

# On the Asymptotics of Stochastic Restrictions

**José A. Hernández**

Department of Applied Economic Analysis, University of Las Palmas de Gran Canaria,  
Spain

Email: joseantonio.hernandez@ulpgc.es

Received 28 March 2016; accepted 31 July 2016; published 3 August 2016

Copyright © 2016 by author and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

**This paper investigates inference methods to introduce prior information in econometric modelling through stochastic restrictions. The goal is to show that stochastic restrictions method estimator can be asymptotically more efficient than the estimator ignoring prior information and can achieve efficiency if prior information grows faster than the sample information in the asymptotics. The set up includes the nonlinear least squares and indirect inference estimators. The paper proposes a new indirect inference estimator that incorporates stochastic equality constraints on the parameters of interest. Finally, the proposed approach is applied to a macroeconomics model where high efficiency gains are shown.**

## Keywords

**Prior Information, Asymptotic Approximation Distribution, Simulation Based Estimation, Nonlinear Models, Capital Stock Estimation, Variable Depreciation Rate**

---

## 1. Introduction

One of the ways in which prior information can be modeled is by the use of the stochastic restrictions approach. The rationale is based on the fact that it brings efficiency gains in the estimators, naturally, subject to the quality of the information available. In some cases, prior information is derived from economic 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 restricted estimator has smaller variance than the non-restricted estimator. In other cases, prior information derives from previous estimations of similar models or samples. This information could be considered an approximation of the unknown parameter, or a range of values that should contain the parameter with some probability. In this case, deterministic restrictions should not be included in the estimation procedure since the restricted estimator will be biased. If this information is not taken into account

despite being valuable, the chance of improving the efficiency of the estimator is wasted. An intermediate solution is to include it with uncertainty. This is the idea behind the stochastic restrictions approach and it is shown to bring efficiency gains (as shown in [1] and [2]) for a linear model under normality of the errors. Nevertheless, stochastic restrictions seem not to have much impact in classic econometrics literature, possibly because of their irrelevance on asymptotic grounds. On the other hand, the Bayesian approach, based on the use of prior information, is increasing its applicability and diffusion in the profession. For instance, related to the mixed logit model (see [3]), the Bayesian approach brings better results in general than the simulated maximum likelihood estimation, mainly due to the prior information that is not considered in the SML method and possibly due to the high variance of the SML estimator resulting from the high number of simulations needed to implement this method ([4] and [5]).

Despite the finite sample efficiency gains of the Theil-Goldberger approach, this result cannot be extended to the asymptotic distributions of the constrained estimator, since the efficiency gain vanishes as sample size increases. This result is proved although not useful for empirical research (see [6]). The reason is that attending at its purposes, asymptotic theory is a tool which provides approximated finite sample distributions of the estimators. If these estimators are constrained and restrictions are correct, then asymptotic theory does not allow the efficient use of all the available information about parameters. Therefore, it would be interesting to extend the finite sample properties of the estimator under stochastic restrictions to the asymptotic context and then to its approximated distribution. In this paper I show that stochastic restrictions could bring asymptotic efficiency gains under some specific assumptions about the asymptotics of prior information.

The main contribution of this paper is the description of a simulation-based estimator in which prior information is taken into account through the stochastic restrictions approach, also, under the same type of assumptions already introduced in the first objective. Simulation-based methods, as the method of simulated moments [7] [8], and the I.I. method of [9] (see also [10] [11], for similar approaches), provide powerful techniques to deal with nonlinear models when traditional methods fail. Nevertheless, there is not a clear way by which prior information can be taken into account in simulation-based estimation methods. In this paper I achieve this goal through the extension of the I.I. method by introducing stochastic restrictions in the initial I.I. criterion, defining the Indirect Inference under Stochastic Restrictions (IIR) estimator and showing efficiency gains when compared to the I.I. estimator.

The structure of the paper is the following: Section 2 provides the motivation for the assumption that supports the main results of this paper, also, I discuss the derived efficiency gains for a traditional estimator in finite sample and asymptotic terms. Section 3 describes the method to combine prior information into the indirect inference criteria through the stochastic restriction approach. Also, asymptotic properties are provided. Section 4 focuses on a macroeconometrics example and numerical evaluation of the efficiency gains of the suggested method, and Section 5 concludes.

## 2. Stochastic Restrictions

The first result to be shown in this paper is that stochastic restrictions yield asymptotic efficiency gains under specific assumptions. Previously, In this section, I provide the definition of stochastic restriction and describe how it behave in asymptotic terms in a standard framework. I show that in order to obtain efficiency gains derived from the introduction of stochastic restrictions it is needed to assume a particular behavior of the prior information in the asymptotic setup. This particular assumption is also motivated in this section.

Consider a general linear model,  $y = x\theta + \varepsilon$ , where the parameters of interest is  $\theta$ , a  $p \times 1$  vector. If prior information is available about  $\theta$ , I could model it as follows:

$$r = G\theta + v \tag{1}$$

which is called a stochastic restriction. In the above equation  $r$  is a  $q \times 1$  vector ( $q < p$ ) containing the values that prior information allocates to a linear combination between parameters,  $G$  is the  $q \times p$  matrix of the parameters coefficients, and  $v$  is a stochastic term that captures the uncertainty about the prior information, for which a distribution is to be assumed.

Let us show how a stochastic restriction could be defined from prior information in a very simple model formed by a Cobb-Douglas production function  $Y = AK^\alpha L^\beta$  using standard macroeconomics notation. Let us assume that available prior information is that “Returns to scale are probably constant”. This means that we expect  $\alpha + \beta$  to be close to one. In this case, the stochastic restriction is the equation  $1 = (\alpha + \beta) + v$ , and  $v$  is a random term whose variance should capture the uncertainty given to the beliefs about the constancy of the

return to scale. In this example  $r = 1$ ,  $q = 1$ , and  $G' = [1, 1]$ . In general, this restriction need not to be linear, and can be denoted as  $G(\theta)$ .

One of the key element of this paper lies on the particular assumption I make about the asymptotics of  $\sigma_v^2$ . In short, I consider an asymptotically decreasing variance of the stochastic restrictions, or, in terms of [12], a parameter sequence. As a result, the relative weights of prior and sample information are preserved in asymptotic terms, which explain efficiency gains as opposed to the standard approach.

This kind of assumption might be considered too strong and, as mentioned in [13], difficult to justify. However, in the context of IV estimation with weak instruments, in [12] and [14] I use a similar assumption, simply justified by the goal of finding better approximations to the finite sample distribution of the estimator of interest. The approximation is derived mainly from standard asymptotic theory, but also, taking into account the extra assumption of a parameter sequence, designed to improve the properties of the considered estimator. Despite of the objection of [13], [12] claims, "... 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<sup>1</sup>. Consequently, there is no need to make such "realistic" assumption ... the quality of the approximation is the only criterion for justifiability". The parameter sequence I choose, as mentioned, is specified on the variance of the stochastic restriction, and using [12] argument, it can also be argued that the rationale lies in the fact that when considered, it makes the asymptotic distribution fit the finite sample distribution better. Yet, in addition, there is a realistic motivation for it: since priors are considered to be obtained from a sample whose size also increases asymptotically, then, it is extended to dynamic terms (defining the asymptotics in terms of both sample sizes) the fact that priors are informative. In other words, our key assumption means that experience matters, which could be considered a natural fact. If priors are informative in static terms, then its quality might increases in the case of additional sampling. Then priors continue to be informative as the size of the sample which generates those increases.

Finally, it is presented an additional argument to motivate the main assumption considered. In a more specific context, the key assumption allows to blend prior and sample information when estimators based on simulation must be used. This is the case of models that generate high nonlinearities in the traditional criterion, making standard methods useless. Generally, the estimators obtained by simulations, despite the fact that are the only solution to estimate some families of models, show high variance, and hence, efficiency gains would be welcomed. The key assumption allows the I.I. procedure becoming a more efficient procedure if stochastic restrictions are correct.

## Efficiency Gains

In this section, I discuss the relevance of taking into account prior information in the estimation of a general nonlinear model. First I remind the properties of the nonlinear least squares estimator ( $\hat{\theta}_{NLS}$ ) and of the estimator that takes into account prior information modelled in the form of stochastic restrictions, called the Nonlinear Least Squares under Stochastic Restriction estimator ( $\hat{\theta}_{SR}$ ). Both estimators are compared in terms of their asymptotic variance covariance matrixes  $AV(\hat{\theta})$ . Second, I show the irrelevance of the prior information when the standard asymptotic analysis is conducted. Nevertheless, I find the opposite result when a more general analysis is carried out, based on assumptions about the asymptotic behavior of  $\sigma_v^2$ , as mentioned in Section 2.

