A spatial stochastic frontier model with spillovers:
Evidence for Italian Regions (1970-1993)

Efthymios G. Tsionas

Lancaster University, UK
m.tsionas@lancaster.ac.uk

Panayotis G. Michaelides[*]

National Technical University of Athens, GR
pmichael@central.ntua.gr

Abstract

Efficiency measurement using stochastic frontier models is well established in applied econometrics. However, no published work seems to be available on efficiency analysis using spatial data dealing with possible spatial dependence between regions. This paper considers a stochastic frontier model with decomposition of inefficiency into an idiosyncratic and a spatial, spillover component. Exact posterior distributions of parameters are derived, and computational schemes based on Gibbs sampling with data augmentation are proposed to conduct simulation-based inference and efficiency measurement. The new method is illustrated using production data for Italian regions (1970-1993). Clearly, further theoretical and empirical research on the subject would be of great interest.

Key Words: Spatial models; efficiency; spillover effects; stochastic frontiers; Bayesian analysis; Gibbs sampling.

JEL Codes: C11, C15, R15.

---

[*] Corresponding author at: Laboratory of Theoretical and Applied Economics; School of Applied Mathematical and Physical Sciences; National Technical University of Athens; Heroon Polytechneiou 9, 157.80; Zografou Campus; Greece; Tel.: +302107721624; Fax: +30 2107721618; E-mail: pmichael@central.ntua.gr (P.G. Michaelides).

A spatial stochastic frontier model with spillovers:
Evidence for Italian Regions (1970-1993)[1]

Abstract

Efficiency measurement using stochastic frontier models is well established in applied econometrics. However, no published work seems to be available on efficiency analysis using spatial data dealing with possible spatial dependence between regions. This paper considers a stochastic frontier model with decomposition of inefficiency into an idiosyncratic and a spatial, spillover component. Exact posterior distributions of parameters are derived, and computational schemes based on Gibbs sampling with data augmentation are proposed to conduct simulation-based inference and efficiency measurement. The new methodis illustrated using production data for Italian regions (1970-1993). Clearly, further theoretical and empirical research on the subject would be of great interest.

Key Words: Spatial models; efficiency; spillover effects; stochastic frontiers; Bayesian analysis; Gibbs sampling.

JEL Codes: C11, C15, R15.

---

1. Introduction

Efficiency measurement using stochastic frontier models is well established in applied econometrics and numerous applications have seen the light of day in agricultural economics, health, marketing *etc*. Stochastic frontiers have been introduced by Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977) and have found numerous applications. For a review of methods and problems see Bauer (1990) and the excellent reviews by Greene (1993, 2001). Extensions of stochastic frontier models have been considered by Greene (1990), Stevenson (1990) and Tsionas (2000, 2002).

Although much work has been done in connection with stochastic frontier models using cross section or panel data, no published work seems to be available on efficiency analysis using spatial data. In fact, countless studieshave been conducted applying stochastic frontier models to regional data, but none – to the best of our knowledge – deals with possible spatial dependence between regions. Exactly this has been the point of departure of our investigation.

Spatial data pose difficult problems because the possibility of spatial correlation must be taken into account (Anselin, 1988 and Anselin and Bera, 1998). This is even more so when one-sided error terms have to be accounted for in order to provide efficiency measures, and the purpose of the present paper is to explore these problems and provide practical methods for likelihood-based inference.

Probably the most appropriate way of addressing spatial dependencein the context of regional inefficiency is through a spatially structured random effects (r.e.) specification. This is actually what the below referenced model is about. In fact, the proposed model sets forth a (latent) random effects vectorthat is specified to follow a spatial autoregressive process.

More precisely, the proposed approach allows for traditional (log-) linear production function relationships between outputs and inputs, while allowing for spatial spillovers in the latent r.e. that measure regional inefficiency. There are a number of papers which adopt this type of spatial regression specification, including Smith &LeSage (2004) who use spatially (autoregressive) structured latent effects in a Probit model; Allenby & McCulloch (2005), and Parent &LeSage (2012) who derive the log-marginal likelihood to be used in a Bayesian modeling averaging context. However, a serious complication that arises when estimating this spatially structured

r.e. model in the context of stochastic frontier analysis is the one-sided nature of the (latent) r.e. vector.

There is a scarcity of works dealing with spatial correlation in the broader context of maximum likelihood estimation or panel data factor models. In this context, see Bai (2009) who considered large panel data models with unobservable multiple interactive effects correlated with the regressors, and Pesaran(2006) whodealt with inference in panel data models with a general multifactor error structure, where the unobserved factors and the individual-specific errors were allowed to follow arbitrary stationary processes. The basic idea there was to filter the individual-specific regressors by means of cross-section averages such that asymptotically as the cross-section dimension tends to infinity, the differential effects of unobserved common factors are eliminated.

Probably, the main advantage of our approach compared to models such as the ones mentioned above is the possibility of mixing different pieces of information (sample information, prior information, etc), in order to construct a model that accounts for the stochastic character of the variables leading to a better approximation of reality. Analytically, the main reason for using a Bayesian approach is that it facilitates representing and taking fuller account of the uncertainties related to model and parameter values. In contrast, most decision analyses based on maximum likelihood or least squares estimation involve fixing the values of parameters that may, in actuality, have an important bearing on the final outcome of the analysis and for which there is considerable uncertainty. One of the major benefits of the Bayesian approach is the ability to incorporate prior information.

Hence, some advantages to using Bayesian analysis include the following: (a) it provides a natural and principled way of combining prior information with data, within a solid decision theoretical framework. In this context, it is possible to incorporate past information about a parameter and form a prior distribution for future analysis. When new observations become available, the previous posterior distribution can be used as a prior; (b) It provides inferences that are conditional on the data, without reliance on asymptotic approximation. Small sample inference proceeds in the same manner as if one had a large sample. Bayesian analysis also can estimate any functions of parameters directly; (c) It obeys the likelihood principle. If two distinct sampling designs yield proportional likelihood functions, then all inferences should be identical from these two designs. Classical inference does not, in

general, obey the likelihood principle; (d) It provides interpretable answers; and (e) It provides a convenient setting for a wide range of models, including spatial econometrics problems. Finally, MCMC, along with other numerical methods, makes computations tractable for virtually all parametric models. See, *inter alia,* Carlin and Lewis (2000); and Robert (2001).

Analytically, the paper develops a model that allows for spatial correlation and spillover effects in inefficiency. The starting point is the usual linear model with a one-sided error term allowing for panel data. The one-sided error term, however, consists of two components, namely an idiosyncratic component, and a spatial or spillover component that relates a region's inefficiency to the inefficiency of neighboring regions. The likelihood function is derived, and Bayesian methods of inference and efficiency measurement are proposed and illustrated in the context of a production function for Italian regions over the period 1970-1993. The production function depends on labor, physical capital, and human capital.

