Biometrika (2010), **97**, 3, *pp.* 621–630
doi: 10.1093/biomet/asq040
Advance Access publication 9 July 2010

Accurate and robust tests for indirect inference

BY VERONIKA CZELLAR

Department of Economics and Decision Sciences, HEC Paris, 1 rue de la Libération, 78351 Jouy en Josas, France
czellarv@hec.fr

AND ELVEZIO RONCHETTI

Department of Econometrics, University of Geneva, Blv. Pont d'Arve 40, 1211 Geneva, Switzerland
elvezio.ronchetti@unige.ch

SUMMARY

In this paper we propose accurate parameter and over-identification tests for indirect inference. Under the null hypothesis the new tests are asymptotically χ^2-distributed with a relative error of order n^{-1}. They exhibit better finite sample accuracy than classical tests for indirect inference, which have the same asymptotic distribution but an absolute error of order $n^{-1/2}$. Robust versions of the tests are also provided. We illustrate their accuracy in nonlinear regression, Poisson regression with overdispersion and diffusion models.

Some key words: Indirect inference; M-estimator; Nonlinear regression; Overdispersion; Parameter test; Robust estimator; Saddlepoint test; Sparsity; Test for over-identification.

1. INTRODUCTION

Statistical models are becoming increasingly complex. Often the likelihood function is not available in closed form and a standard Bayesian or frequentist likelihood analysis is infeasible. In the Bayesian framework this has led to the development of sequential Monte Carlo methods for approximate Bayesian computations; see, for instance, Del Moral et al. (2006), Beaumont et al. (2009), in the statistical literature and Beaumont et al. (2002), Lopes et al. (2009) and references therein in genetics. In the frequentist framework, indirect inference is a broad class of estimators that includes the method of moments and the generalized method of moments estimators (Hansen, 1982) with additively separable orthogonality functions, and simulated method of moments estimators (Lee & Ingram, 1991; Duffie & Singleton, 1993). Indirect inference estimators were introduced by Smith (1993) and Gouriéroux et al. (1993), and have now been applied in a variety of fields, including financial models (Gouriéroux & Monfort, 1996; Billio & Monfort, 2003; Czellar et al., 2007; Calzolari et al., 2008) and in regression models with measurement error (Kuk, 1995; Turnbull et al., 1997; Jiang et al., 1999). Good surveys on indirect inference are provided by Heggland & Frigessi (2004) and Jiang & Turnbull (2004).

In this paper we focus on testing for indirect inference. The procedure can be summarized as follows. Let $y_1, \ldots, y_n$ be a random sample from a distribution F_θ, where $\theta \in \mathbb{R}^p$ is an unknown parameter. If the direct estimation of θ is unfeasible, choose a simpler auxiliary estimator

$\tilde{\mu} \in \mathbb{R}^r (r \geqslant p)$, which can be written as an M-estimator defined by

$$\sum_{i=1}^{n} \Psi(y_i; \tilde{\mu}) = 0, \tag{1}$$

where Ψ is an appropriate function; see, e.g. (10), (11), (13) and (14) below. Then for a given one-to-one function $\mu : \mathbb{R}^p \to \mathbb{R}^r$ and a given symmetric, positive definite matrix $\Omega \in \mathbb{R}^r \times \mathbb{R}^r$, the indirect inference estimator is defined by

$$\hat{\theta} = \arg\min_{\theta} \{\tilde{\mu} - \mu(\theta)\}^{\mathrm{T}} \Omega \{\tilde{\mu} - \mu(\theta)\}, \tag{2}$$

i.e. the value of the parameter θ such that $\mu(\hat{\theta})$ is closest to $\tilde{\mu}$ as measured by the metric defined by the matrix Ω. The function $\mu(\theta)$ is typically chosen as follows.

(i) The function $\mu(\theta) = E_{F_\theta}(\tilde{\mu})$, or $\mathrm{plim}_{F_\theta} \tilde{\mu}$, is deterministic and is often called a binding function (Gouriéroux et al., 1993), a bridge relationship (Turnbull et al., 1997; Jiang et al., 1999) or a target function (Cabrera & Fernholz, 1999).

(ii) The function $\mu(\theta) = m^{-1} \sum_{s=1}^{m} \tilde{\mu}^{(s)}(\theta)$ is obtained by simulation, where for each $s = 1, \ldots, m$, $\tilde{\mu}^{(s)}(\theta)$ is the auxiliary estimate calculated using simulated data $\{y_i^{(s)}(\theta)\}$ generated from F_θ, and $m \geqslant 1$ is fixed.

(iii) An alternative choice obtained by simulation is $\mu(\theta) = \arg\min_\mu m^{-1} \sum_{s=1}^{m} \mathcal{L}[\mu; \{y_i^{(s)}(\theta)\}]$, where $\mathcal{L}$ is the objective function associated with the auxiliary estimator, i.e. $\partial \mathcal{L}/\partial \mu[\tilde{\mu}; \{y_i\}] = 0$.