The purpose of this discussion is to establish formally the setting in which the stochastic restrictions are relevant to explain efficiency gains in the context of a traditional method. This formalization is intended to enhance the understanding of the technical role played by the assumption into  $AV(\hat{\theta})$ , that is, the need for a specific  $T/T^*$  convergence. This allows us in the next section to focus the analysis straight on the construction of the restricted indirect inference estimator once it is known the formal role of the assumption.

We start our discussion by considering a general nonlinear model given by the following equation

$$y_t = f(x_t, \theta) + \varepsilon_t \quad (2)$$

where  $y_t$  is a scalar observable random variable,  $\theta$  is a  $p$ -vector of unknown parameters, and  $\{\varepsilon_t\}$  are i.i.d. unobservable random variables such that  $E(\varepsilon_t) = 0$  and  $V(\varepsilon_t) = \sigma_\varepsilon^2$  for all  $t$ . We assume that  $x_t$  is a vector of exogenous variables. If  $\frac{\partial f(x_t, \theta)}{\partial \theta}$  exist and is continuous, the NLS estimator of  $\theta$  is defined as the value  $\hat{\theta}_{NLS}$

<sup>1</sup>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. Also in our case the approximated distribution depends on further sampling.

that minimizes the criterion function  $\Psi_T(\theta) = \sum_{t=1}^T [y_t - f(x_t; \theta)]^2$  with respect to  $\theta$ . Additional assumptions are to be considered—see, for instance, [15]—to prove the consistency and asymptotic normality of the NLS estimator. This assumptions are, in vague terms, the existence and continuity of  $\frac{\partial f(\cdot)}{\partial \theta}$  and the existence of the limit, as  $T$  goes to infinity, of the second order derivative of  $f(\cdot)$  w.r.t.  $\theta$ . For the sample size  $T$  calling  $F_1' = \frac{\partial f(x, \theta)}{\partial \theta}$  –  $p \times T$  matrix –, the asymptotic distribution of the  $\hat{\theta}_{NLS}$  estimator is,

$$\sqrt{T}(\hat{\theta}_{NLS} - \theta_0) \xrightarrow{d} N\left(0, \sigma_\varepsilon^2 p \lim(F_1' F_1 / T)^{-1}\right). \tag{3}$$

Now I consider a set of  $q$  stochastic restrictions on  $\theta$ ,  $q < p$ , modelled through the equation  $r = G(\theta) + v$ , where  $r$  is a  $q \times 1$  vector and  $v \sim N(0, \sigma_v^2 I_q)$ ,  $v$  independent of  $\theta$  and  $\varepsilon^2$ . Following the Theil-Goldberger approach, the resulting model in which Equation (2) and stochastic restrictions are considered, can be expressed as:

$$\begin{bmatrix} y \\ r \end{bmatrix} = \begin{bmatrix} f(x, \theta) \\ G(\theta) \end{bmatrix} + \begin{bmatrix} \varepsilon \\ v \end{bmatrix}. \tag{4}$$

Since, in general,  $\sigma_v^2 \neq \sigma_\varepsilon^2$ , the model shows heteroscedasticity of known structure. Then, the model can be transformed following the Generalized Least Squares (GLS) by premultiplying the system in (4) by matrix  $P$ , the square root of  $V(e, v)'$ . The resulting model, using a compact notation, becomes:

$$\bar{y} = \bar{f}(\bar{x}, \theta) + \bar{\varepsilon} \tag{5}$$

where  $\bar{y}$  is the transformed  $(T + q)$  dimension vector  $P[y, r]'$ ,  $\bar{f}(\bar{x}, \theta)$  is the transformed matrix  $P[f(\cdot) \ G(\cdot)]'$  of (4) and  $\bar{\varepsilon}$  is the  $(T + q)$ —vector of homoscedastic disturbances, since  $V(\bar{\varepsilon}) = I_{T+q}$ .

Some additional notation should be introduced. Let  $\bar{F} = \frac{\partial \bar{f}}{\partial \theta}$  and  $D = \frac{\partial G(\theta)}{\partial \theta}$  a  $q \times p$  matrix. The Nonlinear Least Squares under Stochastic Restriction (SR) estimator of  $\theta$ ,  $\hat{\theta}_{SR}$ , is the NLS estimator of the model (5). In order to simplify the presentation I will omit details of the proofs of the consistency and asymptotic normality of the SR estimator, since these are the same needed for NLS estimator, although in this case, defined on the relevant variables of the transformed model (5) ( see [15]), for instance. We assume that such required assumptions hold in our model, and hence, asymptotic normality and consistency are obtained. We will call this, *Standard Asymptotic Assumptions* (SAA), explicitly omitted. After some computation, assuming SAA on model (5), it can be easily proved that the asymptotic distribution of the SR estimator is:

$$\sqrt{T}(\hat{\theta}_{SR} - \theta_0) \xrightarrow{d} N\left(0, p \lim\left(\frac{1}{T} \frac{F_1' F_1}{\sigma_\varepsilon^2} + \frac{1}{T} \frac{D' D}{\sigma_v^2}\right)^{-1}\right). \tag{6}$$

Our purpose is to compare the asymptotic variance covariance matrix given in (3) and (6), first in the context of the standard asymptotic theory, and also in a general alternative context to be defined, based on the structure of the variance of the error term  $v$ . We start first by the standard asymptotic setting where the following result is obtained, already provided by [6] for a linear model.

**Proposition 1.** Under SAA, the SR and NLS estimators have the same asymptotic variance covariance matrix.

The proof is immediate. Since  $\sigma_v^2$  is constant,  $p \lim\left(\frac{1}{T} \frac{D' D}{\sigma_v^2}\right) = 0$  and then,

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

<sup>2</sup>Although it is not necessarily, neither realistic that the restrictions should be independent and homoscedastic, assuming that make easier the understanding the effects of stochastic restrictions on efficiency gains.

Hence,  $AV(\hat{\theta}_{SR}) = AV(\hat{\theta}_{NLS})$  as shown in (3)<sup>3</sup>.

Proposition 1 shows that stochastic restrictions bring no asymptotic efficiency gains. The irrelevance of the stochastic restrictions is not a satisfactory result for empirical purposes, where the asymptotic distribution has to be used to approximate the variance of the estimator, especially when the sample size is small. The question that arises is whether or not it would be possible to find a theoretical framework to keep the relevance of the stochastic restrictions in asymptotic terms, as stated in [1] for finite samples and normally distributed error term. Also, it is a matter of interest the nature of the conditions under which this theoretical framework would be built up. We give an affirmative answer to the first question, since I obtain in some cases asymptotic efficiency when using stochastic restrictions. Also, I provide an attempt to motivate our assumptions, and to justify such cases.

The new context is based on the idea that prior information about parameters comes from previous experience. Moreover, experience derives from observations that are taken from a sample of size  $T^*$ . Since  $T^*$  increases, prior information consequently improves, that is, is closer to be correct. Observations are generated by a model that is not essentially related with our model of interest, and hence, the disturbance  $v$  is independent of  $\varepsilon$ . The asymptotic results in this new framework are to be defined as  $T$  and  $T^*$  goes to infinity. Besides, since I want to describe a general analysis, the ratio  $T/T^*$  is allowed to vary from zero to infinity in the limit. The following assumptions are in order to formalize the discussion.

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.

This assumption states that the quality of the prior information increases asymptotically with  $T^*$ . The motivation for these assumptions is based on the idea that experience matters as described in Section 2<sup>4</sup>. It should be noted that A01 implies that the term  $r$  in the stochastic restriction is also a random term and hence it depends on  $T^*$ . In order to capture this feature, it is denoted as  $r^*$ . Since A01 states that prior information improves with the sample size  $T^*$ , in the limit  $r^*$  should equal  $R\beta$  and be true.

The asymptotics is analyzed as  $T$  and  $T^*$  goes to infinite, and different growing rates are allowed for  $T$  and  $T^*$ . The following assumptions formalize the setting where the general analysis is carried out.

Assumption 02 (A02).  $\lim(T/T^*) = 1$ .

Assumption 03 (A03).  $\lim(T/T^*) = \infty$ .

Assumption 04 (A04).  $\lim(T/T^*) = 0$ .

