Indirect inference under stochastic restrictions†

José A. Hernández*
University of Las Palmas de Gran Canaria
Spain

Ignacio Mauleón
University Rey Juan Carlos I
Spain

Abstract

This paper is focused on the stochastic restriction approach, the NLS and the indirect inference estimators.

A methodology to combine sample and prior information is suggested when the indirect inference estimation method is in use. This goal is achieved through stochastic restrictions approach. The resulting estimator is proved to be more efficient than the indirect inference estimator under specific assumptions about the behaviour of the stochastic restriction. As an illustration of the proposed methodology, the capital stock of an economy is estimated through the estimation of its stochastic rate of the depreciation.

Keywords: Indirect inference, prior information, stochastic restrictions, asymptotic distribution, parameter sequence, Monte Carlo experiments, capital stock.

JEL classification code: C10, C11, C15

*Corresponding author: José A. Hernández. Facultad de CC.EE y EE. Campus de Tafira. 35017 Las Palmas de G.C. Spain. Tel. 00 34 928 458 206. Fax 00 34 928 458 183. Email: jhernandez@daea.ulpgc.es

1 Introduction

Stochastic restrictions approach is one of the ways in which prior information could be modelled. The rationale of using it lies on the fact that it brings efficiency gains in the estimation, subject to the quality of the information available. In some cases, prior information comes from the theory, and imposes restrictions among parameters that should hold in exact terms. This prior information could be included in the model as a deterministic restriction, and the estimator that takes it into account has smaller variance than the nonrestricted estimator. In other cases, prior information comes from previous estimations of similar but different models or samples. This information could be considered as a hint, or a range of values that should contain the value of the parameter with some probability. In this case, deterministic restrictions should not be included in the estimation since the restricted estimator will be biased. If this information is not taken into account in the estimation despite being good, then the information will be wasted as well as the chance of improve the efficiency of the estimator. An intermediate solution is to include it with certain predetermined degree of uncertainty. This is the idea behind stochastic restrictions approach, and its interest is that brings efficiency gains, as shown in Theil and Goldberger (1961) and Shiller (1973) for a linear model under normality of the errors. Nevertheless, stochastic restrictions seem not to have much impact in the classic econometric literature, mainly because of its irrelevance to explain asymptotic efficiency gains. On the other side, in the bayesian approach, prior information is in the bases of the methodology, which is increasing its applicability and diffusion in the profession. In some cases, for instance, the mixed logit model (see McFadden and Train, 2000), the bayesian approach dominates the simulated maximum likelihood estimation in terms of efficiency, mainly due to the prior information used in the bayesian approach and not in the SML method and also, due to the high variance of the SML estimator resulting from the high number of simulations needed to implement this method (See Train, 2001 and 2003).

Despite the finite sample efficiency gains, this result cannot be extended to the asymptotic distributions of the estimator obtained using stochastic restrictions, since the efficiency gain vanishes as sample size increases. This result is proved in Lütkepohl (1993), although, as he points out, it is not useful in empirical terms. Therefore, it would be interesting to extend the finite sample properties of the estimation under stochastic restrictions to the asymptotic context and then to its approximated distribution. In this paper we show, as an auxiliary result, that stochastic restrictions yield asymptotic efficiency gains under some specific assumptions about the asymptotics of prior information. Also, we test in a Monte Carlo exercise that our approach provides better approximations to the distribution of the estimator which takes into account stochastic restrictions than the standard approach.

The main goal of this paper focuses in the description of a simulated based estimator in which prior information is taken into account through the stochastic restrictions approach, and also, under the same type of assumptions already introduced in the first objective. The new estimator, based on the indirect inference (I.I.) estimator of Gourieroux et al (1993), is shown to be more efficient than the baseline one. Simulation based methods, as the method of simulated moments of McFadden (1989), Pakes and Pollard (1989), and the I.I. method (see also Smith, 1993, Lee and Ingram, 1991, for similar approaches) provide powerful techniques to deal with nonlinear models when traditional methods fail. Nevertheless, there is no a clear way by which prior information can be taken into account in simulation based estimation methods. In this paper we achieve this goal through the extension of the I.I. method by introducing stochastic restrictions in the initial I.I. criterion. We obtain the Indirect Inference under Stochastic Restrictions (IIR) estimator, discuss its asymptotics properties and efficiency gains compared to the I.I. estimator. A procedure to test the validity of the restrictions is also provided. Finally, we apply the IIR methodology in a macroeconometric example to check its accuracy in the estimation of a endogenous stochastic rate of depreciation of the stock of physical capital. Results show the effectiveness of the proposed methodology in the solution of this kind of problems and at the same time the significant efficiency gain of the IIR with respect to the I.I. method.
The structure of the paper is the following: section 2 provides the motivation for the assumption about the parameter sequence in the priors specification. Section 3 describes the principle of the stochastic restrictions, and discuss its efficiency gains in finite sample and stochastic terms. Section 4 shows the Monte Carlo exercise in which the suggested finite sample approximation is proved to fit better than the conventional. In Section 5 we define the new estimator and provide its asymptotic properties. Section 6 focuses on a macroeconometric example and numerical evaluations of the efficiency gains of the new estimator, and Section 7 concludes.

2 Asymptotic priors

This first result, related with asymptotic efficiency gains, is derived on the bases of a theoretical assumption about the behavior of the variance of the stochastic restriction. More precisely, we consider an asymptotic decreasing variance of the stochastic restrictions. As a result of it, it is preserved the relative weights of prior and sample information in asymptotic terms, which is not the case in the standard approach, and then the resulting variance of the estimator decreases. This assumption makes the resulting asymptotic distribution fit the finite sample distribution better.
This kind of assumption might be considered too strong and, as mentioned in Kadane (1971), difficult

to justify. However, in the context of IV estimation with weak instruments, in Bekker (1994) and Staiger and Stock (1997) we find a similar assumption, justified by the goal of finding better approximations to the finite sample distribution of the estimator of interest. The approximation is derived from standard asymptotic theory mainly, but taking into account the extra assumption of a parameter sequence, designed to make the resulting distribution fit the sample distribution better. Despite of the objection of Kadane (1971), Bekker (1994) claim, "...since the finite sample distribution does not depend on the behavior of observations in the case of further sampling, there is no reason why an approximation should[1]. Consequently, there is no need to make such "realistic" assumption...the quality of the approximation is the only criterion for justifiability". The parameter sequence we choose in this paper lies on the variance of the stochastic restriction. We could also argue that the rationale underlying on it lies is the fact that it makes the asymptotic distribution fit the finite sample distribution better. But, added to that, we find a realistic motivation for it. We consider that the priors are obtained from a sample which size also increase when asymptotics is considered. Then, we extend to dynamic terms (defining the asymptotics in terms of both samples sizes) the property that the priors are informative. In short, we assume that the priors are *consistent,* i.e., when the size of the sample where priors are obtained increases, such priors are more informative. In other words, our key assumption mean that experience matters, and this can be considered to be a natural fact. If priors are informative in static terms, then its quality might increases in the case of additional sampling. That is, priors keep on being informative as increases the size of the sample that generates them.
Finally, the key assumption also allow to blend prior and sample information when estimators based on simulation have to be used. This is the case of models that generates high nonlinearities in the traditional criterion, what makes standard methods not useful. Generally, the estimators obtained by simulations, despite the fact that are the only solution to estimate some family of models, show high variance, and hence, efficiency gains would be welcome. The key assumption, allow to extend the I.I. advantages to a more efficient procedure.

