INDIRECT INFERENCE BASED ON THE SCORE

Peter Fuleky[*]and Eric Zivot[†]

June 22, 2010

Abstract

The Efficient Method of Moments (EMM) estimator popularized by Gallant and Tauchen (1996) is an indirect inference estimator based on the simulated auxiliary score evaluated at the sample estimate of the auxiliary parameters. We study an alternative estimator that uses the sample auxiliary score evaluated at the simulated binding function which maps the structural parameters of interest to the auxiliary parameters. We show that the alternative estimator has the same asymptotic properties as the EMM estimator but in finite samples behaves more like the distance-based indirect inference estimator of Gouriéroux, Monfort and Renault (1993).

Keywords: simulation based estimation, indirect inference, efficient method of moments.

JEL Codes: C13, C15, C22.

---

[*]University of Hawaii Economic Research Organization, 2424 Maile Way, Saunders Hall 540, Honolulu, HI 96822; Email: fuleky@hawaii.edu

[†]Department of Economics, Box 353330, University of Washington, Seattle, WA 98195; Email: ezivot@u.washington.edu

1 Introduction

Indirect inference estimators take advantage of a simplified auxiliary model that is easier to estimate than a proposed structural model. The estimation consists of two stages. First, an auxiliary statistic is calculated from the observed data. Then an analytical or simulated mapping of the structural parameters to the auxiliary statistic is used to calibrate an estimate of the structural parameters. The indirect inference estimators are typically placed into one of two categories: score-based estimators made popular by Gallant and Tauchen (1996), or distance-based estimators proposed by Smith (1993) and refined by Gouriéroux, Monfort and Renault (1993). The simulated score-based estimators have the computational advantage that the auxiliary parameters are estimated from the observed data only once, whereas the distance-based estimators must re-estimate the auxiliary parameters from simulated data as part of the optimization algorithm to estimate the structural parameters. However, many studies have shown (e.g., Michaelides and Ng, 2000; Ghysels, Khalaf and Vodounou, 2003; Duffee and Stanton, 2008) that the computational advantage of the simulated score-based estimators is often offset by poor finite sample properties relative to the distance-based estimators. In this paper we study an alternative score-based estimator that utilizes the sample auxiliary score evaluated with the auxiliary parameters estimated from simulated data. We show that this alternative estimator is asymptotically equivalent to the Gallant and Tauchen (1996) score-based estimator but has finite sample properties that are very close to the distance-based estimators.

The paper is structured as follows. In Section 2, we give an overview of indirect inference estimation, introduce the alternative score-based estimators and derive their asymptotic properties. In Section 3, we compare the finite sample properties of various indirect inference estimators for the parameters of a highly persistent Ornstein-Uhlenbeck process via Monte Carlo. Section 4 concludes. Proofs of all results are given in the Appendix.

2 Review of Indirect Inference

Indirect inference (II) techniques were introduced into the econometrics literature by Smith (1993), Gouriéroux, Monfort, and Renault (1993), Bansal, Gallant, Hussey, and Tauchen (1993, 1995) and Gallant and Tauchen (1996), and are surveyed in Gouriéroux and Monfort (1996) and Jiang and Turnbull (2004). There are four components present in simulation-based II: (1) a true structural model whose parameters θ are one's ultimate interest but are difficult to directly estimate; (2) simulated observations from the structural model for a given θ; (3) an auxiliary approximation to the structural model whose parameters μ are easy to estimate; and (4) a mapping from μ to θ uniquely connecting the parameters of these two models.

To be more specific, assume that a sample of n observations $\{y_t\}_{t=1,\ldots,n}$ are generated from a strictly stationary and ergodic probability model F_θ, $\theta \in \mathbb{R}^p$, with density $p(y_{-m}, \ldots, y_{-1}, y_0; \theta)$ that is difficult or impossible to evaluate analytically[1]. Typical examples are continuous time diffusion models and dynamic stochastic general equilibrium models. Define an auxiliary model $\widetilde{F}_\mu$ in which the parameter $\mu \in \mathbb{R}^r$, with $r \geq p$, is easier to estimate than θ. The auxiliary estimator of μ is defined as

$$\tilde{\mu}_n = \arg\max_{\mu} \widetilde{Q}_n \left(\{y_t\}_{t=1,\ldots,n}, \mu\right). \tag{1}$$

where $\widetilde{Q}_n$ denotes a sample objective function associated with the model $\widetilde{F}_\mu$. For ease of exposition, we consider the case in which the auxiliary estimator is the quasi-maximum likelihood estimator of the model $\widetilde{F}_\mu$, so that $\widetilde{Q}_n$ can be written as

$$\widetilde{Q}_n \left(\{y_t\}_{t=1,\ldots,n}, \mu\right) = \frac{1}{n-m} \sum_{t=m+1}^{n} \tilde{f}(y_t; x_{t-1}, \mu), \tag{2}$$

where $\tilde{f}(y_t; x_{t-1}, \mu)$ is the log density of y_t for the model $\widetilde{F}_\mu$ conditioned on $x_{t-1} = \{y_i\}_{i=t-m,\ldots,t-1}$,

---

[1]For simplicity, we do not consider structural models with additional exogenous variables z_t.

$m \in \mathbb{N}$. We define $\tilde{g}(y_t; x_{t-1}, \mu) = \frac{\partial \tilde{f}(y_t; x_{t-1}, \mu)}{\partial \mu}$ as the $r \times 1$ auxiliary score vector. For more general $\widetilde{Q}_n$, we refer the reader to Gouriéroux and Monfort (1996).

II estimators use the auxiliary model information to obtain estimates of the structural parameters θ. The link between the auxiliary model parameters and the structural parameters is given by the so-called binding function $\mu(\theta)$, which is the functional solution of the asymptotic optimization problem

$$\mu(\theta) = \arg\max_{\mu} E_{F_\theta}[\tilde{f}(y_t; x_{t-1}, \mu)], \tag{3}$$

where $\lim_{n \to \infty} \widetilde{Q}_n(\{y_t\}_{t=1,\ldots,n}, \mu) = E_{F_\theta}[\tilde{f}(y_t; x_{t-1}, \mu)]$, $\tilde{f}(y_t; x_{t-1}, \mu)$ denotes the log density of y_t given $x_{t-1} = (y_{t-1}, \ldots, y_{t-m})$ for the model $\widetilde{F}_\mu$, and $E_{F_\theta}[\cdot]$ means that the expectation is taken with respect to F_θ. In order for $\mu(\theta)$ to define a unique mapping it is assumed that $\mu(\theta)$ is one-to-one and that $\frac{\partial \mu(\theta)}{\partial \theta'}$ has full column rank.

II estimators differ in how they use (3) to define an estimating equation. The distance-based II estimator, originally proposed by Smith (1993) and Gouriéroux, Monfort, and Renault (1993), finds θ to minimize the (weighted) distance between $\mu(\theta)$ and $\tilde{\mu}_n$. The score-based II estimator, made popular by Gallant and Tauchen (1996), finds θ by solving $E_{F_\theta}[\tilde{g}(y_t; x_{t-1}, \tilde{\mu}_n)] = 0$, the first order condition associated with (3)[2]. Typically, the analytic forms of $\mu(\theta)$ and $E_{F_\theta}[\tilde{g}(y_0; x_{-1}, \mu)]$ are not known and simulation-based techniques are used to compute the two types of II estimators.