The purpose of (A02) is to maintain equal weights of prior information and sample information in the limit. Assumption (A03) states that sample information increases more rapidly than prior information, while (A04) set the opposite. We will show that for the three cases, the variance of the SR vary between an inferior bound, given by the variance of the deterministically restricted estimator (I will simply call this as the restricted estimator and denoted it as  $\hat{\theta}_R$ ) and a superior bound given by the variance of the non-restricted estimator, *i.e.*,  $\hat{\theta}_{NLS}$ . These results are shown in the following propositions, proved [Appendix 1](#).

Proposition 2. Under SAA, (A01) and (A02),  $AV(\hat{\theta}_{SR}) < AV(\hat{\theta}_{NLS})$ .

This result shows that stochastic restrictions brings asymptotic efficiency gains with respect to the NLS estimator when sample and prior information increases at the same rate. In other words, Proposition 2 recovers the Theil-Goldberger contribution for asymptotic distributions, and for the resulting approximated finite sample distributions so derived.

Proposition 3. Under SAA, (A01) and (A03),  $AV(\hat{\theta}_{SR}) = AV(\hat{\theta}_{NLS})$ .

This result shows that stochastic constraints do not increase efficiency when sample information increases more rapidly than prior information. In other words, Proposition 3 shows the standard asymptotic conclusion of Proposition 1 as a particular case of the general analysis described by assumption A01.

Proposition 4. Under SAA, (A01) and (A04),  $AV(\hat{\theta}_{SR}) = AV(\hat{\theta}_R)$ .

This result shows that when prior information increases more rapidly than sample information, stochastic

<sup>3</sup>The suggested approximated distribution for finite sample SR estimator is  $\hat{\theta}_{SR} \approx N(\theta_0, \sigma_\varepsilon^2 (F_1'F_1)^{-1})$ .

<sup>4</sup>As is well known, priors in the Theil-Goldberger approach can also be given the interpretation of the posterior mean of a Bayesian estimator. Following this interpretation, we can also justify A01.

constraints increase efficiency to the level of the restricted estimator  $\hat{\theta}_R$ . In other words, in the limit, stochastic constraints reach the maximum level of information and efficiency. Then Proposition 4 shows the standard finite sample conclusion obtained in a deterministic constraint setting as a particular case of the general analysis described by assumption A01. The equation of the  $AV(\hat{\theta}_R)$  is provided, together with the proof, in [Appendix 1](#).

Finally, in the following proposition I show a concluding result for a varying  $\lim T/T^*$  from zero to infinity.

**Proposition 5.** Under (A01), as  $\lim(T/T^*)$  increases from zero to infinity,  $AV(\hat{\theta}_{SR})$  also increase from  $AV(\hat{\theta}_R)$  to  $V(\hat{\theta}_{NLS})$ . That is, if  $\lim(T/T^*) = M$ ,  $0 < M < \infty$ , then,

- 1)  $\frac{d(AV(\hat{\theta}_{SR}))}{dM} > 0$ ,
- 2)  $\lim_{M \rightarrow 0} AV(\hat{\theta}_{SR}) = AV(\hat{\theta}_R)$ ,
- 3)  $\lim_{M \rightarrow \infty} AV(\hat{\theta}_{SR}) = AV(\hat{\theta}_{NLS})$ .

We have established a general setting in which several goals are covered. First I have stated a unique analytical context to explain restricted and non-restricted estimators, in the general terms of the stochastic restrictions approach. In this context, restricted and non-restricted estimators are particular cases of  $\hat{\theta}_{SR}$ , depending on the relative dynamic of sample and prior information. Therefore, when sample information dominates prior information, the general SR estimator leads to the non-restricted estimator. When prior information dominates sample information, the general SR estimator leads to the restricted estimator. The described discussion in terms of the  $T/T^*$  ratio could also be interpreted as a way of looking at the variance of the stochastic restriction. For finite sample size, it is easy to see that the same results could be found when  $\sigma_v^2$  varies from 0 to infinity<sup>5</sup>.

### 3. Indirect Inference under Stochastic Restrictions

The indirect inference method is a simulation-based moment matching estimation procedure. The general idea is to match the moments of the auxiliary model from the simulated data to observed data to obtain the estimates of the structural parameters. The method of Indirect Inference (I.I.) of [9] and the methods of simulated moments of [11] and [16] (see similar methods in [10] and [17]), provide a powerful technique to deal with nonlinear models where traditional methods fail. In spite of the wide applicability of these methods, there is not a methodology to take into account prior information in their implementation (see, for example, [18]). In this section I suggest a way to solve this problem based on the stochastic restriction approach and also on the discussed asymptotic efficiency gain of stochastic restrictions. The analysis will be cast in the framework of the I.I., since this methodology is more general and other simulation-based estimation methods can be viewed as special cases of it. Therefore, notation will follow as closely as possible [9]. The general goal of this section is to provide an example of applicability of the results shown in Section 3, where asymptotic efficiency resulting from stochastic restriction could be theoretically justified. Moreover, this example has empirical implications, since it provides efficiency gains to simulation-based estimators, whose variance is generally high.

First, I define the Indirect Inference under Stochastic Restrictions (IIR) estimation method and provide its distribution. Then, based on the approach introduced in Section 2, I show that the IIR estimator is more efficient than the I.I. method, provided that the stochastic restrictions are asymptotically correct.

In the I.I. approach it is considered a  $p$ -dimension vector of parameters  $\theta$  of a model of interest ( $M$ ), given by a set of  $T$  equations of the form  $y_i = f(x_i, \theta) + \varepsilon_i$  and a  $j$ -dimension vector of parameters  $\beta$ ,  $j \geq p$  of an auxiliary model ( $M^a$ ) given by equations  $y_i = g(x_i, \beta) + u_i$ , which is easier to handle than the model of interest.

Some facts have to be pointed out in order to understand the principle of the I.I. estimation. It is assumed that it is not feasible to estimate  $M$  by mean of a conventional method, due to its complexity or intractability of a conventional criterion for that model. On the other hand,  $M^a$  estimation is feasible by using a traditional method, based on the optimization of a criterion function  $\Psi_T(\beta)$  or auxiliary criterion<sup>6</sup>. As of the optimization on  $\Psi_T(\beta)$  it can be obtained the estimate  $\hat{\beta}$ , denoted as  $\hat{\beta}(\theta)$  since the estimate depends indirectly on the implicit parameters  $\theta$  driving the data generating process. The binding function  $b(\theta)$  is defined as the

<sup>5</sup>This result can be shown on requested to the author.

<sup>6</sup>This function could be, for instance, the likelihood function, as considered in the example described in Section 5.

parametric counterpart of the point estimate  $\hat{\beta}(\theta)$  for a given  $\theta$ . The I.I. estimation focuses on the estimation of the function  $b(\theta)$  from  $\hat{\beta}(\theta)$  for a set of values of  $\theta$ , which generates by simulation independent paths of  $y^s(\theta)$ . The function  $b(\theta)$  is assumed to be twice differentiable with respect to  $\theta$ , and  $D_1 = \left. \frac{\partial b(\theta)}{\partial \theta} \right|_{\theta=\theta_0}$  is of full rank on a neighborhood of  $\theta_0$ . The I.I. estimator of  $\theta$ , following [9], is defined as

$$\tilde{\theta}_{II} = \arg \min \{m'_1 \Omega_1 m_1\}$$

where  $m_1 = \hat{\beta} - \sum_{s=1}^D \hat{\beta}_s(\theta)/S$  and  $\Omega_1$  is a  $j \times j$  symmetric and positive definite matrix to be determined below. Under regular assumptions about the auxiliary criterion  $\Psi_T(\beta)$  and the model—in [Appendix 2](#) these 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

$$W_1(\Omega_1) = (D'_1 \Omega_1 D_1)^{-1} (D'_1 \Omega_1 \Omega_1^{*-1} \Omega_1 D_1) (D'_1 \Omega_1 D_1)^{-1}$$

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 [Appendix 2](#).

The matrix  $\Omega_1$  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  $S \rightarrow \infty$ ) is

$$W_1^* = (D'_1 \Omega_1^* D_1)^{-1}. \tag{7}$$

Since  $\Psi_\infty = p \lim_{T \rightarrow \infty} \Psi_T$ ,  $\Psi_T$  could be used in the place of  $\Psi_\infty$  and a consistent estimator of the asymptotic variance of  $\hat{\beta}$  for  $H_0$ .

We now consider the existence of prior information on the parameters of interest  $\theta$ , what could formally be written as stochastic restrictions:  $r^* = G(\theta) + v$ , where  $E(v) = 0$  and  $V(v) = \Phi_2^*$ ,  $v$  independent of  $\varepsilon$  and  $u$  and the parameters of  $M$  and  $M^a$ . Further properties of  $v$  are to be specified below. The vector  $r^*$  contains the priors about the parameter constraints, and  $V(v)$  is chosen according to the quality of the prior information.

