} g_n(\theta_0) + G \ \sqrt{n}\left(\theta - \theta_0\right) + O_p\left(n^{-\frac{1}{2}}\right)$, where $\sqrt{n}\left(\theta - \theta_0\right)$ is replaced by $-\left(G'V^{-1}G\right)^{-1} G'V^{-1} \ \sqrt{n} g_n(\theta_0)$, see (5), yielding $\sqrt{n} g_n\left(\theta\right) = M \ \sqrt{n} g_n(\theta_0) + O_p\left(n^{-\frac{1}{2}}\right)$. To obtain (50), note that $\sqrt{n}\frac{\rho_2(0)}{\rho_1(0)}\phi = -V^{-1}M \ \sqrt{n} g_n(\theta_0) + O_p\left(n^{-\frac{1}{2}}\right)$, see also (5).

Using (25) and (49), we can demonstrate the asymptotic equivalence of the P_3 and J tests. Indeed, substituting the latter expression into the former, we obtain:

$$P_3 = n g_n\left(\theta\right)' V_n^{-1} B_n \left(B_n' M_n' V_n^{-1} M_n B_n\right)^{-} B_n' V_n^{-1} g_n\left(\theta\right) + o_p(1).$$

Following Lemma 2.2.5d) of Rao and Mitra (1971), $B_n\left(B_n' M_n' V_n^{-1} M_n B_n\right)^{-} B_n'$ is a generalized inverse for $M_n' V_n^{-1} M_n$, since $rk\left(B_n' M_n' V_n^{-1} M_n B_n\right) = rk\left(M_n' V_n^{-1} M_n\right)$. Thus, as V_n is just another generalized inverse for $M_n' V_n^{-1} M_n$, it follows that $P_3 = n g_n\left(\theta\right)' V_n^{-1} g_n\left(\theta\right) + o_p(1) = J_n + o_p(1)$.

Similarly, substituting (50) into (25) and applying the same Lemma of Rao and Mitra (1971), the asymptotic equivalence of P_3 to the Wald statistic (12) (and, hence, to the Pearson statistics P_1 and P_2) is proven: $P_3 = n \left(\frac{\rho_2(0)}{\rho_1(0)}\right)^2 \phi' V_n \phi + o_p(1) = W_n + o_p(1)$.

The asymptotic equivalence of the Pearson-type test to the distance metric test of (11) can be shown by demonstrating the equivalence of the latter to the Wald statistic; see Smith (1997, pp. 510-511) for a proof.

7.3 Constrained GEL estimation

Theorem 5 can be demonstrated as follows. Consider a Taylor expansion of $\sqrt{n} p_i^*$ about $\left(\theta \ 0 \ 0\right)$:

$$\sqrt{n} p_i^* = \sqrt{n} p_i^*\left(\theta \ 0 \ 0\right) + \frac{p_i^*\left(\theta \ 0 \ 0\right)'}{\phi'} \ \sqrt{n}\phi + \frac{p_i^*\left(\theta \ 0 \ 0\right)'}{\psi'} \ \sqrt{n}\psi + O_p\left(n^{-\frac{3}{2}}\right). \tag{51}$$

Define $\rho_j^*(\theta \ \phi \ \psi) \equiv \rho_j^*\left[\phi' g_i(\theta) + \psi' r(\theta)\right]$, $j = 1 \ 2$. As

$$\frac{p_i^*(\theta \ \phi \ \psi)}{\phi'} = \frac{\rho_2^*(\theta \ \phi \ \psi)\, g_i(\theta) \ \ {}_{i=1}^{n} \rho_1^*(\theta \ \phi \ \psi) - \rho_1^*(\theta \ \phi \ \psi) \ \ {}_{i=1}^{n} \rho_2^*(\theta \ \phi \ \psi)\, g_i(\theta)}{\left[\ {}_{i=1}^{n} \rho_1^*(\theta \ \phi \ \psi)\right]^2}$$

and

$$\frac{p_i^*(\theta \ \phi \ \psi)}{\psi'} = \frac{\rho_2^*(\theta \ \phi \ \psi) \ \ {}_{i=1}^{n} \rho_1^*(\theta \ \phi \ \psi) - \rho_1^*(\theta \ \phi \ \psi) \ \ {}_{i=1}^{n} \rho_2^*(\theta \ \phi \ \psi)}{\left[\ {}_{i=1}^{n} \rho_1^*(\theta \ \phi \ \psi)\right]^2} r(\theta),$$

it follows that, since $\rho_j^*(\theta\ \phi\ 0) = \rho_j[\phi' g_i(\theta)]\ (j = 1\ 2)$,

$$\frac{p_i^*\left(\theta\ 0\ 0\right)}{\phi'} = \frac{\rho_2(0)\, g_i\left(\theta\right) \sum_{i=1}^{n} \rho_1(0) - \rho_1(0) \sum_{i=1}^{n} \rho_2(0)\, g_i\left(\theta\right)}{\left[\sum_{i=1}^{n} \rho_1(0)\right]^2} = \frac{\rho_2(0)}{\rho_1(0)} \frac{1}{n} g_i\left(\theta\right)', \tag{52}$$

see (10), and

$$\frac{p_i^*\left(\theta\ 0\ 0\right)}{\psi'} = \frac{\rho_2(0) \sum_{i=1}^{n} \rho_1(0) - \rho_1(0) \sum_{i=1}^{n} \rho_2(0)}{\left[\sum_{i=1}^{n} \rho_1(0)\right]^2} r\left(\theta\right) = 0. \tag{53}$$

Noting that $p_i^*\left(\theta\ 0\ 0\right) = n^{-1}$ and substituting (52) and (53) into (51) yields:

$$\sqrt{n}\left(p_i^* - \frac{1}{n}\right) = \frac{\rho_2(0)}{\rho_1(0)} \frac{1}{n} g_i\left(\theta\right)' \sqrt{n}\phi + O_p\left(n^{-\frac{3}{2}}\right). \tag{54}$$

Finally, subtracting (54) from (10), with θ and θ replaced by θ_0, produces (32).

