CEIS Research Papers

Università di Roma Tor Vergata

ISSN 2610-931X

CEIS Tor Vergata

RESEARCH PAPER SERIES

Vol. 17, Issue 4, No. 459 – May 2019

Estimation of Stochastic Frontier Panel Data Models with Spatial Inefficiency

Federico Belotti, Giuseppe Ilardi and Andrea Piano Mortari

Estimation of stochastic frontier panel data models with spatial inefficiency*

Federico Belotti† Giuseppe Ilardi‡ Andrea Piano Mortari§

Abstract

This paper proposes a stochastic frontier panel data model in which unit-specific inefficiencies are spatially correlated. In particular, this model has simultaneously three important features: i) the total inefficiency of a productive unit depends on its own inefficiency and on the inefficiency of its neighbors; ii) the spatially correlated and time varying inefficiency is disentangled from time invariant unobserved heterogeneity in a panel data model à la Greene (2005); iii) systematic differences in inefficiency can be explained using exogenous determinants. We propose to estimate both the "true" fixed- and random-effects variants of the model using a feasible simulated composite maximum likelihood approach. The finite sample behavior of the proposed estimators are investigated through a set of Monte Carlo experiments. Our simulation results suggest that the estimation approach is consistent, showing good finite sample properties especially in small samples.

Keywords: Stochastic frontiers model; Spatial inefficiency; Panel data, Fixed-effects model.

*The views expressed in this paper are those of the authors and do not involve the responsibility of their institutions. We thank participants at the X North American Workshop on Efficiency and Productivity Analysis for their helpful comments. All remaining errors are ours. First version: May 2018. This version: May 2019.
Corresponding author: Federico Belotti, Department of Economics and Finance, University of Rome Tor Vergata, Via Columbia, 2, 00133 Rome, Italy; e-mail: `federico.belotti@uniroma2.it`.

†University of Rome Tor Vergata, Department of Economics and Finance.

‡Bank of Italy, Directorate General for Economics, Statistics and Research.

§University of Rome Tor Vergata, Centre for Economic and International Studies.

1 Introduction

The analysis of efficiency is an important issue in many economic studies and the Stochastic Frontier (SF) model, originally proposed by Aigner et al. (1977) and Meeusen & van den Broeck (1977), is a popular tool to measure this latent indicator. SF models are regression models with a composite error term given by the difference between a symmetric idiosyncratic error and a one-sided disturbance representing technical inefficiency.[1]

The estimation of these models is typically obtained via likelihood-based methods assuming cross-sectional independence among statistical units. However, as suggested by a well established stream of research investigating the effect of agglomeration on economic activities (see, among many others, Rosenthal & Strange, 2001; Ellison et al., 2010; Martin et al., 2011), this assumption may be difficult to justify since the inefficiency of a productive unit is likely to be affected by the actions of neighboring units. For example, producers can be more efficient by sharing services with neighboring firms or by collectively managing their input purchases (Rosenthal & Strange, 2004). Furthermore, geographical proximity may increase producers efficiency via technological and knowledge spillovers (see Audretsch & Feldman, 2004, and references therein).

In order to capture these externalities, in this work we relax the independence assumption allowing for a specific type of cross-sectional dependence, namely spatial dependence, and attempt to make three contributions to the SF literature. First, within a panel setting, a spatial autoregressive structure is explicitly placed on the inefficiency. Second, we consider a statistical model in which a time varying and spatial auto-regressive inefficiency is disentangled from time-invariant unobserved heterogeneity. Finally, we allow for systematic inefficiency's heterogeneity driven by exogenous determinants. Since both model parameters and inefficiency estimates may be adversely affected when these determinants are neglected, leading to inconsistent estimates and potentially misleading policy conclusions (Wang & Schmidt, 2002), this last feature is crucial from both the methodological and the empirical perspectives.

In particular, we consider the following specification for a stochastic production frontier panel data model

$$
\begin{aligned}
y_{it} &= \alpha_i + \boldsymbol{x}_{it}\boldsymbol{\beta} + v_{it} - u_{it}, && (1)\\
v_{it} &\sim \mathcal{N}(0, \psi^2), && (2)\\
u_{it} &= \rho \boldsymbol{W}\boldsymbol{u}_t + \tilde{u}_{it}, && (3)\\
\tilde{u}_{it} &\sim \mathcal{F}_{\tilde{u}}(\sigma_{it}), && (4)\\
\sigma_{it} &= g(\boldsymbol{z}_{it}\boldsymbol{\delta}), \quad i = 1, \ldots, n, \quad t = 1, \ldots, T, && (5)
\end{aligned}
$$

where, for each unit i and period t, y_{it} represents the output, $\boldsymbol{x}_{it}$ is a $1 \times k$ vector of exogenous inputs and $\boldsymbol{\beta}$ the corresponding $k \times 1$ vector of technological parameters. The composite error term is the difference between the symmetric idiosyncratic error v_{it} and the one-sided disturbance

[1] Notice that when the composite error term is the sum of the two components, the model describes a stochastic cost frontier, and the one-sided disturbance is interpreted as cost inefficiency.

u_{it}, which represents technical inefficiency. The v_{it} disturbance is assumed to be normally distributed with zero mean and variance ψ^2.

The key feature of this statistical model is the presence of spatial interactions in the inefficiency term u_{it}. In particular, the model statement in (3) can rewritten as

$$\boldsymbol{u}_t = (\boldsymbol{I}_n - \rho \boldsymbol{W})^{-1} \tilde{\boldsymbol{u}}_t = \boldsymbol{D} \tilde{\boldsymbol{u}}_t, \quad t = 1, \dots, T, \tag{6}$$

where the disturbance $\tilde{u}_{it}$ is defined as *intrinsic* inefficiency, i.e. the characteristic inefficiency of unit i at time t, $\boldsymbol{W}$ is the $n \times n$ matrix describing the spatial arrangement of the n units, $\boldsymbol{D}$ is the spatial multiplier matrix and $\boldsymbol{I}_n$ is the n-identity matrix. The elements of the intrinsic inefficiency vector $\tilde{\boldsymbol{u}}$ are assumed to be independent between each other as well as from the elements of both $\boldsymbol{v}$ and $\boldsymbol{\alpha}$.[2] Equation (6) can then be interpreted as indicating that the total (or global) inefficiency u_{it} is a function of the intrinsic inefficiency of unit i at time t and of a linear combination of neighboring units' intrinsic inefficiencies at time t scaled by the spatial parameter ρ. In order to ensure that $u_{it} > 0$, we restrict ρ to be in the unit interval. For the similarity with the classical spatial error model, we call model (1)-(5) the Spatial Inefficiency Model (SIM).

We will consider the cases in which $\mathcal{F}_{\tilde{u}}$ is either exponential, with scale parameter σ_{it}, or half normal with standard deviation σ_{it} where the unit- and time-specific scale parameter σ_{it} depends on $g(.)$, a known positive monotonic function, $\boldsymbol{z}_{it}$, a $1 \times s$ vector of exogenous covariates and $\boldsymbol{\delta}$, the corresponding vector of unknown inefficiency effects.

The statistical model is completed by specifying the nature of the unit-specific effects. We consider two different sets of assumptions which lead to different estimation approaches. In the first case, the α's are assumed to be Gaussian random-effects, $\alpha_i \sim \mathcal{N}(0, \tau^2)$, independent from both the exogenous covariates and the idiosyncratic shocks. Following Greene's terminology, we refer to this model as SIM True Random-Effects (SIM-TRE). In the second case, unit-specific effects are treated as parameters to be estimated and no restrictions are placed on the relationship with the exogenous covariates. We refer to this model as SIM True Fixed-Effects (SIM-TFE).

The specification in (1)-(5) is quite general and encompasses a number of models of empirical interest. In particular, the true random- and fixed-effects by Greene (2005) are special cases when $\rho = 0$. When $\alpha_i = \alpha$ and $\rho = 0$ with $\tilde{u}_{it} \sim \mathcal{N}^+(0, \sigma_{it})$, the model can be viewed as a variant of the Battese & Coelli (1995) model in which the scale of the pre-truncated inefficiency is heteroskedastic.

We propose to estimate this model exploiting the simulated likelihood principle. In particular, we derive a computationally feasible simulated composite maximum likelihood estimator that solves two major issues related to the spatial autoregressive structure of the inefficiency: *i)* the multivariate distribution of the composite error term is in general unknown and, *ii)* even when a closed form for the likelihood function is available, the estimation is unfeasible.

We investigate the finite sample properties of the proposed estimators using Monte Carlo simulations. Our results show convincing evidence of consistency when $n \to \infty$ with fixed T.

[2] In what follows, we denote with $\boldsymbol{v}$, $\tilde{\boldsymbol{u}}$ and $\boldsymbol{\alpha}$ the nT-vectors obtained by stacking the corresponding elements.

We apply the proposed inferential procedures to a balanced panel of Indonesian rice farms, documenting the presence of strong spatial inefficiency spillovers. We also find that considering a spatial autoregressive inefficiency can greatly affect the efficiency ranking of productive units.

The remainder of this paper is organized as follows. Section 2 offers a critical review of the literature on spatial SF panel data models, Section 3 describes the derivation of the proposed estimation procedures. Section 4 investigates the small sample properties of the our estimator through a set of Monte Carlo experiments while Section 5 provides an illustration through an application based on the well known balanced panel of Indonesian rice farms.[3] Finally, Section 6 offers some conclusions.

2 Literature Review

The first contribution in the spatial SF literature for panel data dates back to Druska & Horrace (2004) where the authors limit the spatial dependence to the idiosyncratic error and, following Schmidt & Sickles (1984), treat time invariant heterogeneity as if it was inefficiency. Similarly, Glass et al. (2013) consider a spatial autoregressive process for the dependent variable allowing for time-varying inefficiencies *à la* Cornwell et al. (1990).