Function  $G(\theta)$ , that defines constraints between parameters, is differentiable and such that  $D_2 = \frac{\partial G(\theta)}{\partial \theta}$  is a  $q \times p$  matrix of full rank in a neighborhood of  $\theta_0$ .

It is necessary to introduce some additional notation to define the proposed estimation method. Let  $m' = \left(m'_1, (G - r^*)'\right)$ ,  $D = (D_1, D_2)'$ , and  $\Omega$  a block diagonal matrix, with  $\Omega_1$ ,  $\Omega_2$  in their diagonal respectively.

**Definition** The Indirect Inference under Stochastic Restriction (IIR) estimator of  $\theta$  is

$$\tilde{\theta}_{IIR} = \arg \min_{\theta} \{m'(\theta) \Omega m(\theta)\} \tag{8}$$

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 **Appendix 2**.

$$(A8) \quad \frac{\partial G(\theta)}{\partial \theta} = D_2 \text{ is a } q \times p \text{ matrix of full rank in a neighborhood of } \theta_0.$$

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

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

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. Again, the rationale behind (A9) is the intention to maintain a constant relative weight between the sample and prior information asymptotically. The relevance of this assumption lies in the fact, already discussed, that under these 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<sup>7</sup>. The asymptotic properties of the IIR estimator are derived next.

**Proposition 6** Under assumptions (A1) to (A10),  $\tilde{\theta}_{IIR}$ , 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

$$W(\Omega) = [D_1' \Omega_1 D_1 + D_2' \Omega_2 D_2]^{-1} \times [D_1' \Omega_1 \Omega_1^{*-1} \Omega_1 D_1 + D_2' \Omega_2 \Omega_2^{*-1} \Omega_2 D_2] \\ \times [D_1' \Omega_1 D_1 + D_2' \Omega_2 D_2]^{-1}.$$

This result is proved in **Appendix 2**.

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:

$$W^* = (D' \Omega^* D)^{-1} \tag{9}$$

where  $\Omega^*$  is the block diagonal matrix with  $\Omega_1^*$  and  $\Omega_2^*$  in the diagonal.

**Proposition 7** Under assumptions (A1) to (A10)  $\tilde{\theta}_{IIR}$  is asymptotically more efficient than  $\tilde{\theta}_{II}$ .

To proof this result, I compare Equation (7) and Equation (9). 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.

## 4. Empirical Implementation

This section conducts a set of empirical estimations to assess the performance, in terms of bias and efficiency, of the estimation method described in Section 4 compared to the I.I. method.

The model of interest is given by a production function and the perpetual inventory method equation for the capital stock,  $K$ , which depends on the depreciation rate,  $\delta$ , which is considered to be variable. Hence, the considered model consist basically on the main structure of a growth model. A specific feature of the model is that the capital stock in unknown as it depends on the depreciation rate which is an unknown parameter. The

<sup>7</sup>We can also provide in this section a general setting depending of the limit of  $T/T^*$ , as done in Section 3. The arguments and conclusions will be the same. If  $\lim T/T^* = \infty$ ,  $AV(\hat{\theta}_{IIR}) = AV(\hat{\theta}_{II})$  and no efficiency gains are derived. If  $\lim T/T^* = 0$ ,  $AV(\hat{\theta}_{IIR})$  reaches its minimum level.



interest of estimating this model is to obtain as a by-product, estimates of the depreciation rate and hence of the physical capital stock of the economy. Three cases are considered for estimating the depreciation rate, as in [19] where  $\delta$  is function of certain explanatory variables and a purely random term.

To go further into the economic motivation of the model, note that  $K$  is one of the basic economic aggregates, and following the definition provided by the perpetual inventory method, it is given by:

$$K_t = I_t + (1 - \delta)K_{t-1} \quad (10)$$

where  $I$  is investment and  $\delta$  the depreciation rate, which measures the loss in value of the existing capital stock as it ages. Since  $\delta$  is an unknown parameter,  $K$  is not observable and in practice it is usually measured by accounting techniques, which provides not satisfactory figures since, for instance, technological shocks have not effects on the actual value of the net capital stock. One solution to measure the capital stock is by mean of the simultaneous estimation of  $\delta$  jointly with the parameters of a production function,  $Y = F(K, L)$ . That is, the estimation of the production function provides, as a by-product, estimates of a variable depreciation rate, which will allow the measurement of the capital stock. If  $\delta$  is not a constant parameter, then the described methodology to estimate it is not trivial, as can be seen in [19]-[22]. Moreover, since  $\delta$  is assumed to be stochastic, its estimation poses methodological difficulties which are not solvable by standard methods. In this case, simulation-based estimation methods seems to be adequate to solve the high complexity of the model, considerably increased by the presence of a stochastic parameter. As commonly known, the resulting high variance of the estimator seems to be one of the most important *costs* of using these methods, but on the other hand, as previously mentioned, possibly compensated by the availability of prior information. The purpose of this section is to suggest the adequacy of the IIR method to estimate the described model, given the feasibility of priors about  $\delta$ . In particular, prior information about the rate of depreciation is available from other sources (e.g. National Accounts, or estimates deriving from similar models), which could help in the estimation of the production function, since convergence of the algorithms is possibly hard to achieve and estimation is costly in efficiency terms. According to this argument, I use the average estimates of  $\delta$  obtained in similar models as the prior information figures to implement the IIR method.

The theoretical model of interest is given by a Cobb-Douglas production function, and assuming constant returns to scale becomes:

$$y_t = c_0 + \alpha k_t + (1 - \alpha)l_t + \varepsilon_t \quad (11)$$

where  $y$ ,  $l$  and  $k$  are production, labour and capital stock in logs respectively,  $\alpha$  is the elasticity of the capital stock and it is assumed that  $\varepsilon_t \sim N^{iid}(0, \sigma_\varepsilon^2)$ . The capital stock is given by

$$K_t = I_t + (1 - \delta_t)K_{t-1} \quad (12)$$

In the above equation,  $\delta$  is a time dependent parameter and different assumptions are considered about its nature. Three cases are put forward depending on the deterministic part of the depreciation rate: a constant, a dummy variable and the growing rate of  $Y_t$ , in order to follow as closely as possible the empirical models considered in [19] where a similar model is estimated, although in that case  $\delta_t$  has no random term. **Table 1** shows the main characteristics of the depreciation rate stochastic processes. In case I, the random depreciation rate is a constant plus a disturbance. In cases II and III there is an explanatory variable in the rhs of the equation of  $\delta_t$ , following closely the baseline model. The disturbance term is assumed to be  $\varepsilon_t \sim N^{iid}(0, \sigma_\varepsilon^2)$ , independent of  $\varepsilon_t$  and the whole set of explanatory variables.

Besides, the model introduces a prior value  $\tilde{\delta}$  available for the expected sequence  $\delta_t$ . The prior value is taken as the average of the estimated variable depreciation rate in [19], which is  $\tilde{\delta} = 0.06$  for all of the cases. Finally, in order to make the  $\delta_t$  equation consistent with the stochastic restriction, this is expressed as:

$$\tilde{\delta} = \gamma_0 + \gamma_1 \bar{z} + v \quad (13)$$

where  $\bar{z}$  is the sample mean of the explanatory variable of  $\delta_t$  ( $z = 0, D_{2t}$ , or  $\frac{\Delta Y_t}{Y_{t-1}}$  for cases I, II and III respectively) and  $v$  is the error term capturing the uncertainty about  $\tilde{\delta}$ . Specifically, the quality of prior information is determined by the value of  $\sigma_v$ , which is taken to be 0.015, since it leads a plausible approximated interval [0.03, 0.09] for priors on  $\delta$ .