For simulation-based II, it is necessary to be able to easily generate simulated observations from F_θ for a given θ. These simulated observations are typically drawn in two ways. First, a long pseudo-data series of size $S \cdot n$ is simulated giving

$$\{y_t(\theta)\}_{t=1,\ldots,Sn}, \ S \geq 1. \tag{4}$$

---

[2]Gallant and Tauchen (1996) call the score-based II estimator the efficient method of moments (EMM) estimator. Efficiency in the context of EMM refers to the efficiency of the auxiliary model in approximating the structural model, and Gallant and Tauchen (1996, 2004) advocated the use of a particular seminonparametric auxiliary model to achieve efficiency.

Second, S pseudo-data series of size n are simulated giving

$$\{y_t^s(\theta)\}_{t=1,\dots,n}, \ s = 1, \dots, S, \ S \geq 1. \tag{5}$$

Using the simulated observations (4) or (5), the distance-based II estimators are minimum distance estimators defined as

$$\hat{\theta}_S^{Dj}(\widetilde{\Omega}_n) = \arg\min_{\theta} \ J_S^{Dj}(\theta, \widetilde{\Omega}_n) = \arg\min_{\theta} \ \left(\tilde{\mu}_n - \tilde{\mu}_S^j(\theta)\right)' \widetilde{\Omega}_n \left(\tilde{\mu}_n - \tilde{\mu}_S^j(\theta)\right), \ j = A, L, M, \tag{6}$$

where $\widetilde{\Omega}_n$ is a positive definite and symmetric weight matrix which may depend on the data through the auxiliary model, and the simulated binding functions are given by

$$\tilde{\mu}_S^A(\theta) = \arg\max_{\mu} \ S^{-1} \sum_{s=1}^{S} \widetilde{Q}_n\left(\{y_t^s(\theta)\}_{t=1,\dots,n}, \mu\right), \tag{7}$$

$$\tilde{\mu}_S^L(\theta) = \arg\max_{\mu} \ \widetilde{Q}_{Sn}\left(\{y_t(\theta)\}_{t=1,\dots,Sn}, \mu\right), \tag{8}$$

$$\tilde{\mu}_S^M(\theta) = S^{-1} \sum_{s=1}^{S} \arg\max_{\mu} \ \widetilde{Q}_n\left(\{y_t^s(\theta)\}_{t=1,\dots,n}, \mu\right). \tag{9}$$

The superscripts A, L, and M indicate how the binding function is computed from the simulated data: "A" denotes maximizing an aggregation of S objective functions using (5); "L" denotes use of long simulations (4) in the objective function; "M" denotes use of the mean of S estimated binding functions based on (5). The M-type estimator is most commonly used in practice and is more computationally intensive than the A and L-type estimators, but exhibits superior finite sample properties in certain circumstances.

Using (4) or (5), the score-based II estimators are one-step GMM estimators defined as

$$\hat{\theta}_S^{Sj}(\widetilde{\Sigma}_n) = \arg\min_{\theta} \ J_S^{Sj}(\theta) = \arg\min_{\theta} \tilde{g}_S^j(\theta, \tilde{\mu}_n)' \widetilde{\Sigma}_n \tilde{g}_S^j(\theta, \tilde{\mu}_n), \ j = A, L, \tag{10}$$

where $\widetilde{\Sigma}_n$ is a positive definite (pd) and symmetric weight matrix which may depend on the

data through the auxiliary model, and the simulated scores are given by

$$\tilde{g}_S^A(\theta, \tilde{\mu}_n) = S^{-1} \sum_{s=1}^{S} \frac{1}{n-m} \sum_{t=m+1}^{n} \tilde{g}(y_t^s(\theta); x_{t-1}^s(\theta), \tilde{\mu}_n),$$

$$\tilde{g}_S^L(\theta, \tilde{\mu}_n) = \frac{1}{Sn-m} \sum_{t=m+1}^{Sn} \tilde{g}(y_t(\theta); x_{t-1}(\theta), \tilde{\mu}_n).$$

Because (10) fixes the binding function at the sample estimate $\tilde{\mu}_n$ no M-type estimator is available. The implementation of (10) in Gallant and Tauchen (2004) is based on the L-type estimator.

Under regularity conditions described in Gouriéroux and Monfort (1996), the distance-based estimators (6) and score-based estimators (10) are consistent for θ_0 (true parameter vector) and asymptotically normally distributed. The limiting weight matrices that minimize the asymptotic variances of these estimators are $\tilde{\Omega}^* = M_\mu \tilde{\mathcal{I}}^{-1} M_\mu$ and $\tilde{\Sigma}^* = \tilde{\mathcal{I}}^{-1}$, where $\tilde{\mathcal{I}} = \lim_{n \to \infty} \mathrm{var}_{F_\theta}(\sqrt{n} \tilde{g}_n(y_n, \mu(\theta_0)))$, $M_\mu = E_{F_\theta}[\tilde{H}(y_t; x_{t-1}, \mu(\theta_0))]$, $\tilde{g}_n(y_n, \mu(\theta)) = \frac{1}{n-m} \sum_{t=m+1}^{n} \tilde{g}(y_t; x_{t-1}, \mu(\theta))$ and $\tilde{H}(y_t; x_{t-1}, \mu) = \frac{\partial^2 \tilde{f}(y_t; x_{t-1}, \mu)}{\partial \mu \partial \mu'}$. Using consistent estimates of these optimal weight matrices, the distance-based and score-based estimators are asymptotically equivalent with asymptotic variance matrix given by

$$V_S^* = \left(1 + \frac{1}{S}\right) \left(M_\theta' \tilde{\mathcal{I}}^{-1} M_\theta\right)^{-1} = \left(1 + \frac{1}{S}\right) \left(\frac{\partial \mu(\theta_0)'}{\partial \theta'} M_\mu \tilde{\mathcal{I}}^{-1} M_\mu \frac{\partial \mu(\theta_0)}{\partial \theta'}\right)^{-1}, \quad (11)$$

where

$$M_\theta = \left\{ \frac{\partial}{\partial \theta'} E_{F_\theta}[\tilde{g}(y_t; x_{t-1}, \mu)] \right\} \Big|_{\mu=\mu(\theta_0)}.$$

3   Alternative Score-Based II Estimator

Gouriéroux and Monfort (1996, pg. 71) mentioned two alternative II estimators that they claimed are less efficient than the optimal estimators described in the previous section, and referred the reader to Smith (1993) for details. The first one is the simulated quasi-maximum

likelihood (SQML) estimator:

$$\hat{\theta}_S^{\text{SQMLj}} = \arg\max_\theta \widetilde{Q}_n \left(\{y_t\}_{t=1,\ldots,n}, \tilde{\mu}_S^{\text{j}}(\theta) \right), \text{j} = \text{A}, \text{L}, \text{M}. \tag{12}$$

The second one is an alternative score-based estimator of the form

$$\hat{\theta}_S^{\text{Sj2}}(\widetilde{\Sigma}_n) = \arg\min_\theta J^{\text{Sj2}}(\theta, \widetilde{\Sigma}_n) = \arg\min_\theta \ \tilde{g}_n^{\text{j}}(\theta)'\widetilde{\Sigma}_n \tilde{g}_n^{\text{j}}(\theta), \tag{13}$$

where

$$\tilde{g}_n^{\text{j}}(\theta) = \frac{1}{n-m} \sum_{t=m+1}^n \tilde{g}(y_t; x_{t-1}, \tilde{\mu}_S^{\text{j}}(\theta)), \text{j} = \text{A}, \text{L}, \text{M}. \tag{14}$$

In contrast to the score-based estimator (10), the estimator (13) evaluates the auxiliary score with the sample data and a simulated binding function. In this respect, it is more like the distance-based II estimators (6). The estimator (13), however, was not explicitly considered in Smith (1993).