2. The model

In this work, we consider a stochastic frontier model with a decomposition of inefficiency into a spillover component, and an idiosyncratic component. More specifically, consider the specification:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{v} - \mathbf{u} \otimes \mathbf{1}_T$$

$$\mathbf{v} \sim N(0, \sigma_v^2 \mathbf{I}_{Tn})$$

$$\mathbf{u} = \rho \mathbf{W}\mathbf{u} + \boldsymbol{\varepsilon}, \ \mathbf{u} \geq \mathbf{0}$$

$$\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma_\varepsilon^2 \mathbf{I}_n), \ \boldsymbol{\varepsilon} \geq \mathbf{0}.$$

Vector $\mathbf{y}$ is $Tn \times 1$, and $\mathbf{y} = [\mathbf{y}_1, ..., \mathbf{y}_n]'$ where $\mathbf{y}_i$ is the $T \times 1$ vector containing all observations of the dependent variable for the ith region, $\mathbf{X}$ is the $Tn \times k$ matrix of observations for the regressors, $\mathbf{v}$ is $Tn \times 1$ noise, and $\mathbf{u}$ is the $n \times 1$ vector of one-sided error terms. Notice that each block $\mathbf{y}_i$ contains all observations for a given region. This model allows for technical inefficiency. More specifically, inefficiency is decomposed into two components: The first component is idiosyncratic and is given by the one-sided random variable $\boldsymbol{\varepsilon}$. The other component $\rho \mathbf{W}\mathbf{u}$ is spatial and reflects

4

regional spillover effects for any given regional weighting matrix $\mathbf{W}$. Overall inefficiency is the sum of the spatial and idiosyncratic component. The assumption that inefficiency is time invariant is not essential and can be removed. The formulation here is adopted mostly in order to deal in a reasonable and effective way with panel data.

An alternative strategy here would be to follow Parent and LeSage (2012) who simplify things by assuming that $\varepsilon \sim N(0, \sigma^2)$, where $\Lambda \equiv \text{diag}(\lambda_1, \ldots, \lambda_n)$ is introduced to distinguish between the disturbances. Each variance scalar λ_i $i=1,\ldots,n$ is assumed inversely proportional to the number of neighboringregions. This would allow to analytically integrate out the latent r.e. parameters and produce a log-marginal likelihood expression that requires univariate numerical integration over only the spatial dependence parameter ρ. However, despite its eloquent simplicity, this approach seems to be too restrictive here.[2]

Define $\mathbf{\Phi} = \mathbf{I}_n - \rho\mathbf{W}$. Using standard notation and change of variables we have:

$$p(\mathbf{u}) = C_{\mathbf{\Phi}}^{-1} (\tfrac{\pi}{2}\sigma_\varepsilon^2)^{-n/2} \, |\det(\mathbf{\Phi})| \exp\left(-\frac{\mathbf{u}'\mathbf{\Phi}'\mathbf{\Phi}\mathbf{u}}{2\sigma_\varepsilon^2}\right) \cdot I(\mathbf{u} \geq \mathbf{0}, \mathbf{\Phi}\mathbf{u} \geq 0)$$

where $C_{\mathbf{\Phi}} = (\tfrac{\pi}{2}\sigma_\varepsilon^2)^{-n/2} \, |\det(\mathbf{\Phi})| \displaystyle\int_{\{\mathbf{u}:\mathbf{\Phi}\mathbf{u}\geq\mathbf{0},\mathbf{u}\geq\mathbf{0}\}} \exp\left(-\frac{\mathbf{u}'\mathbf{\Phi}'\mathbf{\Phi}\mathbf{u}}{2\sigma_\varepsilon^2}\right) d\mathbf{u}$ is the normalizing

constant of the distribution, $C_{\mathbf{\Phi}}$ depends on $\mathbf{\Phi}$ and thus it cannot be ignored. This constant, however, cannot be computed analytically. Its computation is known as evaluation of multivariate normal probabilities. See Kotz, Balakrishnan, and Johnson (2000, chapter 45, section 5.1).

It has to mentioned that this computation is fairly difficult, and it cannot be reduced to evaluation of multivariate *orthant* probabilities. For orthant probabilities the relatively simple approximation due to Solow (1990) could have been used but in the present case further simplification does not appear possible. The computation and role of $C_{\mathbf{\Phi}}$ will be discussed later on.

Adopting the same prior as before, the kernel posterior distribution is:

---

[2]We would like to thank an anonymous referee for pointing this out.

$$p(\boldsymbol{\beta},\sigma_v,\sigma_\varepsilon,\rho\,|\,\mathbf{y},\mathbf{X}) \propto C_{\boldsymbol{\Phi}}^{-1}\sigma_v^{-(n+\underline{n}+1)}\sigma_\varepsilon^{-(n+\underline{n}+1)}\,|\det(\boldsymbol{\Phi})|\cdot$$

$$\int\limits_{\{\mathbf{u}:\boldsymbol{\Phi}\mathbf{u}\geq 0,\mathbf{u}\geq 0\}} \exp\left[-\frac{(\mathbf{y}+\mathbf{u}\otimes\mathbf{1}_T-\mathbf{X}\boldsymbol{\beta})'(\mathbf{y}+\mathbf{u}\otimes\mathbf{1}_T-\mathbf{X}\boldsymbol{\beta})+\underline{q}_v}{2\sigma_v^2}\right]\cdot\exp\left[-\frac{\mathbf{u}'\boldsymbol{\Phi}'\boldsymbol{\Phi}\mathbf{u}+\underline{q}_u}{2\sigma_\varepsilon^2}\right]d\mathbf{u}$$