Glass et al. (2014, 2016) and Adetutu et al. (2015) extend the standard spatial autoregressive (SAR) model allowing for a compounded error term in which the two components of the error are disentangled using distributional assumptions.[4] In analyzing the efficiency spillovers between African countries, Glass et al. (2017) extend the approach proposed in Glass et al. (2016) by allowing for unobserved heterogeneity in the form of random-effects. Gude et al. (2018) propose a generalized version of the SAR and spatial Durbin frontier models allowing for heterogeneous spatial coefficients in their analysis of Spanish province expenditures. Finally, Ramajo & Hewings (2018) further extend the spatial Durbin stochastic production-frontier panel data model allowing for time varying technical inefficiency. All these works allow for cross-sectional dependence, but none of them explicitly considers a spatial autoregressive structure for the inefficiency. As is well known in the spatial econometric literature (Elhorst, 2010), the spatial Durbin model may be observationally equivalent to a spatial error model under specific parameter configurations. However, in a SF framework, this would imply having the same spatial dependence structure on both the inefficiency and the idiosyncratic error. Furthermore, as also pointed out by Glass et al. (2017), a SAR stochastic frontier model does not define the total inefficiency relative to the estimated best practice frontier. Hence, within such a framework, the analysis of inefficiency has to be performed using the comparison with the best procedure *à la* Schmidt & Sickles (1984). However, while Schmidt & Sickles (1984) estimate inefficiency using fixed-effects, Glass et al. (2016, 2017) propose to post-estimate it using JLMS estimators. As pointed out

[3] The same data have been previously analyzed by several authors among which Lee & Schmidt (1993); Horrace & Schmidt (2000); Druska & Horrace (2004).

[4] In Glass et al. (2016), the authors provide a way to disentangle direct and indirect efficiency *à la* LeSage & Pace (2009).

by Wang & Schmidt (2009), the collection of JLMS estimates is (in a probabilistic sense) a shrinkage of the true latent inefficiencies toward the mean of the distribution. On average, the JLMS estimator will overestimate the smaller realizations of u and underestimate the larger ones. For this reason we argue that, in a comparison with the best setting, substituting the fixed-effects estimator with the JLMS one can be problematic since the bias associated with the JLMS estimator of the best performing unit could potentially contaminate all the other scores.

Mastromarco et al. (2016) propose a traditional SF panel data model accommodating both time and cross sectional dependence.[5] In their model, the total inefficiency is the sum of a homoskedastic time invariant unit-specific component, some unobserved exogenous common factors and a term, obtained from an endogenous threshold efficiency regime selection mechanism, introducing correlation between units. However, the model does not allow explicitly for a spatial autoregressive inefficiency and the consistent estimation of the model's parameters requires that both n and T are sufficiently large.

As mentioned before, a spatial autoregressive structure for the inefficiency implies a multivariate and generally unknown distribution for the composite error term. Nevertheless, when the inefficiency is half-normally distributed, we can exploit the properties of the Closed Skew-Normal (CSN) class of distributions and in particular proposition 13.6.1 in Dominguez-Molina et al. (2004) which shows that the n-dimensional random variable $\boldsymbol{\varepsilon}_t = \boldsymbol{v}_t - \boldsymbol{A}\boldsymbol{u}_t$, for $t = 1, \dots, T$ with $\boldsymbol{v}_t \sim \mathcal{N}_n(\boldsymbol{0}, \boldsymbol{\Psi})$ and $\boldsymbol{u}_t \sim \mathcal{N}_n^+(\boldsymbol{0}, \boldsymbol{\Sigma})$, is distributed as

$$\boldsymbol{\varepsilon}_t \sim CSN_{n,n}(\boldsymbol{0}, \boldsymbol{\Omega}_*, -\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{\Omega}_*^{-1}, \boldsymbol{0}, \boldsymbol{\Lambda}_*), \tag{7}$$

where $\boldsymbol{\Omega}_* = \boldsymbol{\Psi} + \boldsymbol{A}'\boldsymbol{\Sigma}\boldsymbol{A}$ and $\boldsymbol{\Lambda}_* = \boldsymbol{\Sigma} - \boldsymbol{\Sigma}\boldsymbol{A}'(\boldsymbol{\Psi} + \boldsymbol{A}'\boldsymbol{\Sigma}\boldsymbol{A})^{-1}\boldsymbol{A}\boldsymbol{\Sigma}$.[6] Within this framework, two configurations of $\boldsymbol{\Sigma}$, $\boldsymbol{\Psi}$ and $\boldsymbol{A}$ are of interest:

$$\begin{aligned} \boldsymbol{\Psi} &= \psi^2\boldsymbol{I}_n,\ \boldsymbol{A} = \boldsymbol{D} \text{ and } \boldsymbol{\Sigma} = \sigma^2\boldsymbol{I}_n, & (8)\\ \boldsymbol{\Psi} &= \psi^2\boldsymbol{I}_n,\ \boldsymbol{A} = \boldsymbol{I}_n \text{ and } \boldsymbol{\Sigma} = \sigma^2\boldsymbol{D}'\boldsymbol{D}, & (9) \end{aligned}$$

where $\boldsymbol{D}$ is defined in (6).

In the first case, the one considered in this paper, the spatial multiplier $\boldsymbol{D}$ is applied to a vector of independent univariate truncated random variables. The likelihood function associated to this case is given by

$$\begin{aligned} f(\boldsymbol{\epsilon}_t) &= 2^{-n}\phi_n\left(\boldsymbol{\epsilon}_t; \boldsymbol{0}, \psi^2\boldsymbol{I}_n + \sigma^2\boldsymbol{D}'\boldsymbol{D}\right)\\ &\quad \times\Phi_n\left[\sigma^2\boldsymbol{D}'\left(\psi^2\boldsymbol{I}_n + \sigma^2\boldsymbol{D}'\boldsymbol{D}\right)^{-1}\boldsymbol{\epsilon}_t; \boldsymbol{0}, \sigma^2\boldsymbol{I}_n - \sigma^4\boldsymbol{D}'\left(\psi^2\boldsymbol{I}_n + \sigma^2\boldsymbol{D}'\boldsymbol{D}\right)^{-1}\boldsymbol{D}\right]. \end{aligned} \tag{10}$$

In the second case, the spatial transformation $\boldsymbol{D}$ is applied before the truncation operation takes place. In this case the untruncated covariance matrix of the vector $\boldsymbol{u}_t$ is a function of $\boldsymbol{D}$.

[5] The traditional SF panel data models treat time invariant heterogeneity as if it was inefficiency, thus not providing any mechanism to disentangle the former from the latter (Schmidt & Sickles, 1984; Pitt & Lee, 1981; Battese & Coelli, 1988, 1992, 1995; Kumbhakar, 1990).

[6] A detailed presentation of the CSN family and its properties can be found in Gonzalez-Farias et al. (2004).

Therefore, the vector $\boldsymbol{u}_t$ is a realization from a multivariate truncated normal random variable.[7] The likelihood function associated to this case is given by

$$\begin{aligned} f(\boldsymbol{\varepsilon}_t) \;=\; & \Phi_n\left(\mathbf{0};\mathbf{0},\sigma^2\boldsymbol{D}'\boldsymbol{D}\right)^{-1}\phi_n\left(\boldsymbol{\varepsilon}_t;\mathbf{0},\psi^2\boldsymbol{I}_n+\sigma^2\boldsymbol{D}'\boldsymbol{D}\right) \\ & \times\Phi_n\left[\sigma^2\boldsymbol{D}'\boldsymbol{D}\left(\psi^2\boldsymbol{I}_n+\sigma^2\boldsymbol{D}'\boldsymbol{D}\right)^{-1}\boldsymbol{\varepsilon}_t;\mathbf{0},\sigma^2\boldsymbol{D}'\boldsymbol{D}-\sigma^4\boldsymbol{D}'\boldsymbol{D}\left(\psi^2\boldsymbol{I}_n+\sigma^2\boldsymbol{D}'\boldsymbol{D}\right)^{-1}\boldsymbol{D}'\boldsymbol{D}\right]. \end{aligned} \tag{11}$$

In general, these two likelihood functions cannot be factorized as the product of independent univariate densities.[8] Moreover, the maximization of (10) or (11) requires the numerical approximation of several n-dimensional normal integrals for each cross-section in each optimization's step. Regardless of the method used to approximate these integrals, the associated computational burden is hardly manageable, even with moderate cross-sectional dimensions.

These considerations suggest the use of alternative inferential procedures which sidestep the direct evaluation of the n-variate integrals. Schmidt et al. (2009), Areal et al. (2012), Tsionas & Michaelides (2016) and Carvalho (2018) exploit the Bayesian paradigm, thus estimating the inefficiencies alongside the other model parameters through simulation. Schmidt et al. (2009) propose a specification for panel data in which the inefficiencies are modeled using a truncated normal distribution in which the pre-truncated means are heterogeneous and spatially correlated via a conditional autoregressive prior. Hence, this specification allows the inefficiencies to be heteroskedastic and spatially correlated, but it has also some important drawbacks. First, the estimation of this model requires that T is relatively large with respect to n. Moreover, for identification purposes the realizations of the pre-truncated means are centred in zero in each Gibbs iteration, implicitly imposing a relevant characterization of the distribution of the inefficiencies which may not be supported by the data.[9] Finally, this specification does not allow to have efficiency spillovers among the units, since the correlation between the means does not imply a direct effect of the inefficiency of unit i on the one of the neighboring unit j. Areal et al. (2012), Tsionas & Michaelides (2016) and Carvalho (2018) consider special cases of our statistical model (e.g., time invariant and/or homoskedastic inefficiencies). However, the data augmentation step for the intrinsic inefficiencies proposed by Areal et al. (2012) and Carvalho (2018) assumes that the conditional distribution $f(\tilde{\boldsymbol{u}}|\boldsymbol{y},\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is the vector of unknown parameters to be estimated, can be factorized as independent univariate distributions while, even if the intrinsic inefficiencies are *a priori* independent, their conditional distribution is still multivariate. This is because, in presence of spatial dependence, neighboring units are informative about $\tilde{u}_{it}$ since their output is influenced by this latent factor. On the contrary, the *full* conditional distribution of $\tilde{u}_{it}$ given $\boldsymbol{\theta}$, the data and all the other intrinsic inefficiencies is a univariate (unknown) distribution.[10] Finally, concerning the work of Tsionas & Michaelides

[7] The family of truncated normal distributions is not closed to general linear transformation (Horrace, 2005), hence the two spatial specifications imply different distributions for the compounded error term.

[8] As proposed, for instance, by Pavlyuk (2012) and Fusco & Vidoli (2013) in a cross-sectional framework.

[9] The normalization of the location parameters implies that a non-negligible part of the statistical units will have a negative pre-truncation mean. A truncated normal distribution with a negative location parameter has a mode in zero and significantly more mass close to this point compared the half normal distribution.