Smith (1993) showed that (12) is consistent and asymptotically normal with asymptotic variance matrix given by

$$V_S^{SQML} = \left(1 + \frac{1}{S}\right) \left[\frac{\partial\mu(\theta_0)'}{\partial\theta'} M_\mu \frac{\partial\mu(\theta_0)}{\partial\theta'}\right]^{-1} \frac{\partial\mu(\theta_0)'}{\partial\theta'} \widetilde{\mathcal{I}} \frac{\partial\mu(\theta_0)}{\partial\theta'} \left[\frac{\partial\mu(\theta_0)'}{\partial\theta'} M_\mu \frac{\partial\mu(\theta_0)}{\partial\theta'}\right]^{-1},$$

which he showed is strictly greater than (in a matrix sense) the asymptotic variance (11) of the efficient II estimators. He did not derive the asymptotic properties of (13). The following Proposition gives the asymptotic properties of (13).

Proposition 1 *Under the regularity conditions in Gouriéroux and Monfort (1996), the score-based II estimators $\hat{\theta}_S^{Sj2}(\widetilde{\Sigma}_n)$ (j=A,L,M) defined in (13) are consistent and asymptotically normal, when S is fixed and $n \to \infty$:*

$$\sqrt{n}(\hat{\theta}_S^{Sj2}(\widetilde{\Sigma}_n) - \theta_0) \xrightarrow{d} N \left(0, \left(1 + \frac{1}{S}\right) [M_\theta'\Sigma M_\theta]^{-1} \left[M_\theta'\Sigma\widetilde{\mathcal{I}}\Sigma M_\theta\right] [M_\theta'\Sigma M_\theta]^{-1}\right). \tag{15}$$

The proof is given in Appendix B. We make the following remarks:

1. The asymptotic variance of $\hat{\theta}_S^{\text{Sj2}}(\widetilde{\Sigma}_n)$ in (15) is equivalent to the asymptotic variance of Gallant and Tauchen's score based estimator $\hat{\theta}_S^{\text{Sj}}(\widetilde{\Sigma}_n)$, and is equivalent to (11) when $\widetilde{\Sigma}_n$ is a consistent estimate of $\widetilde{\Sigma}^* = \widetilde{\mathcal{I}}^{-1}$. Contrary to the claim in Gouriéroux and Monfort (1996), the alternative score based II estimator is not less efficient than the optimal II estimators.

2. Using the result $\frac{\partial \mu(\theta_0)}{\partial \theta'} = M_\mu^{-1} M_\theta$ (see Gouriéroux and Monfort, 1996 pg. 70) , Smith's SQML estimator (12) is asymptotically equivalent to $\hat{\theta}_S^{\text{Sj2}}(\widetilde{\Sigma}_n)$ when $\widetilde{\Sigma}_n$ is a consistent estimator of M_μ^{-1}. Hence, the SQML is efficient only when $\widetilde{\mathcal{I}} = -M_\mu$.

4   Finite Sample Comparison of II Estimators

In this section, we use Monte Carlo methods to compare the finite sample performance of the alternative score-based estimator (13) to the traditional II estimators (6) and (10) using a simple continuous-time Ornstein-Uhlenbeck (OU) process. Our Monte Carlo design is motivated by Duffee and Stanton (2008) (hereafter, DS). They compared the finite sample properties of EMM and II using highly persistent AR(1) models calibrated to interest rate data and found that EMM is severely biased, has wide confidence intervals, and performs poorly in coefficient and overidentification tests. The OU process we use is the continuous-time analogue of the discrete time AR(1) model, and we calibrate our design so that our results are comparable to those of DS. The analytically tractable OU process also gives us the opportunity to compute non-simulation-based analogues of the simulation-based estimators, and to directly compare the performance of the II estimators to the benchmark maximum likelihood (ML) estimator.

4.1 Model Setup

Assume that the true data generating process is an OU process of the form

$$F_\theta : y = (\theta_0 - \theta_1 y)dt + \theta_2 dW, \quad dW \sim \text{iid N}(0, dt) , \tag{16}$$

where $\theta_0/\theta_1 > 0$ represents the long run (unconditional) mean, $\theta_1 > 0$ captures the speed of mean reversion, and $\theta_2 > 0$ gives the constant volatility of the process. Weekly observations ($\Delta = 1/50$) of annualized interest rates are generated from its exact solution

$$y_t = \frac{\theta_0}{\theta_1}(1 - e^{-\theta_1 \Delta}) + e^{-\theta_1 \Delta} y_{t-\Delta} + \theta_2 \sqrt{\frac{1 - e^{-2\theta_1 \Delta}}{2\theta_1}} \epsilon_t, \quad \epsilon_t \sim \text{iid N}(0, 1) , \tag{17}$$

for the set of true parameters $\theta = (0, 0.6644, 7.1181)'$, which corresponds to a transformation of the AR(1) parameterization of DS (see Appendix A). The parameters θ_0 and θ_2 are assumed to be known by the researcher as previous research, summarized in Phillips and Yu (2009), has shown that these parameters can be very accurately estimated. They are fixed at their true values in the structural model, and only θ_1 is estimated.

A natural auxiliary model is the crude Euler discretization of the OU process

$$\widetilde{F}_\mu : y_t = \mu_0 \Delta + (1 - \mu_1 \Delta)y_{t-\Delta} + \mu_2 \sqrt{\Delta}\xi_{t-\Delta}, \quad \xi_{t-\Delta} \sim \text{iid N}(0, 1) , \tag{18}$$

where all three parameters, μ_0, μ_1, μ_2 are estimated. Thus, θ_1 is over-identified ($r > p$) which gives us the opportunity to evaluate the performance of the estimators in over-identification tests.

For the discretized OU process, the log-likelihood and score vector are given by

$$\tilde{f}_n\left(\{y_t\}_{t=\Delta,\dots,n\Delta},\mu\right) = \frac{1}{n-1}\sum_{t=2\Delta}^{n\Delta}\left[-\frac{1}{2}\ln(2\pi\mu_2^2\Delta) - \frac{1}{2}\frac{\xi_t^2}{\mu_2^2\Delta}\right], \tag{19}$$

$$\tilde{g}(y_t;y_{t-\Delta},\mu) = \frac{\partial\tilde{f}(y_t;y_{t-\Delta},\mu)}{\partial\mu} = \begin{pmatrix} \dfrac{1}{\mu_2^2}\xi_t \\[2mm] -\dfrac{1}{\mu_2^2}\xi_t y_{t-\Delta} \\[2mm] \dfrac{1}{\mu_2^2}\left(\dfrac{\xi_t^2}{\mu_2^2\Delta}-1\right) \end{pmatrix}, \tag{20}$$

where $\xi_t = y_t - \mu_0\Delta - (1-\mu_1\Delta)y_{t-\Delta}$. The unconstrained auxiliary estimator $\tilde{\mu}_n$ found by maximizing (19) is the least squares estimator.