Define $\mathbf{e}=\mathbf{X}\boldsymbol{\beta}-\mathbf{y}$, and $\boldsymbol{\Sigma}^{-1}=\boldsymbol{\Phi}'\boldsymbol{\Phi}$. The integral above is evaluated as follows:

$$\int\limits_{\{\mathbf{u}:\boldsymbol{\Phi}\mathbf{u}\geq 0,\mathbf{u}\geq 0\}} \exp\left[-\frac{(\mathbf{y}+\mathbf{u}\otimes\mathbf{1}_T-\mathbf{X}\boldsymbol{\beta})'(\mathbf{y}+\mathbf{u}\otimes\mathbf{1}_T-\mathbf{X}\boldsymbol{\beta})}{2\sigma_v^2}\right]\cdot\exp\left[-\frac{\mathbf{u}'\boldsymbol{\Phi}'\boldsymbol{\Phi}\mathbf{u}}{2\sigma_\varepsilon^2}\right]d\mathbf{u}=$$

$$\int\limits_{\{\mathbf{u}:\boldsymbol{\Phi}\mathbf{u}\geq 0,\mathbf{u}\geq 0\}} \exp\left[-\tfrac{1}{2}\left\{\frac{(\mathbf{u}\otimes\mathbf{1}_T-\mathbf{e})'(\mathbf{u}\otimes\mathbf{1}_T-\mathbf{e})}{2\sigma_\varepsilon^2}+\frac{\mathbf{u}'\boldsymbol{\Sigma}^{-1}\mathbf{u}}{2\sigma_\varepsilon^2}\right\}\right]d\mathbf{u}=$$

$$\int\limits_{\{\mathbf{u}:\boldsymbol{\Phi}\mathbf{u}\geq 0,\mathbf{u}\geq 0\}} \exp\left[-\tfrac{1}{2}\frac{\mathbf{u}'\left(T\sigma_\varepsilon^2\mathbf{I}_n+\sigma_v^2\boldsymbol{\Sigma}^{-1}\right)\mathbf{u}-2\sigma_\varepsilon^2\mathbf{u}'\sum_{t=1}^{T}\mathbf{e}_t+\sigma_\varepsilon^2\mathbf{e}'\mathbf{e}}{\sigma_v^2\sigma_\varepsilon^2}\right]d\mathbf{u}$$

Define $\boldsymbol{\Omega}=\sigma_v^2\sigma_\varepsilon^2(T\sigma_\varepsilon^2\mathbf{I}_n+\sigma_v^2\boldsymbol{\Sigma}^{-1})^{-1}$, and $\boldsymbol{\mu}_*=\boldsymbol{\Omega}\bar{\mathbf{e}}/(T\sigma_v^2)$, with $\bar{\mathbf{e}}=T^{-1}\sum_{t=1}^{T}\mathbf{e}_t$ or in matrix form $\bar{\mathbf{e}}=T^{-1}(\mathbf{I}_n\otimes\mathbf{1}_T')\mathbf{e}$. The above integral reduces to:

$$\int\limits_{\{\mathbf{u}:\boldsymbol{\Phi}\mathbf{u}\geq 0,\mathbf{u}\geq 0\}} \exp\left[-\tfrac{1}{2}\left\{(\mathbf{u}-\boldsymbol{\mu}_*)'\boldsymbol{\Omega}^{-1}(\mathbf{u}-\boldsymbol{\mu}^*)-\boldsymbol{\mu}_*'\boldsymbol{\Omega}^{-1}\boldsymbol{\mu}_*+\frac{\mathbf{e}'\mathbf{e}}{\sigma_v^2}\right\}\right]d\mathbf{u}\propto$$

$$\left[\det(\boldsymbol{\Omega})\right]^{1/2}\exp\left[-\tfrac{1}{2}\left(\frac{\mathbf{e}'\mathbf{e}}{\sigma_v^2}-\boldsymbol{\mu}_*'\boldsymbol{\Omega}^{-1}\boldsymbol{\mu}_*\right)\right]$$

Therefore, the kernel posterior distribution is

$$p(\boldsymbol{\beta},\sigma_v,\sigma_\varepsilon,\rho\,|\,\mathbf{y},\mathbf{X})\propto$$

$$\sigma_v^{-(Tn+\underline{n}+1)}\sigma_\varepsilon^{-(n+\underline{n}+1)}C_{\boldsymbol{\Phi}}^{-1}\,|\det(\boldsymbol{\Phi})|\cdot\left[\det(\boldsymbol{\Omega})\right]^{1/2}\exp\left[-\tfrac{1}{2}\left(\frac{\mathbf{e}'\mathbf{e}}{\sigma_v^2}-\boldsymbol{\mu}_*'\boldsymbol{\Omega}^{-1}\boldsymbol{\mu}_*\right)\right]\cdot\exp\left[-\tfrac{1}{2}\left(\frac{\underline{q}_v}{\sigma_v^2}+\frac{\underline{q}_\varepsilon}{2\sigma_\varepsilon^2}\right)\right]$$

We have $\mathbf{e}'\mathbf{e}=(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})=(n-k)s^2+(\boldsymbol{\beta}-\mathbf{b})\mathbf{X}'\mathbf{X}(\boldsymbol{\beta}-\mathbf{b})$, where $\mathbf{b}=(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$, and $(Tn-k)s^2=(\mathbf{y}-\mathbf{Xb})'(\mathbf{y}-\mathbf{Xb})$ are the usual least squares quantities. Note that s^2 does not depend on parameter ρ. Next, we have

$$p(\sigma_v,\sigma_\varepsilon,\rho\,|\,\mathbf{y},\mathbf{X})\propto\sigma_v^{-(n-k+\underline{n}+1)}\sigma_\varepsilon^{-(n+\underline{n}+1)}C_{\boldsymbol{\Phi}}^{-1}\,|\det(\boldsymbol{\Phi})|\cdot\left[\det(\boldsymbol{\Omega})\right]^{1/2}$$

$$\exp\left[-\tfrac{1}{2}\left(\frac{(n-k)s^2+\underline{q}_v}{\sigma_v^2}-\boldsymbol{\mu}_*'\boldsymbol{\Omega}^{-1}\boldsymbol{\mu}_*\right)\right]\cdot\exp\left[-\tfrac{1}{2}\left(\frac{\underline{q}_\varepsilon}{\sigma_\varepsilon^2}\right)\right]$$

Integrating with respect to $\boldsymbol{\beta}$ using properties of the multivariate normal distribution, we obtain:

$$p(\sigma_v, \sigma_\varepsilon, \rho \,|\, \mathbf{y}, \mathbf{X}) \propto$$

$$\sigma_v^{-(Tn-k+\underline{n}+1)} \sigma_\varepsilon^{-(n+\underline{n}+1)} C_\Phi^{-1} \,|\det(\boldsymbol{\Phi})| \cdot \left[\det(\boldsymbol{\Omega})\right]^{1/2} \exp\left[-\frac{1}{2}\left(\frac{(Tn-k)s^2 + \underline{q}_v}{\sigma_v^2} - \boldsymbol{\mu}_*'\boldsymbol{\Omega}^{-1}\boldsymbol{\mu}_*\right)\right] \cdot \exp\left(-\frac{1}{2}\frac{\underline{q}_\varepsilon}{\sigma_\varepsilon^2}\right)$$

This kernel posterior is three-dimensional, and could be maximized to obtain a quasi-Bayes estimator. Integrating with respect to the variance parameters using numerical integration in two dimensions would produce the marginal posterior distribution of ρ namely $p(\rho \,|\, \mathbf{y}, \mathbf{X})$. The posterior is, however, highly nonlinear and conducting inferences using this posterior may involve awkward numerical problem. Hence, the three-dimensional conditional distribution could be simplified to only two dimensions, by fixing the variance ratio $\sigma_v/\sigma_\varepsilon$ as is often the case and robustness of the results with respect to various assumed ratios could be checked. However, in what follows we adopt a relevant approach which overcomes the need for fixing the aforementioned variance ratio.