Assume that the parameter θ is partitioned into $(\theta_1^{\mathrm{T}}, \theta_2^{\mathrm{T}})^{\mathrm{T}}$, $\theta_1 \in \mathbb{R}^{p_1}$ and $\theta_2 \in \mathbb{R}^{p-p_1}$ and consider the null hypothesis $H_0 : \theta_1 = \theta_{10}$. In the indirect inference setting, likelihood ratio-type tests are based on the optimal value of the objective function in (2), i.e.

$$W(\hat{\theta}, \hat{\theta}^c, \Omega^*) = \{\tilde{\mu} - \mu(\hat{\theta}^c)\}^{\mathrm{T}} \Omega^* \{\tilde{\mu} - \mu(\hat{\theta}^c)\} - \{\tilde{\mu} - \mu(\hat{\theta})\}^{\mathrm{T}} \Omega^* \{\tilde{\mu} - \mu(\hat{\theta})\}, \tag{3}$$

where Ω^* is the inverse of the asymptotic variance-covariance matrix of the auxiliary estimator under the true data-generating model as defined in Gouriéroux et al. (1993, p. S91), and $\hat{\theta}^c$ is the constrained indirect inference estimator under H_0. In contrast with standard likelihood ratio tests, the factor 2 does not appear in front. Under the null hypothesis $nW(\hat{\theta}, \hat{\theta}^c, \Omega^*)$ is asymptotically $\chi^2_{p_1}$-distributed. The classical test for the validity of the over-identifying conditions ($r > p$) is given by the J test:

$$J(\hat{\theta}, \Omega^*) = \{\tilde{\mu} - \mu(\hat{\theta})\}^{\mathrm{T}} \Omega^* \{\tilde{\mu} - \mu(\hat{\theta})\},$$

where $nJ(\hat{\theta})$ is asymptotically χ^2_{r-p}-distributed. For the simulation-based indirect inference, the W and J statistics are premultiplied by $m/(m+1)$. In general, the optimal weighting matrix Ω^* is unknown and needs to be estimated by a consistent estimator, for instance, $\widetilde{\Omega}^* = M_\mu I^{-1} M_\mu$, with $M_\mu = n^{-1} \sum_{i=1}^{n} \partial \Psi / \partial \mu^{\mathrm{T}}(y_i; \tilde{\mu})$ and $I = n^{-1} \sum_{i=1}^{n} \{\Psi(y_i; \tilde{\mu}) \Psi^{\mathrm{T}}(y_i; \tilde{\mu})\}$. The χ^2-approximations to the distributions of the W and J tests have an absolute error of order $n^{-1/2}$ and their finite sample accuracy can be very poor and can lead to misleading inference; see, e.g. Michaelides & Ng (2000), Duffee & Stanton (2008), and Table 1 and Fig. 1 below.

To improve this, we introduce saddlepoint tests based on indirect inference estimators, which are also asymptotically χ^2-distributed under the null but with a relative error of order n^{-1}. Another advantage of the saddlepoint tests is that they do not require the knowledge of Ω^*. We show by simulation that saddlepoint tests are very accurate down to small sample sizes. They outperform in accuracy classical tests with both non-simulation-based and simulation-based indirect inference estimators. We also introduce robust versions of the saddlepoint tests which provide accurate inference in the presence of local deviations from the stochastic assumptions of the model. In some cases they outperform nonrobust saddlepoint and classical tests even when there are no deviations from the assumed model.

Our paper is related to several earlier contributions, in particular work by Robinson et al. (2003), Lô & Ronchetti (2009), Kitamura & Stutzer (1997) and Sowell in an unpublished 2009 Carnegie Mellon University technical report. In § 3 we clarify the connection between these papers and our proposal.

2. Saddlepoint tests for indirect inference

2·1. *Just-identified case*

The following proposition introduces a saddlepoint parameter test for just-identified ($r = p$) indirect inference estimators.

Proposition 1. *Consider the null hypothesis* $H_0 : \theta_1 = \theta_{10} \in \mathbb{R}^{p_1}$, *where* $\theta = (\theta_1^{\mathrm{T}}, \theta_2^{\mathrm{T}})^{\mathrm{T}}$. *Denote by* $\hat{\theta} = (\hat{\theta}_1^{\mathrm{T}}, \hat{\theta}_2^{\mathrm{T}})^{\mathrm{T}}$ *the indirect inference estimator given by* (2) *and make the following assumptions.*

Assumption 1. *The dimension of the auxiliary estimator* $\tilde{\mu}$ *is* $r = p$.

Assumption 2. *The indirect inference estimator* $\hat{\theta}$ *satisfies* $\tilde{\mu} = \mu(\hat{\theta})$.

Assumption 3. *The density of* $\hat{\theta}$ *exists and has the saddlepoint approximation*

$$f_{\hat{\theta}}(t) = (2\pi/n)^{p/2} \exp[nK_{\psi}\{\lambda(t); t\}] |B(t)| |\Sigma(t)|^{-1/2} \{1 + \mathcal{O}(n^{-1})\},$$

where $\psi(y_i; \theta) \equiv \Psi\{y_i; \mu(\theta)\}$, $K_{\psi}(\lambda; t) = \log E_{F_\theta}[\exp\{\lambda^{\mathrm{T}} \psi(y_i; t)\}]$, $\lambda(t)$ *is the saddlepoint satisfying* $\partial K_{\psi}/\partial\lambda(\lambda; t) = 0$, $|\cdot|$ *denotes the determinant,*

$$B(t) = \exp[-K_{\psi}\{\lambda(t); t\}] E_{F_\theta}[\partial\psi/\partial t(y_i; t) \exp\{\lambda(t)^{\mathrm{T}} \psi(y_i; t)\}]$$

and

$$\Sigma(t) = \exp[-K_{\psi}\{\lambda(t); t\}] E_{F_\theta}[\psi(y_i; t)\psi^{\mathrm{T}}(y_i; t) \exp\{\lambda(t)^{\mathrm{T}} \psi(y_i; t)\}].$$

Define the saddlepoint test statistic by

$$S(\hat{\theta}_1) = \inf_{\theta_2} \sup_{\lambda} \{-\log E_{F_{(\theta_{10}, \theta_2)}}(\exp[\lambda^{\mathrm{T}} \Psi\{y_i; \mu(\hat{\theta}_1, \theta_2)\}])\}. \tag{4}$$

Then, under H_0, $2nS(\hat{\theta}_1)$ *is asymptotically* $\chi^2_{p_1}$*-distributed with relative error of order* n^{-1}.