Comparing (17) with (18) shows that $\tilde{\mu}_n$ is a biased estimator of θ, and that the binding function $\mu(\theta)$ has the form[3]

$$\mu_0(\theta) = \frac{\theta_0}{\theta_1\Delta}\left(1-e^{-\theta_1\Delta}\right), \ \ \mu_1(\theta) = \frac{1}{\Delta}\left(1-e^{-\theta_1\Delta}\right), \ \ \mu_2(\theta) = \theta_2\sqrt{\frac{1-e^{-\theta_1\Delta}}{2\theta_1\Delta}}. \tag{21}$$

Given (20) and that the expected score vector $E_{F_\theta}[\tilde{g}(y_t;y_{t-\Delta},\mu)]$ has a closed form expression, non-simulation-based versions of distance-based and score-based II estimators are available. We denote these estimators IN, SN1 and SN2, respectively.

For the Monte Carlo analysis, $n = 1000$ observations with $\Delta = 1/50$ are simulated from (17) with $\theta = (0, 0.6644, 7.1181)'$ and are treated as the observed sample in each Monte Carlo run. For the simulations (4) and (5) used to compute the simulation-based II estimators, we set $S = 20$ so that the simulation-based estimators have a 95% asymptotic efficiency relative to the non-simulation-based estimators (see (11)), and use the same random number seed for all values of θ during the optimizations. When simulating from (17), the stability constraint $\theta_1 > 0$ is imposed and simulations are started from the long run mean of the process θ_0/θ_1. When estimating the auxiliary model parameters, the stability constraint $\mu_1 > 0$ is imposed.

---

[3]The bias associated with estimating the crude Euler approximation was first shown by Lo (1988) and is called the discretization bias.

4.2 Objective Functions and Confidence Intervals

Figure 1 illustrates the LR-type statistics for testing $H_0 : \theta_1 = \theta_1^0$ as functions of θ_1^0 for the II estimators based on a single representative sample. The widths of 95% confidence intervals are obtained by inverting the LR statistics. The 95% confidence intervals contain values of θ_1^0 such that the value of the LR statistic lies below the 95% quantile of the chi-square distribution with 1 degree of freedom.

Table 1 summarizes the point estimates and confidence intervals for each of the estimators. The distance-based (D) II estimates and the alternative score-based (S2) estimates are very similar, and are much smaller than the Gallant-Tauchen score-based (S1) estimates. In general, for each estimator, the A-type and L-type simulation-based estimates are close to the non-simulation based N-type estimate. For the D and S2 estimators, the M-type estimates are slightly smaller than the other estimates.

Figure 1 shows that the shapes of the LR statistics for the non-simulation-based and simulation-based estimators are very similar. The LR statistics for the two score-based estimators, however, have very different shapes. As noted by DS, the shape of LR^{S1} is highly asymmetric due to the scaling of some sample moments by the population variance. It is relatively flat for θ_1^0 values above $\hat{\theta}_1^{\text{S1}}$ because the population variance approaches zero as the mean reversion of the process increases. However, it peaks sharply as θ_1^0 approaches zero because the population variance diverges to infinity as the process becomes more persistent. In contrast, the shapes of the LR functions for the S2 and D estimators are almost identical and are roughly symmetric in θ_1^0. This occurs because they are scaled by the variance of the observed sample which is constant for any θ_1^0.

The S1 confidence interval covers a wide range of θ_1 above the point estimate, but only little of the range below the point estimate, and it does not contain the true value $\theta_1 = 0.6644$. While the shape of the S1 criterion function puts a high penalty for θ_1 close to the boundary of stationarity, it causes point estimates above the true value to be rejected with high probability. The point estimates differ across estimators using the same amount

of information, and similar to the results reported in DS, LR^{S1} has its minimum the farthest away from the true value $\theta_1 = 0.6644$. The M-type LR-type statistics are shifted toward the true value reflecting the different finite sample properties of the M-type estimators in comparison to the N, L and A-type ones.

4.3   Computational Issues

The S2 estimator can be considered a hybrid estimator consisting of two steps. In the first step the simulation-based binding function $\tilde{\mu}_S(\theta)$ is calculated. In distance-based II this simulated binding function is directly compared to the auxiliary estimate $\tilde{\mu}$. In the S2 estimator the mean score evaluated with $\tilde{\mu}_S(\theta)$ is compared to the mean score evaluated with $\tilde{\mu}$, where the latter is equal to zero by construction. Because the score function is evaluated with the observed data, a fixed input, all the variability of the S2 objective function can be attributed to the simulated binding function $\tilde{\mu}_S(\theta)$, just like in the case of the D estimators' objective function. Therefore the objective functions of the simulation-based S2 and D estimators will also look similar. [4]

Gallant and Tauchen (2002) criticize distance-based II for its computational inefficiency, because it potentially involves two nested optimizations: the estimator of the simulated binding function is embedded within the D estimator. This may lead to numerical instability if the auxiliary estimator does not have a closed form analytical expression but instead relies on an optimizer. The inner (binding function) optimization, which is computed within a tolerance, will cause some jitter, and render the outer (structural) optimization problem non-smooth. However, if a simple auxiliary model is chosen such that the auxiliary estimator has a closed form analytical solution, the speed and stability of the D estimator becomes much improved. Because the S2 estimator also uses the simulated binding function, similar issues have to be weighed when the auxiliary model for S2 is chosen. Interestingly, Table 2 indicates that the average computation time associated with the SL1 and SA1 estimators

---

[4]The shape of the objective function is equivalent to the shape of the LR statistic except for a level shift.

are actually slightly higher than the L and A-type estimators that use a simulated binding function. This occurs because of the irregular shapes of the SL1 and SA1 objective functions.

4.4 Bias and RMSE

Figure 2 and Table 3 summarize the empirical distributions of the ML and II estimators of θ_1. The distributions are based on 1000 Monte Carlo simulations. Gouriéroux and Monfort (1996, pg. 66) note that the score-based and distance-based II estimates should be very close in a just identified setting. However, Figure 2 shows that the distributions of these estimators in an over-identified setting can be very different. The S1 estimators are extremely biased (seven times that of the MLE), a confirmation of DS's finding. In contrast, the corresponding S2 and II estimators (N, L and A-type) have a slightly lower bias than the MLE, and their distributions closely resemble that of the MLE.

While the ML and II estimators are not subject to the discretization bias [5], they are affected by a finite sample bias (Ball and Torous, 1996; Phillips and Yu, 2009) due to the highly persistent nature of the adopted parameterization of the OU process. The M-type estimator has been shown to correct this finite sample bias in a just-identified setting (Gouriéroux, Renault, and Touzi, 2006; Gouriéroux, Phillips, and Yu, 2008, Phillips and Yu, 2009), but the results of Table 3 shows this not the case in an over-identified setting. Whereas the N, A and L-type estimators show a positive bias, the M-type estimator shows a negative finite sample bias. [6]

---

[5]MLE is based on the transition density of the continuous time structural model, and the indirect estimators correct the discretization bias of the auxiliary estimator.

[6]In a just identified setting where the θ_0 and θ_2 parameters are assumed to be unknown, and are being estimated along with θ_1, the mean estimate, bias and RMSE of $\hat{\theta}_1^{EM2}$ is 0.6813, (0.0169), and [0.3662] respectively, and the mean estimate, bias and RMSE of $\hat{\theta}_1^{IM}$ is 0.6810, (0.0167), and [0.3656] respectively. These results represent a 90% reduction in bias compared to MLE and confirm the finite sample bias correcting properties of the M-type estimators in just identified models.