We consider the augmented posterior:

$$p(\boldsymbol{\beta}, \sigma_v, \sigma_\varepsilon, \rho, \mathbf{u} \,|\, \mathbf{y}, \mathbf{X}) \propto \sigma_v^{-(Tn+\underline{n}+1)} \sigma_\varepsilon^{-(n+\underline{n}+1)} \,|\det(\boldsymbol{\Phi})| \cdot C_\Phi^{-1} \cdot$$

$$\exp\left[-\frac{(\mathbf{y} + \mathbf{u} \otimes \mathbf{1}_T - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} + \mathbf{u} \otimes \mathbf{1}_T - \mathbf{X}\boldsymbol{\beta}) + \underline{q}_v}{2\sigma_v^2}\right] \cdot \exp\left[-\frac{\mathbf{u}'\boldsymbol{\Phi}'\boldsymbol{\Phi}\mathbf{u} + \underline{q}_u}{2\sigma_\varepsilon^2}\right]$$

The following conditional distributions required for Gibbs sampling can be derived:

$$\boldsymbol{\beta} \,|\, \sigma_v, \sigma_\varepsilon, \rho, \mathbf{u}, \mathbf{y}, \mathbf{X} \sim N\left((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{y} + \mathbf{u} \otimes \mathbf{1}_T), \sigma_v^2(\mathbf{X}'\mathbf{X})^{-1}\right),$$

$$\frac{(\mathbf{y} + \mathbf{u} \otimes \mathbf{1}_T - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} + \mathbf{u} \otimes \mathbf{1}_T - \mathbf{X}\boldsymbol{\beta}) + \underline{q}_v}{\sigma_v^2} \,|\, \boldsymbol{\beta}, \sigma_\varepsilon, \rho, \mathbf{u}, \mathbf{y}, \mathbf{X} \sim \chi^2(Tn+\underline{n}),$$

$$\frac{\mathbf{u}'\boldsymbol{\Phi}'\boldsymbol{\Phi}\mathbf{u} + \underline{q}_u}{\sigma_\varepsilon^2} \,|\, \boldsymbol{\beta}, \sigma_v, \rho, \mathbf{u}, \mathbf{y}, \mathbf{X} \sim \chi^2(n+\underline{n}).$$

Drawing from these distributions involves standard operations. With respect to $\boldsymbol{\beta}$ we impose the constraints that input elasticities are positive at the means of data and L>0 other points symmetrically placed around the means, within the range of the data.

With L=10 all computed output elasticities turned out positive when evaluated at the observed data.

The conditional posterior distribution of ρ is given by:

$$p(\rho \mid \boldsymbol{\beta}, \sigma_v, \sigma_\varepsilon, \mathbf{u}, \mathbf{y}, \mathbf{X}) \propto \mid \det(\boldsymbol{\Phi}) \mid C_{\boldsymbol{\Phi}}^{-1} \exp\left[-\frac{\mathbf{u}'\boldsymbol{\Phi}'\boldsymbol{\Phi}\mathbf{u} + \underline{q}_u}{2\sigma_\varepsilon^2}\right] p(\rho),$$

truncated to the interval given by the inverse of minimum and the maximum eigenvalues of the spatial matrix to ensure stability of the system. Since the conditional posterior is non-standard, we resort to griddy Gibbs sampling to generate random numbers from this distribution. The troublesome part in these operations is the evaluation of the multivariate normal probability involved in $C_{\boldsymbol{\Phi}}^{-1}$. One alternative is to simply ignore this term, in which case, however, results will be approximate. The simplest alternative is to compute this term using Monte Carlo techniques.

In other words, we generate a number of vectors $\mathbf{u}$ from $N_n(\mathbf{0}, \boldsymbol{\Phi}'\boldsymbol{\Phi})$ and count how many times the restrictions $\mathbf{u} \geq \mathbf{0}, \boldsymbol{\Phi}\mathbf{u} \geq \mathbf{0}$ are satisfied. One important property of this problem is that matrix $\boldsymbol{\Phi}$ depends only on ρ for any given regional weighting matrix $\mathbf{W}$. Therefore, a simplification would be to compute $C_{\boldsymbol{\Phi}}^{-1}$ over a range of ρ values in advance of any other computations, and then use interpolation to compute the value of $C_{\boldsymbol{\Phi}}^{-1}$ at the intermediate values that will be required during the course of computations. This is the strategy adopted in the present paper.

The conditional posterior distribution of latent inefficiency variables is:

$$p(\mathbf{u} \mid \boldsymbol{\beta}, \sigma_v, \sigma_\varepsilon, \rho, \mathbf{u}, \mathbf{y}, \mathbf{X}) \propto \exp\left[-\frac{(\mathbf{y} + \mathbf{u} \otimes \mathbf{1}_T - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} + \mathbf{u} \otimes \mathbf{1}_T - \mathbf{X}\boldsymbol{\beta}) + \underline{q}_v}{2\sigma_v^2}\right] \cdot \exp\left[-\frac{\mathbf{u}'\boldsymbol{\Phi}'\boldsymbol{\Phi}\mathbf{u} + \underline{q}_u}{2\sigma_\varepsilon^2}\right]$$

If we combine the terms in the two exponentials, and after some algebra, it can be shown that:

$$\mathbf{u} \mid \boldsymbol{\beta}, \sigma_v, \sigma_\varepsilon, \rho, \mathbf{y}, \mathbf{X} \sim N_n(\boldsymbol{\mu}_*, \boldsymbol{\Omega}), \quad \mathbf{u} \geq \mathbf{0}, \quad \boldsymbol{\Phi}\mathbf{u} \geq \mathbf{0},$$

where $\boldsymbol{\Omega} = \sigma_v^2 \sigma_\varepsilon^2 \left(T\sigma_\varepsilon^2 \mathbf{I}_n + \sigma_v^2 \boldsymbol{\Phi}'\boldsymbol{\Phi}\right)^{-1}$, $\boldsymbol{\mu}_* = \boldsymbol{\Omega}\bar{\mathbf{e}}/(T\sigma_v^2)$, $\mathbf{e} = \mathbf{X}\boldsymbol{\beta} - \mathbf{y}$.