[10] Updating unit intrinsic inefficiencies via Random-walk Metropolis produces a highly autocorrelated Gibbs

(2016) we note that, while the authors describe a statistical model equivalent to (8) (see p.246 of Tsionas & Michaelides, 2016), they end up estimating a specification similar to the one reported in (9).[11] It is worth emphasizing that since spatial interactions occur before truncation, it is not possible to distinguish between intrinsic and total inefficiency.

3 Estimation

In what follows, we describe two inferential procedures allowing the estimation of the parameters of model (1)-(5), both based on the strategy of augmenting the likelihood function by simulating the inefficiency vector. In particular, depending on the assumptions made about time invariant unobserved heterogeneity, we derive a simulated likelihood estimator that is computationally feasible even in presence of a spatial autoregressive structure of the inefficiency.

3.1 Estimation of SIM-TRE

In this section, we present an estimation strategy for the SIM-TRE. Given the independence assumption between $\boldsymbol{v}$, $\boldsymbol{\alpha}$ and $\tilde{\boldsymbol{u}}$, the likelihood function can be defined in general terms as

$$\begin{aligned} L(\boldsymbol{\theta}) &= \int f(\boldsymbol{v}, \tilde{\boldsymbol{u}}, \boldsymbol{\alpha} | \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{\theta}) d\boldsymbol{\alpha} d\tilde{\boldsymbol{u}} \\ &= \int f(\boldsymbol{y} | \boldsymbol{X}, \boldsymbol{\alpha}, \tilde{\boldsymbol{u}}, \boldsymbol{W}, \rho, \boldsymbol{\beta}, \psi^2, \tau^2) f(\tilde{\boldsymbol{u}} | \boldsymbol{Z}, \boldsymbol{\delta}) f(\boldsymbol{\alpha} | \tau^2) d\boldsymbol{\alpha} d\tilde{\boldsymbol{u}}, \end{aligned} \tag{12}$$

where $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{\delta}', \sigma, \tau, \psi, \rho)$. The derivation of the likelihood function hides two challenges: i) the marginalization of $\tilde{\boldsymbol{u}}$ and $\boldsymbol{\alpha}$ from $f(\boldsymbol{v}, \tilde{\boldsymbol{u}}, \boldsymbol{\alpha} | \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{\theta})$ may not lead to a closed-form expression; ii) even when an analytical expression is available (e.g., normal-half normal models, see Section 2), its maximization involves the direct evaluation of n-variate Gaussian integrals, thus leading to a computationally unfeasible estimator, even in presence of moderate sample sizes.

If the intrinsic inefficiency is assumed to belong to a one-parameter family of distributions (e.g., exponential or half normal), the marginalization in equation (12) can be performed by simulation. The log-likelihood can be expressed in term of its simulated counterpart as

$$\begin{aligned} \ell(\boldsymbol{\theta}) &= \log \left\{ \int \prod_{i=1}^{n} \phi_T \left[\boldsymbol{y}_i - \boldsymbol{X}_i \boldsymbol{\beta} + (\boldsymbol{I}_T \otimes \boldsymbol{d}_i)(\boldsymbol{\sigma} \odot \tilde{\boldsymbol{u}}); \boldsymbol{0}, \boldsymbol{I}_T \psi^2 + \boldsymbol{\iota}_T \boldsymbol{\iota}_T' \tau^2 \right] d\tilde{\boldsymbol{u}} \right\} \\ &= \log \left(E_{\tilde{\boldsymbol{u}}} \left\{ \prod_{i=1}^{n} \phi_T \left[\boldsymbol{y}_i - \boldsymbol{X}_i \boldsymbol{\beta} + (\boldsymbol{I}_T \otimes \boldsymbol{d}_i)(\boldsymbol{\sigma} \odot \tilde{\boldsymbol{u}}); \boldsymbol{0}, \boldsymbol{I}_T \psi^2 + \boldsymbol{\iota}_T \boldsymbol{\iota}_T' \tau^2 \right] \right\} \right) \\ &\approx \log \left\{ \frac{1}{R} \sum_{r=1}^{R} \prod_{i=1}^{n} \phi_T \left[\boldsymbol{y}_i - \boldsymbol{X}_i \boldsymbol{\beta} + (\boldsymbol{I}_T \otimes \boldsymbol{d}_i)(\boldsymbol{\sigma} \odot \tilde{\boldsymbol{u}}_r); \boldsymbol{0}, \boldsymbol{I}_T \psi^2 + \boldsymbol{\iota}_T \boldsymbol{\iota}_T' \tau^2 \right] \right\}, \end{aligned} \tag{13}$$

chain for all parameters. In the case of Carvalho (2018), 30000 MCMC draws from the posterior distribution are equivalent to less then 100 independent draws. This issue is further exacerbated when all remaining intrinsic inefficiencies are correctly included in the conditioning set of the n full conditional distributions.

[11] The ambiguity around the statistical model is also reflected in the fact that the authors added an "additional" inequality constraint in the distribution of the inefficiency which is not needed (see Tsionas & Michaelides, 2016, p. 250).

where $\boldsymbol{X}_i$ is the $T \times k$ matrix of regressors for unit i, $\phi_T(.)$ is the T-variate Gaussian density, $\boldsymbol{\iota}_T$ is a T-vector of ones, $\boldsymbol{d}_i$ is the i-th row of the matrix $\boldsymbol{D}$, $\tilde{\boldsymbol{u}}_r = (\tilde{u}_{11r}, ..., \tilde{u}_{nTr})'$ is the $nT \times 1$ vector of simulated draws from the one sided distribution $\mathcal{F}_{\tilde{u}}$ with scale parameter equal to 1, $\boldsymbol{\sigma} = g(\boldsymbol{Z\delta})$ is a $nT \times 1$ vector, R is the number of draws and, finally, $\otimes$ and $\odot$ denote, respectively, the Kronecker and the element-wise products. However, evaluating the terms $\prod_{i=1}^n \phi_T(.)$ is computationally challenging because the standard statistical packages tend to approximate them to zero. In the double digit approximation, the closest numbers to 0 without being approximated to 0 are $\pm 10^{-323}$ and, even in moderate sample sizes (i.e., $n = 100$ and $T = 5$), the large majority of the R addenda in (13) fall within the interval defined by these two thresholds.

Given that the log-likelihood in (13) cannot be maximized, we propose to use the simulation approach for independently approximating the likelihood contribution of each statistical unit as

$$\begin{aligned}\tilde{\ell}_i(\boldsymbol{\theta}) &= \log L_i(\boldsymbol{\theta}) \\ &\approx \log\left[\frac{1}{R}\sum_{r=1}^{R}\phi_T\left(\boldsymbol{y}_i - \boldsymbol{X}_i\boldsymbol{\beta} + (\boldsymbol{I}_T \otimes \boldsymbol{d}_i)(\boldsymbol{\sigma} \odot \tilde{\boldsymbol{u}}_r); \boldsymbol{0}, \boldsymbol{I}_T\psi_v^2 + \boldsymbol{\iota}_T\boldsymbol{\iota}_T'\tau^2\right)\right]. \qquad (14)\end{aligned}$$

The marginalization of the inefficiencies via simulations is now a simpler task for two reasons. First the simulation of $\tilde{\boldsymbol{u}}_r$ is straightforward given that its elements are i.i.d. and thus we have to draw from a univariate distribution. Second the evaluation of equation (14) only requires the calculation of T-variate normal densities.

If we combine the n simulated log-likelihood contributions $\tilde{\ell}_i(\boldsymbol{\theta})$ as

$$\tilde{\ell}(\boldsymbol{\theta}) = \sum_{i=1}^{n} \tilde{\ell}_i(\boldsymbol{\theta}), \qquad (15)$$

the resulting maximizer $\tilde{\boldsymbol{\theta}}$ can be viewed as a simulated composite likelihood estimator (SCLE) for the whole sample in which part of the dependence among the statistical units is voluntarily misspecified.[12] This approach is similar to the partial maximum likelihood approach described in Wang et al. (2013), but instead of working with the bivariate marginal likelihood contributions we work with the univariate marginal simulated ones.

The simulated approach requires that the number of simulations draws R is sufficiently large to guarantee that the simulated average in (14) is a good approximation of the corresponding expectation. Since the dimension of the integration problem is relatively high, a particular attention has to be paid on how to obtain the draws efficiently while guaranteeing a good multidimensional coverage. On this point, we found that the shuffled Halton sequences significantly reduce the computational burden in this context, while the use of traditional simulations techniques, such as pseudo-uniform random draws or standard Halton sequences, is not advisable due to their high correlation.[13]

[12] A detailed introduction to composite likelihood methods can be found in Varin et al. (2011).

[13] See Hess et al. (2003) for a detailed discussion about shuffled Halton sequences.

For inference purposes, we have to recognize that the SCLE belongs to the class of M-estimators, hence its asymptotic variance is equal to

$$\boldsymbol{A}_0^{-1}\boldsymbol{B}_0\boldsymbol{A}_0^{-1}, \tag{16}$$

where

$$\boldsymbol{A}_0 = -\mathbb{E}\left[\nabla_{\boldsymbol{\theta\theta}}\tilde{\ell}_i(\boldsymbol{\theta}_0)\right] \text{ and } \boldsymbol{B}_0 = \mathbb{E}\left[\nabla_{\boldsymbol{\theta}}\tilde{\ell}_i(\boldsymbol{\theta}_0)\nabla_{\boldsymbol{\theta}}\tilde{\ell}_i(\boldsymbol{\theta}_0)'\right], \tag{17}$$

with $\nabla_{\boldsymbol{\theta}}$ and $\nabla_{\boldsymbol{\theta\theta}}$ denoting the vector of first and second derivatives of the objective function respectively, and $\boldsymbol{\theta}_0$ is the true parameters vector.[14]

3.2 Estimation of SIM-TFE

The most appropriate estimation procedure for the SIM-TFE depends on the length of the panel. In long panels ($T > 10$), where the incidental parameters problem becomes negligible (Belotti & Ilardi, 2018), the unit-specific intercepts can be treated as parameters to be estimated as proposed by Greene (2005). On the other hand, when the panel is short, the incidental parameters problem can be avoided using a data transformation.