4.5 Test Statistics

Table 4 shows the empirical rejection rates of nominal 5% overidentification tests and LR-type coefficient tests of $\theta_1 = \theta_1^0$ based on 1000 Monte Carlo simulations. The tests, especially the LR-type coefficient tests, based on the S1 estimators are extremely oversized. The extreme right skewness in the finite sample distributions of the S1 estimators combined with the asymmetry of the S1 objective functions, causes a high rejection rate of the LR-type coefficient tests. In contrast, the rejection rates of the other estimators are approximately equal and closer to the nominal level. Fuleky (2009) shows that the higher rejection rates of the LR-type tests based the M-type estimators is caused by the over-identification restrictions in conjunction with the nonlinearity of the binding function in small samples. [7]

5   Conclusion

In this paper we study the asymptotic and finite sample properties of an alternative score-based II estimator that uses the sample auxiliary score evaluated at the simulated binding function. We show that this estimator is asymptotically equivalent to Gallant and Tauchen's simulated score estimator, but in finite samples behaves much more like the distance-based II estimators. For estimating the mean reversion parameter of a highly persistent OU process, we show that the alternative score-based estimator does not exhibit the poor finite sample properties of the simulated score estimator. Our results counter some of the criticisms of the score-based II estimators raised by Duffee and Stanton (2008).

---

[7]In a just identified setting where the θ_0 and θ_2 parameters are assumed to be unknown, and are being estimated along with θ_1, the empirical size of the simple LR-type test for testing $H_0 : \theta_1 = 0.6644$ is 0.079 for EM2 and 0.083 for IM respectively. These results indicate that the M-type estimators have improved inference properties in just identified models.

References

Ball, C. and Torous, W. (1996). Unit roots and the estimation of interest rate dynamics. *Journal of Empirical Finance*, 3(2):215-238.

Duffee, G. and Stanton, R. (2008). Evidence on Simulation Inference for Near Unit-Root Processes with Implications for Term Structure Estimation. *Journal of Financial Econometrics*, 6(1):108-150.

Fuleky, P. (2009). Simulation-based estimation of continuous-time models. Phd Thesis, Department of Economics, University of Washington.

Gallant, A. and Tauchen, G. (1996). Which Moments to Match? *Econometric Theory*, 12(4):657-81.

Gallant, A. and Tauchen, G. (1997). Estimation of Continuous-Time Models for Stock Returns and Interest Rates. *Macroeconomic Dynamics*, 1(1):135-68.

Gallant, A. and Tauchen, G. (2002). Simulated Score Methods and Indirect Inference for Continuous Time Models. *Handbook of Financial Econometrics*, North Holland, Amsterdam.

Gallant, A. and Tauchen, G. (2004). EMM: A Program for Effcient Method ofMoments Estimation. Unpublished manuscript, Department of Economics, Duke University.

Ghysels, E., Khalaf, L. and Vodounou, C. (2003), "Simulation Based Inference in Moving Average Models", *Annales D'Économie et de Statistique*, **69**, 85-99.

Gouriéroux, C. and Monfort, A. (1996). Simulation-Based Econometric Methods. Oxford University Press, USA.

Gouriéroux, C., Monfort, A., and Renault, E. (1993). Indirect Inference. *Journal of Applied Econometrics*, 8:S85-S118.

Gouriéroux, C., Phillips, P., and Yu, J. (2008). Indirect Inference for Dynamic Panel Models. *Journal of Econometrics.*

Gouriéroux, C., Renault, E., and Touzi, N. (2000). Calibration by simulation for small sample bias correction. *Simulation-Based Inference in Econometrics, Methods and Applications.* Cambridge University Press, Cambridge.

Jiang, W. and Turnbull, B. (2004). The Indirect Method: Inference Based on Intermediate Statistics|A Synthesis and Examples. *Statistical Science*, 19(2):239-263.

Lo, A. (1988). Maximum Likelihood Estimation of a Generalized Ito Process with Discretely Sampled Data. *Econometric Theory*, 4(2):231-247.

Michaelides, A. and Ng, S. (2000), "Estimating the Rational Expectations Model of Speculative Storage: A Monte Carlo Comparison of Three Simulation Estimators", *Journal of Econometrics*, **96**, 231-266.

Phillips, P. and Yu, J. (2007). Maximum Likelihood and Gaussian Estimation of Continuous Time Models in Finance. *Handbook of Financial Time Series*, North-Holland.

Smith Jr, A. (1993). Estimating Nonlinear Time-Series Models Using Simulated Vector Autoregressions. *Journal of Applied Econometrics*, 8:S63-S84.

6 Appendix A: OU Parameterization

The parameters for the OU structural model (16) are derived from the following AR(1) parameterization used by DS

$$y_t = \theta_0^{DS} + \theta_1^{DS} y_{t-\Delta} + \theta_2^{DS} \epsilon_t, \quad \epsilon_t \sim \text{iid } N(0,1). \tag{22}$$

DS assumed weekly observations, used $\Delta = 1$ to represent the observation interval and set $\theta_0^{DS} = 0, \theta_1^{DS} = 0.9868$, and $\theta_2^{DS} = 1$. The AR(1) parameters in (22) can be mapped to the OU parameters in (16) by inverting the following correspondence between (17) and (22)

$$\theta_0^{DS} = \frac{\theta_0}{\theta_1}(1 - e^{-\theta_1 \Delta}), \quad \theta_1^{DS} = e^{-\theta_1 \Delta}, \quad \theta_2^{DS} = \theta_2 \sqrt{\frac{1 - e^{-2\theta_1 \Delta}}{2\theta_1}}.$$

Thus, we obtain the second set of parameters from the transformation: $\theta_1 = -\Delta^{-1} \log \theta_1^{DS} = -50 \log 0.9868 = 0.6644$, $\theta_0 = \theta_0^{DS} \theta_1 / (1 - \theta_1^{DS}) = 0 \cdot 0.6644 / (1 - 0.9868) = 0$, $\theta_2 = \theta_2^{DS} \sqrt{2\theta_1 / (1 - (\theta_1^{DS})^2)} = 1\sqrt{2 \cdot 0.6644 / (1 - 0.9868^2)} = 7.1181$. Here, θ_1 can be interpreted as the annualized mean reversion toward the long run mean of zero, and θ_2 as the annualized volatility of the OU process. The value $\theta_1 = 0.6644$ implies that the half-life of a shock to interest rates is approximately one year. We consider a time horizon of 20 years, which corresponds to 1000 observations.

7 Appendix B: Proof of Proposition 1

The regularity conditions from Gouriéroux and Monfort (1996, Appendix 4A) are:

(A1) $\tilde{f}_n(y_n, \mu) = \frac{1}{n-m} \sum_{t=m+1}^{n} \tilde{f}(y_t; x_{t-1}, \mu) \xrightarrow{p} \tilde{f}_E(\theta, \mu) = E_{F_\theta}[\tilde{f}(y_t; x_{t-1}, \mu)]$ uniformly in (θ, μ) as $n \to \infty$.

(A2) $\tilde{f}_E(\theta, \mu)$ has a unique maximum with respect to μ : $\mu(\theta) = \arg\max_\mu \tilde{f}_E(\theta, \mu)$.