7.4 P_1^{pr}, P_{2a}^{pr} and P_{2b}^{pr} tests

Smith (2000) derived the following minimum chi-square test statistic of parametric restrictions: $MC_n = n\left(\frac{\rho_2(0)}{\rho_1(0)}\right)^2 \left(\phi - \phi\right)' V_n \left(\phi - \phi\right)$. We now show that all tests derived in section 4.2 are asymptotically equivalent to that statistic and, thus, appropriate for testing (28).

The proof for P_1^{pr} is straightforward. Indeed, from (32), it follows immediately that

$$\begin{aligned}
n\left(p_i - p_i^*\right)^2 &= \left[\frac{\rho_2(0)}{\rho_1(0)}\right]^2 \frac{1}{n}\left(\phi - \phi\right)' g_i(\theta_0)\, g_i(\theta_0)' \left(\phi - \phi\right) + o_p(1) \\
\sum_{i=1}^{n} \left(np_i - np_i^*\right)^2 &= n\left[\frac{\rho_2(0)}{\rho_1(0)}\right]^2 \left(\phi - \phi\right)' \frac{1}{n} \sum_{i=1}^{n} g_i(\theta_0)\, g_i(\theta_0)' \left(\phi - \phi\right) + o_p(1) \\
P_1^{pr} &= n\left[\frac{\rho_2(0)}{\rho_1(0)}\right]^2 \left(\phi - \phi\right)' V_n \left(\phi - \phi\right) + o_p(1) \\
&= MC_n + o_p(1).
\end{aligned}$$

The proof for P_{2a}^{pr} and P_{2b}^{pr} is similar since, from (10) and (54), we have $np_i = 1 + O_p\left(n^{-\frac{1}{2}}\right)$ and $np_i^* = 1 + O_p\left(n^{-\frac{1}{2}}\right)$, respectively.

7.5 P_3^{pr} test

The P_3^{pr} statistic is asymptotically equivalent to Smith's (1997) MC_n statistic, as we show next. Since $\sqrt{n}\left(F_{gel} - F_{gel}^*\right) = \frac{\rho_2(0)}{\rho_1(0)} B' \sqrt{n}\left(\phi - \phi\right) + O_p\left(n^{-\frac{1}{2}}\right)$, see (38), P_3^{pr} (43) can be written as:

$$P_3^{pr} = n\left[\frac{\rho_2(0)}{\rho_1(0)}\right]^2 \left(\phi - \phi\right)' B\left[B'V^{-1}G\Omega\left(I - P\right)G'V^{-1}B\right]^- B'\left(\phi - \phi\right) + o_p(1).$$

Following Lemma 2.2.5d) of Rao and Mitra (1971), we know that $B\left[B'V^{-1}G\Omega\left(I-P\right)G'V^{-1}B\right]^{-}B'$ is a generalized inverse for $V^{-1}G\Omega\left(I-P\right)G'V^{-1}$, since $rk \quad B\left[B'V^{-1}G\Omega\left(I-P\right)G'V^{-1}B\right]^{-}B' \quad = rk\left[V^{-1}G\Omega\left(I-P\right)G'V^{-1}\right]$. Thus, as V is just another generalized inverse for $V^{-1}G\Omega\left(I-P\right)G'V^{-1}$, it follows that $P_3 = n \quad \frac{\rho_2(0)}{\rho_1(0)} \quad ^2 \left(\phi-\phi\right)' V_n \left(\phi-\phi\right) + o_p\left(1\right) = MC_n + o_p\left(1\right)$.

References

Andrews, D. W. K. (1988): "Chi-square diagnostic tests for econometric models: theory", *Econometrica*, 56(6), 1419-1453.

Back, K. and D. P. Brown (1993): "Implied probabilities in GMM estimators", *Econometrica*, 61(4), 971-975.

Hansen, L. P. (1982): "Large sample properties of generalised method of moments estimators", *Econometrica*, 50(4), 1029-1054.

Hansen, L. P., J. Heaton and A. Yaron (1996): "Finite-sample properties of some alternative GMM estimators", *Journal of Business & Economic Statistics*, 14(3), 262-280.

Imbens, G. W. (1997): "One-step estimators for over-identified generalised method of moments models", *Review of Economic Studies*, 64, 359-383.

Imbens, G. W., R. H. Spady and P. Johnson (1998): "Information theoretic approaches to inference in moment condition models", *Econometrica*, 66(2), 333-357.

Kitamura, Y. and M. Stutzer (1997): "An information-theoretic alternative to generalised method of moments estimation", *Econometrica*, 65(4), 861-874.

Newey, W. K., J. J. S. Ramalho, and R. J. Smith, (2001): "Asymptotic bias for GMM and GEL estimators with estimated nuisance parameters", mimeo.

Newey, W. K. and R. J. Smith (2001): "Higher order properties of GMM and generalized empirical likelihood estimators", mimeo.

Qin, J. and J. Lawless (1994): "Empirical likelihood and general estimating equations", *Annals of Statistics*, 22(1), 300-325.

Rao, C. R. and S. K. Mitra (1971): *Generalized inverse of matrices and its applications*, Wiley.

Smith, R. J. (1997), "Alternative semi-parametric likelihood approaches to generalised method of moments estimation", *Economic Journal*, 107(441), 503-519.