In the first case, we propose to exploit a simulated composite likelihood dummy variables approach. The log-likelihood contribution for unit i in period t can be expressed in terms of its simulated counterpart as

$$\begin{aligned}
\tilde{\ell}_{it}(\boldsymbol{\theta}) &= \log L_{it}(\boldsymbol{\theta}) \\
&= \log\left[\int f(y_{it}|\boldsymbol{x}_{it}, \alpha_i, \tilde{\boldsymbol{u}}, \boldsymbol{W}, \rho, \boldsymbol{\beta}, \psi^2) f(\tilde{\boldsymbol{u}}|\boldsymbol{Z}, \boldsymbol{\delta}) d\tilde{\boldsymbol{u}}\right] \\
&= \log E_{\tilde{\boldsymbol{u}}}\left[\phi\left(y_{it} - \alpha_i - \boldsymbol{x}_{it}\boldsymbol{\beta} + \boldsymbol{d}_i(\boldsymbol{\sigma}_t \odot \tilde{\boldsymbol{u}}_t); 0, \psi^2\right)\right] \\
&\approx \log\frac{1}{R}\sum_{r=1}^{R}\left[\phi\left(y_{it} - \alpha_i - \boldsymbol{x}_{it}\boldsymbol{\beta} + \boldsymbol{d}_i(\boldsymbol{\sigma}_t \odot \tilde{\boldsymbol{u}}_{tr}); 0, \psi^2\right)\right].
\end{aligned} \tag{18}$$

By exploiting the block diagonal structure of the Hessian matrix, the maximization of the simulated composite log-likelihood function based on (18) is computationally feasible also in presence of a large number of nuisance parameters.

In presence of short panels, a strategy to avoid the incidental parameters problem consists in eliminating the unit-specific effects through a data transformation. Following Belotti & Ilardi (2018), we can rewrite the model (1)-(5) in first-differences as

$$\Delta\boldsymbol{y}_i = \Delta\boldsymbol{X}_i\boldsymbol{\beta} - \Delta\boldsymbol{\eta}_i + \Delta\boldsymbol{v}_i, \tag{19}$$

[14]Even if we do not formally establish the asymptotic properties of the SCLE, it is worth noting that our Monte Carlo simulations show consistency of the proposed estimators and that the ratio between the average standard errors obtained using the sample analog of (16) and the standard deviations over replications of the estimated coefficients is close to one.

where $\Delta \boldsymbol{y}_i = (\Delta y_{i2}, ..., \Delta y_{iT})'$, $\Delta \boldsymbol{v}_i = (\Delta v_{i2}, ..., \Delta v_{iT})'$, $\Delta \boldsymbol{\eta}_i = (\Delta \eta_{i2}, ..., \Delta \eta_{iT})'$ and $\Delta \eta_{it} = \boldsymbol{d}_i \Delta \tilde{\boldsymbol{u}}_t$ with Δ the first-difference operator, i.e. $\Delta y_{it} = y_{it} - y_{it-1}$.[15]

The marginal log-likelihood contribution for the model in first-differences can be expressed in terms of its simulated counterpart as

$$\begin{aligned} \ell_i^*(\boldsymbol{\theta}) &= \log \left[\int f(\Delta \boldsymbol{y}_i | \boldsymbol{\theta}, \Delta \boldsymbol{X}_i, \Delta \boldsymbol{\eta}_i) f(\Delta \tilde{\boldsymbol{u}}_i | \boldsymbol{\sigma}) d\Delta \tilde{\boldsymbol{u}}_i \right] && (20) \\ &= \log \left\{ \mathbb{E}_{\Delta \tilde{\boldsymbol{u}}} \left[\phi_{T-1} \left(\Delta \boldsymbol{y}_i - \Delta \boldsymbol{X}_i \boldsymbol{\beta} + \Delta \boldsymbol{\eta}_i ; \boldsymbol{0}, \Psi_{T-1} \right) \right] \right\} && (21) \\ &\approx \log \left[\frac{1}{R} \sum_{r=1}^{R} \phi_{T-1} \left(\Delta \boldsymbol{y}_i - \Delta \boldsymbol{X}_i \boldsymbol{\beta} + \Delta \boldsymbol{\eta}_{ir} ; \boldsymbol{0}, \Psi_{T-1} \right) \right], && (22) \end{aligned}$$

where

$$\Psi_{T-1} = \psi^2 \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ 0 & -1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & -1 \\ 0 & 0 & \dots & -1 & 2 \end{pmatrix}, \tag{23}$$

and $\Delta \boldsymbol{\eta}_{ir} = (\Delta \eta_{i2r}, ..., \Delta \eta_{iTr})'$ with $\Delta \eta_{itr} = \boldsymbol{d_i}(\boldsymbol{\sigma}_t \odot \tilde{\boldsymbol{u}}_{rt} - \boldsymbol{\sigma}_{t-1} \odot \tilde{\boldsymbol{u}}_{rt-1})$ and $\tilde{\boldsymbol{u}}_{rt} = (\tilde{u}_{1tr}, ..., \tilde{u}_{ntr})'$. The maximization can then proceed by combining the n simulated log-likelihood contributions as in (15).[16]

3.3 Efficiency analysis

A fundamental feature of SF models is the estimation of technical (cost) inefficiency. The standard approach is to post-estimate the inefficiency of unit i in period t by exploiting the mean of the conditional distribution of u_{it} given ε_{it} (Jondrow et al., 1982, JLMS). However, given the spatial autoregressive structure of u_{it} in model (1)-(5), its post-estimation has to be based on the multivariate conditional distribution of $\tilde{\boldsymbol{u}} | \boldsymbol{\varepsilon}$.

[15] It is important to remark that unless the inefficiency can be decomposed additively in a time invariant and a time varying components, the TFE specification can be used to successfully separate unobserved heterogeneity from inefficiency. As discussed by Belotti & Ilardi (2018) for the non spatial TFE, despite the use of a fixed-effects killing transformation, the distributional assumptions allow to correctly identify the parameters associated with time invariant inefficiency factors. Moreover, other transformations can be used to remove the fixed-effects such as the within-group transformation or a more general unit-specific detrending transformation such as the one used in Atella et al. (2014) and Kutlu et al. (2019).

[16] Also in this case, the asymptotic variance has a sandwich structure as reported in equation (16) for the SIM-TRE.

In particular, the conditional mean of $\tilde{\boldsymbol{u}}|\boldsymbol{\varepsilon}$ can be expressed as

$$\begin{aligned}\mathbb{E}(\tilde{\boldsymbol{u}}|\boldsymbol{\varepsilon}) &= \int \tilde{\boldsymbol{u}} f(\tilde{\boldsymbol{u}}|\boldsymbol{\varepsilon}) d\tilde{\boldsymbol{u}} \\ &= \frac{1}{f(\boldsymbol{\varepsilon})} \int \tilde{\boldsymbol{u}} f(\tilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}) d\tilde{\boldsymbol{u}} \\ &= \frac{1}{f(\boldsymbol{\varepsilon})} \int \tilde{\boldsymbol{u}} f(\boldsymbol{\varepsilon}|\tilde{\boldsymbol{u}}) f(\tilde{\boldsymbol{u}}) d\tilde{\boldsymbol{u}} \\ &= \frac{\int \tilde{\boldsymbol{u}} f(\boldsymbol{\varepsilon}|\tilde{\boldsymbol{u}}) f(\tilde{\boldsymbol{u}}) d\tilde{\boldsymbol{u}}}{\int f(\boldsymbol{\varepsilon}|\tilde{\boldsymbol{u}}) f(\tilde{\boldsymbol{u}}) d\tilde{\boldsymbol{u}}}. \end{aligned} \tag{24}$$

Given that the integral in equation (24) is intractable, we propose to approximate it using simulations. Hence, the JLMS estimator can be expressed as

$$\begin{aligned}\mathbb{E}(\tilde{\boldsymbol{u}}|\hat{\boldsymbol{\varepsilon}}) &\approx \frac{\sum_{r=1}^{R} \hat{\boldsymbol{\sigma}} \odot \tilde{\boldsymbol{u}}_r f(\hat{\boldsymbol{\varepsilon}}_r|\tilde{\boldsymbol{u}}_r)}{\sum_{r=1}^{R} f(\hat{\boldsymbol{\varepsilon}}_r|\tilde{\boldsymbol{u}}_r)} \\ &= \sum_{r=1}^{R} \hat{w}_r \; \hat{\boldsymbol{\sigma}} \odot \tilde{\boldsymbol{u}}_r, \end{aligned} \tag{25}$$

where $\tilde{\boldsymbol{u}}_r = (\tilde{u}_{11r}, ..., \tilde{u}_{nTr})'$ is the $nT \times 1$ vector of simulated draws from the one sided distribution $\mathcal{F}_{\tilde{u}}$ with scale parameter equal to 1, and the weight $\hat{w}_r$ is defined, for the SIM-TRE, as

$$\hat{w}_r = \frac{\prod_{i=1}^{n} \phi_T \left(\boldsymbol{y}_i - \boldsymbol{X}_i \hat{\boldsymbol{\beta}} + (\boldsymbol{I}_T \otimes \hat{\boldsymbol{d}}_i)(\hat{\boldsymbol{\sigma}} \odot \tilde{\boldsymbol{u}}_r); \boldsymbol{0}, \boldsymbol{I}_T \hat{\psi}_v^2 + \boldsymbol{\iota}_T \boldsymbol{\iota}_T' \hat{\tau}^2 \right)}{\sum_{r=1}^{R} \prod_{i=1}^{n} \phi_T \left(\boldsymbol{y}_i - \boldsymbol{X}_i \hat{\boldsymbol{\beta}} + (\boldsymbol{I}_T \otimes \hat{\boldsymbol{d}}_i)(\hat{\boldsymbol{\sigma}} \odot \tilde{\boldsymbol{u}}_r); \boldsymbol{0}, \boldsymbol{I}_T \hat{\psi}_v^2 + \boldsymbol{\iota}_T \boldsymbol{\iota}_T' \hat{\tau}^2 \right)}, \tag{26}$$

while, for the SIM-TFE, we have that

$$\hat{w}_r = \frac{\prod_{i=1}^{n} \prod_{t=1}^{T} \phi \left(y_{it} - \hat{\alpha}_i - \boldsymbol{x}_{it} \hat{\boldsymbol{\beta}} + \hat{\boldsymbol{d}}_i (\hat{\boldsymbol{\sigma}}_t \odot \tilde{\boldsymbol{u}}_{tr}); 0, \hat{\psi}^2 \right)}{\sum_{r=1}^{R} \prod_{i=1}^{n} \prod_{t=1}^{T} \phi \left(y_{it} - \hat{\alpha}_i - \boldsymbol{x}_{it} \hat{\boldsymbol{\beta}} + \hat{\boldsymbol{d}}_i (\hat{\boldsymbol{\sigma}}_t \odot \tilde{\boldsymbol{u}}_{tr}); 0, \hat{\psi}^2 \right)}. \tag{27}$$