**Table 1.** Patterns for the depreciation rate.

Case	I. Constant	II. Dummy variable	III. Growing Rate of Y
$\delta$ equation	$\delta_t = \gamma_0 + \epsilon_t$	$\delta_t = \gamma_0 + \gamma_1 D_{2t} + \epsilon_t$	$\delta_t = \gamma_0 + \gamma_1 \left( \frac{\Delta Y_t}{Y_{t-1}} \right) + \epsilon_t$
Stochastic restriction	$\tilde{\delta} = \gamma_0 + v$	$\tilde{\delta} = \gamma_0 + \gamma_1 \bar{D}_2 + v$	$\tilde{\delta} = \gamma_0 + \gamma_1 \bar{Y} + v$

### IIR Estimation

The empirical model of interest is formed by the following equations:

$$\begin{aligned}
 (I) \quad & y_t = c_0 + c_1 D_{1t} + \alpha k_t + (1 - \alpha) l_t + \epsilon_t \\
 (II) \quad & K_t = I_t + (1 - \delta_t) K_{t-1} \\
 (III) \quad & \delta_t = \begin{cases} \gamma_0 + \epsilon_t & \text{(Case I)} \\ \gamma_0 + \gamma_1 D_{2t} + \epsilon_t & \text{(Case II)} \\ \gamma_0 + \gamma_1 \left( \frac{\Delta Y_t}{Y_{t-1}} \right) + \epsilon_t & \text{(Case III)} \end{cases} \\
 (IV) \quad & \tilde{\delta} = \gamma_0 + \gamma_1 \bar{z} + v
 \end{aligned} \tag{14}$$

where data requires  $\epsilon_t$  to follow an AR(1) process, modelled as  $\epsilon_t = \rho \epsilon_{t-1} + \varpi_t$ ,  $\varpi_t \sim N(0, \sigma_\varpi^2)$ . The autoregressive structure  $\epsilon_t$  could be understood as a result of total factor productivity shock<sup>8</sup>. It should be noted that data allows for a structural change for the intercept, captured by  $c_1$ . The breaking period is the one which leads to the best fit among all possible dates. As in the baseline estimation (see [19]),  $D_{1t} = 0$  if  $t \leq 1986$  and  $D_{1t} = 1$  elsewhere. For Case II  $D_2$  is a dummy variable, which becomes empirically relevant, and given by  $D_{2t} = 0$  if  $t \leq 1989$  and  $D_{2t} = 1$  elsewhere. The breaking period is picking selected the best fit among all possible dates. For the Case III equation, the rate explanatory variable  $\frac{\Delta Y_t}{Y_{t-1}}$  is considered in order to check the role of the intensiveness in using the capital stock in the depreciation pattern. Finally, it is assumed that  $\epsilon_t \sim N(0, \sigma_\epsilon^2)$ .

The parameter vector of model (14) is  $\theta = (c_0, c_1, \alpha, \rho, \sigma_\epsilon^2, \gamma_0, \gamma_1, \sigma_\epsilon^2)$ , —  $\gamma_1$  only appears in Cases II and III— which is estimated by IIR using data of the variables  $y$ ,  $l$ ,  $I$ ,  $D_1$ ,  $D_2$  and  $z$  for the sample period 1970-1997. Non-residential investment is considered for the estimation. The data are taken from the Spanish National Statistical Institute and are measured at 1986 prices.

The auxiliary criterion is maximum likelihood and the auxiliary model for the IIR estimation is exactly the same model considered in [19], which is much closed to the model of interest, being in this case the depreciation rate deterministic and no restrictions imposed by the existence of returns to scale into the production function. Henceforth, the equations of the auxiliary empirical model are:

$$\begin{aligned}
 (I') \quad & y_t = \beta_1 + \beta_2 D_{1t} + \beta_3 k_t + \beta_4 l_t + u_t \\
 (II') \quad & K_t = I_t + (1 - \delta_t) K_{t-1} \\
 (III') \quad & \delta_t = \begin{cases} \beta_7 & \text{(Case I)} \\ \beta_7 + \beta_8 D_{2t} & \text{(Case II)} \\ \beta_7 + \beta_8 \left( \frac{\Delta Y_t}{Y_{t-1}} \right) & \text{(Case III)} \end{cases}
 \end{aligned} \tag{15}$$

<sup>8</sup>Following real business cycle models, a standard assumption for technology consists of a time trend and an innovation which follows an AR(1) process. In this case the error in the production function follows an AR(1), what supports our results.

where  $u_t$  follows an AR(1) process, given by  $u_t = \beta_3 u_{t-1} + w_t$ ,  $w_t \sim N^{iid}(0, \sigma_w^2)$  and  $V(u_t) = \beta_6$ . Again the random error is considered to follow an AR(1), capturing the total factor productivity dynamics and yielding more accurate estimates. Finally, the parameter vector of the auxiliary model (15) is  $\beta = (\beta_1, \dots, \beta_8)$ , being  $\beta_8$  only estimated for cases II and III.

The motivation for structure of the auxiliary model relies on the fact that it is a more simple model, since no random term is considered in the equation of the variable rate of depreciation, and, on the other hand, it is a more general model, since no constant returns to scale are imposed in the production function. Very little can be said in priors grounds about the adequacy of one specific model to be the best auxiliary model for I.I. estimating. Nevertheless, it is in general admitted that the model should be similar, and if possible, more general. Both of this characteristics are considered in the selection of the model considered, which is also supported by the empirical results.

As defined in the previous section, the IIR estimator of  $\theta$  is given by:

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

being  $Q_T$ , for the specification of the models given in (14) and (15):

$$Q_T = \left[ \hat{\beta} - \sum \hat{\beta}_s(\theta)/S \right]' \Omega_1^* \left[ \hat{\beta} - \sum \hat{\beta}_s(\theta)/S \right] + \frac{\left[ E(\tilde{\delta}) - \tilde{\delta} \right]^2}{\sigma_v^2}$$

where  $\hat{\beta}$  and  $\hat{\beta}_s$  are the ML estimators of the auxiliary model obtained from original and simulated data respectively,  $S$  is the number of simulations, taken to be 100,  $\tilde{\delta}$  is the prior value of  $\delta$ ,  $\sigma_v^2$  the figure capturing the uncertainty about the prior information and  $\Omega_1^*$  is the optimal distance matrix, also needed for computing  $AV(\tilde{\theta}_{IIR})$ . The term  $E(\tilde{\delta})$  is given by the specific equation of the stochastic restriction in each case and  $\tilde{\delta}$  and  $\sigma_v$  are taken to be 0.06 and 0.0015 respectively, as proved in [Appendix 2](#),

$$AV(\tilde{\theta}_{IIR}) = \left( 1 + \frac{1}{S} \right) \left[ D_1' \Omega_1^* D_1 + D_2' \Omega_2^* D_2 \right]^{-1}$$

where  $D_1$ , and the matrices  $J_0$  and  $\bar{H}_0$  included in  $\Omega_1^*$  are computed using the numerical derivatives involved in the definition, that is  $\frac{\partial \beta}{\partial \theta}$ , the Hessian of the auxiliary criterion and the outer product of gradient.

The ratio  $T/T^*$  was taken to be equal to one. The reason for this is that since it is numerically equivalent to consider different  $\sigma_v$  or  $T/T^*$ , to test the sensibility of the results I simply considered alternative values of  $\sigma_v$ . Naturally, the estimates of  $\gamma_0$  obtained were closed to  $\tilde{\delta}$  as  $\sigma_v$  decreases, although  $\sigma_v < 0.015$  leads to priors for  $\delta$  weakly supported from previous estimates.

**Table 2** shows the estimates results obtained for all of the cases. Each one of the models has been estimated simultaneously by I.I. and IIR using the same simulation path, in order to test for the efficiency gains and consistency of the results in a more direct way. In all of the cases and for both methods, **Table 2** points estimates of the intercept, capital elasticity, coefficient of the AR(1) error term and the variance of the error are fairly closed to those found in the baseline model estimates of [19], although not always statistically significant.

The point estimate of 0.3 for  $\alpha$  is very general admitted for production function estimates. The constant depreciation rate is estimated at 4.5% and 4.0%, for IIR and II respectively, which is higher than the constant and non-stochastic depreciation rate estimate of 3.7% found in the baseline deterministic model. This value is still low when compared to conventional values and more complex specifications seem in order. The remaining results pursue this point. In Case II, columns 4 and 5 give the results obtained allowing for a dummy variable,  $D_{2t}$  which is statistically significant and points to an increased depreciation rate in the second sub sample, yielding an average value for the whole sample close to 6.5%. These are more reasonable results, as discussed above, and again larger vales estimates than in the deterministic  $\delta$  model.

Columns 6 and 7 give the results for Case III, and the coefficient  $\gamma_1$  is positive for both methods and significant only for IIR method. This would mean that an increase in the GDP growth rate increases the depreciation rate, which is consistent result with the fact that an increase in aggregate demand explains an intensive use of capital and then an increase in its depreciation rate.