3 Stochastic Restrictions and Efficiency Gains

In this section we discuss the relevance of taking into account prior information in the Nonlinear Least Squares (NLS) estimation. We define the Nonlinear Least Squares under Stochastic Restriction (SR) method, obtain its asymptotic distribution, and compare it with the NLS distribution under standard

[1] It is important to note that in section 3, under standard assumptions, the finite sample efficiency gains due to stochastic restrictions vanishes as the sample size increases. Then, also in our case the approximated distribution depends on further sampling.

asymptotic theory (this approach does not ascribe to SR any advantages against NLS) and under our key assumption. The goal of this section is showing that stochastic restrictions are able to be modelled asymptotically and bring efficiency gains. This result will be used in the next section to redefine the I.I. estimator when stochastic restrictions are taking into account in the estimation.

Consider a general nonlinear model given by the equation $y_t = f(z_t, \theta) + \varepsilon_t$. The NLS estimation of θ is defined as the value $\hat{\theta}_{NLS}$ that minimizes the criterium function $\Psi_T(\theta) = \sum_{t=1}^{T} [y_t - f(z_t; \theta)]^2$ with respect to θ. Calling $F_1' = \frac{\partial f}{\partial \theta}$ – px1 vector, we have $\frac{1}{T}\left[\frac{\partial^2 \Psi_T}{\partial\theta\partial\theta'}(\theta_0)\right] = \frac{F_1' F_1}{T}$ and the NLS asymptotic distribution is

$$\sqrt{T}(\hat{\theta}_{NLS} - \theta_0) \xrightarrow{d} N\left(0, \sigma_\varepsilon^2 p\lim\left(F_1' F_1/T\right)^{-1}\right) \tag{1}$$

We consider now q stochastic restrictions on θ, $q < p$, modelled through the equation $r = G(\theta) + v$, where r is a qx1 vector and $v \sim N(0, \sigma_v^2 I_q)$, v independent of θ and ε. The resulting model is given by:

$$\begin{bmatrix} y \\ r \end{bmatrix} = \begin{bmatrix} f(x,\theta) \\ G(\theta) \end{bmatrix} + \begin{bmatrix} \varepsilon \\ v \end{bmatrix}$$

In order to reach homoscedasticity, the model can be transformed and expressed is matrix form as:

$$\bar{y} = \bar{f}(\bar{x}, \theta) + \bar{\varepsilon}$$

with $V(\bar{\varepsilon}) = I_{T+q}$. Some additional notation should be introduced. Let $\bar{F} = \frac{\partial \bar{f}}{\partial \theta}$ and $D = \frac{\partial G(\theta)}{\partial \theta}$ a qxp matrix. After some computation we reach the asymptotic distribution of the SR estimator:

$$\sqrt{T}(\hat{\theta}_{SR} - \theta_0) \xrightarrow{d} N\left(0, p\lim\left(\frac{1}{T}\frac{F_1' F_1}{\sigma_\varepsilon^2} + \frac{1}{T}\frac{D'D}{\sigma_v^2}\right)^{-1}\right) \tag{2}$$

We now compare the asymptotic distributions given in (1) and (2) under different scenarios. The first one corresponding to the standard asymptotic theory, while the alternative one is given by a particular assumption made on the behavior of the variance of the error term v.

Proposition 1. Under standard asymptotic analysis, the asymptotic distribution of the estimator SR is identical to the asymptotic distribution of the NLS estimator.

Proof. Since σ_v^2 is constant, $p\lim\left(\frac{1}{T}\frac{D'D}{\sigma_v^2}\right) = 0$. Then, the second term in the asymptotic variance covariance matrix vanishes and

$$\sqrt{T}(\hat{\theta}_{SR} - \theta_0) \xrightarrow{d} N\left(0, p\lim\left(\frac{1}{T}\frac{F_1' F_1}{\sigma_\varepsilon^2}\right)^{-1}\right)$$

which coincides with the asymptotic distribution of the NLS estimator given in equation (1)[2]. Therefore, asymptotic analysis leads to the same distribution for $\hat{\theta}_{SR}$ and $\hat{\theta}_{NLS}$, and stochastic restrictions brings no efficiency gains. Nevertheless, the previous result is not satisfactory when the asymptotic distribution has to be used to approximate the variance of the estimator, specially when the sample size is small. Because of this reason we consider a specific framework in which the analysis is developed. We consider that prior information about parameters comes from previous experience and such experience derives from observations that are taken from a sample of size T^*. This observations comes from a model that is not essentially related with our model of interest, and hence, the disturbance v is independent of ε. The asymptotic results in this new framework are defined as T and T^* goes to infinity. The following assumptions are in order.

Assumption 01 (A01). The variance of v, the error term of the stochastic restriction, is $\sigma_v^{*2} = \sigma_v^2/T^*$ where T^* is the sample size of the model generating the prior information.

Assumption 02 (A02) $T/T^* = 1 + o_p(1)$.

The purpose of A01 is to maintain the relative weights of prior information when taking limits. This principle will be developed in a similar way in a more developed assumption in section 5.

Proposition 2. Under (A01), (A02) and regular conditions on the regressors and errors of the model, the SR estimator is asymptotically more efficient than the NLS estimator.

Proof. Under A01, the asymptotic distribution (2) becomes

$$\sqrt{T}(\hat{\theta}_{SR} - \theta_0) \xrightarrow{d} N\left(0, p\lim\left(\frac{1}{T}\frac{F_1'F_1}{\sigma_\varepsilon^2} + \frac{1}{T}\frac{D'D}{\sigma_v^{*2}}\right)^{-1}\right)$$

and, from A02

$$p\lim\left(\frac{1}{T}\frac{D'D}{\sigma_v^{*2}}\right) = p\lim\left(\frac{D'D}{\sigma_v^2}\right)$$

Then,

$$\left\{p\lim\left(\frac{1}{T}\frac{F_1'F_1}{\sigma_\varepsilon^2} + \frac{1}{T}\frac{D'D}{\sigma_v^{*2}}\right)\right\}^{-1} - \left\{p\lim\left(\frac{1}{T}\frac{F_1'F_1}{\sigma_\varepsilon^2}\right)\right\}^{-1}$$

is a definite negative matrix, and $\hat{\theta}_{SR}$ is asymptotically more efficient than $\hat{\theta}_{NLS}$, i.e., than the SR estimator obtained under the standard analysis. When normality is not assumed, the result is also extended to the approximated distributions[3]. An additional result is shown below in which we justify the inclusion of assumption (A01).