Here we face a computational problem similar to the one described in Section 3: the evaluation of equation (26) involves the product of n T-variate densities (or nT univariate Gaussian densities in equation 27). The problem is that, even with moderate sample size, most of the weights are approximated to zero. For this reason, and in analogy with the strategy adopted for estimation, we propose to use the following estimator for the SIM-TRE case

$$\mathbb{E}(\tilde{\boldsymbol{u}}_i|\hat{\boldsymbol{\varepsilon}}_i) \approx \sum_{r=1}^{R} \hat{\omega}_{ir} \; \hat{\boldsymbol{\sigma}}_i \odot \tilde{\boldsymbol{u}}_{ir}, \quad \forall i = 1, \dots, n, \tag{28}$$

where

$$\hat{\omega}_{ir} = \frac{\phi_T \left(\boldsymbol{y}_i - \boldsymbol{X}_i \hat{\boldsymbol{\beta}} + (\boldsymbol{I}_T \otimes \hat{\boldsymbol{d}}_i)(\hat{\boldsymbol{\sigma}} \odot \tilde{\boldsymbol{u}}_r); \boldsymbol{0}, \boldsymbol{I}_T \hat{\psi}_v^2 + \boldsymbol{\iota}_T \boldsymbol{\iota}_T' \hat{\tau}^2 \right)}{\sum_{r=1}^{R} \phi_T \left(\boldsymbol{y}_i - \boldsymbol{X}_i \hat{\boldsymbol{\beta}} + (\boldsymbol{I}_T \otimes \hat{\boldsymbol{d}}_i)(\hat{\boldsymbol{\sigma}} \odot \tilde{\boldsymbol{u}}_r); \boldsymbol{0}, \boldsymbol{I}_T \hat{\psi}_v^2 + \boldsymbol{\iota}_T \boldsymbol{\iota}_T' \hat{\tau}^2 \right)}, \tag{29}$$

Similarly, for the SIM-TFE, we can use

$$\mathbb{E}(\tilde{u}_{it}|\hat{\varepsilon}_{it}) \approx \sum_{r=1}^{R} \hat{\omega}_{itr}\ \hat{\sigma}_{it}\tilde{u}_{itr}, \quad \forall i = 1, \dots, n \quad t = 1, \dots, T, \tag{30}$$

where $\hat{\omega}_{itr}$ is defined as

$$\hat{\omega}_{itr} = \frac{\phi\left(y_{it} - \hat{\alpha}_i - \boldsymbol{x}_{it}\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{d}}_i(\hat{\boldsymbol{\sigma}}_t \odot \tilde{\boldsymbol{u}}_{tr}); 0, \hat{\psi}^2\right)}{\sum_{r=1}^{R} \phi\left(y_{it} - \hat{\alpha}_i - \boldsymbol{x}_{it}\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{d}}_i(\hat{\boldsymbol{\sigma}}_t \odot \tilde{\boldsymbol{u}}_{tr}); 0, \hat{\psi}^2\right)}. \tag{31}$$

The evaluation of equation (31) requires, for each unit i, the estimation of α_i. When the model is estimated exploiting the simulated function (22), the fixed-effects are ruled out from the parameter space, hence their estimation has to be performed in a second stage. The $n \times 1$ vector of fixed-effects can be obtained as

$$\hat{\boldsymbol{\alpha}} = \frac{1}{T}\sum_{t=1}^{T}\left(\boldsymbol{y}_t - \boldsymbol{X}_t\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{c}}_t\right), \tag{32}$$

where $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{c}}_t = \mathbb{E}(\tilde{\boldsymbol{u}}_t|\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\sigma}}_t)$ are consistent estimates. In particular, $\hat{\boldsymbol{c}}_t = \hat{\boldsymbol{D}}\hat{\boldsymbol{\sigma}}_t$ when $\tilde{u}_{it} \sim \mathcal{E}(\sigma_{it})$ and $\hat{\boldsymbol{c}}_t = \sqrt{2\pi^{-1}}\hat{\boldsymbol{D}}\hat{\boldsymbol{\sigma}}_t$ when $\tilde{u}_{it} \sim \mathcal{N}^+(0, \sigma_{it}^2)$ ($\hat{\boldsymbol{\sigma}}_t = \hat{\sigma}$ in the homoskedastic case).[17] This estimator is equivalent to the mean-adjusted estimator of $\boldsymbol{\alpha}$ in the fixed-effects linear model.

In all cases, the global inefficiency can be computed using

$$\hat{\boldsymbol{u}} = \left[(\boldsymbol{I}_n - \hat{\rho}\boldsymbol{W})^{-1} \otimes \boldsymbol{I}_T\right]\hat{\boldsymbol{\xi}}, \tag{33}$$

where $\hat{\boldsymbol{\xi}} = \mathbb{E}(\tilde{\boldsymbol{u}}|\hat{\boldsymbol{\varepsilon}})$. A measure of global technical efficiency (TE) for unit i in period t is given by

$$TE_{it} = \exp(-\hat{u}_{it}). \tag{34}$$

Under stationarity, the Leontief expansion of $\boldsymbol{D}$ is $\boldsymbol{D} = \boldsymbol{I} + \rho\boldsymbol{W} + \rho^2\boldsymbol{W}^2 + \dots$, thus the vector of global efficiency can be approximately decomposed as

$$TE_t = \exp(-\hat{\boldsymbol{D}}\hat{\boldsymbol{\xi}}_t) = \underbrace{\exp(-\hat{\boldsymbol{\xi}}_t)}_{\text{Intrinsic efficiency}} \times \underbrace{\exp(-\hat{\rho}\boldsymbol{W}\hat{\boldsymbol{\xi}}_t) \times \dots \times \exp(-\hat{\rho}^H\boldsymbol{W}^H\hat{\boldsymbol{\xi}}_t)}_{\text{Spillover efficiency}}, \tag{35}$$

where the total efficiency can be obtained as the product of intrinsic and spillover efficiency.[18]

In the spirit of LeSage & Pace (2009), the $n \times 1$ vector of global (total) inefficiencies $\hat{\boldsymbol{u}}_t$, can

[17] Equation (32) refers to the case of production frontiers. For cost frontiers, the $\hat{\boldsymbol{c}}_t$ term enters the expression with a minus sign.

[18] We follow Waugh (1950) for selecting H in the Leontief expansion by setting the approximation error to 10^{-7}.

be decomposed into direct ($\hat{\boldsymbol{u}}_t^{Dir}$) and indirect ($\hat{\boldsymbol{u}}_t^{Indir}$) inefficiencies as [19]

$$\hat{\boldsymbol{u}}_t = \hat{\boldsymbol{D}} \begin{pmatrix} \hat{\xi}_{1t} \\ \hat{\xi}_{2t} \\ . \\ . \\ \hat{\xi}_{nt} \end{pmatrix} = \begin{pmatrix} \hat{u}_{11t}^{Dir} & + & \hat{u}_{12t}^{Indir} & + & . & . & + & \hat{u}_{1nt}^{Indir} \\ \hat{u}_{21t}^{Indir} & + & \hat{u}_{22t}^{Dir} & + & . & . & + & \hat{u}_{2nt}^{Indir} \\ . & + & . & + & . & . & + & . \\ . & + & . & + & . & . & + & . \\ \hat{u}_{n1t}^{Indir} & + & \hat{u}_{n2t}^{Indir} & + & . & . & + & \hat{u}_{nnt}^{Dir} \end{pmatrix}. \quad (36)$$

The diagonal elements of (36) univocally define the unit-specific direct inefficiencies, $\hat{u}_{ii,t}^{Dir}$, as the sum of the intrinsic inefficiency of unit i plus the feedback effects, the latter being the impacts passing through neighboring units and back to the unit itself. As for the indirect inefficiency, it is possible to construct two different measures: $\hat{u}_{i,t}^{Indir} = \sum_{j \neq i} \hat{u}_{ij,t}^{Indir}$, which refers to the inefficiency spillovers to unit i from all the j units and, $\hat{u}_{j,t}^{Indir} = \sum_{i \neq j} \hat{u}_{ji,t}^{Indir}$, which represents the spillovers from unit j to all i units. Since $\hat{u}_{i,t}^{Indir}$ can be thought as the portion of total inefficiency imported from other units while $\hat{u}_{j,t}^{Indir}$ represent the inefficiency that is exported to other units, this decomposition can be useful for the identification of the most relevant productive units in generating the spillover effects.

4 Monte Carlo evidence

In this Section we study the finite sample properties of the SCLE for a heteroskedastic SIM via numerical simulations.[20] In all the experiments we consider the following heteroskedastic normal-half normal SIM

$$\begin{aligned} y_{it} &= 1 + \alpha_i + \beta x_{it} + v_{it} - u_{it} && (37) \\ v_{it} &\sim \mathcal{N}(0, \psi^2) && (38) \\ u_{it} &= \rho \sum_{j=1}^{n} w_{ij} u_{jt} + \tilde{u}_{it} && (39) \\ \tilde{u}_{it} &\sim \mathcal{N}^+(0, \sigma_{it}^2) && (40) \\ \sigma_{it} &= \exp(\delta_0 + 0.5 z_{it}) && (41) \end{aligned}$$

where $z_{it} \sim \mathcal{N}(0,1)$, $\beta = 1$, $\psi = \tau = 0.25$. We investigate the effect of different sample sizes ($n = 121, 256$), different (average) signal-to-noise ratios $\bar{\lambda} = \frac{\bar{\sigma}}{\psi} = 1$ or 2, obtained by setting $\delta_0 = (-1.5, -0.8)$ and different degrees of spatial interaction ($\rho = 0.3, 0.7$). Following Shi & Lee (2017), the spatial weights matrix W is generated from as rook matrix. Individual units are arranged row by row on a $n \times n$ chessboard where neighbors are defined as those who share

[19] In our case this decomposition derives naturally from the statistical model, while models with a spatial autoregressive structure for the dependent variable, e.g. SAR or SDM specifications, require a reduced form. This implies that the point estimates of global inefficiency are not comparable with those obtained from a non-spatial SF model since they are not relative to the estimated best practice frontier.