(A3) $\tilde{f}_n(y_n, \mu)$ and $\tilde{f}_E(\theta, \mu)$ are differentiable with respect to μ, and $\tilde{g}_E(\theta, \mu) = \frac{\partial \tilde{f}_E(\theta, \mu)}{\partial \mu}$
$= \lim_{n \to \infty} \frac{\partial \tilde{f}_n(y_n, \mu)}{\partial \mu}$.

(A4) The only solution to the asymptotic first order conditions is $\mu(\theta) : \tilde{g}_E(\theta, \mu) = 0 \Rightarrow \mu = \mu(\theta)$.

(A5) The equation $\mu = \mu(\theta)$ admits a unique solution in θ.

(A6) $p \lim_{n \to \infty} \frac{\partial^2 \tilde{f}_n(y_n, \mu(\theta))}{\partial \mu \partial \mu'} = E_{F_\theta}[\widetilde{H}(y_t; x_{t-1}, \mu(\theta_0))] = M_\mu$

(A7) $\sqrt{n}\tilde{g}_n(y_n, \mu(\theta_0)) = \sqrt{n}\frac{\partial \tilde{f}_n(y_n, \mu(\theta_0))}{\partial \mu} \xrightarrow{d} N(0, \widetilde{\mathcal{I}})$ as $n \to \infty$.

For ease of exposition, we only give the proof for $\hat{\theta}_S^{\mathrm{SL2}}(\widetilde{\Sigma}_n) = \hat{\theta}_S^{\mathrm{L}}$ which follows closely the proof from Gouriéroux and Monfort (1996, Appendix 4A). The results for the other estimators are similar. For consistency, first note that for fixed S and as $n \to \infty$

$$\tilde{g}_n(y_n, \mu(\theta)) \xrightarrow{p} \tilde{g}_E(\theta_0, \mu(\theta)),$$

$$\tilde{\mu}_S^{\mathrm{L}}(\theta) = \arg\max_\mu \tilde{f}_{Sn}(y_{Sn}(\theta), \mu) \xrightarrow{p} \arg\max_\mu S\tilde{f}_E(\theta, \mu) = \mu(\theta).$$

Then $\hat{\theta}_S^{\mathrm{L}} \xrightarrow{p} \arg\min_\theta \tilde{g}_E(\theta_0, \mu(\theta))'\Sigma\tilde{g}_E(\theta_0, \mu(\theta))$ which, by A4, is uniquely minimized at $\theta = \theta_0$. Hence, $\hat{\theta}_S^{\mathrm{L}} \xrightarrow{p} \theta_0$.

For asymptotic normality, the first order condition of the optimization problem in (13) is

$$\frac{\partial \tilde{g}_n(y_n, \tilde{\mu}_S^{\mathrm{L}}(\hat{\theta}_S^{\mathrm{L}}))'}{\partial \theta'}\widetilde{\Sigma}_n\tilde{g}_n(y_n, \tilde{\mu}_S^{\mathrm{L}}(\hat{\theta}_S^{\mathrm{L}})) = 0. \tag{23}$$

Taking a mean value expansion (MVE) of $\tilde{g}_n(y_n, \tilde{\mu}_S^{\mathrm{L}}(\hat{\theta}_S^{\mathrm{L}}))$ around θ_0 and plugging it into (23) gives

$$\frac{\partial \tilde{g}_n(y_n, \tilde{\mu}_S^{\mathrm{L}}(\hat{\theta}_S^{\mathrm{L}}))'}{\partial \theta'}\widetilde{\Sigma}_n \left[\tilde{g}_n(y_n, \tilde{\mu}_S^{\mathrm{L}}(\theta_0)) + \frac{\partial \tilde{g}_n(y_n, \tilde{\mu}_S^{\mathrm{L}}(\bar{\theta}))}{\partial \mu'}\frac{\partial \tilde{\mu}_S^{\mathrm{L}}(\bar{\theta})}{\partial \theta'}(\hat{\theta}_S^{\mathrm{L}} - \theta_0) \right] = 0 , \tag{24}$$

where $\bar{\theta}$ represents the vector of intermediate values. Using the results

$$\frac{\partial \tilde{g}_n(y_n, \tilde{\mu}_S^{\mathrm{L}}(\hat{\theta}_S^{\mathrm{L}}))'}{\partial \theta'} = \frac{\partial \tilde{\mu}_S^{\mathrm{L}}(\hat{\theta}_S^{\mathrm{L}}))'}{\partial \theta'} \frac{\partial \tilde{g}_n(y_n, \tilde{\mu}_S^{\mathrm{L}}(\hat{\theta}_S^{\mathrm{L}}))'}{\partial \mu'} \xrightarrow{p} \frac{\partial \mu(\theta_0)'}{\partial \theta'} \frac{\partial \tilde{g}_E(\theta_0, \mu(\theta_0))'}{\partial \mu'} = M_\theta' \ ,$$

$$\frac{\partial \tilde{g}_n(y_n, \tilde{\mu}_S^{\mathrm{L}}(\bar{\theta}))}{\partial \mu'} \frac{\partial \tilde{\mu}_S^{\mathrm{L}}(\bar{\theta}))}{\partial \theta'} \xrightarrow{p} \frac{\partial \tilde{g}_E(\theta_0, \mu(\theta_0))}{\partial \mu'} \frac{\partial \mu(\theta_0)}{\partial \theta'} = M_\theta,$$

and re-arranging (24) then gives

$$\sqrt{n}(\hat{\theta}_S^{\mathrm{L}} - \theta_0) = -\left[M_\theta' \Sigma M_\theta\right]^{-1} M_\theta' \Sigma \sqrt{n} \tilde{g}_n(y_n, \tilde{\mu}_S^{\mathrm{L}}(\theta_0)) + o_p(1). \tag{25}$$

Next, use a MVE of $\tilde{g}_n(y_n, \tilde{\mu}_S^{\mathrm{L}}(\theta_0))$ around $\tilde{\mu}$ to give

$$\sqrt{n} \tilde{g}_n(y_n, \tilde{\mu}_S^{\mathrm{L}}(\theta_0)) = \sqrt{n} \tilde{g}_n(y_n, \tilde{\mu}) + \frac{\partial \tilde{g}_n(y_n, \bar{\mu})}{\partial \mu'} \sqrt{n}(\tilde{\mu}_S^{\mathrm{L}}(\theta_0) - \tilde{\mu}) \tag{26}$$

$$= \sqrt{n} \tilde{g}_n(y_n, \tilde{\mu}) + M_\mu \sqrt{n}(\tilde{\mu}_S^{\mathrm{L}}(\theta_0) - \tilde{\mu}) + o_p(1),$$

and another MVE of $\tilde{g}_n(y_n, \tilde{\mu}) = 0$ around $\mu(\theta_0)$ to give

$$\sqrt{n} \tilde{g}_n(y_n, \tilde{\mu}) = \sqrt{n} \tilde{g}_n(y_n, \mu(\theta_0)) + \frac{\partial \tilde{g}_n(y_n, \bar{\bar{\mu}})}{\partial \mu'} \sqrt{n}(\tilde{\mu} - \mu(\theta_0)) = 0,$$

so that

$$\sqrt{n}(\tilde{\mu} - \mu(\theta_0)) = -M_\mu^{-1} \sqrt{n} \tilde{g}_n(y_n, \mu(\theta_0)) + o_p(1). \tag{27}$$