---

[2] The suggested approximated distribution for finite sample SR estimator is $\hat{\theta}_{SR} \approx N(\theta_0, \sigma_\varepsilon^2(F_1'F_1)^{-1})$

[3] The approximate distribution for finite sample size is $\hat{\theta}_{SR} \approx N\left(\theta_0, \left(\frac{F_1'F_1}{\sigma_\varepsilon^2} + \frac{D'D}{\sigma_v^2}\right)^{-1}\right)$

One additional result could be checked through a simulation exercise: when sample size is small, the analytical expression of the approximated variance of the SR estimator found under (A01) and (A02) is a better approximation of the *observed* variance of the SR estimator than the expression of the approximated variance of the SR estimator derived from the standard asymptotic theory. This result is shown in the next section.

4 Monte Carlo exercise about SR estimation

We consider the following model:

$$\begin{aligned} y_t &= bz_t + u_t \\ u_t &= \alpha u_{t-1} + \varepsilon_t \end{aligned} \tag{3}$$

where $\varepsilon \sim iid\ N(0, \sigma_\varepsilon)$ and suppose that there is strong prior in favor of zero autocorrelation of the errors. In terms of stochastic restrictions we have the equation $\tilde{\alpha} = \alpha + v$, $v \sim N(0, \sigma_v)$, being σ_v small and $\tilde{\alpha}$ close to zero since prior information is *correct* and so that close to $\alpha = 0$. We can write the previous model as follows

$$\begin{aligned} y_t &= \beta_1 y_{t-1} + \beta_2 z_t + \beta_3 z_{t-1} + \varepsilon_t \\ &= x_t'\beta + \varepsilon_t \end{aligned} \tag{4}$$

where $x_t' = (y_{t-1}, z_t, z_{t-1})$, and $\beta = (\beta_1, \beta_2, \beta_3)$. Note that $\beta_1 = \alpha$; $\beta_2 = b$ and $\beta_3 = -\alpha b$. Calling $\theta = (\alpha, b)'$ to the original model parameters (equations (3)), we can write $\beta = \beta(\theta)$, the vector of parameters of the auxiliary model (equation (4)). Taking into account the stochastic restriction the model is

$$\begin{bmatrix} Y \\ \alpha^h \end{bmatrix} = \begin{bmatrix} X \\ R \end{bmatrix} \beta(\theta) + \begin{bmatrix} \varepsilon \\ v \end{bmatrix} \tag{5}$$

where $R = (1, 0, 0)$, α^h is the prior information available about α and h is the index of simulation to be explained below.

The purpose of this exercise is to test the accuracy to approximate the variance of the SR estimator by means of the analytical expression deriving from the standard approach – which coincides with the NLS estimation variance – and from the alternative approach suggested under (A01) and (A02). In order to check this, we first simulate and estimate model (5) by SR method, compute the observed variance of the estimator across simulations and then compare it with alternative expressions to find which one approximates it better. In other words, we try to check in this example whether or not the stochastic restrictions are relevant or not in the approximated distributions resulting from alternative approaches to measure it.

Sample size in the Monte Carlo experiment is $T = 40$ and the number of simulations is $H = 5000$. In each simulation prior information about the values of α is extracted from a simulated distribution $N(0, \sigma_v)$ and values are denoted with index h, $h = 1, ..., H$. Different values of σ_v are considered to extend the discussion, reflecting different levels of certainty of prior information[4]. True values of the parameters are $\theta_0 = (\alpha_0, b_0) = (0, 1.5)$. Results are shown in Table 1 and should be considered as the *true* properties of the estimator or, at least, close approximations since the number of simulations is high.

In Table 2 are computed the alternative variances approximations: the one suggested by asymptotic theory, in which stochastic restrictions result to be irrelevant, given by

$$V_1(\widehat{\tilde{\theta}_{NLS}}) \simeq \left(D' \left(p \lim(\frac{X'X}{T}) \frac{1}{\sigma_\varepsilon^2} \right) D \right)^{-1}$$

and the alternative approximation resulting from assumptions (A01) and (A02) given by

$$V_2(\widehat{\tilde{\theta}_{SR}}) \simeq \left(D' \left(p \lim(\frac{X'X}{T}) \frac{1}{\sigma_\varepsilon^2} + \frac{R'R}{\sigma_\alpha^2} \right) D \right)^{-1}$$

The following conclusions are reached. First, as the alternative approach suggests, stochastic restrictions bring efficiency gains, as observed when column I is compared with columns II, III and IV in Table 1. Also, efficiency increases with certainty about the prior information. Second, the analytical form of the approximated variance obtained under (A01) and (A02) is closer to the *real* one than the one suggested by the standard approach.

5 Indirect Inference under Stochastic Restrictions

The method of Indirect Inference (I.I.) of Gourieroux et al (1993) (GMR93) and the methods of simulated moments of Lee and Ingram (1991) and Duffie and Singleton (1993), (see similar methods in Smith, 1993 and Gallant and Tauchen, 1994), provide a powerful technique to deal with nonlinear models where traditional methods fail. In spite of the wide applicability of these methods, there is

[4] Additional properties of the exercise are available from the authors on request.

not a methodology to take into account prior information in their implementation (see, for example, Canova, 1994). In this section we suggest a way to solve this problem based on the stochastic restriction approach. The analysis will be cast in the framework of the I.I., since this methodology is more general and other simulated based estimation methods can be viewed as special cases.

First we define the new estimation method of Indirect Inference under Stochastic Restrictions (IIR) and provide its distribution. Then, based on the approach introduced in section 2 we show that the IIR estimator is more efficient than the I.I. method. Finally we suggest a method to test the validity of the restrictions.

In the I.I. approach it is considered a $p-$dimension vector of parameters θ of a true model (M) and a $j-$ dimension vector of parameters β, $j \geq p$ of the auxiliary model. The binding function $b(\theta)$ could be estimated from $\hat{\beta}(\theta)$, a consistent criterion for β given a value of θ. The $b(\theta)$ function is twice differentiable with respect to θ, where $\partial b(\theta)/\partial\theta = D_1$, is of full rank.