This is a multivariate truncated normal distribution with the additional inequality constraints $\boldsymbol{\Phi u} \geq \mathbf{0}$. Generating random draws from this distribution is accomplished by using the conditional distributions $u_i \mid \boldsymbol{\beta}, \sigma_v, \sigma_\varepsilon, \rho, u_{j, j\neq i}, \mathbf{y}, \mathbf{X}$, for each $i = 1, \ldots n$. Using standard results for the multivariate normal distribution, we have

$$u_i \mid \boldsymbol{\beta}, \sigma_v, \sigma_\varepsilon, \rho, u_{j, j\neq i}, \mathbf{y}, \mathbf{X} \sim N(m_i, e_{ii}^{-1}), \text{ for each } i = 1, \ldots n,$$

where $m_i = \mu_{*i} - \sum_{j \neq i} (u_j - \boldsymbol{\mu}_{*j}) e_{ji} / e_{ii}$, $\boldsymbol{\Omega}^{-1} = [e_{ij}]$

subject to the constraints $u_i \geq 0$, $\boldsymbol{\Phi u} \geq \mathbf{0}$. The constraints can be converted to a single constraint of the form $A_i \leq u_i \leq B_i$ as follows:

For each $j = 1, .., n$,

If $\phi_{ji} > 0$, set $\overline{A}_j = -\phi_{ji}^{-1} \sum_{s \neq i} \phi_{js} u_s$, if $\phi_{ji} < 0$, set $\overline{B}_j = -\phi_{ji}^{-1} \sum_{s \neq i} \phi_{js} u_s$, else if $\phi_{ji} = 0$ set $\overline{A}_j = -\infty$ and $\overline{B}_j = \infty$. Next define $A_i = \max(0, \max_{j=1,..,n} \overline{A}_j)$ and $B_i = \min_{j=1,..,n} \overline{B}_j$. Therefore, the full conditional distribution is:

$$u_i \mid \boldsymbol{\beta}, \sigma_v, \sigma_\varepsilon, \rho, u_{j, j\neq i}, \mathbf{y}, \mathbf{X} \sim N(m_i, e_{ii}^{-1}), \text{ subject to } A_i \leq u_i \leq B_i, \text{ for each } i = 1, .., n.$$

This involves drawing random numbers from a truncated normal distribution. For this task specialized algorithms are available.

Results concerning the finite integrability of the posterior for this model, follow directly from Fernandez, Osiewalski and Steel (1997). In informal terms, a stochastic frontier model with panel data $(T > 1)$ for which the marginal prior distribution $p(\mathbf{u})$ is proper, yields a proper posterior.

Efficiency measurement in this model involves marginal posterior distributions of the idiosyncratic component, ε, and the spatial component, $\mathbf{u}$. Since $\boldsymbol{\varepsilon} \sim N(0, \sigma_\varepsilon^2 \mathbf{I}_n)$, $\boldsymbol{\varepsilon} \geq 0$, it is clear that:

$$p_\varepsilon(\varepsilon \mid y, X) = \int_0^\infty \left(\frac{\pi}{2} \sigma_\varepsilon^2\right)^{-1/2} \exp\left(-\frac{\varepsilon^2}{\sigma_\varepsilon^2}\right) \cdot p(\sigma_\varepsilon \mid \mathbf{y}, \mathbf{X}) d\sigma_\varepsilon, \; \varepsilon \geq 0.$$

In terms of efficiency, let $r = \exp(-\varepsilon)$. Then we have:

$$p_r(r \mid \mathbf{y}, \mathbf{X}) = p_\varepsilon(-\ln r \mid \mathbf{y}, \mathbf{X}) r^{-1} = \int_0^\infty \left(\frac{\pi}{2}\sigma_\varepsilon^2\right)^{-1/2} \exp\left(-\frac{(\ln r)^2}{\sigma_\varepsilon^2}\right) \cdot p(\sigma_\varepsilon \mid \mathbf{y}, \mathbf{X}) d\sigma_\varepsilon .$$

For the spatial or spillover component, $\mathbf{u}$, we have:

$$p_u(\mathbf{u} \mid y, X) = \int_{1/\omega_{max}}^{1/\omega_{min}} \int_0^\infty \left(\frac{\pi}{2}\sigma_\varepsilon^2\right)^{-n/2} \mid \det(\Phi) \mid \exp\left(-\frac{1}{2\sigma_\varepsilon^2}\mathbf{u}'\Phi'\Phi\mathbf{u}\right) \cdot p(\sigma_\varepsilon, \rho \mid y, X) d\sigma_\varepsilon d\rho ,$$

$$\mathbf{u} \geq \mathbf{0} .$$

with the understanding that $p(\sigma_\varepsilon, \rho \mid \mathbf{y}, \mathbf{X})$ includes its integrating constant. Efficiency measurement using this distribution is not possible since the integral is not available analytically. However, the conditional distribution for region i is given by $u_i \mid \sigma_\varepsilon, \rho, \mathbf{y}, \mathbf{X} \sim N(0, V_{ii})$, $u_i \geq 0$, where $\mathbf{V} = \sigma_\varepsilon^2(\Phi'\Phi)^{-1}$. Therefore, the distribution

$$p(u_i \mid \mathbf{y}, \mathbf{X}) = \int\int p(u_i \mid \sigma_\varepsilon, \rho, \mathbf{y}, \mathbf{X}) p(\sigma_\varepsilon, \rho \mid \mathbf{y}, \mathbf{X}) d\sigma_\varepsilon d\rho \approx S^{-1} \sum_{i=1}^S p(u_i \mid \sigma_\varepsilon^{(i)}, \rho^{(i)}, \mathbf{y}, \mathbf{X}) .$$

This Monte Carlo approximation based on the available Gibbs draws may be used for posterior predictive efficiency measurement. A different concept of efficiency is based on posterior region-specific efficiency measurement. It utilizes the conditional distribution $\mathbf{u} \mid \boldsymbol{\beta}, \sigma_v, \sigma_\varepsilon, \rho, \mathbf{y}, \mathbf{X} \sim N_n(\boldsymbol{\mu}_*, \boldsymbol{\Omega})$, $\mathbf{u} \geq \mathbf{0}$, $\Phi\mathbf{u} \geq \mathbf{0}$. If we had the distribution $\int p(\mathbf{u} \mid \boldsymbol{\beta}, \sigma_v, \sigma_\varepsilon, \rho, \mathbf{y}, \mathbf{X}) d\boldsymbol{\beta} d\sigma_v d\sigma_\varepsilon d\rho$ efficiency measurement would have been easy. However, we do not even have the normalizing constant of the distribution because the mean is not zero and the constraints present an impediment to the computation of the normalizing constant or moments of the distribution. For this reason, to compute moments of the form:

$$\int_{\Theta \times \{\mathbf{u}:\mathbf{u} \geq \mathbf{0}, \Phi\mathbf{u} \geq \mathbf{0}\}} \mathbf{u}^q p(\mathbf{u} \mid \boldsymbol{\beta}, \sigma_v, \sigma_\varepsilon, \rho, \mathbf{y}, \mathbf{X}) d\boldsymbol{\beta} d\sigma_v d\sigma_\varepsilon d\rho d\mathbf{u}$$

(where Θ is the parameter space) we resort to a four step algorithm. See Appendix.