Table 1: Monte Carlo estimated sizes for J, W and DM tests of overidentifying moment conditions: asset-pricing model (10 000 replications)

n	Size	J									W						DM	
		2s(n)	ri(n)	cu(n)	et(n)	et(s)	et(r)	el(n)	el(s)	el(r)	et(n)	et(s)	et(r)	el(n)	el(s)	el(r)	et	el
100	20.0%	26.7	26.1	24.0	25.5	29.8	29.2	25.9	28.3	28.5	29.9	25.3	26.9	28.1	28.3	24.8	27.2	27.9
	10.0%	17.5	16.7	12.8	15.9	19.6	20.1	16.4	17.6	19.3	19.8	16.1	14.8	18.7	17.6	15.3	16.9	17.0
	5.0%	12.2	11.3	7.2	10.6	13.7	14.6	11.2	11.2	14.0	13.8	11.0	8.3	13.6	11.2	10.4	11.0	11.1
	2.5%	9.5	8.5	4.5	7.8	9.9	11.5	8.3	7.4	10.8	10.3	8.3	4.6	10.4	7.4	7.6	7.7	7.3
	1.0%	6.9	5.9	2.5	5.2	6.9	8.7	5.7	4.3	8.1	7.3	5.8	2.2	7.6	4.3	5.3	5.0	4.1
	0.5%	5.7	4.4	1.6	3.7	5.4	7.2	4.2	2.9	7.0	5.6	4.4	1.3	6.1	2.9	4.0	3.6	2.8
	0.1%	3.9	2.3	0.7	1.7	3.3	4.7	2.1	1.1	4.6	3.5	2.4	0.4	4.0	1.1	2.7	1.8	1.2
200	20.0%	25.3	25.1	24.2	24.8	28.0	27.7	25.0	26.8	27.2	28.1	24.7	25.7	26.0	26.8	23.6	26.0	26.8
	10.0%	15.0	14.7	13.3	14.4	17.3	18.0	14.7	15.7	17.3	17.4	14.5	13.6	16.5	15.7	13.9	15.2	15.5
	5.0%	9.9	9.5	7.6	9.2	11.2	12.5	9.5	9.4	11.9	11.3	9.5	7.1	11.2	9.4	8.9	9.3	9.1
	2.5%	6.8	6.6	4.5	6.2	7.8	9.1	6.5	5.8	8.5	7.8	6.5	3.8	8.1	5.8	6.6	5.8	5.6
	1.0%	4.6	4.3	2.4	4.0	5.0	6.2	4.3	2.9	6.0	5.0	4.3	1.5	5.9	2.9	4.6	3.6	3.0
	0.5%	3.5	3.1	1.4	2.9	3.5	4.9	3.1	1.9	4.8	3.6	3.1	0.9	4.8	1.9	3.6	2.5	1.9
	0.1%	1.9	1.5	0.4	1.3	2.0	3.0	1.5	0.7	3.0	2.1	1.6	0.2	3.2	0.7	2.1	1.1	0.7
500	20.0%	23.1	23.0	22.7	22.8	25.6	25.4	22.9	24.7	25.1	25.7	22.9	23.7	24.2	24.7	22.1	24.0	24.4
	10.0%	13.1	13.0	12.4	12.8	15.0	15.6	13.0	13.5	15.0	15.0	13.0	12.0	14.4	13.5	12.4	13.4	13.6
	5.0%	8.0	7.9	7.3	7.6	9.3	10.1	7.8	7.9	9.7	9.4	7.8	6.3	9.3	7.9	7.9	7.8	7.7
	2.5%	5.0	4.9	4.1	4.7	6.2	6.9	4.9	4.6	6.6	6.2	4.9	3.4	6.6	4.6	5.4	4.7	4.6
	1.0%	3.0	3.0	2.3	2.9	3.7	4.4	3.0	2.4	4.2	3.7	3.0	1.3	4.5	2.4	3.7	2.5	2.2
	0.5%	2.1	2.1	1.5	1.9	2.4	3.3	2.1	1.4	3.2	2.5	2.0	0.6	3.7	1.4	2.8	1.6	1.3
	0.1%	0.8	0.9	0.4	0.7	1.1	1.7	0.8	0.5	1.8	1.2	0.8	0.1	2.4	0.5	1.8	0.5	0.4
1000	20.0%	21.8	21.8	21.6	21.7	23.5	23.5	21.9	22.6	23.2	23.5	21.8	22.0	22.3	22.6	20.8	22.5	22.7
	10.0%	11.9	11.8	11.6	11.8	13.2	13.9	11.9	12.1	13.6	13.3	11.9	11.2	12.5	12.1	11.2	12.3	12.3
	5.0%	6.7	6.8	6.5	6.7	8.0	8.5	6.8	7.1	8.3	8.1	6.8	5.9	8.0	7.1	6.6	6.9	6.8
	2.5%	4.4	4.3	4.1	4.3	4.8	5.5	4.4	3.9	5.4	5.0	4.4	3.0	5.2	3.9	4.3	4.1	3.9
	1.0%	2.4	2.4	2.3	2.4	2.5	3.5	2.4	1.8	3.4	2.6	2.4	1.2	3.4	1.8	2.5	2.2	1.9
	0.5%	1.7	1.7	1.6	1.7	1.7	2.4	1.7	1.0	2.5	1.7	1.7	0.6	2.4	1.0	1.9	1.3	1.0
	0.1%	0.7	0.7	0.6	0.6	0.7	1.2	0.6	0.3	1.3	0.7	0.7	0.1	1.5	0.3	1.0	0.5	0.3

Note: The actual size closest to the nominal size of all tests contained in Tables 1 and 2 is underlined.