The I.I. estimator of θ, following GMR93, is defined as

$$\tilde{\theta}_{II} = \arg\min \left\{ m_1' \Omega_1 m_1 \right\}$$

where $m_1 = \hat{\beta} - \sum\limits_{s=1}^{S} \hat{\beta}(\theta)/S$ and Ω_1 is a jxj symmetric and positive definite matrix to be determined below. Under regular assumptions on auxiliary criterion $\Psi_T(\beta)$ and the model, - in the Appendix this assumptions are shown - the asymptotic distribution of the I.I. estimator is

$$\sqrt{T}(\tilde{\theta}_{II} - \theta_0) \xrightarrow{d} N\left(0, \left(1 + \frac{1}{S}\right) W_1(\Omega_1)\right)$$

where

$$W_1(\Omega_1) = \left(D_1'\Omega_1 D_1'\right)^{-1} \left(D_1'\Omega_1\Omega_1^{*-1}\Omega_1 D_1\right) \left(D_1'\Omega_1 D_1'\right)^{-1}$$

and $\Omega_1^* = J_0 \bar{H}_0^{-1} J_0$ being $J_0 = p\lim -\frac{\partial^2 \Psi_T(b(\theta))}{\partial\beta\partial\beta'}$, $\bar{H}_0 = H_0 - K_0$, being H_0 and K_0 are matrixes related with properties of the variance-covariance matrix of $\sqrt{T}\frac{\partial \Psi_T}{\partial \beta}$ and specified in assumptions (A6) and (A7) in the Appendix.

The matrix Ω_1 is chosen according to the optimality criterion, and then taken as $\Omega_1 = \Omega_1^*$. In this case, the asymptotic variance-covariance matrix of $\tilde{\theta}_{II}$ (taking $S \longrightarrow \infty$) is

$$W_1^* = \left(D_1'\Omega_1^* D_1\right)^{-1} \tag{6}$$

Since $\Psi_\infty = \underset{T\rightarrow\infty}{p\lim}\Psi_T$, Ψ_T could be used in the place of Ψ_∞ and a consistent estimator of the asymptotic variance of $\hat{\beta}$ for H_0[5].

We now consider the existence of prior information on the parameters of interest θ, what could formally by written as stochastic restrictions: $G(\theta) = r + v$, where $E(v) = 0$ and $V(v) = \Phi_2^*$, v independent of the error term of the model. Further properties of v are to be specified below. Vector r contains the known values that verifies the prior information, and the variance covariance matrix is chosen according to the quality of the prior information. Function $G(\theta)$ is differentiable and such that $\partial G/\partial\theta = D_2$ is a qxp matrix of full rank in a neighborhood of θ_0.

It is necessary to introduce some additional notation to define the new estimation method. Let $m' = (m_1', (G - r)')$, $D = (D_1, D_2)'$, and Ω a block diagonal matrix, with Ω_1, Ω_2 in their diagonal respectively.

Definition The indirect inference under stochastic restriction estimator of θ is

$$\tilde{\theta}_{IIR} = \arg\min_{\theta} \left\{ m'(\theta)\Omega m(\theta) \right\} \tag{7}$$

where

$$m'\Omega m = \left\{ m_1'(\theta)\Omega_1 m_1(\theta) + (G(\theta) - r)'\Omega_2(G(\theta) - r) \right\}$$

Some additional assumptions are in order to derive the asymptotic behavior of the IIR estimator.

A1)- A7). Are the regular conditions needed to obtain the asymptotic distributions if the I.I. estimator, shown in the Appendix.

A8) $\frac{\partial G(\theta)}{\partial\theta} = D_2$ is a qxp matrix of full rank in a neighborhood of θ_0.

A9) $\sqrt{T^*}(G(\theta) - r) \xrightarrow{d} N(0, \Omega_2^{*-1})$

A10) $\frac{T}{T^*} = o_p(1)$

Assumption (A9) describes the asymptotic properties of the stochastic restrictions, and it leads to the approximate distribution

$$G(\theta) - r \approx N(0, T^{*-1}\Omega_2^{*-1})$$

and hence similar to assumption (A01) introduced in section 3. The rationale behind this assumption is the intention to maintain a constant relative weight between the sample and prior information asymptotically. The relevance of this assumption lies on the fact, already discussed, that under this

[5] In order to estimate H, the semiparametric procedure suggested by Andrews and Monahan (1992) could be applied. The idea is to regress $\frac{1}{\sqrt{T}} \frac{\partial \Psi_T(\beta)}{\partial \beta}$ to an Arma vector and to estimate its variance covariance matrix from residuals in a non parametric way (see, for instance, Newey and West (1987)). Another possibility is to estimate a VAR from $\frac{1}{\sqrt{T}} \frac{\partial \Psi_T(\beta)}{\partial \beta}$ (see Den Haan, et. al. (1995)).

hypotheses, the approximate distribution for small sample size of the resulting estimator is closer to the observed distribution of the estimator. Note that (A9) implies consistency of the random variable r. Again, (A9) will bring efficiency gains in the restricted estimator. The asymptotic properties of the IIR estimator are derived next.

Proposition 3 Under assumptions (A1) to (A10), the Indirect Inference under Stochastic Restrictions estimator, $\tilde{\theta}_{IIR}$, is consistent, asymptotically normal and has the asymptotic distribution

$$\sqrt{T}(\tilde{\theta}_{IIR} - \theta_0) \xrightarrow{d} N\left(0, \left(1 + \frac{1}{S}\right) W(\Omega)\right)$$

where

$$\begin{aligned} W(\Omega) \quad = \quad & \left[D_1'\Omega_1^{-1}D_1 + D_2'\Omega_2^{-1}D_2\right]^{-1} \\ & \text{x}[D_1'\Omega_1\Omega_1^{*-1}\Omega_1 D_1 + D_2'\Omega_2\Omega_2^{*-1}\Omega_2 D_2] \\ & \text{x}\left[D_1'\Omega_1 D_1 + D_2'\Omega_2 D_2\right]^{-1} \end{aligned}$$

This result is proved in the Appendix.

For the optimal matrix $\Omega_1 = \Omega_1^* = J_0\bar{H}_0^{-1}J_0$ and $\Omega_2 = \Omega_2^*$ the variance-covariance matrix reduces to:

$$W^* = (D'\Omega^* D)^{-1} \tag{8}$$

where Ω^* is the block diagonal matrix with Ω_1^* and Ω_2^* in the diagonal.

Proposition 4 Under assumptions (A1) to (A10) $\tilde{\theta}_{IIR}$ is asymptotically more efficient than $\tilde{\theta}_{II}$

To proof this result, we compare equations (8) and (6). The difference

$$W^* - W_1^* = (D'\Omega^* D)^{-1} - (D_1'\Omega_1^* D_1)^{-1}$$

is a negative definite matrix since

$$D'\Omega^* D - D_1'\Omega_1^* D_1 = D_2'\Omega_2^* D_2$$

is a positive definite matrix.