3. Empirical Analysis

In this section we consider an application to Italian regions (1970-1993). Efficiency measurement must necessarily deal with the fact that regions could be different in terms of structure, as well as in terms of temporal performance. In other words, it will be necessary to include regional effects and time effects in the model. Failing to account for these effects would lead to potentially misguided inferences regarding efficiency measurement. With this modification, the spatial lag model with both regional and time effects is:

$$\mathbf{y} = (\mathbf{I}_T \otimes \mathbf{1}_n)\boldsymbol{\delta} + (\mathbf{1}_T \otimes \mathbf{I}_n)\boldsymbol{\theta} + \mathbf{X}\boldsymbol{\beta} + \mathbf{v} - \mathbf{u} \otimes \mathbf{1}_T$$

where $\boldsymbol{\delta}$ are coefficients related to time effects, and $\boldsymbol{\theta}$ are coefficients related to regional effects. The basic model used in this application is a Translog production function (e.g. Koop et *al.* 2000) of the form:

$$
\begin{aligned}
lnY_{it} = \mathbf{z}'_{it}\boldsymbol{\varphi} &+ \beta_\kappa lnK_{it} + \beta_L lnL_{it} + \beta_H lnH_{it} + \frac{1}{2}\beta_{KK} lnK_{it}{}^2 + \frac{1}{2}\beta_{LL} lnL_{it}{}^2 \\
&+ \frac{1}{2}\beta_{HH} lnH_{it}{}^2 + \beta_{KL} lnK_{it} lnL_{it} + \beta_{KH} lnK_{it} lnH_{it} \\
&+ \beta_{LH} lnL_{it} lnH_{it} + u_{it}
\end{aligned}
$$

where $\mathbf{z}_{it}$ represents the vector of regional and time dummy variables which are interacted with ln$K$, ln$L$ and ln$Y$, $\boldsymbol{\varphi}$ represents the respective coefficients (effects), Y is output, K is physical capital, L is labor, and H represents human capital.

The data is obtained from the CRENoS database. See Marrocu, Paci, and Pala (2000), Paci and Saba (1997), and Paci and Pasceddu (2000). Output is measured using value added, labor is measured using units of labor at sub-sectoral level, human capital is measured using number of registrations in high school, and capital is measured using total gross capital stock in agriculture, industry, market services, and non-market services. The capital stock is available for the period 1970-1993, but output and labor are available over a longer period, namely 1960-1996 so the availability of capital stock data restricts the sample to 1970-93.

As discussed earlier, in the aforementioned Translog specification we impose the so-called regularity conditions, namely that elasticities have to be positive.

Gibbs sampling has been implemented using 30,000 iterations, 10,000 of which have been discarded to mitigate the impact of start-up effects. We use a flat prior on ρ, $\underline{q}_v = \underline{q}_\varepsilon = 10^{-4}$, and $\underline{n} = 0$.

The posterior results are reported in Table 1. Reported are posterior means of parameters and posterior standard deviations. The posterior mean of ρ is positive with a small posterior standard deviation indicating positive externalities in inefficiency implying that when a region improves in terms of efficiency, neighboring regions can benefit from this reduction and reduce their own inefficiency level. Marginal posterior distributions of all parameters are reported in Figure 3. The other important finding is that the marginal posterior of ρ is practically incompatible with values near zero so the extent of positive externalities in inefficiency is significant.

Since the posterior for the spatial parameter ρ does not have a recognizable functional form we implemented a random walk Metropolis-Hasting procedure. Furthermore, we have imposed the spatial parameter ρ to be in the range of the inverse of minimum and the maximum eigenvalues of the spatial matrix to ensure stability of the system. The acceptance rate of the random walk Metropolis-Hastings algorithm was 26.1%.

Finally, posterior predictive efficiency distributions are reported in Figure 4. The straight line corresponds to the posterior predictive efficiency distribution of the spillover component (u_i): This is in fact the average of posterior predictive distributions across all nineteen u_i distributions ($i = 1,..,19$). The dashed line corresponds to the posterior predictive efficiency distribution of the idiosyncratic component. Region-specific means of posterior predictive efficiency computed using the procedure described previously are presented in Table 2. These measures range from 0.832 for Trantino and 0.832 Apulia to almost perfect efficiency for Umbria.

The joint marginal posterior distribution of σ_ε and ρ (the two important parameters of the model) is reported[3] in Figure 5. The conclusion that the marginal posterior of ρ is practically incompatible with values near zero is easily seen in the contours of the joint posterior, where it is also apparent that the two parameters are

---

[3] The bivariate distribution is estimated using kernel techniques. The density was normalized to be equal to one at the mode.

not strongly correlated. The same holds true for every bivariate combination of β_j, σ_v, σ_ε, and ρ. Essentially this implies that the Gibbs sampler is a highly efficient computational scheme for Bayesian inference, at least for the particular data set.

4. Concluding Remarks

The purpose of the paper was to propose a stochastic frontier model that accounts for the spatial component of regional data. The proposed model allows for a spillover effect in inefficiency, as well as idiosyncratic efficiency components. The likelihood function is exceedingly complex but it was shown that Monte Carlo methods may be used to perform Bayesian inference and provide efficiency measures. The model is applicable to contexts different than the spatial context to which the present paper has been limited. Spillover effects are quite common in the empirical economics literature, for example there are R&D spillovers, contagion effects in financial economics etc. For particular specifications of the spatial weight matrix, the model considered in the present paper could be used in these contexts as well. Clearly, further theoretical and empirical research on the subject would be of great interest.

References

Aigner D, Knox-Lovell KA, van den Broeck J, Koop G, Osiewalski J, Steel MFJ. 1994. Stochastic frontier models: A Bayesian perspective, *Journal of Econometrics***61**: 273-303.

Anselin, L., 1988, *Spatial econometrics: Methods and models*, Kluwer, Dordrecht.