The proof is based on Robinson et al.'s (2003) test for M-estimators and is given in the online Supplementary Material.

Remark 1. Sufficient conditions for Assumption 3 are provided by Theorem 2 in Almudevar et al. (2000), p. 290.

Remark 2. The S test does not depend on Ω. This is an advantage when compared to the W test which is asymptotically $\chi^2_{p_1}$-distributed only for an optimal weighting matrix.

Remark 3. By the same arguments as in Lô & Ronchetti (2009, p. 2128), Proposition 1 still holds when the observations are independent but not identically distributed and the auxiliary M-estimator is defined by $\sum_{i=1}^{n} \Psi_i(y_i; \tilde{\mu}) = 0$. In this case, denote by F^i the distribution of y_i

and equation (4) becomes

$$S(\hat\theta_1) = \inf_{\theta_2} \sup_{\lambda} \left\{ -n^{-1} \sum_{i=1}^{n} \log E_{F^i_{(\theta_{10},\theta_2)}} (\exp[\lambda^{\mathrm{T}} \Psi_i\{y_i; \mu(\hat\theta_1, \theta_2)\}]) \right\}.$$

Remark 4. In the case when the cumulant generating function does not have a closed form, we can apply Robinson et al.'s (2003) empirical version of the test. It is defined by

$$\hat S(\hat\theta_1) = \inf_{\theta_2} \sup_{\lambda} \left[-K_0^w\{\lambda; (\hat\theta_1, \theta_2)\} \right],$$

$$K_0^w(\lambda; \theta) = \log \left(\sum_{i=1}^{n} w_i \exp[\lambda^{\mathrm{T}} \Psi\{y_i; \mu(\theta)\}] \right),$$

based on the empirical distribution $\hat F_0 = (w_1, \ldots, w_n)$, which satisfies the null hypothesis and is closest to $\hat F = (1/n, \ldots, 1/n)$ with respect to the backward Kullback–Leibler distance $d(\hat F_0, \hat F) = \sum_{i=1}^{n} w_i \log\{w_i/(1/n)\}$, i.e.

$$w_i = \exp[\lambda(\theta^*)^{\mathrm{T}} \Psi\{y_i; \mu(\theta^*)\}] \bigg/ \sum_{j=1}^{n} \exp[\lambda(\theta^*)^{\mathrm{T}} \Psi\{y_j; \mu(\theta^*)\}], \tag{5}$$

where $\theta^* = (\theta_{10}^{\mathrm{T}}, \theta_2^{*\mathrm{T}})^{\mathrm{T}}$, $\theta_2^* = \arg\min_{\theta_2}[-\kappa\{\lambda(\theta_{10}, \theta_2); (\theta_{10}, \theta_2)\}]$, $\kappa(\lambda; \theta) = \log(1/n \sum_{i=1}^{n} \exp[\lambda^{\mathrm{T}} \Psi\{y_i; \mu(\theta)\}])$ and $\lambda(\theta) = \arg\max_{\lambda}\{-\kappa(\lambda; \theta)\}$.

2·2. *Extension to the over-identified case*

We consider the over-identified case, $r > p$, which plays an important role in certain applications, in particular in econometrics. The goal is to construct a saddlepoint test that can be used to test hypotheses both on parameters of the model and on the validity of over-identifying conditions.

We make the following assumptions:

Assumption 4. The function $\mu(\theta)$ is continuously differentiable in $\theta \in \Theta \subset \mathbb{R}^p$.

Assumption 5. The matrix $\partial \mu^{\mathrm{T}}(\theta)/\partial\theta$ has full column rank for all $\theta \in \Theta \subset \mathbb{R}^p$.

From the first-order conditions of the minimization problem (2), $\hat\theta$ satisfies

$$\frac{\partial \mu^{\mathrm{T}}(\hat\theta)}{\partial\theta} \Omega\{\tilde\mu - \mu(\hat\theta)\} = \bar M(\hat\theta)^{\mathrm{T}} \Omega^{1/2} \{\tilde\mu - \mu(\hat\theta)\} = 0, \tag{6}$$

where $\bar M(\theta) = \Omega^{1/2} \partial\mu(\theta)/\partial\theta^{\mathrm{T}}$ and $\Omega^{1/2}$ is a Cholesky decomposition of Ω, i.e. $\Omega = (\Omega^{1/2})^{\mathrm{T}} \Omega^{1/2}$. Define a p-dimensional manifold in $\mathbb{R}^r$ as $\mathcal{M} = \{\Omega^{1/2}\mu(\theta) \in \mathbb{R}^r \mid \theta \in \Theta\}$. Consistently with the generalized method of moments literature, we refer to the p-dimensional tangent space $T_{\Omega^{1/2}\mu(\theta)}\mathcal{M}$ as the identifying space (Sowell, 1996), and to the $(r-p)$-dimensional orthogonal complement $N_{\Omega^{1/2}\mu(\theta)}\mathcal{M}$ as the over-identifying space (Hansen, 1982). We use Sowell's (1996) spectral decomposition of the idempotent matrix

$$\bar M(\theta)\{\bar M(\theta)^{\mathrm{T}} \bar M(\theta)\}^{-1} \bar M(\theta)^{\mathrm{T}} = C(\theta) \begin{pmatrix} I_p & 0 \\ 0 & 0_{(r-p)\times(r-p)} \end{pmatrix} C(\theta)^{\mathrm{T}}$$

such that $C(\theta)^{\mathrm{T}} C(\theta) = I_r$, $C(\theta) = \{C_1(\theta), C_2(\theta)\}$ and write the identifying and over-identifying spaces as

$$T_{\Omega^{1/2}\mu(\theta)}\mathcal{M} = \{\bar M(\theta) x \mid x \in \mathbb{R}^p\} = \{C_1(\theta) x \mid x \in \mathbb{R}^p\},$$

$$N_{\Omega^{1/2}\mu(\theta)}\mathcal{M} = \{y \in \mathbb{R}^r \mid x^{\mathrm{T}} \bar M(\theta)^{\mathrm{T}} y = 0, \text{ for all } x \in \mathbb{R}^p\} = \{C_2(\theta)\xi \mid \xi \in \mathbb{R}^{r-p}\}.$$