Finally, the validity of the restrictions could be tested from the approximated distribution of $G(\tilde{\theta})$ under $H_0 : G(\theta) - r = v$, taking $v \sim N(0, \Omega_2^{*-1})$. A Taylor approximation of $G(\theta)$ evaluated in any θ,

$\tilde{\theta}$ such that $\sqrt{T}(\tilde{\theta}-\theta_0) \xrightarrow{d} N(0,W)$, yields

$$G(\tilde{\theta}) = G(\theta) + \left.\frac{\partial G(\theta)}{\partial \theta}\right|_{\theta=\tilde{\theta}} (\tilde{\theta}-\theta)$$

Since $\tilde{\theta} \approx N(\theta, T^{-1}W)$, from (**??**), under H_0, and assuming that $\tilde{\theta}$ and v are independent, we have

$$G(\tilde{\theta}) \approx N(0, T^{-1}(D_2' W D_2 + \Omega_2^{*-1}))$$

where D_2 denotes the same matrix as in previous section.

6 A macroeconometric example

The stock of physical capital of an economy is one of the basic economic aggregates. However it is not observable since it depends on the rate of depreciation. Nevertheless it could be estimated jointly with the parameters of a production function. In most cases prior information on the rate of depreciation is available from other sources (e.g. National Accounts). We consider here a simplified production function $y_t = \alpha k_t + \varepsilon_t$, where y and k are production and capital stock in logs respectively and α is the elasticity of the capital stock. From the perpetual inventory method, capital stock is given by $K_t = I + (1-\delta_t)K_{t-1}$ and different assumptions are made about the pattern δ_t. In the auxiliary model, production is explained by current and lagged production and investment.

First we describe several cases in which the rate of depreciation is modelled thorough three different ways: stochastic liner with normal error, uniform and Autorregresive. More specifically, the main characteristics of the stochastic processes, in each one of the cases, are:

i) **Case I**. The rate of depreciation is given by the process

$$\delta_t = \gamma_0 + \gamma_1 z_t + v_t \tag{9}$$

where $z_t = (I_t - I_{t-1})/I_{t-1}$, as in Hernández and Mauleón (2002a)[6], $v \sim N(0, \sigma_v^2)$. Then,

$$\delta \sim N(\gamma_0 + \gamma_1 \bar{z}, \sqrt{\gamma_1^2 \sigma_z^2 + \sigma_v^2})$$

[6]The economic idea behind this assumption is that technological progress is incorporated in the production through new investment. Such technological progress explain an increasing productivity in the new capital assets relative to the vintage capital, and this causes obsolescence. This argument explain the positive relation between the rate of depreciation and the rate at which investment varies in time, and hence that $\delta_1 > 0$. The part of the depreciation associated with the use of capital is captured in parameter δ_0.

The stochastic restriction equation should be consistent with the process of the rate of depreciation. Hence, an additional equation should be taken into account:.

$$\tilde{\delta} = \gamma_0 + \gamma_1 \overline{z} + \epsilon \tag{10}$$

i) **Case II**. The rate of depreciation is given by a uniform distribution. Then,

$$\delta = v \sim U[\gamma_0; \gamma_1] \tag{11}$$

and the parameters to be estimated are the mean and the variance of the distribution, univocaly related with d_L and d_U. In this case, prior information consist in a pair of values, $\tilde{\gamma}_0$, and $\tilde{\gamma}_1$, and could be modelled as

$$\tilde{\delta} = \epsilon \sim U[\tilde{\gamma}_0; \tilde{\gamma}_1] \tag{12}$$

where $\tilde{\gamma}_h$ are values close to parameters γ_h, $h = 0, 1$.

iii) **Case III**. The rate of depreciation follows an autorregresive process. Then,

$$\delta_t = \gamma_0 + \gamma_1 \delta_{t-1} + v_t \tag{13}$$

where $v \sim N(0, \sigma_v^2)$, $|\gamma_1| < 1$. The equation about prior information to be taken into account should be such that the prior value of δ was close to the expected value of the process of δ. Then, since the process is stationary, prior information could be modelled as

$$\tilde{\delta} = \frac{\gamma_0}{1 - \gamma_1} + \epsilon \tag{14}$$

The equations above define the processes and the stochastic restrictions equations related with the rate of depreciation of the capital stock. The model contain also some additional equations, as the capital stock equation,

$$K_t = I_t + (1 - \delta_t) K_{t-1}$$

and finally, the production function equation,

$$\log Y_t = \alpha \log(K_t) + \varepsilon_t$$

where ε_t is the white noise disturbance.

We have simulated series for the rate of depreciation, capital stock and production. Then, we use the only observable variables Y_t and I_t, generated by the vector of parameters $\theta = (\alpha, \delta_0, \delta_1, \sigma_v)$, (values are shown in Table 5). The equations contained in the three models are, then, the equation of the production function, of the capital stock and equations (9) and (10) in Case I, (11) and (12) in Case II and (13) and (14) in Case III. The mentioned equations are considered to estimate each one of the models through the restricted indirect inference estimator as we describe next.

6.1 IIR estimation

The IIR estimation is carried out taking priors of δ that are close with the true mean of the parameter sequence, and with small variance. The auxiliary model are chosen according to their adequacy to the main IIR criterion[7]. In such models production is explained by lagged investment and production. The IIR estimator of θ is

$$\tilde{\theta}_{IIR} = \arg\min \{Q_T\}$$

where Q_T, taking OLS as the auxiliary criterion, is

$$Q_T = \left[\hat{\beta} - \sum \hat{\beta}_s(\theta)/S\right]' (X'X/\sigma_\varepsilon^2) \left[\hat{\beta} - \sum \hat{\beta}_s(\theta)/S\right] + \frac{(\tilde{\delta} - E(\tilde{\delta}))^2}{\sigma_\epsilon^2}$$

and $\hat{\beta}$ the OLS estimator of the auxiliary model from the original data and $\hat{\beta}_s$ the OLS estimator obtained with simulated data in simulation s. The value $\tilde{\delta}$ is the prior about δ. $E(\tilde{\delta})$ equation and σ_ϵ value will depend on the case that is considered, as described in Table 5. In general σ_ε^2 is also unknown, but we have considered it as known to simplify the estimation problem, although the estimation of such parameter would not bring further difficulties.

In the simulation exercise, the number of replications and estimations of the model is 1000. Table 5 shows the results obtained for the different cases, and the Case III, where $\delta \sim$ AR(1) results shows bias in the mean and the variance of the estimators. Nevertheless, the conclusions that we reach tells us that the method IIR performs well in general, and that could be an adequate tool to estimate a stochastic rate of depreciation in macro models.