Anselin, L., and A.K. Bera, 1998, Spatial dependence in linear regression models with an introduction to spatial econometrics, in A. Ullah and D.E.A. Giles, *Handbook of applied economic statistics*, New York, Marcel Dekker, 237-289

Bai, J. (2009), Panel Data Models with Interactive Fixed Effects, *Econometrica*, 77, 1229-1279.

Bauer PW. 1990. Recent developments in the econometric estimation of frontiers. *Journal ofEconometrics***46**: 39-56.

Carlin, B. P. and Louis, T. A., 2000, *Bayes and Empirical Bayes Methods for Data Analysis*, Second Edition, London: Chapman & Hall.

Casella G, George E. 1992. Explaining the Gibbs sampler.*The American Statistician***46**: 167-174.

Fernandez C, Osiewalski J, Steel MFJ. 1997. On the use of panel data in stochastic frontier models. *Journal of Econometrics***79**: 169-193

Gelfand AE, Smith AFM. 1990. Sampling based approaches to calculating marginal densities. *Journal of the American Statistical Association***85**: 398-409.

Geweke J. 1996. Simulation-based Bayesian inference for economic time series, Research Department, Federal Reserve Bank of Minneapolis, Working Paper 570.

Greene WH. 1990. A gamma-distributed stochastic frontier model.*Journal of Econometrics* **46**: 141-163.

Greene, W., 1993, The econometric approach to efficiency analysis, in H. O. Fried, C. A. Knox Lovell, and S. S. Schmidt (eds.), *The measurement of productive efficiency: Techniques and Applications* (Oxford: Oxford University Press).

Greene, W., 2001, New developments in the estimation of stochastic frontier models with panel data, Department of Economics, New York University.

Koop G, Osiewalski J, Steel MFJ. 1997. Bayesian efficiency analysis through individuals effects: Hospital cost frontiers. Journal of Econometrics **76**: 77-105.

Koop G, Steel MFJ. 2001. Bayesian analysis of stochastic frontier models. In *A Companion to Theoretical Econometrics*, Baltagi B (ed). Blackwell, 520-573.

Koop G, Steel MFJ, Osiewalski J., 1995, Posterior analysis of stochastic frontier models using Gibbs sampling.*Computational Statistics***10**: 353-373.

Koop, G., Osiewalski, J., & Steel, M., 2000, Modelling the sources ofoutput growth in a panel of countries, *Journal of Business and Economic Statistics*, 18(3), 284-299.

Kotz, S., N. Balakrishnan, and N.L. Johnson, 2000, *Continuous multivariate distributions*, volume 1: Methods and applications, Wiley, New York.

LeSage, James P. and R. Kelley Pace, 2009, *Introduction to Spatial Econometrics*, CRC Press, Taylor &Francis Group: Boca Raton, FL, January.

Marrocu M., Paci R. and Pala R., 2000, Estimation of total factor productivity for regions and sectors in Italy. A panel cointegration approach, *Contributi di RicercaCRENoS, 00/16*

Meeusen W, van den Broeck J., 1977, Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review***8**: 435-444.

Paci R. and Saba A. (1997), The empirics of regional economic growth in Italy. 1951-93. *RivistaInternazionale di ScienzeEconomiche e Commerciali*

Paci R. andPusceddu N. (2000) Stimadello stock di capitalenelleregioniitaliane,*RassegnaEconomica, Quaderni di Ricerca, 97-118.*

Parent, Olivier and James P. LeSage, 2012, Determinants of Knowledge Production and Their Effects on Regional Economic Growth, *Journal of Regional Science*, Volume 62, Number 2, pp.256-284.

Pesaran, M. H., 2006, Estimation and inference in large heterogeneouspanels with a multifactor error structure, *Econometrica*, 74, 967-1012.

Rao, C.R., 1965, *Linear statistical inference and its applications*, Wiley: New York.

Robert, C. P., 2001, *The Bayesian Choice*, Second Edition, New York: Springer-Verlag.

Roberts GO, Smith AFM 1994. Simple conditions for the convergence of the Gibbs sampler and Metropolis-Hastings algorithms. *Stochastic Processes and their Applications***49**: 207-216.

Schmidt P., 1977, Formulation and estimation of stochastic frontier production function models,*Journal of Econometrics***6**: 21-37.

Smith AFM, Roberts CO, 1993, Bayesian computation via the Gibbs sampler and related Markov Chain Monte Carlo methods. *Journal of the Royal Statistical Society***B55**: 3-23.

Smith, Tony E. and James P. LeSage, 2004, A Bayesian Probit Model with Spatial Dependencies, in *Advances in Econometrics*, Volume 18: Spatial and Spatiotemporal

Econometrics, (Oxford: Elsevier Ltd), JamesP. LeSage and R. Kelley Pace (eds.), , pp. 127-160.

Solow, A.R., 1990, A method for approximating multivariate normal orthant probabilities, *Journal of Statistical Computation and Simulation***37**, 225-229

Stevenson RE. 1990. Likelihood functions for generalized stochastic frontier estimation. *Journal of Econometrics***13**: 57-66.

Tanner MA, Wong W-H. 1987. The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association***82**: 528-550.

Tsionas E.G. 2000.Full likelihood inference in normal-gamma stochastic frontier models.*Journalof Productivity Analysis***13**: 179-201.

Tsionas E.G. 2002.Stochastic Frontier Models with Random Coefficients, *Journal of Applied Econometrics*, **17** (2): 121-147.

Zellner A. 1971. *Introduction to Bayesian inference in econometrics*.Wiley: New York.

Table 1. Posterior results for Italian regions

	Posterior mean	Posterior s.d.
e_K	0.301	0.044
e_L	0.598	0.041
e_H	0.249	0.037
σ_v	0.1914	0.0163
σ_ε	0.2212	0.0172
ρ	0.3140	0.0443
VI	0.5156	0.1316

Notes: *VI* is the portion of variance of the frontier attributed to the idiosyncratic component. "e_K, e_L, e_H" denote capital elasticities of capital, labor and human capital evaluated at the means of data. Reported results are averaged out for parameter uncertainty.