A construction of the function $C(\theta)$ is available in the online Supplementary Material. From the first-order conditions (6) for any $x \in \mathbb{R}^p$, $x^{\mathrm{T}}\bar{M}(\hat{\theta})^{\mathrm{T}}\Omega^{1/2}\{\tilde{\mu} - \mu(\hat{\theta})\} = 0$, which implies that $\Omega^{1/2}\{\tilde{\mu} - \mu(\hat{\theta})\} \in N_{\Omega^{1/2}\mu(\hat{\theta})}\mathcal{M}$. Choosing the associated $\hat{\xi} = C_2(\hat{\theta})^{\mathrm{T}}\Omega^{1/2}\{\tilde{\mu} - \mu(\hat{\theta})\}$, we have $\tilde{\mu} = \mu(\hat{\theta}) + \Omega^{-1/2}C_2(\hat{\theta})\hat{\xi}$. Define the augmented parameter $\eta = (\theta^{\mathrm{T}}, \xi^{\mathrm{T}})^{\mathrm{T}}$ and the function $g(\eta) = \mu(\theta) + \Omega^{-1/2}C_2(\theta)\xi$. Since the auxiliary estimator $\tilde{\mu}$ is an M-estimator defined by (1), we can write the augmented indirect inference estimator $\hat{\eta} = (\hat{\theta}^{\mathrm{T}}, \hat{\xi}^{\mathrm{T}})^{\mathrm{T}}$ as an M-estimator:

$$\sum_{i=1}^{n} \Psi\{y_i; g(\hat{\eta})\} = 0, \tag{7}$$

and we can define the $\hat{S}$ test based on (7) to test the hypothesis $H_0 : \eta_1 = \eta_{10}$ in $\mathbb{R}^{r_1}$.

Specifically, the $\hat{S}$ test statistic can be used to perform the following tests. We define a parameter test for the hypothesis $H_0 : \theta_1 = \theta_{10} \in \mathbb{R}^{p_1}$ by marginalizing over θ_2 and ξ:

$$\hat{S}^{\mathrm{par}}(\hat{\theta}_1) = \inf_{\theta_2,\xi} \sup_{\lambda} \left[-K_0^{\omega}\{\lambda; (\hat{\theta}_1, \theta_2, \xi)\}\right],$$

where $2n\hat{S}^{\mathrm{par}}(\hat{\theta}_1)$ is asymptotically $\chi^2_{p_1}$-distributed. We define an over-identification test for the hypothesis $H_0 : \xi = 0$ in $\mathbb{R}^{r-p}$ by marginalizing over θ:

$$\hat{S}^{\mathrm{over}}(\hat{\xi}) = \inf_{\theta} \sup_{\lambda} \left[-K_0^{\omega}\{\lambda; (\theta, \hat{\xi})\}\right], \tag{8}$$

where $2n\hat{S}^{\mathrm{over}}(\hat{\xi})$ is asymptotically χ^2_{r-p}-distributed.

More general hypotheses of the form $H_0 : h(\eta) = 0 \in \mathbb{R}^{r_1}, r_1 \leqslant r$ can be tested using (8), where inf is taken over $\{\eta : h(\eta) = h(\hat{\eta})\}$. This includes, for instance, testing the validity of a single orthogonality condition.

3. Connections to the literature

Robinson et al. (2003) introduced a saddlepoint test statistic for M-estimators, that is asymptotically χ^2-distributed with a relative error of order n^{-1} under the null hypothesis. We extend their methodology in two directions by constructing saddlepoint tests for hypotheses on the parameters in indirect inference and the validity of the over-identifying conditions.

Just-identified ($r = p$) indirect inference estimators for θ are M-estimators with score function $\Psi\{y; \mu(\theta)\}$. Therefore, this case can be reinterpreted in the generalized linear models set-up when the binding function $\mu(\cdot)$ is a deterministic and known inverse link function such as the $\exp(\cdot)$ for Poisson regression. A saddlepoint test for generalized linear models in its parametric version was developed by Lô & Ronchetti (2009) and our proposal can be viewed as an extension to the nonparametric set-up and to the case where the link function is not deterministic but must be simulated.

The nonparametric over-identification test corresponds to the proposal by Kitamura & Stutzer (1997). However, our derivation in § 2·2 using a nonparametric version of Robinson et al.'s (2003) original test explains the good second-order properties of this test beyond the first-order equivalence. This test can be viewed as an empirical likelihood procedure, where the weights w_i in (5) are obtained by minimizing the backward Kullback–Leibler distance $d(\hat{F}_0, \hat{F})$, in contrast with minimizing the forward Kullback–Leibler distance $d(\hat{F}, \hat{F}_0)$, which leads to the weights of Owen's original empirical likelihood (Owen, 2001). However, the latter does not enjoy the second-order properties of the former. For a comparison in the framework of M-estimators, see Monti & Ronchetti (1993).