Also, as an additional exercise, estimations have been made by the I.I. and the IIR methods with the same observations, in order to test the efficiency gains of the proposed method. In Table 6 we show the results for the Cases I and II under the same true parameters vector than in the previous exercise. For the Case II we have replicated the exercise for different values of σ_ϵ, the variance of the stochastic

[7] More information about the auxiliary model and criteria could be requested to authors.

restrictions, in order to test the sensibility of the efficiency gains to different scenarios of uncertainty. We also show in Table 6 that efficiency gains of the IIR estimator with respect to I.I. are important, specially in Case I, and they depend positively, on the quality of prior information.

Finally, results in Table 6 show that IIR approach behaves properly and could be an useful approach to estimate models in which it is necessary to estimate by simulation[8] and in which prior information on parameters is available.

7 Conclusion

In this paper we first prove that stochastic restrictions could bring efficiency gain in asymptotic distributions, and also, in a particular example we check that the resulting distribution explain better the distribution of the NLS estimator. Second, we propose a procedure to combine the method of Indirect Inference (I.I.) of Gourieroux et al. (1993) with stochastic restrictions approach (Theil and Goldberger (1961), Shiller (1972) and Litterman (1986)). The proposed method of Indirect Inference under Stochastic Restriction (IIR) also suggest a way to extend any simulated based estimation method to the context of stochastic restrictions, since these should be included into the criterion in the same way. The IIR method is intended to provide a technique to incorporate prior information into highly complex econometric procedures developed recently, in order to improve the estimation results. We provide the distribution of the IIR estimator and a test for the validity of the restrictions. We show that IIR estimator is more efficient than II estimator. Finally, a macroeconometric application is suggested where IIR is implemented to estimate a stochastic endogenous rate of depreciation of the capital stock of an economy.

[8] See Hernández and Mauleón (2002) for empirical implementation of a simpler method.

Tables

TABLE 1

Properties of the SR estimator

	σ_α^2					
	0.02		0.04		0.06	
	b	α	b	α	b	α
Mean	1.499580	0.001056	1.4999353	-1.392e-05	1.500986	0.000873
Variance	0.025580	0.011769	0.025772	0.016660	0.025591	0.018979
MCE	0.025580	0.011770	0.025772	0.016660	0.025592	0.018979

TABLE 2

Alternative estimators of $V(\tilde{\theta}_{SR})$

$\widehat{V_1(\tilde{\theta}_{SR})}$		$\widehat{V_2(\tilde{\theta}_{SR})}$					
I		**II** $(\sigma_\alpha^2 = 0.02)$		**III** $\sigma_\alpha^2 = 0.04$		**IV** $\sigma_\alpha^2 = 0.06$	
0.025000	-1.393e-19	0.011111	-6.191e-20	0.015384	-8.573e-20	0.017647	-9.834e-20
	0.025097		0.025097		0.025097		0.025097

TABLE 3

Patterns for the Rate of Depreciation

Case	Linear dependence	Uniform Distribution	Autorregresive
Structure	$\delta_t = \gamma_0 + \gamma_1 z_t + v_t$	$\delta = v$	$\delta_t = \gamma_0 + \gamma_1 \delta_{t-1} + v_t$
Error	$v \sim N(0, \sigma_v^2)$	$v \sim U[\gamma_0; \gamma_1]$	$v \sim N(0, \sigma_v^2),\ \lvert\gamma_1\rvert < 1$
Stochastic Restriction	$\tilde{\delta} = \delta_0 + \delta_1 \bar{z} + \epsilon$	$\tilde{\gamma}_s = \gamma_s + \epsilon;\ s = 0, 1$	$\tilde{\delta} = \frac{\gamma_0}{1-\gamma_1} + \epsilon$

TABLE 4

Mean and standard deviation of IIR estimates

	Patterns for δ		
	$\delta_t = \gamma_0 + \gamma_1 z_t + v_t$	$\delta = v \sim U[\gamma_0; \gamma_1]$	$\delta_t = \gamma_0 + \gamma_1 \delta_{t-1} + v_t$
θ	$(\alpha, \gamma_0, \gamma_1, \sigma_v)$	$(\alpha, \gamma_0, \gamma_1)$	$(\alpha, \gamma_0, \gamma_1, \sigma_v)$
θ_0	(0.4,0.06,0.01,0.004)	(0.4,0.03,0.12)	(0.4, 0.06, 0.2, 0.02)
α	0.400	0.403	0.3998
	(0.002)	(0.012)	(0.0038)
γ_0	0.060	0.038	0.0413
	(0.003)	(0.016)	(0.0335)
γ_1	0.009	0.118	0.4499
	($5.4\text{x}10^{-4}$)	(0.046)	(0.4465)
σ_v	0.004	-	0.0266
	($8.7\text{x}10^{-5}$)	-	(0.0408)

TABLE 5

Properties of the IIR estimator

	Case I	Case II	Case III
	$\delta_t = \delta_0 + \delta_1 z_t + v_t$	$\delta = v \sim U[\delta_L; \delta_U]$	$\delta_t = \gamma_0 + \gamma_1 \delta_{t-1} + v_t$
θ_0	(0.4,0.06,0.01,0.004)	(0.4,0.03,0.12)	(0.4, 0.06, 0.2, 0.02)
α	0.400	0.4014	0.3998
	(0.002)	(0.0047)	(0.0038)
δ_0	0.060	0.0374	0.0413
	(0.003)	(0.0065)	(0.0335)
δ_1	0.009	0.1174	0.4499
	(5.4x10^{-4})	(0.0066)	(0.4465)
σ_v^2	0.004	-	0.0266
	(8.7x10^{-5})	-	(0.0408)

TABLE 6

Efficiency gains of IIR estimator

Case	I: $\delta_t = \delta_0 + \delta_1 z_t + v_t$		II: $\delta = v \sim U[\delta_L; \delta_U]$			
θ_0	(0.4,0.06,0.01,0.004)		(0.4,0.03,0.12)			
	I.I.	*IISR*	*I.I.*		*IISR*	
σ_ϵ	-	*0.005*	-	*0.003*	*0.004*	*0.005*
α	0.0399	0.3948	0.4024	0.4020	0.4013	0.4012
$std(\alpha)$	0.0084	(5.5x10^{-4})	(0.0132)	(0.0065)	(0.0074)	(0.0112)
δ_0	0.0633	0.0628	0.0491	0.0334	0.0325	0.0323
$std(\delta_0)$	0.0714	(1.7x10^{-5})	(0.2416)	(0.0094)	(0.0110)	(0.0163)
δ_1	0.0058	0.0093	0.0976	0.1232	0.1222	0.1220
$std(\delta_1)$	0.0771	(0.0013)	(0.2454)	(0.0096)	(0.0111)	(0.0162)
σ_v	0.0065	0.0037	-	-	-	-
$std(\sigma_v)$	0.0107	(0.0048)				

Appendix: Derivation of the asymptotic distribution of IIR estimator

Here we develop similar proofs to the used on the asymptotic properties of the I.I. estimator. To show the asymptotic distribution of IIR estimator we need several regularity conditions, as for the I.I. distribution. The most important are

A1) The general auxiliary criterion function $\psi_T(y_T, z_T; \beta)$ converges to a deterministic limit denoted by $\psi_\infty(\theta, \beta)$ when T goes to infinity.