Table 2: Monte Carlo estimated sizes for Pearson-type tests of overidentifying moment conditions: asset-pricing model (10 000 replications)

n	Size	P1		P2		P3 (L=8)						P3 (L=16)					
		et	el	et	el	et(n)	et(s)	et(r)	el(n)	el(s)	el(r)	et(n)	et(s)	et(r)	el(n)	el(s)	el(r)
100	20.0%	26.7	28.6	30.4	28.3	25.2	22.4	20.2	25.8	22.9	21.8	25.1	22.5	20.2	25.4	22.9	21.2
	10.0%	17.0	19.3	20.4	17.6	14.3	15.4	9.3	14.0	14.7	8.3	14.2	15.6	9.7	13.7	15.0	8.1
	5.0%	11.8	14.0	14.6	11.2	8.4	10.9	3.8	7.3	10.2	2.1	8.6	11.1	4.6	7.3	10.4	2.3
	2.5%	8.9	10.8	10.9	7.4	5.0	7.9	1.1	3.8	6.9	0.3	5.3	8.2	1.7	3.7	7.5	0.4
	1.0%	6.4	8.1	7.9	4.3	2.4	4.7	0.1	1.4	4.0	0.0	2.7	5.2	0.3	1.6	4.5	0.0
	0.5%	5.0	7.0	6.2	2.9	1.4	2.7	0.0	0.6	2.3	0.0	1.6	3.5	0.1	0.8	2.8	0.0
	0.1%	2.8	4.6	4.3	1.1	0.3	0.6	0.0	0.1	0.5	0.0	0.6	1.1	0.0	0.1	0.9	0.0
200	20.0%	25.5	27.2	28.5	26.8	25.4	21.6	21.5	25.5	23.3	22.6	25.0	21.5	21.2	25.3	22.8	22.3
	10.0%	15.4	17.3	18.1	15.7	14.3	14.2	10.2	14.1	13.3	10.5	14.1	14.3	10.1	13.7	13.3	9.9
	5.0%	10.0	11.9	12.0	9.4	8.1	10.0	4.8	7.6	8.8	4.3	8.1	10.1	4.8	7.3	8.8	3.9
	2.5%	7.0	8.5	8.7	5.8	4.7	7.1	2.4	4.3	6.0	1.3	4.7	7.2	2.5	4.0	5.9	1.1
	1.0%	4.7	6.0	5.7	2.9	2.5	4.7	0.8	1.9	3.6	0.3	2.5	4.8	0.9	1.6	3.7	0.3
	0.5%	3.5	4.8	4.3	1.9	1.5	3.3	0.3	1.0	2.6	0.1	1.5	3.4	0.4	0.8	2.4	0.1
	0.1%	1.9	3.0	2.7	0.7	0.5	1.3	0.0	0.3	1.0	0.0	0.5	1.5	0.0	0.3	1.0	0.0
500	20.0%	23.6	25.1	26.0	24.7	24.2	21.0	21.6	24.3	22.8	22.2	24.0	21.0	21.3	24.0	22.5	22.0
	10.0%	13.5	15.0	15.6	13.5	13.5	12.4	10.4	13.3	12.6	11.0	13.4	12.3	10.4	12.8	12.3	10.6
	5.0%	8.3	9.7	9.9	7.9	7.7	8.1	5.1	7.7	7.2	5.6	7.5	8.1	4.9	7.4	7.0	5.2
	2.5%	5.3	6.6	6.8	4.6	4.5	5.6	2.5	4.6	4.5	2.6	4.4	5.5	2.5	4.1	4.4	2.4
	1.0%	3.3	4.2	4.2	2.4	2.4	3.6	1.0	2.3	2.7	1.0	2.3	3.6	1.0	2.0	2.6	0.7
	0.5%	2.3	3.2	3.1	1.4	1.4	2.6	0.4	1.4	1.8	0.4	1.3	2.6	0.5	1.1	1.7	0.3
	0.1%	1.0	1.8	1.7	0.5	0.4	1.2	0.1	0.5	0.7	0.1	0.4	1.1	0.1	0.4	0.6	0.0
1000	20.0%	22.1	23.2	23.8	22.6	22.9	20.8	21.1	22.6	22.0	21.2	22.9	20.7	20.9	22.5	21.7	21.1
	10.0%	12.3	13.6	13.6	12.1	12.6	11.5	10.2	12.3	11.9	10.5	12.4	11.5	10.1	12.1	11.7	10.2
	5.0%	7.2	8.3	8.4	7.1	7.1	7.1	5.2	7.0	6.5	5.6	7.0	7.0	5.2	6.9	6.3	5.3
	2.5%	4.6	5.4	5.4	3.9	4.2	4.8	2.7	4.0	4.0	2.6	4.0	4.7	2.7	3.8	4.0	2.5
	1.0%	2.6	3.4	3.0	1.8	2.2	3.0	1.1	2.0	2.3	1.0	2.1	3.0	1.0	1.8	2.2	0.9
	0.5%	1.8	2.5	2.1	1.0	1.2	2.1	0.6	1.2	1.5	0.5	1.2	2.0	0.5	1.0	1.4	0.4
	0.1%	0.7	1.3	1.0	0.3	0.4	1.0	0.1	0.4	0.5	0.1	0.4	1.0	0.1	0.3	0.5	0.1

Note: The actual size closest to the nominal size of all tests contained in Tables 1 and 2 is underlined.

Table 3: Monte Carlo estimated sizes for J, W and DM tests of overidentifying moment conditions: chi-squared moments model (10 000 replications)