In a more general setting, **Table 2** contains several key findings. First, both I.I. and IIR generate estimates

**Table 2.** IIR and I.I. estimates<sup>1</sup>.

Case	I: $\delta_t = \gamma_0 + \epsilon_t$		II: $\delta_t = \gamma_0 + \gamma_1 D_{2t} + \epsilon_t$		III: $\delta_t = \gamma_0 + \gamma_1 \frac{\Delta Y_t}{Y_{t-1}} + \epsilon_t$	
Method	IIR	I.I.	IIR	I.I.	IIR	I.I.
$c_0$	5.035 (4.83)	5.022 (4.03)	5.199 (3.14)	5.116 (3.40)	5.166 (3.73)	5.167 (3.21)
$c_1$	0.040 (1.69)	0.040 (1.45)	0.0403 (1.97)	0.039 (1.76)	0.041 (1.73)	0.041 (1.58)
$\alpha$	0.291 (5.04)	0.300 (4.30)	0.2975 (6.22)	0.298 (6.22)	0.297 (1.72)	0.291 (1.20)
$\rho$	0.200 (1.75)	0.200 (1.62)	0.201 (1.50)	0.204 (1.45)	0.206 (2.42)	0.206 (2.20)
$\sigma_\epsilon^2$	$1.1 \times 10^{-4}$ (9.3)	$1.1 \times 10^{-4}$ (9.3)	$1.1 \times 10^{-4}$ (9.5)	$1.1 \times 10^{-4}$ (9.5)	$1.0 \times 10^{-4}$ (10.01)	$1.0 \times 10^{-4}$ (9.1)
$\gamma_0$	0.045 (1.94)	0.040 (1.68)	0.065 (2.65)	0.071 (1.50)	0.067 (2.05)	0.070 (1.70)
$\gamma_1$	-	-	0.010 (2.53)	0.010 (2.47)	0.009 (2.18)	0.008 (1.38)
$\sigma_c^2$	0.017 (0.42)	0.016 (0.03)	0.016 (0.15)	0.017 (0.15)	0.015 (0.55)	0.015 (0.35)

<sup>1</sup>t-values in brackets.

with very little difference from the baseline model estimates which contains no stochastic depreciation rate. This result allows for confidence in terms of bias and adequacy of the simulation-based methods for the estimation of this specific model, although not significant differences are found for the estimates of the parameters underlying the variable depreciation rate.

Second, IIR is more efficient than I.I., which is shown for the parameter for which prior information is available. In fact, efficiency losses are small provided that I use conservative choices for the variance of the stochastic restriction. Alternative estimations were conducted for different quality levels of the prior information, confirming that efficiency losses are inversely related to the quality of the prior information.

Third, in the implementation of the IIR method, convergence is achieved faster than for the I.I. estimation, which shows that the proposed methodology is a practical way to mix prior and sample information in a simulation-based estimation method. On the other hand, preliminary results suggest that by reducing the number of simulations (say, to 50), it will be possible to reduce the computation time of IIR without adversely affecting its finite sample properties.

## 5. Conclusions

This paper formalizes some intuitions about the role of prior information on asymptotic rules of inference. In particular, the natural idea that despite prior information is asymptotically irrelevant when modeled through stochastic restrictions, this theoretical result may not avoid using accurate prior information for empirical purposes. Nevertheless, so far there is no any contribution in the literature providing ground for it.

Asymptotic theory is a tool that provides approximate figures for the mean and the variance-covariance matrix of estimators that in general may have an empirical interest, that is, may be one of the few practical solutions to estimate a model of interest. Nevertheless, if prior information is irrelevant in asymptotic terms, it will be so in the derived finite sample approximation of the variance of such estimator. This result of course is not

helpful and leads to discard any use of prior information even knowing that prior information in general may be relevant if accurate-in terms of efficiency. This paper is intended to provide an insight in the previous discussion in the sense that if prior information is proved to be asymptotically relevant, then it will also be for the finite sample approximation and thus will bring efficiency gains on empirical ground. This previous discussion is the motivation of this paper and the solution I provide may be understood as a contribution oriented to enhance the usefulness of any estimator as in asymptotic terms there is no room for using prior information in the form of stochastic restrictions.

On the other hand it is worth it to recall the large variance of the I.I. estimator (as well as of others simulation based estimators). This additional setup provides specific motivation to face the challenge of providing theoretical ground for the asymptotic efficiency gains due to stochastic restrictions.

The main contribution, which is the formulation of a new estimator (the IIR estimator), more efficient than the baseline estimator is achieved through the introduction of one specific assumption, which in short is that prior information increases with sample size. This idea, the cornerstone of the suggested approach, is intended to be taken as a potential contribution for the large family of simulation based estimators in the sense that they are now allowed to mix sample and prior information to achieve efficiency gains.

As expected, this discussion is open for future research as empirical results that may be found for testing this insight, may support it or not.

## References

- [1] Theil, H. and Goldberger, A. (1961) On Pure and Mixed Statistical Estimation in Economics. *International Economic Review*, **2**, 65-78. <http://dx.doi.org/10.2307/2525589>
- [2] Shiller, R. (1973) A Distributed Lag Estimator Derived from Smoothness Priors. *Econometrica*, **41**, 775-778. <http://dx.doi.org/10.2307/1914096>
- [3] McFadden, D. and Train, K. (2000) Mixed MNL Models for Discrete Response. *Journal of Applied Econometrics*, **15**, 447-470. [http://dx.doi.org/10.1002/1099-1255\(200009/10\)15:5<447::AID-JAE570>3.0.CO;2-1](http://dx.doi.org/10.1002/1099-1255(200009/10)15:5<447::AID-JAE570>3.0.CO;2-1)
- [4] Train, K. (2001) A Comparison of Hierarchical Bayes and Maximum Simulated Likelihood for Mixed Logit. Working Paper, Department of Economics, University of California, Berkeley.
- [5] Train, K. (2003) Discrete Choice Methods with Simulations. Cambridge University Press, Cambridge. <http://dx.doi.org/10.1017/CBO9780511753930>
- [6] Lütkepohl, H. (1993) Introduction to Multiple Time Series Analysis. Springer Verlag, Berlin. <http://dx.doi.org/10.1007/978-3-642-61695-2>
- [7] McFadden, D. (1989) A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration. *Econometrica*, **57**, 995-1026. <http://dx.doi.org/10.2307/1913621>
- [8] Pakes, A. and Pollard, D. (1989) Simulation and the Asymptotics of Optimization Estimators. *Econometrica*, **57**, 1027-1057. <http://dx.doi.org/10.2307/1913622>
- [9] Gourieroux, C., Montfort, A. and Renault, A. (1993) Indirect Inference. *Journal of Applied Econometrics*, **8**, 85-118. <http://dx.doi.org/10.1002/jae.3950080507>
- [10] Smith, A. (1993) Estimating Nonlinear Time Series Models Using Simulated Vector Autoregressions. *Journal of Applied Econometrics*, **8**, 63-85. <http://dx.doi.org/10.1002/jae.3950080506>
- [11] Lee, B. and Ingram, B. (1991) Simulation Estimation of Time Series Models. *Journal of Econometrics*, **47**, 197-207. [http://dx.doi.org/10.1016/0304-4076\(91\)90098-X](http://dx.doi.org/10.1016/0304-4076(91)90098-X)
- [12] Bekker, P. (1994) Alternative Approximations to the Distributions of Instrumental Variable Estimators. *Econometrica*, **62**, 657-681. <http://dx.doi.org/10.2307/2951662>
- [13] Kadane, J. (1971) Comparison of k-Class Estimator when the Disturbances Are Small. *Econometrica*, **39**, 723-737. <http://dx.doi.org/10.2307/1909575>
- [14] Staiger, D. and Stock, J. (1997) Instrumental Variables Regression with Weak Instruments. *Econometrica*, **65**, 557-586. <http://dx.doi.org/10.2307/2171753>
- [15] Amemiya, T. (1985). *Advanced Econometrics*. Harvard University Press, Cambridge.
- [16] Duffie, D. and Singleton, K. (1993) Simulated Moments Estimation of Markov Models of Asset Prices. *Econometrica*, **61**, 929-1052. <http://dx.doi.org/10.2307/2951768>
- [17] Gallant, R. and Tauchen, G. (1996) Which Moments to Match? *Econometric Theory*, **12**, 657-81. <http://dx.doi.org/10.1017/S026646660006976>

- [18] Canova, F. (1994) Statistical Inference in Calibrated Models. *Journal of Applied Econometrics*, **9**, 123-145. <http://dx.doi.org/10.1002/jae.3950090508>
- [19] Hernández, J.A. and Mauleón, I. (2005) Econometric Estimation of a Variable Rate of Depreciation of the Capital Stock. *Empirical Economics*, **3**, 575-595. <http://dx.doi.org/10.1007/s00181-004-0234-4>
- [20] Prucha, I.R. (1995) On the Econometric Estimation of a Constant Rate of Depreciation. *Empirical Economics*, **20**, 299-302. <http://dx.doi.org/10.1007/BF01205440>
- [21] Nadiri, M.I. and Prucha, I. (1993) Estimation of the Depreciation Rate of Physical and R&D Capital in the US Total Manufacturing Sector. NBER, WP 4591.
- [22] Prucha, I.R. and Nadiri, M.I. (1996) Endogenous Capital Utilization and Productivity Measurement in Dynamic Factor Demand Models. Theory and Application to the U.S. Electrical Machinery Industry. *Journal of Econometrics*, **71**, 343-379. [http://dx.doi.org/10.1016/0304-4076\(94\)01709-3](http://dx.doi.org/10.1016/0304-4076(94)01709-3)
- [23] Golub, G.H. and Van Loan, C.F. (1989) *Matrix Computations*. Johns Hopkins University Press, Baltimore.