A2) This limit function has a unique maximum with respect to β, and this maximum is $b(\theta)$. That is, $b(\theta) = \arg\max_\beta\{\psi_\infty(\theta, \beta)\}$

A3) ψ_T and ψ_∞ are differentiable with respect to β and $\lim \frac{\partial \psi_T}{\partial \beta} = \frac{\partial \psi_\infty}{\partial \beta}$

A4) The solution of the asymptotic first order condition $\frac{\partial \psi_\infty(\theta,\beta)}{\partial \beta} = 0$ is well defined in θ and β

A5) $plim - \frac{\partial^2 \Psi_T(b(\theta))}{\partial \beta \partial \beta'} = -\frac{\partial^2 \Psi_\infty(b(\theta_0))}{\partial \beta \partial \beta'} = J_0$ and J_0^{-1} exist

A6) $\sqrt{T}\frac{\partial \Psi_T(b(\theta_0))}{\partial \beta} \xrightarrow{d} N(0, H_0)$

A7) $\lim_{T\longrightarrow\infty} cov\left[\sqrt{T}\frac{\partial \Psi_T(y_T^{s_1})}{\partial \beta}, \sqrt{T}\frac{\partial \Psi_T(y_T^{s_2})}{\partial \beta}\right] = K_0$ for $s_1 \neq s_2$.

A8) $\frac{\partial G(\theta)}{\partial \theta} = D_2$ is of full rank

A9) $\sqrt{T^*}(G(\theta_0) - r) \xrightarrow{d} v \sim N(0, \Omega_2^{*-1})$

A10) $\frac{T}{T^*} = 1 + o_p(1)$

Let us first prove the consistency of the IIR estimator. Under assumptions (A1) to (A4), following Gourieroux et al. (1993) it is proved that the intermediate estimators $\hat{\beta}$ and $\frac{1}{S}\sum_{s=1}^{S} \hat{\beta}_S(\theta)(\equiv \hat{\beta}_{ST}(\theta)$ to simplify notation) converge to $b(\theta_0)$ and $b(\theta)$ respectively. Also, from (A9), $r \longrightarrow G(\theta_0)$. Then,

$$
\begin{aligned}
\hat{\theta}_{IIR} &= \arg\min_\theta \left\{m'(\theta)\Omega m(\theta)\right\} \\
&= \arg\min_\theta\{[\hat{\beta} - \hat{\beta}_{ST}(\theta)]'\Omega_1[\hat{\beta} - \hat{\beta}_{ST}(\theta)] \\
&\qquad +(G(\theta) - r)'\Omega_2(G(\theta) - r)\} \\
&\longrightarrow \{[b(\theta_0) - b(\theta)]'\Omega_1[\hat{\beta} - \hat{\beta}_{ST}(\theta)] \\
&\qquad [G(\theta) - G(\theta_0)]'\Omega_2[G(\theta) - G(\theta_0)]\} \\
&= \{\theta : b(\theta) = b(\theta_0),\ G(\theta) = G(\theta_0)\} \text{ (since } \Omega_1,\ \Omega_2 \text{ are positive definite)} \\
&= \theta_0 \text{ (since } b \text{ and } G \text{ are well defined)}
\end{aligned}
$$

Let us now find the asymptotic distribution of $\hat{\theta}_{IIR}$. Under assumptions (A1) to (A7), asymptotic expansions of $\hat{\beta}_T$ and $\frac{1}{S}\sum_{s=1}^{S} \hat{\beta}_S(\theta_0)$ are deduced from the first order condition. We have, that (following

Gourieroux et al. (1993))

$$\sqrt{T}(\hat{\beta}_T - b(\theta_0)) = J_0^{-1}\sqrt{T}\frac{\partial \Psi_T(b(\theta_0))}{\partial \beta} + o_p(1) \tag{15}$$

and

$$\sqrt{T}(\hat{\beta}_{ST}(\theta_0) - b(\theta_0)) = \frac{J_0^{-1}}{S}\sqrt{T}\sum_{s=1}^{S}\frac{\partial \Psi_T(y_T^s, b(\theta_0))}{\partial \beta} + o_p(1) \tag{16}$$

The asymptotic expansion of $\hat{\theta}_{IIR}(\Omega)$ is deduced as follows. The first order condition for $\hat{\theta}_{IIR}$ ($\hat{\theta}$ for short) from the criterion (7) is:

$$-\frac{\partial \hat{\beta}'_{ST}[\hat{\theta}(\Omega_1)]}{\partial \theta}\Omega_1[\hat{\beta}_T - \hat{\beta}_{ST}(\hat{\theta}(\Omega_1)] + \frac{\partial G'(\hat{\theta})}{\partial \theta}\Omega_2[G(\hat{\theta}) - r] = 0$$

An expansion around the limit value θ_0 gives

$$\begin{aligned} & -\frac{\partial \hat{\beta}'_{ST}(\theta_0)}{\partial \theta}\Omega_1\sqrt{T}[\hat{\beta}_T - \hat{\beta}_{ST}(\theta_0)] + \frac{\partial G'(\theta_0)}{\partial \theta}\Omega_2\sqrt{T}[G(\theta_0) - r] \\ & +\sqrt{T}\left[\frac{\partial \hat{\beta}'_{ST}(\theta_0)}{\partial \theta}\Omega_1\frac{\partial \hat{\beta}_{ST}(\theta_0)}{\partial \theta} + \frac{\partial G'(\theta_0)}{\partial \theta}\Omega_2\frac{\partial G(\theta_0)}{\partial \theta}\right](\hat{\theta}_{IIR}(\Omega) - \theta_0) \\ = \;& o_p(1) = S_1 + S_2 + S_3 + S_4 \text{ (say)} \end{aligned}$$

since $S_1 + S_3 = o_p(1)$, as shown in the asymptotic properties of the $\hat{\theta}_{II}$ estimator under the considered assumptions. From (A9) and the consistency of $\hat{\theta}_{IIR}$, it follows that $S_2 + S_4$ is also $o_p(1)$. Rewriting the above equation in the limit, and calling $D_1 = \frac{\partial b(\theta_0)}{\partial \theta}$, $D_2 = \frac{\partial G(\theta_0)}{\partial \theta}$

$$\begin{aligned} \sqrt{T}(\hat{\theta}_{IIR}(\Omega) - \theta_0) &= \left[D_1'\Omega_1 D_1 + D_2'\Omega_2 D_2\right]^{-1} D_1'\Omega_1\sqrt{T}[\hat{\beta}_T - \hat{\beta}_{ST}(\theta_0)] \\ &\quad - \left[D_1'\Omega_1 D_1 + D_2'\Omega_2 D_2\right]^{-1} D_2'\Omega_2\frac{\sqrt{T}}{\sqrt{T^*}}\sqrt{T^*}[G(\theta_0) - r] + o_p(1) \end{aligned} \tag{17}$$