Sowell's 2009 technical report treats the empirical saddlepoint approximation of the density of an M-estimator (Ronchetti & Welsh, 1994) as an empirical likelihood function, defines an estimator as the maximum of this function, and derives the corresponding empirical likelihood ratio test. In the just-identified case, this is the analytic counterpart of bootstrap likelihoods as defined in Davison et al. (1995), where the saddlepoint approximation is replaced by resampling. In the set-up of generalized method of moments, Sowell shows that the new estimator is first-order equivalent to the standard generalized method of moments estimator, but exhibits smaller higher-order bias. The derivation of more general second-order properties for this estimator and test seems to be an open problem.

Finally, note that the results derived in § 2·2 connect to the modern theory of sparsity; see, for instance, Meier et al. (2009) and references therein, and it would be fruitful to deepen this link.

4. ILLUSTRATION OF THE ACCURACY OF SADDLEPOINT TESTS

4·1. *Non-simulation-based indirect inference: a nonlinear regression*

In this subsection we focus on non-simulation-based indirect inference estimators defined by a binding function and we compare the accuracy of classical and saddlepoint tests. To illustrate this point, consider data $y_1, \ldots, y_n$ generated from the model

$$y_i = \exp(x_i\theta) + \epsilon_i, \qquad (9)$$

where $x_i \in \mathbb{R}$, $\theta \in \mathbb{R}$ is an unknown parameter and $\epsilon_1, \ldots, \epsilon_n$ are independent and identically distributed standard normal errors. We choose the auxiliary linear model obtained by Taylor approximation of the right-hand side of (9),

$$y_i = 1 + \tilde{x}_i^{\mathrm{T}}\mu + u_i,$$

where $\tilde{x}_i = \{x_i, (x_i)^2, \ldots, (x_i)^r\}^{\mathrm{T}}$, $\mu = (\mu_1, \ldots, \mu_r)^{\mathrm{T}}$ and $u_1, \ldots, u_n$ are independent and identically distributed standard normal errors. Consider the indirect inference estimator with the auxiliary least-squares estimator $\tilde{\mu}$, characterized by the Ψ_i and $\mu(\theta)$ functions:

$$\Psi_i(y_i;\mu) = (y_i - 1 - \tilde{x}_i^{\mathrm{T}}\mu)\tilde{x}_i, \quad \mu(\theta) = E_{F_\theta}(\tilde{\mu}) = \left(\sum_{i=1}^{n} \tilde{x}_i\tilde{x}_i^{\mathrm{T}}\right)^{-1} \sum_{i=1}^{n} \tilde{x}_i\{\exp(x_i\theta) - 1\}. \qquad (10)$$

The resulting indirect inference estimator $\hat{\theta}$ is a generalized method of moments estimator. We generate 10 000 samples for each of the sample sizes $n = 40, 100, 200, 400$, from model (9) with $\theta = 0{\cdot}5$, and $\{x_{ij}\}$ from a uniform distribution on $[-1, 1]$. We consider $r = 1, 2, 3$ and use the weighting matrix $\widetilde{\Omega}^*$. For $r = 1$ we compute the W, S, which has an analytic form in this case as is shown in the Supplementary Material, and $\hat{S}$ statistics to test the hypothesis $H_0 : \theta = 0{\cdot}5$. For $r = 2, 3$, we compute the W, $\hat{S}^{\mathrm{par}}$ to test the hypothesis $H_0 : \theta = 0{\cdot}5$ and the J and $\hat{S}^{\mathrm{over}}$ to test the validity of the over-identifying conditions. Table 1 shows the empirical rejection frequencies of the W, S, $\hat{S}^{\mathrm{par}}$, J and $\hat{S}^{\mathrm{over}}$ tests for a 5% nominal size. In the just-identified case, the S test is accurate even for sample sizes as small as 40 and, starting from sample sizes of 100, the $\hat{S}$ test has a similar accuracy as the S test. Moreover, in each case the $\hat{S}$ tests greatly improve the accuracy of their classical counterparts and the gain is particularly important for small sample sizes and increased number of auxiliary parameters.

4·2. *The effect of simulation: a Poisson regression with overdispersion*

In this section we apply the $\hat{S}$ test to Poisson regression with overdispersion and study its accuracy when it is computed with non-simulation-based and simulation-based indirect inference estimators. Consider the Poisson regression model with overdispersion:

$$y_i \mid (z_i, \varepsilon_i) \sim \mathcal{P}\{\exp(\alpha + \beta z_i + \sigma\varepsilon_i)\},$$

Table 1. *Actual size* (%) *of classical and saddlepoint tests with nominal size* 5%

	$r=1$			$r=2$				$r=3$			
	W	$\hat{S}$	S	W	$\hat{S}^{\text{par}}$	J	$\hat{S}^{\text{over}}$	W	$\hat{S}^{\text{par}}$	J	$\hat{S}^{\text{over}}$
$n=40$	6·8	6·3	4·9	10·2	6·9	9·1	6·2	15·6	5·0	15·0	4·4
$n=100$	5·6	5·4	4·9	6·7	5·9	6·7	5·7	7·9	6·9	8·0	5·3
$n=200$	5·5	5·4	5·2	5·9	5·5	6·3	5·8	6·5	6·2	6·4	5·7
$n=400$	5·3	5·2	5·2	5·9	5·5	5·4	5·2	6·0	5·7	6·0	5·5

where z_i are Bernoulli variables with probability 0·5 and ε_i are independent and identically distributed standard normal random variables. This model was estimated through non-simulation-based indirect inference by Jiang & Turnbull (2004) by using the auxiliary estimator $\tilde{\mu}$:

$$(\tilde{\mu}_1, \tilde{\mu}_2)^{\mathrm{T}} = \arg\max_{(\mu_1,\mu_2)} \sum_{i=1}^{n} \{y_i(\mu_1 + z_i\mu_2) - \exp(\mu_1 + z_i\mu_2)\}, \quad \tilde{\mu}_3 = n^{-1}\sum_{i=1}^{n} y_i^2.$$

This estimator is characterized by

$$\Psi(y_i, z_i; \mu) = \{y_i - \exp(\mu_1 + z_i\mu_2), y_i z_i - \exp(\mu_1 + z_i\mu_2) z_i, y_i^2 - \mu_3\}^{\mathrm{T}} \tag{11}$$

and

$$\mu(\theta) = (\alpha + \sigma^2/2, \beta, [\{1 + \exp(\beta)\}\exp(\alpha + \sigma^2/2) + \{1 + \exp(2\beta)\}\exp\{2(\alpha + \sigma^2)\}]/2)^{\mathrm{T}}.$$

In § 4 and in Fig. 1 of the Supplementary Material, we describe how to compute a simulation-based indirect inference estimator $\hat{\theta}^{(m)}$ using the second procedure for $\mu(\theta)$ described in § 1.