In addition, use a MVE of the simulated score $\tilde{g}_{Sn}(y_{Sn}(\theta_0), \tilde{\mu}_S^{\mathrm{L}}(\theta_0))$ around $\mu(\theta_0)$

$$\sqrt{n} \tilde{g}_{Sn}(y_{Sn}(\theta_0), \tilde{\mu}_S^{\mathrm{L}}(\theta_0)) = \sqrt{n} \tilde{g}_{Sn}(y_{Sn}(\theta_0), \mu(\theta_0)) + \frac{\partial \tilde{g}_{Sn}(y_{Sn}(\theta_0), \bar{\bar{\mu}})}{\partial \mu'} \sqrt{n}(\tilde{\mu}_S^{\mathrm{L}}(\theta_0) - \mu(\theta_0)) = 0,$$

so that

$$\sqrt{n}(\tilde{\mu}_S^{\mathrm{L}}(\theta_0) - \mu(\theta_0)) = -\left[\frac{\partial \tilde{g}_{Sn}(y_{Sn}(\theta_0), \bar{\bar{\mu}})}{\partial \mu'}\right]^{-1} \sqrt{n}\tilde{g}_{Sn}(y_{Sn}(\theta_0), \mu(\theta_0)) \qquad (28)$$

$$= -S^{-1}M_\mu^{-1}\sqrt{n}\sum_{s=1}^{S}\tilde{g}_n(y_n^s(\theta_0), \mu(\theta_0)) + o_p(1),$$

since $\tilde{g}_{Sn}(y_{Sn}(\theta_0), \mu(\theta_0)) = \sum_{s=1}^{S}\tilde{g}_n(y_n^s(\theta_0), \mu(\theta_0))$ and so

$$\frac{\partial \tilde{g}_{Sn}(y_{Sn}(\theta_0), \bar{\bar{\mu}})}{\partial \mu'} = \sum_{s=1}^{S}\frac{\partial \tilde{g}_n(y_n^s(\theta_0), \bar{\bar{\mu}})}{\partial \mu'} \xrightarrow{p} S \cdot M_\mu.$$

By subtracting (27) from (28) we get

$$\sqrt{n}(\tilde{\mu}_S^{\mathrm{L}}(\theta_0) - \tilde{\mu}) = M_\mu^{-1}\sqrt{n}\left[\tilde{g}_n(y_n, \mu(\theta_0)) - S^{-1}\sum_{s=1}^{S}\tilde{g}_n(y_n^s(\theta_0), \mu(\theta_0))\right]. \qquad (29)$$

Using (29) and $\tilde{g}_n(y_n, \tilde{\mu}) = 0$, (26) can be rewritten as

$$\sqrt{n}\tilde{g}_n(y_n, \tilde{\mu}_S^{\mathrm{L}}(\theta_0)) = \sqrt{n}\left[\tilde{g}_n(y_n, \mu(\theta_0)) - S^{-1}\sum_{s=1}^{S}\tilde{g}_n(y_n^s(\theta_0), \mu(\theta_0))\right], \qquad (30)$$

Because y_n and $y_n^s(\theta_0)$ $(s = 1, \ldots, S)$ are independent it follows that

$$\mathrm{AsyVar}[\sqrt{n}\tilde{g}_n(y_n, \tilde{\mu}_S^{\mathrm{L}}(\theta_0))] =$$

$$\mathrm{AsyVar}[\sqrt{n}\tilde{g}_n(y_n, \mu(\theta_0))] + S^{-2}\sum_{s=1}^{S}\mathrm{AsyVar}[\sqrt{n}\tilde{g}_n(y_n, \mu(\theta_0))] = \left(1 + \frac{1}{S}\right)\mathcal{I},$$

so that

$$\sqrt{n}\tilde{g}_n(y_n, \tilde{\mu}_S^{\mathrm{L}}(\theta_0)) \xrightarrow{d} N\left(0, \left(1 + \frac{1}{S}\right)\mathcal{I}\right). \qquad (31)$$

Plugging (31) into (25) gives the desired result.

Estimator	$\hat{\theta}_1$	lower	upper	length
SN1	1.6649	1.0015	∞	∞
SL1	1.9695	1.3253	∞	∞
SA1	1.9551	1.2728	∞	∞
SN2	1.0827	0.1421	1.3325	1.1905
SL2	1.035	0.0733	1.3275	1.2542
SA2	1.0275	0.0739	1.3241	1.2502
SM2	0.8246	0.0121	1.2513	1.2391
DN	1.0866	0.1812	1.3268	1.1456
DL	1.0470	0.1215	1.3160	1.1945
DA	1.0394	0.1212	1.3126	1.1914
DM	0.8382	0.0240	1.2387	1.2147

Table 1: Point estimates and asymptotic 95 percent confidence intervals for $\theta_1 = 0.6644$ from a representative simulation of the OU process (16).

SN1	SL1	SA1	SN2	SL2	SA2	SM2	DN	DL	DA	DM
0.02	0.63	0.88	0.03	0.52	0.87	1.12	0.01	0.51	0.85	1.11

Table 2: Average estimation time in seconds based on 1000 Monte Carlo experiments. Estimation was performed in R.2.8 on a Dell Poweredge 1850 Server (3GHz, 2GB RAM).

Estimator	Mean	Bias	RMSE
MLE	0.7723	0.1079	0.3178
SN1	1.3761	0.7117	1.6274
SL1	1.3860	0.7216	1.6871
SA1	1.3910	0.7266	1.6920
SN2	0.7697	0.1053	0.3165
SL2	0.7617	0.0973	0.3178
SA2	0.7615	0.0971	0.3184
SM2	0.5473	-0.1171	0.3355
DN	0.7741	0.1097	0.3193
DL	0.7664	0.1020	0.3208
DA	0.7661	0.1017	0.3214
DM	0.5520	-0.1124	0.3356

Table 3: Finite sample properties of estimators of $\theta_1 = 0.6644$ based on 1000 Monte Carlo simulations.

Estimator	J-test	LR test
SN1	0.130	0.379
SL1	0.121	0.363
SA1	0.121	0.354
SN2	0.078	0.051
SL2	0.074	0.053
SA2	0.073	0.052
SM2	0.078	0.129
DN	0.082	0.049
DL	0.070	0.056
DA	0.070	0.054
DM	0.076	0.123