n	Size	J									W						DM	
		2s(n)	ri(n)	cu(n)	et(n)	et(s)	et(r)	el(n)	el(s)	el(r)	et(n)	et(s)	et(r)	el(n)	el(s)	el(r)	et	el
100	20.0%	34.6	34.6	34.6	34.7	38.0	38.0	35.0	36.5	37.2	37.7	34.1	35.1	33.8	36.5	31.2	35.7	36.3
	10.0%	27.0	26.9	26.9	27.2	28.0	30.1	27.5	25.9	29.5	27.6	27.0	23.4	25.4	25.9	22.8	26.5	26.0
	5.0%	22.3	22.3	22.3	22.6	21.5	25.2	23.1	19.3	24.9	21.3	22.3	16.9	20.3	19.3	17.8	20.7	19.3
	2.5%	18.8	18.8	18.8	19.1	17.6	21.7	19.8	14.8	21.8	17.6	19.3	12.7	16.7	14.8	14.4	17.1	15.5
	1.0%	15.5	15.5	15.5	15.8	13.4	18.2	16.7	11.1	18.9	13.5	16.2	9.6	13.4	11.1	11.3	13.9	11.5
	0.5%	13.4	13.4	13.4	13.8	11.5	16.4	14.7	9.0	17.3	11.7	14.2	8.0	11.8	9.0	9.5	12.0	9.9
	0.1%	9.8	9.8	9.8	10.2	8.1	13.0	11.3	5.9	14.3	8.6	10.8	5.4	8.9	5.9	7.3	8.9	6.8
200	20.0%	29.0	29.0	29.0	29.0	30.9	31.9	29.1	29.6	31.2	30.5	28.5	28.2	26.0	29.6	23.1	29.6	30.0
	10.0%	20.9	20.9	20.9	21.0	21.1	23.8	21.1	19.5	22.9	20.8	20.8	16.8	17.4	19.5	14.7	20.2	19.8
	5.0%	16.4	16.4	16.4	16.5	14.6	19.0	16.8	12.9	18.4	14.4	16.5	10.5	12.8	12.9	10.8	15.1	13.6
	2.5%	13.8	13.8	13.8	13.8	11.0	15.8	14.1	8.9	15.7	10.9	13.9	6.9	9.8	8.9	8.5	11.7	9.7
	1.0%	10.4	10.4	10.4	10.6	7.6	12.7	11.0	5.8	13.1	7.6	10.6	4.4	7.8	5.8	6.3	8.7	6.5
	0.5%	9.1	9.0	9.0	9.2	6.1	10.8	9.5	4.1	11.3	6.1	9.2	3.1	6.5	4.1	5.3	7.2	5.1
	0.1%	6.3	6.3	6.3	6.5	3.5	8.4	6.9	2.1	9.0	3.5	6.6	1.7	4.7	2.1	3.8	4.7	2.9
500	20.0%	25.4	25.4	25.4	25.4	26.3	27.9	25.5	25.3	27.3	26.1	25.2	24.0	21.7	25.3	19.1	26.4	26.1
	10.0%	16.4	16.4	16.4	16.5	15.6	18.8	16.5	14.4	18.3	15.4	16.3	12.5	13.1	14.4	11.3	15.7	14.8
	5.0%	11.5	11.5	11.5	11.5	9.8	13.3	11.6	8.7	13.0	9.7	11.4	6.9	9.0	8.7	7.8	10.2	9.1
	2.5%	8.6	8.6	8.6	8.6	6.6	10.3	8.6	5.3	10.2	6.5	8.5	3.7	6.7	5.3	5.8	7.1	5.9
	1.0%	6.3	6.3	6.3	6.4	3.7	7.5	6.4	2.7	7.7	3.6	6.3	1.7	4.9	2.7	4.0	5.0	3.3
	0.5%	5.2	5.2	5.2	5.2	2.6	6.2	5.3	1.8	6.5	2.6	5.2	1.0	4.0	1.8	3.2	3.6	2.3
	0.1%	3.0	3.0	3.0	3.1	1.2	4.2	3.2	0.7	4.7	1.2	3.1	0.3	2.7	0.7	2.0	2.0	1.0
1000	20.0%	23.2	23.2	23.2	23.2	24.4	25.1	23.3	23.6	24.6	24.2	23.1	22.5	20.8	23.6	18.8	23.8	23.5
	10.0%	14.0	14.0	14.0	14.1	13.9	16.0	14.1	12.9	15.7	13.7	14.0	11.6	12.0	12.9	10.7	13.9	13.2
	5.0%	9.1	9.1	9.1	9.1	8.4	10.8	9.1	7.4	10.3	8.3	9.1	6.1	8.0	7.4	7.0	8.5	7.9
	2.5%	6.5	6.5	6.5	6.5	5.0	7.9	6.5	4.1	7.7	4.8	6.4	2.9	5.8	4.1	5.2	5.6	4.5
	1.0%	4.3	4.3	4.3	4.3	2.7	5.3	4.3	2.0	5.5	2.7	4.2	1.1	4.2	2.0	3.6	3.2	2.2
	0.5%	3.2	3.2	3.2	3.2	1.7	4.1	3.2	1.2	4.3	1.7	3.2	0.6	3.4	1.2	2.8	2.1	1.3
	0.1%	1.7	1.7	1.7	1.7	0.7	2.2	1.7	0.4	2.5	0.7	1.7	0.1	2.1	0.4	1.6	0.9	0.4