As an illustration, we choose parameter values $\alpha = -1$, $\beta = 1$ and $\sigma = 0{\cdot}5$ and a sample size $n = 250$. We consider the simple hypothesis test $H_0 : (\alpha, \beta, \sigma)^{\mathrm{T}} = (-1, 1, 0{\cdot}5)^{\mathrm{T}}$ at the nominal level 5%. The empirical rejection frequencies of the non-simulation-based W and $\hat{S}$ are, respectively, 11·3% and 7·7%. It appears that for small simulation sizes the χ^2_3-distribution is not an accurate approximation of the simulation-based $\hat{S}$ test. For $m < 10$, the W tests are more accurate. For $m \geqslant 10$, the simulation-based $\hat{S}$ test overperforms the W tests and for $m \geqslant 20$, the simulation-based and non-simulation-based $\hat{S}$ tests have similar good accuracy.

4·3. *Robust saddlepoint tests: the case of diffusion models*

An additional problem that arises when estimating and testing through indirect inference is model misspecification, which can lead to biased estimators and misleading test results. Dridi et al. (2007) provide a complete discussion of this problem. Here we focus on the local robustness properties of indirect inference estimators and tests when the observations $\{y_i\}$ are generated from a neighbourhood of the structural model of the form $(1 - \varepsilon)F_\theta + \varepsilon G$, where G is an arbitrary distribution. Local robustness is achieved if the influence function of the indirect inference estimator is bounded. Since this influence function is proportional to the influence function of the auxiliary estimator (Genton & de Luna, 2000), and this is an M-estimator defined by (1), boundedness of Ψ in (1) implies local robustness for the indirect inference estimator. Hence we construct a robust $\hat{S}$ test by (4) with an auxiliary M-estimator based on a bounded Ψ function. This is an alternative to the W test proposed by Genton & Ronchetti (2003) which is based on the objective function (3) with robust $\tilde{\mu}$ estimators.

To illustrate this point, consider a geometric Brownian motion with drift

$$dy_t = \beta y_t dt + \sigma y_t dW_t, \tag{12}$$

where W_t is a standard Wiener process. Let $\theta = (\beta, \sigma)^{\mathrm{T}}$ and consider the auxiliary model $y_t = (1 + \mu_1)y_{t-1} + \mu_2 y_{t-1}\epsilon_t$, obtained through a crude Euler discretization of (12), where