Table 4: Rejection frequencies of nominal 5 percent tests. J-test denotes the test for overidentification restrictions and has an asymptotic chi-square distribution with 2 degrees of freedom. LR test denotes the LR-type coefficient test of $\theta_1 = 0.6644$, and has an asymptotic chi-square distribution with 1 degree of freedom. Rejection frequencies are based on 1000 Monte Carlo simulations.
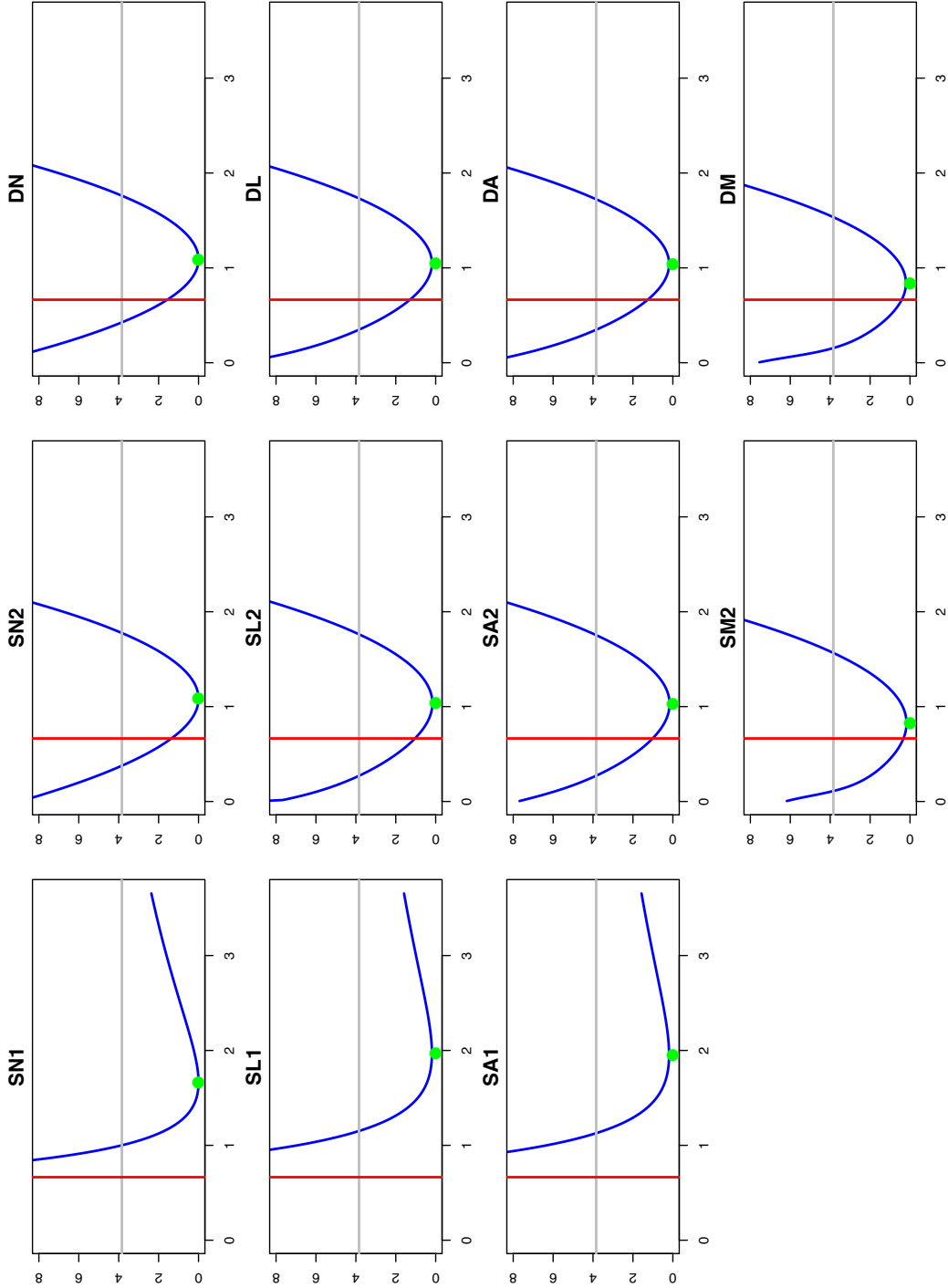
Figure 1: Representative LR-type statistics for testing $H_0 : \theta_1 = \theta_1^0$ as functions of θ_1^0. The underlying model is described in Section 4.1. The horizontal grey line, the vertical red line, and the green dot represent the $\chi^2(1)$ critical value, the true value of $\theta_1 = 0.6644$, and the structural estimate respectively.
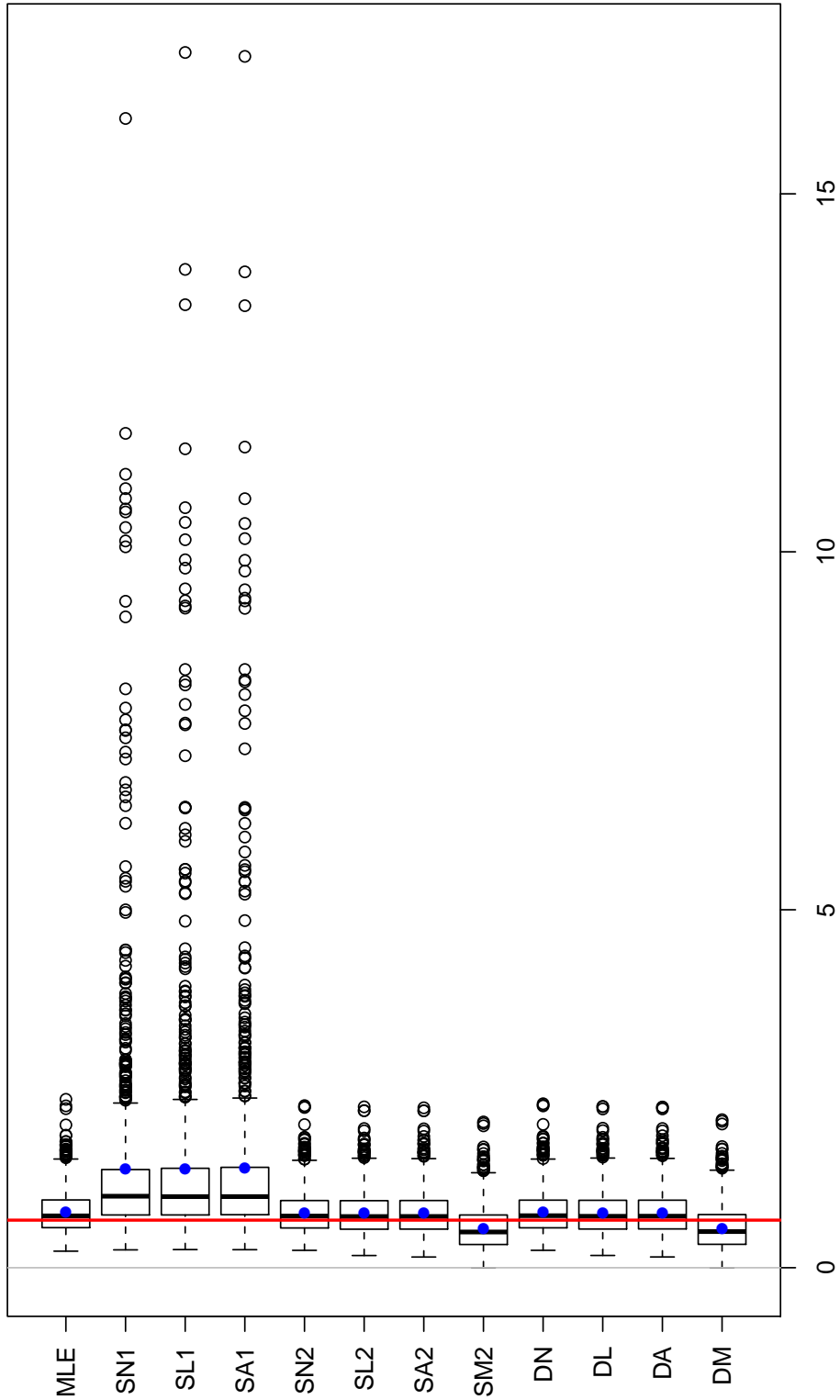
Figure 2: Distribution of θ_1 estimates in the model described in Section 4.1. The boxplot is augmented with a red line representing the true value of $\theta_1 = 0.6644$ and a blue dot representing the mean of the estimates.