[20] The Stata code implementing the methods described in this paper is available upon request.

a common border. Units in the interior of the chessboard have 4 neighbors, and units on the border and corner have respectively 3 and 2 neighbors.[21] Define the $n \times n$ matrix $\tilde{M}_n$ such that $\tilde{M}_{n,ij} = 1$ if and only if units i and j are neighbors on the chessboard, and $\tilde{M}_{n,ij} = 0$ otherwise. The spatial weights matrix W_n is defined as a row normalized $\tilde{M}_n$.[22] All simulation designs have a common base: the unit-specific parameters $\alpha_1, ..., \alpha_n$ are drawn from a standard Gaussian random variable, just one explanatory variable is used $x_{it} \sim \mathcal{N}(0,1)$ when the model is SIM-TRE and $x_{it} = 0.5\alpha_i + \sqrt{0.5^2} w_{it}$ with $w_{it} \sim \mathcal{N}(0,1)$ when the model is SIM-TFE. Finally, the number of replications is 1000.

Simulation results are summarized for each set of simulations by reporting the average bias and Mean Squared Error (MSE) of the estimates, together with the linear ($r_{u,\hat{u}}$) and the Spearman rank correlation coefficients between the (true) simulated inefficiencies and the estimated ones. The inefficiency's bias and MSE are reported for both intrinsic and global inefficiency scores and are computed for each replication over the $N = n \times T$ observations, and then these quantities are averaged over replications, e.g., MSE($\hat{u}_{it}$)$= R^{-1} \sum_{r=1}^{R} (NT)^{-1} \sum_{i=1}^{n} \sum_{t=1}^{T} (E(u_{it}|\hat{\varepsilon}_{it}) - u_{it}^0)^2$, where E$(u_{it}|\hat{\varepsilon}_{it})$ is the JLMS estimate and u_{it}^0 is the simulated (true) inefficiency.

4.1 Random-effects estimator

The first simulation exercise considers the SIM in the random-effects case. Table 1 and 2 summarize the simulation results showing evidence of consistency of the SCLE: an increase in the cross-sectional dimension produces significant reductions in both bias and MSE. In the "low" spatial correlation case ($\rho = 0.3$), all parameters are accurately estimated regardless of the value of $\bar{\lambda}$. In particular, the bias is negligible, in the order of the third decimal digit, except for γ_0 and γ_1 for which, in any case, the maximum absolute percentage bias is around 8 percent ($n = 121$, $T = 5$, $\bar{\lambda} = 1$).

In the "high" spatial correlation case ($\rho = 0.7$), we still observe evidence of consistency, although when $\bar{\lambda} = 2$, the estimator of β_0 is less accurate. Overall, we find that the spatial correlation parameter ρ is always very well estimated.

The inefficiencies are accurately estimated in all the scenarios, with the correlation between u and $\hat{u}$ ranging from 76% to 88%. We do not find significant improvements when the cross-sectional dimension increases, implying that inefficiencies are correctly estimated even in small samples. Likewise, the results remain substantially unaffected by changes in ρ. On the other hand, as expected, the performance of the inefficiency estimator are greatly affected by the signal-to-noise ratio (Wang & Schmidt, 2009).

[21] This design of the spatial weights matrix is motivated by the observation that regions in most observed regional structures have similar connectivity as units in a rook matrix.

[22] Note that a row-normalized $\boldsymbol{W}$ is consistently used in Monte Carlo experiments in the spatial econometrics literature (e.g., Shi & Lee, 2017; LeSage & Pace, 2018).

4.2 Fixed-effects estimator

Table 3 and 4 summarize the simulation results for the SCLE, using the same structure adopted before. Overall the results show that the SCLE does not suffer the incidental parameters problem, especially in the estimation of ψ. In the "low" spatial correlation case all parameters are accurately estimated with significant reductions in both bias and MSE when n gets larger. Similarly to the TRE case, the bias is always negligible except for γ_0 for which the maximum absolute percentage bias is around 13 percent ($n = 121$, $T = 5$, $\bar{\lambda} = 1$).

When ρ gets larger, although the estimator still shows consistency, γ_0 has a moderate bias when $n = 121$ and $\bar{\lambda} = 2$. Nevertheless, it is worth emphasizing that the bias is halved when the cross-sectional dimension increases, going from around -0.11 to -0.053. Overall, we find that the performances of the fixed-effects SCLE are comparable with those of the random-effects case.

As for the TRE case, the inefficiency scores are accurately estimated in all the scenarios, with the correlation between u and $\hat{u}$ ranging from 75% to 86%. As expected, the performances of the JLMS estimator remains substantially the same in the estimation of intrinsic inefficiencies, while they improve when ρ gets larger, especially the Spearman rank correlation. It is worth noting that we do not observe noticeable differences in the performance of the JLMS estimator compared to the random-effects case. This evidence suggests that, even in short panels, the estimation of the inefficiency scores is not affected by the additional stage required to estimate the fixed-effects.

5 Empirical application: Indonesian Rice Farming

In this section we apply the proposed estimators to a balanced panel of Indonesian rice farms. This data set has been analyzed by several authors (Erwidodo, 1990; Lee & Schmidt, 1993; Horrace & Schmidt, 1996, 2000; Druska & Horrace, 2004) and is described in Section 5.1. We estimate both the true random- and fixed-effects versions of the Spatial Inefficiency Model (denoted SIM-TRE and SIM-TFE, respectively). We also estimate the corresponding non-spatial versions of the true random- and fixed-effects SF models (denoted NS-TRE and NS-TFE, respectively) together with the pooled non-spatial SF model (PSF). Except for the PSF model, all other specifications account for time invariant unobserved heterogeneity and, the SIM-TRE and SIM-TFE account for global spatial interactions via a spatial auto-regressive inefficiency.

Following Druska & Horrace (2004), we employ the following Cobb-Douglas specification for the frontier function:

$$y_{it} = \alpha_i + \boldsymbol{g}_{it}\boldsymbol{\beta} + \boldsymbol{s}_{it}\boldsymbol{\delta} + \eta_1 t + \eta_2 t^2 + v_{it} - u_{it} \tag{42}$$

$$u_{it} = \rho \sum_{j=1}^{171} w_{ij} u_{jt} + \tilde{u}_{it} \tag{43}$$

$$\tilde{u}_{it} \sim \mathcal{N}^+(0, \sigma_{it}^2) \tag{44}$$

$$\sigma_{it}^2 = \exp(\boldsymbol{z}_{it}\boldsymbol{\gamma}) \tag{45}$$

where y_{it} is output of the i-*th* farm at time t and $\boldsymbol{g}_{it}$ is a vector of inputs (in logs). The remaining variables in equation (42) are: *i*) a vector of dummy variables to shift the frontier technology $\boldsymbol{s}_{it}$; *ii*) a second order polynomial in time to account for Hicks-neutral technological change (η_1, η_2). In the case of NS-TRE and SIM-TRE the last element of $\boldsymbol{g}_{it}$ is the intercept. Equation (45) includes a set of variables to model the variance of the intrinsic inefficiency $\boldsymbol{z}_{it}$. Finally, $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{\delta}', \eta_1, \eta_2, \rho, \boldsymbol{\gamma}', \nu_1, \nu_2)$ is the parameter vector to be estimated.

5.1 Data and spatial weights matrix

The data originates from a survey, dated 1977, of 171 rice farms located in the production area of the Cimanuk River Basin in West Java, Indonesia. The survey aimed at investigating their farming practices over six (three wet and three dry) growing seasons.[23] The summary statistics are reported in Table 5. In our application, kilograms of rice produced by farm i at time t is the output. Inputs include: seed (kg), urea (kg), trisodium phosphate (kg), labor (labor-hours), and land (hectares). We include a set of dummy variables that shift the production frontier and that are, respectively, equal to one if pesticide are used, if high yielding varieties are planted, if mixed varieties are planted or if it was a wet season. Furthermore, we interact the wet season and the pesticide dummy because more bugs may be present during the wet season. Finally, we control for a quadratic time trend.

As determinant of inefficiency equation we include the share of family labor over total labor, a dummy variable equal to one if the farmer doesn't share the field with others and an dummy equal to one if the wage payed in that farm is above the average wage in the period. All variables are in logs except shares and dummies.

As for the spatial weighting matrix $\boldsymbol{W}$ we follow Druska & Horrace (2004) and construct a weighting matrix based on the the inverse of geographical distance between individual farms. Since we only have information on the villages where the farms are located we use geographical coordinates of the villages to determine physical distances between producing units while the individual distance between farms within the same village are arbitrarily set to 10 km. In order to maintain the symmetry of $\boldsymbol{W}$ we follow Ord (1975) and use $\boldsymbol{W} = \boldsymbol{M}^{-1/2}\boldsymbol{W}_0\boldsymbol{M}^{-1/2}$, where $\boldsymbol{W}_0$ is the non-normalized spatial weighting matrix and $\boldsymbol{M}$ is a diagonal matrix containing the row sums of $\boldsymbol{W}_0$.

5.2 Results

The empirical results for the five estimated models are reported in Table 6. In terms of estimated frontier parameters, most of the results are quite similar. As expected all the inputs (labor, land, seed, urea and phosphate) are productive with positive and statistically significant elasticities, and returns-to-scale are constant or slightly decreasing. These findings are consistent with previous analyses of this data (Lee & Schmidt, 1993; Druska & Horrace, 2004). High-yielding and mixed rice varieties are significantly more productive than traditional varieties, and productivity

[23] For a detailed discussion of the data see Erwidodo (1990).

is higher during the wet season than in the dry season. We do not find a significant effects of pesticides adoption even if we explicitly control for their usage in the wet season, which could be characterized by more water and insects. As for the coefficient estimates of the second order polynomial in time, they describe a U-shaped technological trend.

As for the determinants of technical inefficiency, we note that the negative effect associated with paying a wage above the average becomes statistically significant when we control for time-invariant unobserved heterogeneity. Furthermore, not sharing the plot with somebody else increases the variance of inefficiency but only when we add a spatial autoregressive structure for the latter. Across all models we find no significant effect for the share of family labour.