## Appendix 1: Asymptotic Variance Covariance Matrix of the SR Estimator

We set the following assumptions to prove Propositions 2 to 5.

(A01).  $V(v) = \sigma_v^{*2} = \sigma_v^2/T^*$  .

(A02).  $p \lim(T/T^*) = 1$ .

(A03).  $p \lim(T/T^*) = \infty$ .

(A04).  $p \lim(T/T^*) = 0$ .

**Proposition 2.** Under SAA, (A01) and (A02),  $AV(\hat{\theta}_{SR}) < AV(\hat{\theta}_{NLS})$ .

**Proof.** Under (A01), by construction, the asymptotic distribution of  $\hat{\theta}_{SR}$  (6) is now given by

$$\sqrt{T}(\hat{\theta}_{SR} - \theta_0) \xrightarrow{d} N\left(0, p \lim\left(\frac{1}{T} \frac{F_1'F_1}{\sigma_\varepsilon^2} + \frac{1}{T} \frac{D'D}{\sigma_v^{*2}}\right)^{-1}\right). \quad (16)$$

From (A02),  $p \lim(T/T^*) = 1$ , then,

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

and

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

is a definite negative matrix, and  $AV(\hat{\theta}_{SR}) < AV(\hat{\theta}_{NLS})$ . The resulting approximated distribution for finite sample size is

$$\hat{\theta}_{SR} \approx N\left(\theta_0, \left(\frac{F_1'F_1}{\sigma_\varepsilon^2} + \frac{D'D}{\sigma_v^2}\right)^{-1}\right)$$

what means that efficiency gains are also extended to finite sample distributions.

**Proposition 3.** Under SAA, (A01) and (A03),  $AV(\hat{\theta}_{SR}) = AV(\hat{\theta}_{NLS})$ .

**Proof.** To prove this proposition I use the general form of the *Sherman-Morrison-Woodbury* formula (see [23]), which is

$$(A + BCB')^{-1} = A^{-1} - A^{-1}B(C^{-1} + B'A^{-1}B)^{-1}B'A^{-1} \quad (17)$$

where  $A$  and  $C$  are  $k \times k$  matrixes such that  $A^{-1}$ ,  $C^{-1}$  exist, and  $B$  is a  $k \times q$  matrix. Now, taking

$$A = \frac{1}{\sigma_\varepsilon^2} \frac{F_1'F_1}{T},$$

$$B = R' \text{ and}$$

$$C = I_q \frac{1}{T} \frac{1}{\sigma_v^{*2}} \stackrel{(A01)}{=} I_q \frac{T^*}{T} \frac{1}{\sigma_v^2}.$$

From the distribution given in (16) and taking into account Equation (17), I can rewrite  $AV(\hat{\theta}_{SR})$  as

$$\begin{aligned} AV(\hat{\theta}_{SR}) &= p \lim \left\{ \left( \frac{F_1'F_1}{T} \frac{1}{\sigma_\varepsilon^2} \right)^{-1} - \left( \frac{F_1'F_1}{T} \frac{1}{\sigma_\varepsilon^2} \right)^{-1} R' \left[ I_q \sigma_v^2 \frac{T}{T^*} + R \left( \frac{F_1'F_1}{T} \frac{1}{\sigma_\varepsilon^2} \right)^{-1} R' \right]^{-1} R \left( \frac{F_1'F_1}{T} \frac{1}{\sigma_\varepsilon^2} \right)^{-1} \right\} \\ &= p \lim \left( \frac{F_1'F_1}{\sigma_\varepsilon^2} \right)^{-1} = AV(\hat{\theta}_{NLS}) \end{aligned}$$

since by (A03)  $p \lim \frac{T}{T^*} = \infty$ , and then,  $\lim \left[ I_q \sigma_v^2 \frac{T}{T^*} + R \left( \frac{F_1' F_1}{T} \frac{1}{\sigma_\varepsilon^2} \right)^{-1} R' \right]^{-1} = 0$ .

**Proposition 4.** Under SAA, (A01), and (A04),  $AV(\hat{\theta}_{SR}) = AV(\hat{\theta}_R)$ , where  $\hat{\theta}_R$  is the restricted nonlinear estimator of the model.

*Proof.* From the rewritten equation of  $AV(\hat{\theta}_{SR})$ ,

$$AV(\hat{\theta}_{SR}) = p \lim \left[ \left( \frac{F_1' F_1}{T} \frac{1}{\sigma_\varepsilon^2} \right)^{-1} - \left( \frac{F_1' F_1}{T} \frac{1}{\sigma_\varepsilon^2} \right)^{-1} R' \left[ I_q \sigma_v^2 \frac{T}{T^*} + R \left( \frac{F_1' F_1}{T} \frac{1}{\sigma_\varepsilon^2} \right)^{-1} R' \right]^{-1} R \left( \frac{F_1' F_1}{T} \frac{1}{\sigma_\varepsilon^2} \right)^{-1} \right]. \quad (18)$$

Since (A04) states that  $p \lim T/T^* = 0$ , then,

$$\begin{aligned} p \lim \left[ I_q \sigma_v^2 \frac{T}{T^*} + R \left( \frac{F_1' F_1}{T} \frac{1}{\sigma_\varepsilon^2} \right)^{-1} R' \right]^{-1} &= \left[ p \lim \left( I_q \sigma_v^2 \frac{T}{T^*} + R \left( \frac{F_1' F_1}{T} \frac{1}{\sigma_\varepsilon^2} \right)^{-1} R' \right) \right]^{-1} \\ &= \left[ R p \lim \left( \frac{F_1' F_1}{T} \frac{1}{\sigma_\varepsilon^2} \right)^{-1} R' \right]^{-1} \end{aligned}$$

and, by substituting the above equation into the equation of  $AV(\hat{\theta}_{SR})$ , I have

$$\begin{aligned} AV(\hat{\theta}_{SR}) &= p \lim \left[ \left( \frac{F_1' F_1}{T} \frac{1}{\sigma_\varepsilon^2} \right)^{-1} - \left( \frac{F_1' F_1}{T} \frac{1}{\sigma_\varepsilon^2} \right)^{-1} R' \left[ R \left( \frac{F_1' F_1}{T} \frac{1}{\sigma_\varepsilon^2} \right)^{-1} R' \right]^{-1} R \left( \frac{F_1' F_1}{T} \frac{1}{\sigma_\varepsilon^2} \right)^{-1} \right] \\ &= p \lim \left[ \left( \frac{F_1' F_1}{T} \frac{1}{\sigma_\varepsilon^2} \right)^{-1} - \left( \frac{F_1' F_1}{T} \frac{1}{\sigma_\varepsilon^2} \right)^{-1} R' \left[ R \left( \frac{F_1' F_1}{T} \right)^{-1} R' \right]^{-1} R \left( \frac{F_1' F_1}{T} \right)^{-1} \right] \end{aligned}$$

which is easily checked to be the asymptotic variance covariance matrix of the NLS estimator of the model

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

which is the restricted model.

**Proposition 5.** Under (A01), and as  $\lim(T/T^*)$  increases from zero to infinity,  $AV(\hat{\theta}_{SR})$  also increases from  $AV(\hat{\theta}_R)$  to  $V(\hat{\theta}_{NLS})$ . That is, if  $\lim(T/T^*) = M$ ,  $0 < M < \infty$ , then,

- 1)  $\frac{d(AV(\hat{\theta}_{SR}))}{dM} > 0$ ,
- 2)  $\lim_{M \rightarrow 0} AV(\hat{\theta}_{SR}) = AV(\hat{\theta}_R)$ ,
- 3)  $\lim_{M \rightarrow \infty} AV(\hat{\theta}_{SR}) = AV(\hat{\theta}_{NLS})$ .