From (16), (15), we get

$$\sqrt{T}(\hat{\beta}_T - \hat{\beta}_{ST}(\theta_0)) = J_0^{-1}\sqrt{T}\left[\frac{\partial \Psi_T}{\partial \beta} - \frac{1}{S}\sum_{s=1}^{S}\frac{\partial \Psi_T(y_T^s)}{\partial \beta}\right]$$

and using (A6), (A7),

$$\sqrt{T}(\hat{\beta}_T - \hat{\beta}_{ST}(\theta_0)) \xrightarrow{d} N\left(0, \left(1 + \frac{1}{S}\right)\Omega_1^{*-1}\right)$$

where $\Omega_1^* = J_0\bar{H}_0 J_0$, and $\bar{H}_0 = H_0 - K_0$.

Finally, using assumptions (A8), (A9) and (A10):

$$\sqrt{T}(\hat{\theta}_{IIR}(\Omega) - \theta_0) \xrightarrow{d} N\left(0, \left(1 + \frac{1}{S}\right) W(\Omega)\right)$$

where

$$\begin{aligned} W(\Omega) &= \left[D_1'\Omega_1 D_1 + D_2'\Omega_2 D_2\right]^{-1} D_1'\Omega_1\Omega_1^{*-1}\Omega_1 D_1 \\ &\quad x \left[D_1'\Omega_1 D_1 + D_2'\Omega_2 D_2\right]^{-1} \\ &\quad + \left[D_1'\Omega_1 D_1 + D_2'\Omega_2 D_2\right]^{-1} D_2'\Omega_2\Omega_2^{*-1}\Omega_2 D_2 \\ &\quad x \left[D_2'\Omega_2 D_2 + D_2'\Omega_2 D_2\right]^{-1} \\ &= \left[D_1'\Omega_1 D_1 + D_2'\Omega_2 D_2\right]^{-1} \\ &\quad x[D_1'\Omega_1\Omega_1^{*-1}\Omega_1 D_1 + D_2'\Omega_2\Omega_2^{*-1}\Omega_2 D_2] \\ &\quad x \left[D_1'\Omega_1 D_1 + D_2'\Omega_2 D_2\right]^{-1} \end{aligned}$$

The optimal matrixes Ω_i are respectively, Ω_i^* as a consequence of theGauss-Markov theorem and the asymptotic variance-covariance matrix of the IIR estimator

$$W^* = \left[D_1'\Omega_1^* D_1 + D_2'\Omega_2^* D_2\right]^{-1} = \left[D'\Omega^* D\right]^{-1}$$

also taking $S \longrightarrow \infty$.

[†]Financial Support from the University of Las Palmas de Gran Canaria project ULPGC - UNI2001/4 is gratefully acknowledged. The paper has benefited from comments by Juan Mora and Chihwa Kao, although authors assume full responsibility from any remaining error.

References

Andrews, D. and J. Monahan, 1992, An Improved Heteroskedasticity and Autocorrelation Consistent Matrix Estimator, Econometrica, 60, 953-966.

Bekker, P., 1994, Alternative approximations to the distributions of instrumental variable estimators, Econometrica, 62, 657-681.

Canova, F., 1994, Statistical Inference in Calibrated Models, Journal of Applied Econometrics, 9, 123-145.

Duffie, D. and K. Singleton, 1993, Simulated Moments Estimation of Markov Models of Asset Prices, Econometrica, 61, 929-1052.

Gallant, R. and G. Tauchen, 1996, Which Moments to Match?, Econometric Theory, 12, 657-81.

Gourieroux, C., A. Montfort, A. and A. Renault, 1993, Indirect Inference, Journal of Applied Econometrics, 8, 85-118.

Lee, B. and B. Ingram, 1991, Simulation Estimation of Time Series Models, Journal of Econometrics, 47, 197-207.

Kadane, J. B., 1971, Comparison of k-class Estimator when the Disturbances are Small, Econometrica, 39, 723-737.

Litterman, R., 1986, Forecasting with Bayesian Vector Autorregression, Journal of Business and Economic Statistics, 4, 25-38.

Lütkepohl, H., 1993, Introduction to Multiple Time Series Analysis (Springer Verlag, Berlin).

Hernández, J. and Mauleón, I., 2002, On the econometric estimation of a variable rate of depreciation, IVIE Working Paper WP-AD, 2002/07.

McFadden, D., 1989, A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration, Econometrica, 57, 995-1026.

Nadiri, M.I. and I. Prucha, 1993, Estimation of the Depreciation Rate of Physical and R&D Capital in the US Total Manufacturing Sector, National Bureau of Economic Research, WP 4591.

Pakes, A. and D. Pollard, 1989, Simulaton and the Asymptotics of Optimization Estimators, Econometrica, 57, 1027-1057.

Prucha, I., 1995, On the Econometric Estimation of a Constant Rate of Depreciation, Empirical Economics, 20, 299-302.

Schmidt, P., 1976, Econometrics (Marcel Dekker, New York).

Shiller, R., 1973, A Distributed Lag Estimator Derived from Smoothness Priors, Econometrica, 41, 775-778.

Smith, A. (1993) Estimating Nonlinear time Series Models using Simulated Vector Autoregressions", Journal of Applied Econometrics, 8, 63-85.

Staiger, D. and J. Stock, 1997, Instrumental Variables Regression with Weak Instruments, Econometrica, 65, 557-586.

Theil, H and A. Goldberger, 1961, On Pure and Mixed Statistical Estimation in Economics, International Economic Review, 2, 65-78.

Train, K., 2001, A comparison of hierarchical Bayes and maximul simulated likelihood for mixed logit, Working Paper, Department of Economics, University of California, Berkeley.

Train, K., 2003, Discrete Choice Methods with Simulations (Cambridge University Press, Cambridge).

McFadden, D. and K. Train, 2000, Mixed MNL models of discrete response, Journal of Applied Econometrics, 15, 447–470.