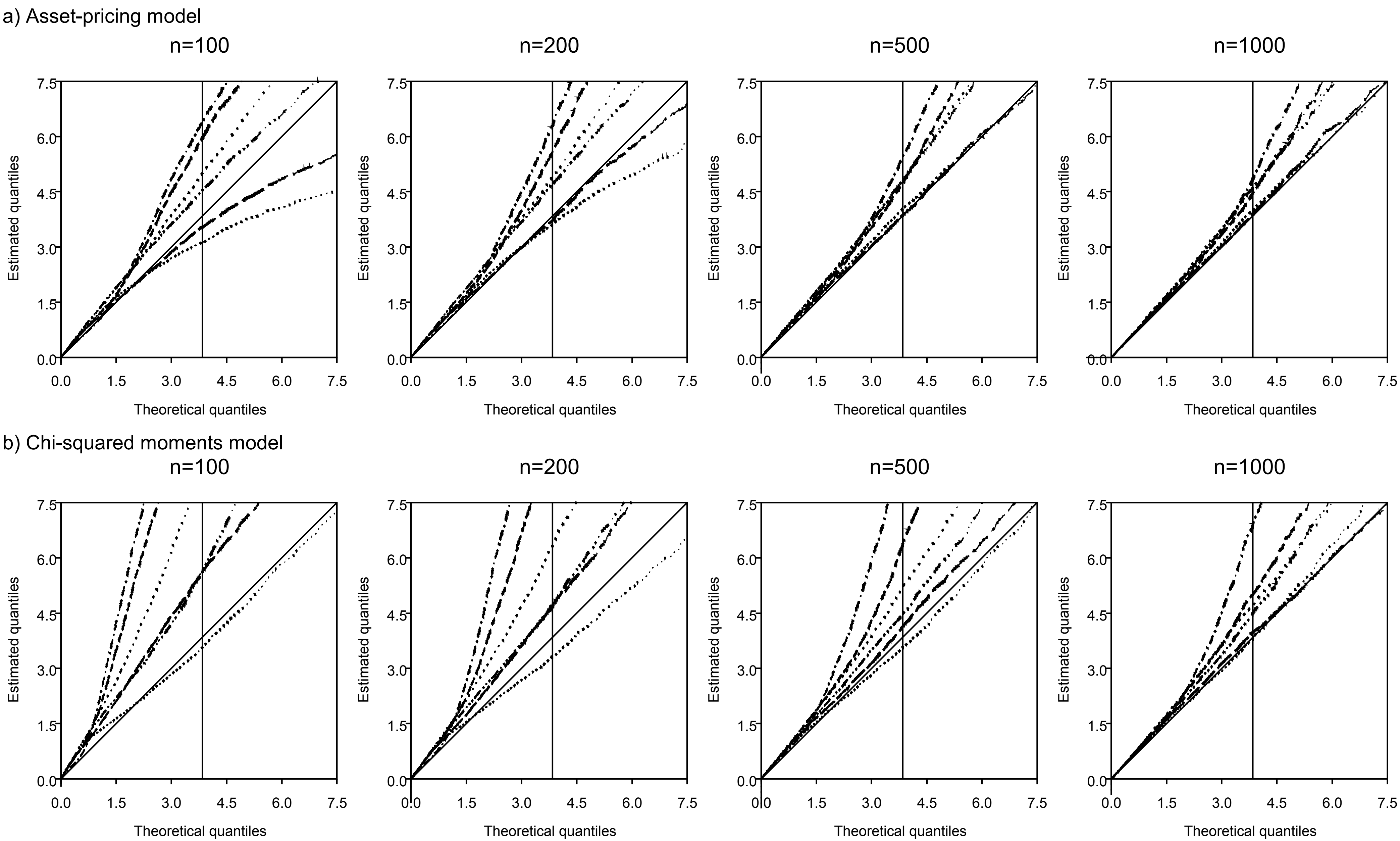
Note: The actual size closest to the nominal size of all tests contained in Tables 3 and 4 is underlined.

Table 4: Monte Carlo estimated sizes for Pearson-type tests of overidentifying moment conditions: chi-squared moments model (10 000 replications)

n	Size	P1		P2		P3 (L=8)						P3 (L=16)					
		et	el	et	el	et(n)	et(s)	et(r)	el(n)	el(s)	el(r)	et(n)	et(s)	et(r)	el(n)	el(s)	el(r)
100	20.0%	35.6	37.2	38.4	36.5	32.8	33.7	27.9	30.3	32.7	22.1	34.3	34.9	29.8	34.3	33.7	29.9
	10.0%	27.9	29.5	28.5	25.9	21.6	28.0	17.1	16.7	25.6	9.2	24.4	29.7	20.6	21.8	28.4	11.5
	5.0%	23.5	24.9	22.1	19.3	15.3	23.4	10.5	9.8	20.6	4.3	18.0	25.6	14.6	13.5	24.1	4.6
	2.5%	20.0	21.8	18.2	14.8	11.0	19.6	6.3	6.1	16.6	2.1	14.3	22.2	9.9	8.6	20.7	2.3
	1.0%	17.1	18.9	14.2	11.1	7.6	15.3	3.1	3.7	12.4	0.9	10.4	18.6	5.3	5.1	17.3	1.1
	0.5%	15.2	17.3	12.2	9.0	5.7	12.6	1.8	2.7	9.6	0.5	8.4	16.3	3.0	3.6	15.0	0.6
	0.1%	12.0	14.3	9.0	5.9	3.1	6.1	0.1	1.4	2.1	0.0	4.8	12.1	0.5	2.0	10.9	0.0
200	20.0%	29.8	31.3	31.3	29.6	28.0	27.7	23.5	26.1	27.6	20.4	29.0	28.6	24.9	28.7	27.8	25.9
	10.0%	21.5	22.9	21.8	19.5	17.6	21.8	13.2	14.1	19.1	8.2	19.2	23.2	15.8	17.5	21.1	11.4
	5.0%	17.2	18.4	15.5	12.9	11.1	17.8	7.6	7.6	14.8	3.5	13.4	19.4	10.4	10.9	17.4	4.6
	2.5%	14.4	15.7	11.7	8.9	7.4	15.0	4.4	4.5	11.5	1.6	9.6	16.6	7.1	6.7	14.8	1.5
	1.0%	11.5	13.1	8.4	5.8	4.6	11.7	2.1	2.4	8.4	0.6	6.6	13.9	4.1	3.3	11.8	0.4
	0.5%	9.8	11.3	6.8	4.1	3.1	9.7	1.3	1.5	6.7	0.3	5.2	11.8	2.6	2.1	9.9	0.1
	0.1%	7.4	9.0	4.1	2.1	1.5	6.7	0.3	0.5	3.9	0.1	2.8	8.9	1.0	0.8	7.2	0.0
500	20.0%	26.0	27.3	26.8	25.3	25.4	23.8	22.6	23.4	25.0	20.5	26.0	24.0	23.2	24.8	24.9	22.2
	10.0%	17.2	18.3	16.1	14.4	14.1	17.0	11.1	12.6	14.7	9.2	15.2	17.7	11.8	13.8	15.8	10.6
	5.0%	12.0	13.0	10.5	8.7	8.5	12.7	5.8	7.0	9.4	4.2	9.5	13.8	6.9	8.1	11.0	5.0
	2.5%	9.1	10.2	7.2	5.3	5.3	9.9	3.2	4.0	6.8	2.1	6.2	11.0	4.4	4.8	8.4	2.0
	1.0%	6.6	7.7	4.5	2.7	2.8	7.3	1.3	1.8	4.7	0.9	3.8	8.4	2.3	2.2	6.1	0.7
	0.5%	5.6	6.5	3.1	1.8	1.8	6.0	0.7	1.3	3.5	0.5	2.5	6.9	1.4	1.3	5.1	0.5
	0.1%	3.6	4.7	1.6	0.7	0.6	3.9	0.2	0.5	1.8	0.1	1.2	4.9	0.4	0.5	3.1	0.1
1000	20.0%	23.6	24.6	24.7	23.6	23.5	21.8	21.6	22.3	23.1	20.3	23.7	22.0	21.9	23.1	23.1	21.2
	10.0%	14.6	15.7	14.4	12.9	13.1	13.9	10.7	12.1	13.0	9.8	13.6	14.4	11.2	12.6	13.5	10.3
	5.0%	9.5	10.3	9.0	7.4	7.7	9.8	5.4	6.7	7.7	4.9	8.2	10.5	6.0	7.3	8.5	5.2
	2.5%	6.8	7.7	5.5	4.1	4.3	7.5	2.4	3.9	5.0	2.7	4.9	8.1	3.2	4.1	6.0	2.5
	1.0%	4.6	5.5	3.1	2.0	2.1	5.4	1.0	2.1	2.7	1.3	2.5	6.0	1.5	2.0	3.8	1.0
	0.5%	3.5	4.3	2.1	1.2	1.2	4.1	0.6	1.4	1.9	0.7	1.5	4.7	0.9	1.3	2.8	0.5
	0.1%	1.9	2.5	1.0	0.4	0.4	2.2	0.1	0.5	0.8	0.2	0.5	2.7	0.4	0.4	1.3	0.1