Table 2. Means of posterior predictive region-specific efficiency

| Region	Efficiency	Region	Efficiency
Piedmont	0.981	Umbria	0.997
Val d'Aosta	0.983	Marches	0.902
Lombardy	0.933	Latium	0.904
Trantino	0.832	Abruzzo	0.973
Veneto	0.935	Molise	0.881
Friuti	0.931	Campania	0.901
Liguria	0.922	Apulia	0.833
Emilia Rom.	0.951	Basilicata	0.872
Tuscany	0.057	Calabria	0.894
		Sicily	0.888

**Figure 3. Marginal posterior distributions of parameters for spatial frontier:
Italian regional data**
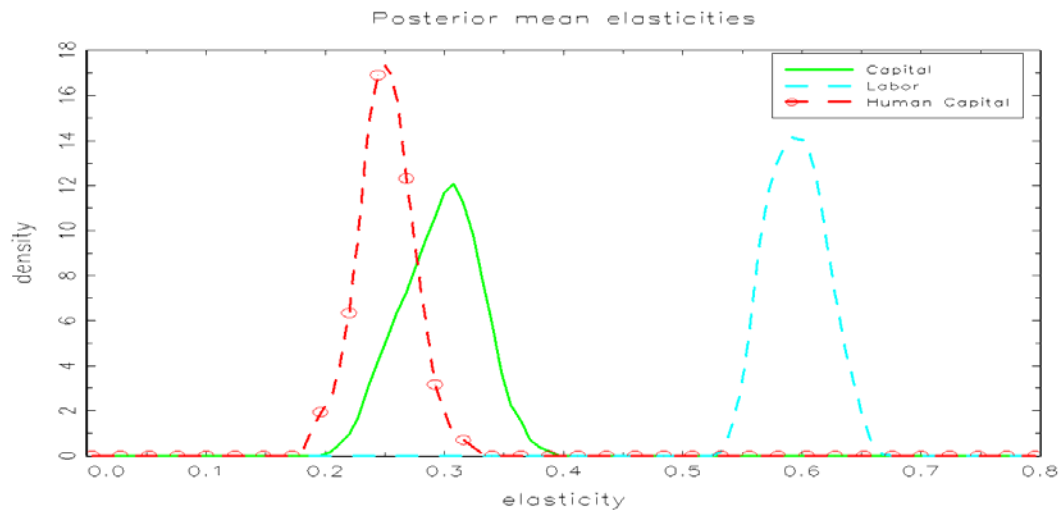
Figure 4. Posterior predictive distributions for idiosyncratic and spillover efficiency: Italian regional data.
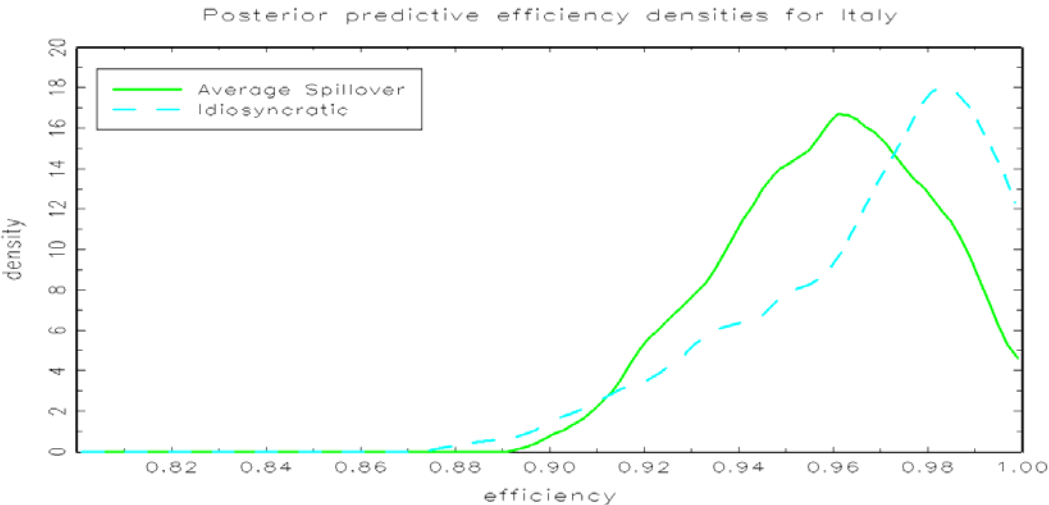
Figure 5. Joint posterior distribution of σ_ε and ρ

Joint posterior of ρ and σ_e
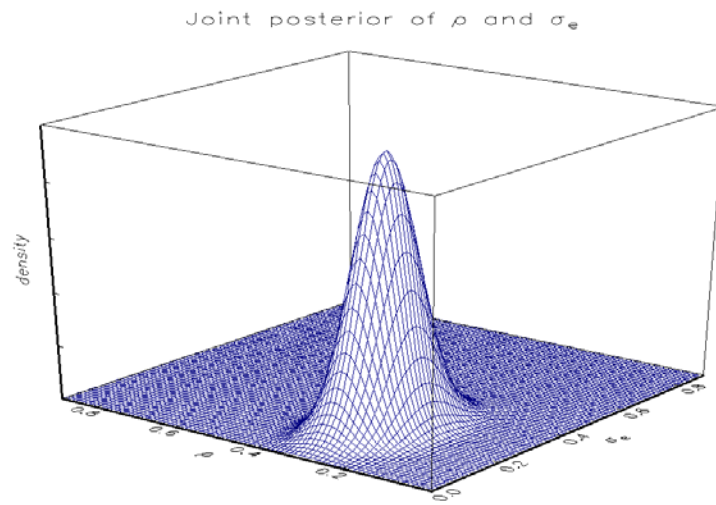
APPENDIX

Computational Algorithm

Step 1: For each Gibbs draw $\boldsymbol{\beta}^{(i)}, \sigma_v^{(i)}, \sigma_\varepsilon^{(i)}, \rho^{(i)}$, $i = 1,..,S$

Step 2: Draw vectors $\mathbf{u}^{(1)}, \mathbf{u}^{(2)},..., \mathbf{u}^{(N)}$ from the distribution $N_n(\boldsymbol{\mu}_*, \boldsymbol{\Omega})$ subject to $\mathbf{u}^{(j)} \geq 0$, $\boldsymbol{\Phi}\mathbf{u}^{(j)} \geq 0$.

Step 3: Compute the moments in the usual way, i.e.

$$E\left[\mathbf{u}^q \mid \boldsymbol{\beta}^{(i)}, \sigma_v^{(i)}, \sigma_\varepsilon^{(i)}, \rho^{(i)}, \mathbf{y}, \mathbf{X}\right] = N^{-1} \sum_{j=1}^{N} \mathbf{u}^{q,(j)}$$

Step 4: Compute moments: $E\left[\mathbf{u}^q \mid \mathbf{y}, \mathbf{X}\right] = S^{-1} \sum_{i=1}^{S} E\left[\mathbf{u}^q \mid \boldsymbol{\beta}^{(i)}, \sigma_v^{(i)}, \sigma_\varepsilon^{(i)}, \rho^{(i)}, \mathbf{y}, \mathbf{X}\right]$.

In this application we use $N = 50$ simulations to compute moments.