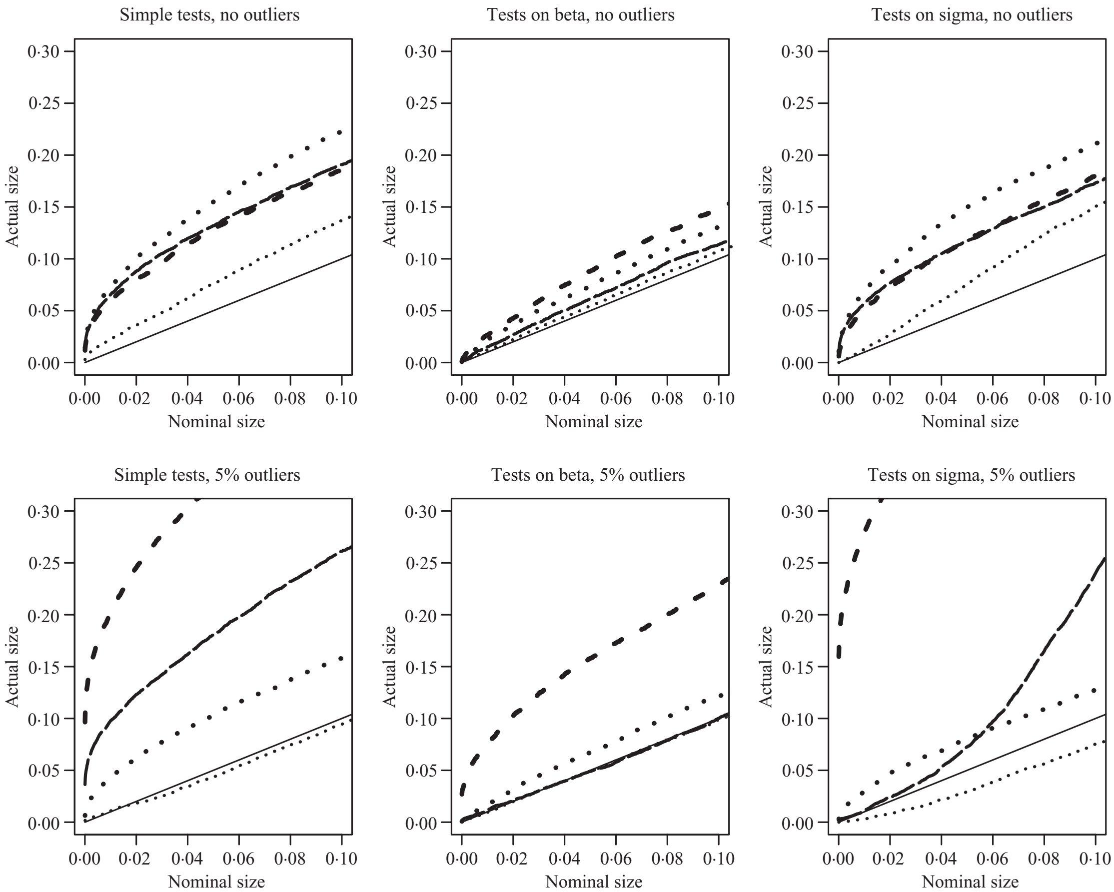

Fig. 1. p-value plots of simulation-based W (long dashes), $\hat{S}$ (small dashes), robust W (large dots) and robust $\hat{S}$ (small dots) tests for the model $dy_t = \beta y_t dt + \sigma y_t dW_t$ with $\beta = -0{\cdot}05$, $\sigma = 0{\cdot}2$ and $n = 40$ when testing $H_0 : (\beta, \sigma)^{\mathrm{T}} = (-0{\cdot}05, 0{\cdot}2)^{\mathrm{T}}$ (first column), $H_0 : \beta = -0{\cdot}05$ (second column) and $\sigma = 0{\cdot}2$ (third column). The upper panels present tests when the original data is generated under the model without contamination while the lower panels present tests with 5% additive outliers. The simulation size is $m = 20$.

ϵ_t are independent and identically distributed standard normal variables. The auxiliary maximum likelihood estimator is defined by $\sum_{t=1}^{n} \Psi(r_t; \tilde{\mu}) = 0$, where $r_t = y_t/y_{t-1}$ and

$$\Psi(r_t; \mu) = \{r_t - 1 - \mu_1, (r_t - \mu_1 - 1)^2 - \mu_2^2\}^{\mathrm{T}}. \tag{13}$$

Let $\hat{\theta}$ be the simulation-based indirect inference estimator of type (iii) as described in § 1 with $\mathcal{L}[\mu; \{r_t^{(s)}(\theta)\}] = [\sum_t \Psi\{r_t^{(s)}(\theta); \mu\}]^{\mathrm{T}}[\sum_t \Psi\{r_t^{(s)}(\theta); \mu\}]$, where $r_t^{(s)}(\theta)$ is generated from a fine Euler discretization with 20 subintervals between $y_{t-1}(\theta)$ and $y_t(\theta)$.

The structural model has the exact discretization $r_t \sim \log \mathcal{N}(\beta - \sigma^2/2, \sigma^2)$. We simulate samples according to this distribution and contaminated samples that contain 5% additive outliers generated from $\log \mathcal{N}(\beta - \sigma^2/2, \sigma^2\tau^2)$, with $\tau = 5$. This particular distribution is used only to generate the data for the purpose of the simulation study, but not to compute the estimators or test statistics. This is to simulate the practical situation in which the contamination is unknown and inference is carried out assuming the structural model (12).

Our test is based on Genton & Ronchetti's (2003) robust indirect inference estimator $\hat{\theta}_R$, where the auxiliary estimator $\tilde{\mu}_R$ is defined by the M-estimator with bounded score function

$$\Psi(r_t;\mu) = \left\{\psi_c\left(\frac{r_t - 1 - \mu_1}{\mu_2}\right), \chi_c\left(\frac{r_t - 1 - \mu_1}{\mu_2}\right)\right\}^{\mathrm{T}}, \tag{14}$$

where $\psi_c(z) = \min\{c, \max(-c, z)\}$ is the Huber function, $\chi_c(z) = \psi_c^2(z) - E_\Phi(\psi_c^2)$, Φ is the cumulative function of the standard normal distribution and $c = 1{\cdot}345$. In more complex models, such simple robust indirect inference estimators may not be available. Instead, indirect robust generalized method of moments estimators can be used, as developed in Czellar et al. (2007). As an illustration, we choose parameter values $\beta = -0{\cdot}05$ and $\sigma = 0{\cdot}2$, we consider three hypotheses for the parameters: $H_0 : (\beta, \sigma)^{\mathrm{T}} = (-0{\cdot}05, 0{\cdot}2)^{\mathrm{T}}$, $H_0 : \beta = -0{\cdot}05$ and $H_0 : \sigma = 0{\cdot}2$, and we study the accuracy of $\hat{S}$ and robust $\hat{S}$ tests at the model and in the presence of contamination. Figure 1 shows p-value plots comparing classical W, robust W, $\hat{S}$ and robust $\hat{S}$ tests for sample size $n = 40$ and simulation size $m = 20$. Under no contamination, the nonrobust and robust $\hat{S}$ tests provide a more accurate inference than the corresponding W tests in all cases but one. Under contamination, the nonrobust $\hat{S}$ test is strongly oversized, even more than the classical W. On the other hand, the robust $\hat{S}$ provides very accurate inference in both contaminated and noncontaminated cases. The Supplementary Material shows the relative errors of the p-values of the χ_2^2-approximation for the test $H_0 : (\beta, \sigma)^{\mathrm{T}} = (-0{\cdot}05, 0{\cdot}2)^{\mathrm{T}}$ for sample sizes $n = 40$ and $n = 240$ without contamination. The robust $\hat{S}$ tests are the most accurate tests even in the case without contamination, and the gain is particularly important in the tails of the distributions.

Acknowledgement

The authors would like to thank the editor, two referees, Laurent Calvet, Christophe Croux, René Garcia, Christian Gouriéroux, Eric Renault, Enrique Sentana, Fallaw Sowell, Eric Zivot and seminar participants at the University of Washington, University of Geneva and Université Libre de Bruxelles for helpful comments. Elvezio Ronchetti thanks the Swiss National Science Foundation Pro*Doc Program for partial financial support.