As mentioned in Section 1, our spatial inefficiency random- and fixed-effects models (SIM-TRE and SIM-TFE) nest both their non-spatial version (NS-TRE and NS-TFE) and the pooled one (PSF). Hence, a set of appropriate Wald and Hausman tests can be used to select the best model. First, we note that the $\hat{\tau}$ parameter, i.e. the estimated standard deviation of $\boldsymbol{\alpha}$ in the "true" random-effects variant of the fitted SF models, is strongly statistically significant.[24] This can be interpreted as evidence of the presence of time-invariant unobserved heterogeneity and, consequently, the NS-TRE is preferred to the PSF model.[25] Secondly, the estimated spatial autoregressive parameter $\hat{\rho}$ in the SIM-TRE model is 0.83 and strongly statistically significant, signalling the presence of strong spatial correlation in the inefficiency component.[26] Based on this evidence, we conclude that a SF model controlling simultaneously for both time-invariant unobserved heterogeneity and a spatial autoregressive inefficiency better fits the data at hand. Finally, in order to choose between a fixed- or a random-effects specification, we rely on the Hausman test.[27] The p-value for the aforementioned test reported in Table 6 suggest that we cannot reject the random-effects assumption at the 1% level. Given the moderate sample size, we interpret this result as evidence in favor of the SIM-TRE model.

As shown in Figure 1, the estimated intrinsic technical inefficiency scores from SIM-TRE is very similar to the scores obtained from the NS-TRE model while the global scores are substantially higher. Since the global inefficiency is also a function of neighboring unit inefficiencies and given the very high estimated spatial correlation coefficient, we are not surprised by this finding. A similar evidence can be depicted from the bottom panel of Table 6 where the summary statistics for the estimated inefficiency scores are reported. We find that the maximum value estimated for the NS-TRE inefficiency is equal to the minimum value obtained for the global inefficiency from the SIM-TRE.

Explicitly introducing the spatial dimension also has a high impact on the ranking, in terms

[24]Testing the statistical significance of τ is a non standard problem since under the null hypothesis the parameter lies on the boundary of the parameter space (Molenberghs & Verbeke, 2007).

[25]The same conclusion can be obtained in this case using a non standard LR test.

[26]Applying a different statistical model on the same data, Druska & Horrace (2004) find similar values for the spatial correlation coefficient. We also estimated the heteroskedastic version of the SDM SF model proposed by Glass et al. (2017) finding a small and statistically insignificant estimate for the spatial correlation parameter.

[27]The test has been implemented by assessing the joint statistical significance of the Mundlak terms in a correlated SIM-TRE.

of efficiency, of the productive units. This evidence is clear by looking at Table 7 where the pairwise Spearman correlation rank between different models is reported. We note that there is a very high correlation (about 0.93) between the rankings obtained from PSF and NS-TRE while their correlation with our preferred spatial inefficiency model drops below 70%.

Finally, Figure 2 shows the evolution over time of direct ($\hat{u}_{i,t}^{dir}$), indirect ($\hat{u}_{i,t}^{Indir}$) and global technical inefficiency for a selected farm. For example, between the first and the second cropping seasons, while the direct inefficiency remains almost constant, the global inefficiency decreases. This is the result of some kind of positive externalities coming from neighboring units (the indirect inefficiency decreases). Similarly, between seasons four and five, the increase of the considered unit's direct inefficiency is offset by a positive spillover, leading to a decrease in the global inefficiency. The same evidence is shown in Figure 3, where the global efficiency is the product between intrinsic and spillover efficiency.

6 Concluding remarks

In this work we propose a statistical model in which time invariant unobserved heterogeneity may be disentangled from time-varying heteroskedatic inefficiency while allowing for weak cross-sectional dependence, namely spatial dependence. This model nests the "true" fixed- and random-effects models proposed by Greene (2005) as well as other well known SF models for panel data. Compared to SAR SF models for panel data, our model features a measure of total inefficiency that is relative to the estimated best practice frontier.

Regardless of the assumptions on the relationship between the explanatory variables and the time-invariant unobserved heterogeneity, a spatial autoregressive inefficiency component makes the standard maximum likelihood (ML) estimation of the parameters unfeasible since it implies the approximation of n-variate Gaussian integrals. Furthermore we find that, even with moderate cross-sectional dimensions, the simulated ML approach cannot be applied since the vast majority of simulated likelihood contributions is approximated to zero being below the minimum allowed number in the standard double digit approximation. We originally solve this problem by exploiting a simulated first-order composite likelihood approach, proposing both a fixed- and a random-effects version of our inferential procedure. We propose a simulation-based approach in order to obtain point estimates of inefficiency. In the spirit of LeSage & Pace (2009), the proposed procedure allows to disentangle unit specific inefficiency from system-wide spillovers and study their dynamic over time.

We study the finite sample behaviour of the proposed inferential procedures using Monte Carlo simulations finding convincing evidence of consistency when n gets larger with fixed T. We illustrate the usefulness of the new approach using a balanced panel of Indonesian rice farms, documenting the presence of strong spatial inefficiency spillovers. Furthermore, we show that considering a spatial autoregressive inefficiency can greatly influence the ranking of productive units and that this modelling framework may be useful to capture the spillover effects described by the agglomeration economics literature.

References

Adetutu, M., Glass, A., Kenjegalieva, K., & Sickles, R. (2015). The effects of efficiency and tfp growth on pollution in europe: A multistage spatial analysis. *J. Product. Anal.*, (pp. 307–326).

Aigner, D., Lovell, C., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, *6*, 21–37.

Areal, F. J., Balcombe, K., & Tiffin, R. (2012). Integrating spatial dependence into stochastic frontier analysis. *Australian Journal of Agricultural and Resource Economics*, *56*, 521–541.

Atella, V., Belotti, F., Depalo, D., & Piano Mortari, A. (2014). Measuring spatial effects in the presence of institutional constraints: The case of italian local health authority expenditure. *Regional Science and Urban Economics*, *49*, 232–241.

Audretsch, D. B., & Feldman, M. P. (2004). Chapter 61 - knowledge spillovers and the geography of innovation. In J. V. Henderson, & J.-F. Thisse (Eds.), *Cities and Geography* (pp. 2713 – 2739). Elsevier volume 4 of *Handbook of Regional and Urban Economics*.

Battese, G., & Coelli, T. (1988). Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. *Journal of Econometrics*, *38*, 387–399.

Battese, G., & Coelli, T. (1992). Frontier production functions, technical efficiency and panel data: with application to paddy farmers in india. *Journal of Productivity Analysis*, *3*, 153–169.

Battese, G., & Coelli, T. (1995). A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics*, *20*, 325–332.

Belotti, F., & Ilardi, G. (2018). Consistent inference in fixed-effects stochastic frontier models. *Journal of Econometrics*, *202*, 161 – 177.

Carvalho, A. (2018). Efficiency spillovers in bayesian stochastic frontier models: application to electricity distribution in new zealand. *Spatial Economic Analysis*, *13*, 171–190.

Cornwell, C., Schmidt, P., & Sickles, R. (1990). Production frontiers with cross-sectional and time series variation in efficiency levels. *Journal of Econometrics*, *46*, 185–200.

Dominguez-Molina, J., Gonzalez-Farias, G., & Ramos-Quiroga, R. (2004). Skew-normality in stochastic frontier analysis. In M. Genton (Ed.), *Skew Elliptical Distributions and their Applications: A Journey beyond Normality* chapter 13. (pp. 235–253). Boca Raton, Florida. Chapman and Hall / CRC.

Druska, V., & Horrace, W. C. (2004). Generalized moments estimation for spatial panel data: Indonesian rice farming. *American Journal of Agricultural Economics*, *86*, 185–198.

Elhorst, J. P. (2010). Applied spatial econometrics: raising the bar. *Spatial Economic Analysis*, *5*, 9–28.

Ellison, G., Glaeser, E. L., & Kerr, W. R. (2010). What causes industry agglomeration? evidence from coagglomeration patterns. *American Economic Review*, *100*, 1195–1213.

Erwidodo (1990). Panel data analysis on farm-level efficiency, input demand and output supply of rice farming in west java, indonesia. Ph.D. dissertation, Department of Agricultural Economics, Michigan State University.

Fusco, E., & Vidoli, F. (2013). Spatial stochastic frontier models: Controlling spatial global and local heterogeneity. *International Review of Applied Economics*, *27*, 679–694.

Glass, A., Kenjegalieva, K., & Paez-Farrell, J. (2013). Productivity growth decomposition using a spatial autoregressive frontier model. *Econom. Lett.*, (pp. 291–295).

Glass, A., Kenjegalieva, K., & Sickles, R. (2014). Estimating efficiency spillovers with state level evidence for manufacturing in the us. *Econom. Lett.*, (pp. 154–159).

Glass, A., Kenjegalieva, K., & Sickles, R. (2016). A spatial autoregressive stochastic frontier model for panel data with asymmetric efficiency spillovers. *Journal of Econometrics*, (pp. 289–300).

Glass, A. J., Kenjegalieva, K., Ajayi, V., Adetutu, M., & Sickles, R. C. (2017). Relative winners and losers from efficiency spillovers in africa with policy implications for regional integration. Nottingham Trent University, Discussion Papers in Economics.

Gonzalez-Farias, G., Dominguez-Molina, J., & Gupta, A. (2004). The closed skew normal distribution. In M. Genton (Ed.), *Skew Elliptical Distributions and their Applications: A Journey beyond Normality* chapter 2. (pp. 40–57). Boca Raton, Florida. Chapman and Hall / CRC.

Greene, W. (2005). Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *Journal of Econometrics*, *126*, 269–303.

Gude, A., Alvarez, I., & Orea, L. (2018). Heterogeneous spillovers among spanish provinces: a generalized spatial stochastic frontier model. *Journal of Productivity Analysis*, *50*, 155–173.

Hess, S., Polak, J. W., & Daly, A. (2003). On the performance of shuffled halton sequences in the estimation of discrete choice models. In "Proceedings of the European Transport Conference 2003 (Strasbourg, France)".

Horrace, W. C. (2005). Some results on the multivariate truncated normal distribution. *Journal of Multivariate Analysis*, *94*, 209–221.