Note: The actual size closest to the nominal size of all tests contained in Tables 3 and 4 is underlined.

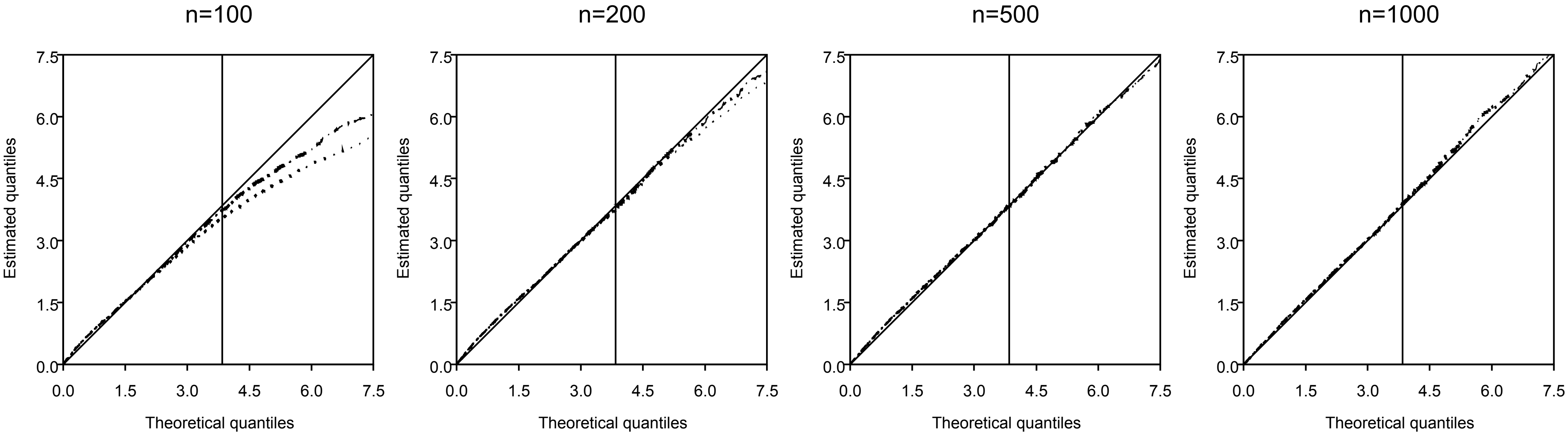
Figure 1: QQ-plots of P3 tests of overidentifying moment conditions (L=8; 10 000 replications)



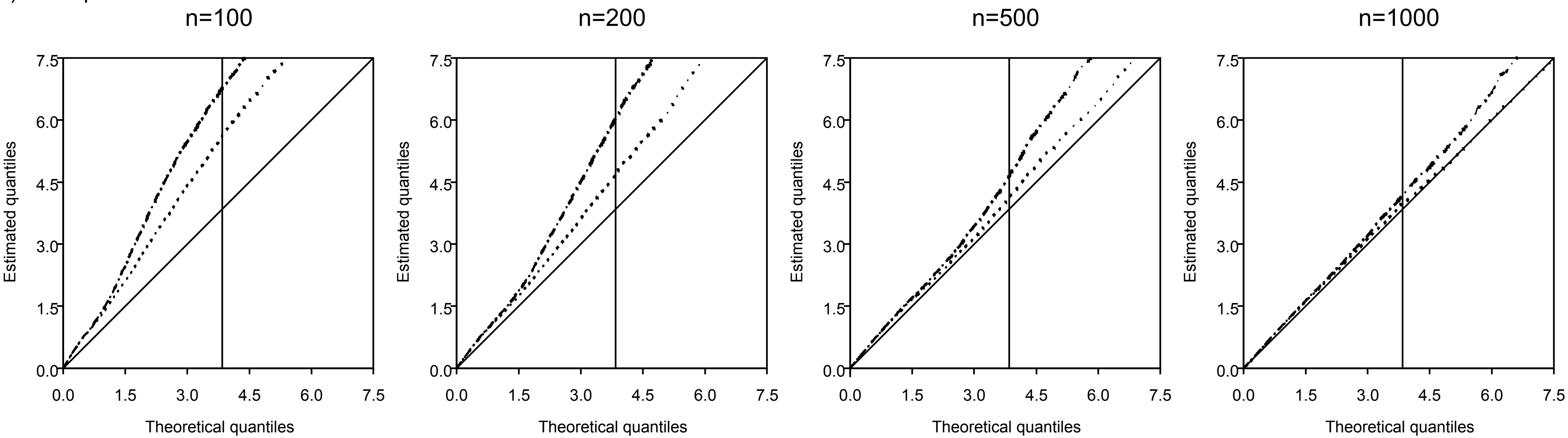
Notes: P3et(n) (dotted line), P3et(s) (dot-dashed line), P3et(r) (dashed line), P3el(n) (three-dot-dashed line), P3el(s) (two-dashed line), P3el(r) (frequent-dotted line).