*Proof.* Taking limits in the term where  $T/T^*$  appears into the Equation (18) I have

$$\left[ I_q \sigma_v^2 M + R p \lim \left( \frac{F_1' F_1}{T} \frac{1}{\sigma_\varepsilon^2} \right)^{-1} R' \right]^{-1} = H(M) \text{ (say)} \quad (19)$$

and going back to the Equation (18), I have

$$AV(\hat{\theta}_{SR}) = p \lim \left( \frac{F_1' F_1}{T} \frac{1}{\sigma_\varepsilon^2} \right)^{-1} - p \lim \left( \frac{F_1' F_1}{T} \frac{1}{\sigma_\varepsilon^2} \right)^{-1} R' H(M) R^{-1} R p \lim \left( \frac{F_1' F_1}{T} \frac{1}{\sigma_\varepsilon^2} \right)^{-1}.$$



Since  $\frac{dH(M)}{dM} < 0$ , then  $\frac{dAV(\hat{\theta}_{SR})}{dM} > 0$  what proves 1).

From (19),

$$\lim_{M \rightarrow 0} H(M) = \left[ Rp \lim \left( \frac{F_1' F_1}{T \sigma_\varepsilon^2} \right)^{-1} R' \right]^{-1}$$

$$\lim_{M \rightarrow \infty} H(M) = 0$$

then, by substituting the above results into (18), I obtain  $\lim_{M \rightarrow 0} AV(\hat{\theta}_{SR}) = AV(\hat{\theta}_R)$ , and

$\lim_{M \rightarrow \infty} AV(\hat{\theta}_{SR}) = AV(\hat{\theta}_{NLS})$ , what proves 2) and 3).

## Appendix 2: Asymptotic Distribution of IIR Estimator 3 mm

Here I develop similar proofs to the used on the asymptotic properties of the I.I. estimator. To show the asymptotic distribution of IIR estimator I 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  $\beta$ , and this maximum is  $b(\theta)$ . That is,  $b(\theta) = \arg \max_\beta \{\psi_\infty(\theta, \beta)\}$

A3)  $\psi_T$  and  $\psi_\infty$  are differentiable with respect to  $\beta$  and  $p \lim \frac{\partial \psi_T}{\partial \beta} = \frac{\partial \psi_\infty}{\partial \beta}$ .

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

A5)  $p \lim -\frac{\partial^2 \psi_T(b(\theta))}{\partial \beta \partial \beta'} = -\frac{\partial^2 \psi_\infty(b(\theta))}{\partial \beta \partial \beta'} = J_0$  and  $J_0^{-1}$  exist.

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

A7)  $\lim_{T \rightarrow \infty} cov \left[ \sqrt{T} \frac{\partial \psi_T(y_T^{s_1})}{\partial \beta}, \sqrt{T} \frac{\partial \psi_T(y_T^{s_2})}{\partial \beta} \right] = K_0$  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)  $\lim_{T, T^* \rightarrow \infty} \frac{T}{T^*} = 1$ .

Let us first prove the consistency of the IIR estimator. Under assumptions (A1) to (A4), following [9] 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),  $r^* \rightarrow G(\theta_0)$ . Then,

$$\hat{\theta}_{IIR} = \arg \min_\theta \{m'(\theta) \Omega m(\theta)\} = \arg \min_\theta \left\{ \left[ \hat{\beta} - \hat{\beta}_{ST}(\theta) \right]' \Omega_1 \left[ \hat{\beta} - \hat{\beta}_{ST}(\theta) \right] + (G(\theta) - r^*)' \Omega_2 (G(\theta) - r^*) \right\}$$

$$\rightarrow \left\{ \left[ b(\theta_0) - b(\theta) \right]' \Omega_1 \left[ \hat{\beta} - \hat{\beta}_{ST}(\theta) \right] \left[ G(\theta) - G(\theta_0) \right]' \Omega_2 \left[ G(\theta) - G(\theta_0) \right] \right\}$$

$$= \left\{ \theta : b(\theta) = b(\theta_0), G(\theta) = G(\theta_0) \right\} \text{ (since } \Omega_1, \Omega_2 \text{ are positive definite)}$$

$$= \theta_0 \text{ (since } b \text{ and } G \text{ are well defined).}$$

Let us now find the asymptotic distribution of  $\hat{\theta}_{HR}$ . 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 [9])

$$\sqrt{T}(\hat{\beta}_T - b(\theta_0)) = J_0^{-1} \sqrt{T} \frac{\partial \Psi_T(b(\theta_0))}{\partial \beta} + o_p(1) \tag{20}$$

and

$$\sqrt{T}(\hat{\beta}_{ST}(\theta_0) - b(\theta_0)) = \frac{J_0^{-1}}{S} \sqrt{T} \sum_{s=1}^S \frac{\partial \Psi_T(y_T^s, b(\theta_0))}{\partial \beta} + o_p(1). \tag{21}$$

The asymptotic expansion of  $\hat{\theta}_{HR}(\Omega)$  is deduced as follows. The first order condition for  $\hat{\theta}_{HR}$  ( $\hat{\theta}$  for short) from the criterion (8) is:

$$\begin{aligned} & -\frac{\partial \hat{\beta}'_{ST}[\hat{\theta}(\Omega_1)]}{\partial \theta} \Omega_1 [\hat{\beta}_T - \hat{\beta}_{ST}(\hat{\theta}(\Omega_1))] \\ & + \frac{\partial G'(\hat{\theta})}{\partial \theta} \Omega_2 [G(\hat{\theta}) - r^*] \\ & = 0. \end{aligned}$$

An expansion around the limit value  $\theta_0$  gives

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

since  $S_1 = o_p(1)$ , as shown in the asymptotic properties of the  $\hat{\theta}_H$  estimator under the considered assumptions. From (A9) and the consistency of  $\hat{\theta}_{HR}$ , it follows that  $S_2 + S_3$  is also  $o_p(1)$ . Rewriting the above equation in

the limit, and calling  $D_1 = \left. \frac{\partial b(\theta)}{\partial \theta} \right|_{\theta=\theta_0}$  and  $D_2 = \frac{\partial G(\theta_0)}{\partial \theta}$ ,

$$\begin{aligned} \sqrt{T}(\hat{\theta}_{HR}(\Omega) - \theta_0) &= [D_1' \Omega_1 D_1 + D_2' \Omega_2 D_2]^{-1} D_1' \Omega_1 \sqrt{T} [\hat{\beta}_T - \hat{\beta}_{ST}(\theta_0)] \\ & - [D_1' \Omega_1 D_1 + D_2' \Omega_2 D_2]^{-1} D_2' \Omega_2 \frac{\sqrt{T}}{\sqrt{T^*}} \sqrt{T^*} [G(\theta_0) - r^*] + o_p(1). \end{aligned} \tag{22}$$

From (21), (20), I 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):

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

where

$$\begin{aligned}
W(\Omega) &= [D_1'\Omega_1D_1 + D'\Omega_2D_2]^{-1} D_1'\Omega_1\Omega_1^{*-1}\Omega_1D_1 \times [D_1'\Omega_1D_1 + D'\Omega_2D_2]^{-1} \\
&\quad + [D_1'\Omega_1D_1 + D'\Omega_2D_2]^{-1} D_2'\Omega_2\Omega_2^{*-1}\Omega_2D_2 \times [D_2'\Omega_2D_2 + D'\Omega_2D_2]^{-1} \\
&= [D_1'\Omega_1D_1 + D'\Omega_2D_2]^{-1} \times [D_1'\Omega_1\Omega_1^{*-1}\Omega_1D_1 + D_2'\Omega_2\Omega_2^{*-1}\Omega_2D_2] \\
&\quad \times [D_1'\Omega_1D_1 + D'\Omega_2D_2]^{-1}.
\end{aligned}$$

The optimal  $\Omega_i$  is  $\Omega_i^*$  by the Gauss-Markov theorem. The AVC matrix of the IIR estimator, taking  $S \rightarrow \infty$ , is:

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



Scientific Research Publishing

**Submit or recommend next manuscript to SCIRP and we will provide best service for you:**

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.

A wide selection of journals (inclusive of 9 subjects, more than 200 journals)

Providing 24-hour high-quality service

User-friendly online submission system

Fair and swift peer-review system

Efficient typesetting and proofreading procedure

Display of the result of downloads and visits, as well as the number of cited articles

Maximum dissemination of your research work

Submit your manuscript at: <http://papersubmission.scirp.org/>