Supplementary material

Supplementary material is available at *Biometrika* online.

References

Almudevar, A., Field, C. & Robinson, J. (2000). The density of multivariate M-estimates. *Ann. Statist.* **28**, 275–97.

Beaumont, M., Cornuet, J.-M., Marin, J.-M. & Robert, C. (2009). Adaptive approximate Bayesian computation. *Biometrika* **96**, 983–90.

Beaumont, M., Zhang, W. & Balding, D. (2002). Approximate Bayesian computation in population genetics. *Genetics* **162**, 2025–35.

Billio, M. & Monfort, A. (2003). Kernel-based indirect inference. *J. Finan. Economet.* **1**, 297–326.

Cabrera, J. & Fernholz, L. (1999). Target estimation for bias and mean square error reduction. *Ann. Statist* **27**, 1080–104.

Calzolari, G., Fiorentini, G. & Sentana, E. (2008). Indirect estimation of large conditionally heteroskedastic factor models, with an application to the Dow 30 stocks. *J. Economet.* **146**, 10–25.

Czellar, V., Karolyi, G. A. & Ronchetti, E. (2007). Indirect robust estimation of the short-term interest rate process. *J. Empirical Finan.* **14**, 546–63.

Davison, A., Hinkley, D. V. & Worton, B. (1995). Accurate and efficient construction of bootstrap likelihoods. *Statist. Comp.* **5**, 257–64.

Del Moral, P., Doucet, A. & Jasra, A. (2006). Sequential Monte Carlo samplers. *J. R. Statist. Soc.* B **68**, 411–36.

DRIDI, R., GUAY, A. & RENAULT, E. (2007). Indirect inference and calibration of dynamic stochastic general equilibrium models. *J. Economet.* **136**, 397–430.

DUFFEE, G. R. & STANTON, R. (2008). Evidence on simulation inference for near unit-root processes with implications for term structure estimation. *J. Finan. Economet.* **6**, 108–42.

DUFFIE, D. & SINGLETON, K. J. (1993). Simulated moments estimation of Markov models of asset prices. *Econometrica* **61**, 929–52.

GENTON, M. G. & DE LUNA, X. (2000). Robust simulation-based estimation. *Statist. Prob. Lett.* **48**, 253–59.

GENTON, M. G. & RONCHETTI, E. (2003). Robust indirect inference. *J. Am. Statist. Assoc.* **98**, 67–76.

GOURIÉROUX, C. & MONFORT, A. (1996). *Simulation-Based Econometric Methods*. New York: Oxford University Press.

GOURIÉROUX, C., MONFORT, A. & RENAULT, E. (1993). Indirect inference. *J. Appl. Economet.* **8**, S85–S118.

HANSEN, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica* **50**, 1029–54.

HEGGLAND, K. & FRIGESSI, A. (2004). Estimating functions in indirect inference. *J. R. Statist. Soc.* B **66**, 447–62.

JIANG, W. & TURNBULL, B. (2004). The indirect method: inference based on intermediate statistics—a synthesis and examples. *Statist. Sci.* **19**, 239–63.

JIANG, W., TURNBULL, B. & CLARK, L. C. (1999). Semiparametric regression models for repeated events with random effects and measurement error. *J. Am. Statist. Assoc.* **94**, 111–24.

KITAMURA, Y. & STUTZER, M. (1997). An information-theoretic alternative to generalized method of moments estimation. *Econometrica* **65**, 861–74.

KUK, A. Y. C. (1995). Asymptotically unbiased estimation in generalised linear models with random effects. *J. R. Statist. Soc.* B **57**, 395–407.

LEE, B. & INGRAM, B. (1991). Simulation estimation of time-series models. *J. Economet.* **47**, 197–205.

LÔ, S. & RONCHETTI, E. (2009). Robust and accurate inference for generalized linear models. *J. Mult. Anal.* **100**, 2126–36.

LOPES, S., BALDING, D. & BEAUMONT, M. (2009). PopABC: a program to infer historical demographic parameters. *Bioinformatics* **25**, 2747–9.

MEIER, L., VAN DE GEER, S. & BÜHLMANN, P. (2009). High-dimensional additive modeling. *Ann. Statist.* **37**, 3779–821.

MICHAELIDES, A. & NG, S. (2000). Estimating the rational expectations model of speculative storage: a Monte Carlo comparison of three simulation estimators. *J. Economet.* **96**, 231–66.

MONTI, A. & RONCHETTI, E. (1993). On the relationship between empirical likelihood and empirical saddlepoint approximation for multivariate M-estimators. *Biometrika* **80**, 329–38.

OWEN, A. (2001). *Empirical Likelihood*. Boca Raton: Chapman & Hall/CRC.

ROBINSON, J., RONCHETTI, E. & YOUNG, G. (2003). Saddlepoint approximations and tests based on multivariate M-estimates. *Ann. Statist.* **31**, 1154–69.

RONCHETTI, E. & WELSH, A. (1994). Empirical saddlepoint approximations for multivariate M-estimators. *J. R. Statist. Soc.* B **56**, 313–26.

SMITH, A. A. (1993). Estimating nonlinear time series models using simulated vector autoregressions. *J. Appl. Economet.* **8**, S63–S84.

SOWELL, F. B. (1996). Optimal tests for parameter instability in the generalized method of moments framework. *Econometrica* **64**, 1085–107.

TURNBULL, B. W., JIANG, W. & CLARK, L. C. (1997). Regression models for recurrent event data: parametric random effects models with measurement error. *Statist. Med.* **16**, 853–64.

[*Received June* 2009. *Revised March* 2010]