Figure 2: QQ-plots of et(r) P3 tests of overidentifying moment conditions (10 000 replications)

a) Asset-pricing model



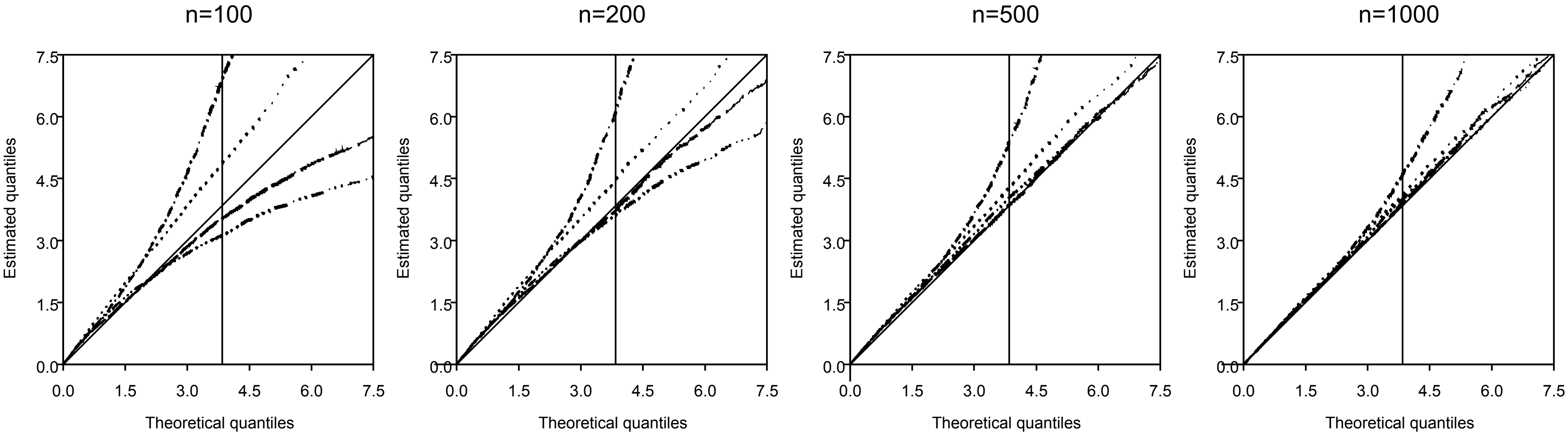
b) Chi-squared moments model



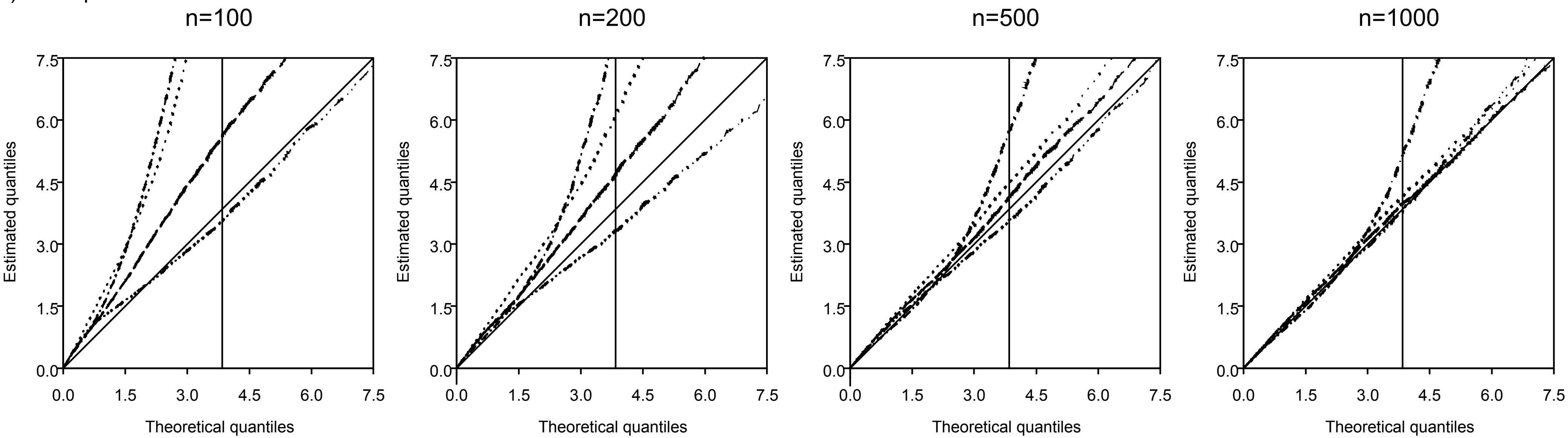
Notes: L=8 (dotted line), L=16 (dot-dashed line).

Figure 3: QQ-plots of robust forms of Wald and P3 (L=8) tests of overidentifying moment conditions (10 000 replications)

a) Asset-pricing model



b) Chi-squared moments model



Notes: Wet(r) (dotted line), Wel(r) (dot-dashed line), P3et(r) (dashed line), P3el(r) (three-dot-dashed line).