Horrace, W. C., & Schmidt, P. (1996). Confidence statements for efficiency estimates from stochastic frontier models. *Journal of Productivity Analysis*, *7*, 257–282.

Horrace, W. C., & Schmidt, P. (2000). Multiple comparisons with the best, with economic applications. *Journal of Applied Econometrics*, *15*, 1–26.

Jondrow, J., Knox Lovell, C. A., Materov, I. S., & Schmidt, P. (1982). On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics*, *19*, 233–238.

Kumbhakar, S. (1990). Production frontiers, panel data and time-varying technical inefficiency. *Journal of Econometrics*, *46*, 201–212.

Kutlu, L., Tran, K., & Tsionas, E. (2019). Time-varying true individual effects model with endogenous regressors. *Journal of Econometrics (forthcoming)*, .

Lee, Y., & Schmidt, P. (1993). A production frontier model with flexible temporal variation in technical inefficiency. In H. Fried, C. Lovell, & S. Schmidt (Eds.), *The measurement of productive efficiency: techniques and applications*. Oxford University Press.

LeSage, J. P., & Pace, R. K. (2009). *Introduction to Spatial Econometrics*. Taylor & Francis.

LeSage, J. P., & Pace, R. K. (2018). Spatial econometric monte carlo studies: raising the bar. *Empirical Economics*, *55*, 17–34.

Martin, P., Mayer, T., & Mayneris, F. (2011). Spatial concentration and plant-level productivity in france. *Journal of Urban Economics*, *69*, 182 – 195.

Mastromarco, C., Serlenga, L., & Shin, Y. (2016). Modelling technical efficiency in cross sectionally dependent stochastic frontier panels. *Journal of Applied Econometrics*, *31*, 281–297.

Meeusen, W., & van den Broeck, J. (1977). Efficiency estimation from cobb-douglas production function with composed errors. *International Economic Review*, *18*, 435–444.

Molenberghs, G., & Verbeke, G. (2007). Likelihood ratio, score, and wald tests in a constrained parameter space. *The American Statistician*, *61*, 22–27.

Ord, K. (1975). Estimation methods for models of spatial interaction. *Journal of the American Statistical Association*, *70*, 120–126.

Pavlyuk, D. (2012). Maximum likelihood estimator for spatial stochastic frontier models. In "Proceedings of the 12th International Conference Reliability and Statistics in Transportation and Communication".

Pitt, M., & Lee, L. (1981). The measurement and sources of technical inefficiency in the indonesian weaving industry. *Journal of Development Economics*, *9*, 43–64.

Ramajo, J., & Hewings, G. J. D. (2018). Modelling regional productivity performance across western europe. *Regional Studies*, *52*, 1372–1387.

Rosenthal, S. S., & Strange, W. C. (2001). The determinants of agglomeration. *Journal of Urban Economics*, *50*, 191 – 229.

Rosenthal, S. S., & Strange, W. C. (2004). Chapter 49 - evidence on the nature and sources of agglomeration economies. In J. V. Henderson, & J.-F. Thisse (Eds.), *Cities and Geography* (pp. 2119 – 2171). Elsevier volume 4 of *Handbook of Regional and Urban Economics*.

Schmidt, A. M., Moreira, A. R. B., Helfand, S. M., & Fonseca, T. C. O. (2009). Spatial stochastic frontier models: Accounting for unobserved local determinants of inefficiency. *Journal of Productivity Analysis*, *31*, 101–112.

Schmidt, P., & Sickles, R. C. (1984). Production frontiers and panel data. *Journal of Business & Economic Statistics*, *2*, 367–74.

Shi, W., & Lee, L.-F. (2017). Spatial dynamic panel data models with interactive fixed effects. *Journal of Econometrics*, *197*, 323 – 347.

Tsionas, E. G., & Michaelides, P. G. (2016). A Spatial Stochastic Frontier Model with Spillovers: Evidence for Italian Regions. *Scottish Journal of Political Economy*, *63*, 243–257.

Varin, C., Reid, N., & Firth, D. (2011). An overview of composite likelihood methods. *Statistica Sinica*, *21*, 5–42.

Wang, H., Iglesias, E. M., & Wooldridge, J. M. (2013). Partial maximum likelihood estimation of spatial probit models. *Journal of Econometrics*, *172*, 77 – 89.

Wang, H., & Schmidt, P. (2002). One-step and two-step estimation of the effects of exogenous variables on technical efficiency levels. *Journal of Productivity Analysis*, *18*, 129–144.

Wang, W. S., & Schmidt, P. (2009). On the distribution of estimated technical efficiency in stochastic frontier models. *Journal of Econometrics*, *148*, 36 – 45.

Waugh, F. (1950). Inversion of the leontief matrix by power series. *Econometrica*, *18*, 142–154.

Table 1: Simulation results for the SIM-TRE. The Spearman rank correlation coefficient is in parentheses ($\bar{\lambda}$=1).

(a) $\rho = 0.3$

$n = 121$ $T = 5$

	Bias	MSE
β_0	0.0081	0.0098
β_1	-0.0008	0.0002
γ_0	-0.0279	0.0739
γ_1	0.0207	0.0167
ψ	-0.0054	0.0004
σ_a	-0.0035	0.0004
ρ	-0.0286	0.0387
E(u|ε)	0.0049	0.0270
$\mathrm{r}_{u,\hat{u}}$	0.7603	
	(0.6008)	
E($\tilde{u}$|ε)	-0.0010	0.0194
$\mathrm{r}_{\tilde{u},\hat{\tilde{u}}}$	0.7595	
	(0.5796)	

$n = 256$ $T = 5$

	Bias	MSE
β_0	0.0017	0.0046
β_1	-0.0004	0.0001
γ_0	-0.0300	0.0373
γ_1	0.0135	0.0080
ψ	-0.0033	0.0002
σ_a	0.0004	0.0002
ρ	-0.0136	0.0207
E(u|ε)	-0.0018	0.0229
$\mathrm{r}_{u,\hat{u}}$	0.7650	
	(0.6102)	
E($\tilde{u}$|ε)	-0.0045	0.0182
$\mathrm{r}_{\tilde{u},\hat{\tilde{u}}}$	0.7611	
	(0.5853)	

(b) $\rho = 0.7$

$n = 121$ $T = 5$

	Bias	MSE
β_0	-0.0054	0.0488
β_1	-0.0008	0.0002
γ_0	-0.0628	0.1223
γ_1	0.0392	0.0253
ψ	-0.0038	0.0008
σ_a	0.0028	0.0005
ρ	-0.0256	0.0148
E(u|ε)	-0.0129	0.0749
$\mathrm{r}_{u,\hat{u}}$	0.8038	
	(0.7189)	
E($\tilde{u}$|ε)	-0.0037	0.0189
$\mathrm{r}_{\tilde{u},\hat{\tilde{u}}}$	0.7725	
	(0.5922)	

$n = 256$ $T = 5$

	Bias	MSE
β_0	0.0043	0.0257
β_1	-0.0007	0.0001
γ_0	-0.0289	0.0358
γ_1	0.0121	0.0073
ψ	-0.0031	0.0004
σ_a	0.0048	0.0002
ρ	-0.0086	0.0066
E(u|ε)	-0.0041	0.0514
$\mathrm{r}_{u,\hat{u}}$	0.8132	
	(0.7332)	
E($\tilde{u}$|ε)	-0.0044	0.0174
$\mathrm{r}_{\tilde{u},\hat{\tilde{u}}}$	0.7755	
	(0.6003)	

Table 2: Simulation results for the SIM-TRE. The Spearman rank correlation coefficient is in parentheses ($\bar{\lambda}$=2).

(a) $\rho = 0.3$

$n = 121$ $T = 5$

	Bias	MSE
β_0	-0.0161	0.0115
β_1	-0.0012	0.0003
γ_0	-0.0335	0.0159
γ_1	0.0264	0.0063
ψ	0.0029	0.0006
σ_a	-0.0045	0.0006
ρ	-0.0243	0.0162
E(u|ε)	-0.0212	0.0535
$r_{u,\hat{u}}$	0.8742	
	(0.7431)	
E($\tilde{u}$|ε)	-0.0118	0.0425
$r_{\tilde{u},\hat{\tilde{u}}}$	0.8713	
	(0.7125)	

$n = 256$ $T = 5$

	Bias	MSE
β_0	-0.0166	0.0061
β_1	-0.0011	0.0001
γ_0	-0.0276	0.0084
γ_1	0.0168	0.0027
ψ	0.0030	0.0003
σ_a	-0.0003	0.0003
ρ	-0.0184	0.0096
E(u|ε)	-0.0228	0.0483
$r_{u,\hat{u}}$	0.8761	
	(0.7472)	
E($\tilde{u}$|ε)	-0.0119	0.0412
$r_{\tilde{u},\hat{\tilde{u}}}$	0.8718	
	(0.7161)	

(b) $\rho = 0.7$

$n = 121$ $T = 5$

	Bias	MSE
β_0	-0.1359	0.0853
β_1	-0.0012	0.0004
γ_0	-0.0683	0.0419
γ_1	0.0551	0.0123
ψ	0.0307	0.0024
σ_a	0.0171	0.0013
ρ	-0.0359	0.0099
E(u|ε)	-0.1482	0.1635
$r_{u,\hat{u}}$	0.8785	
	(0.8139)	
E($\tilde{u}$|ε)	-0.0166	0.0493
$r_{\tilde{u},\hat{\tilde{u}}}$	0.8573	
	(0.6942)	

$n = 256$ $T = 5$

	Bias	MSE
β_0	-0.0802	0.0508
β_1	-0.0011	0.0002
γ_0	-0.0464	0.0177
γ_1	0.0275	0.0043
ψ	0.0203	0.0012
σ_a	0.0145	0.0007
ρ	-0.0171	0.0047
E(u|ε)	-0.0948	0.1234
$r_{u,\hat{u}}$	0.8819	
	(0.8200)	
E($\tilde{u}$|ε)	-0.0159	0.0459
$r_{\tilde{u},\hat{\tilde{u}}}$	0.8599	
	(0.7013)	

Table 3: Simulation results for the SIM-TFE. The Spearman rank correlation coefficient is in parentheses ($\bar{\lambda}$=1).

(a) $\rho = 0.3$

$n = 121$ $T = 5$

	Bias	MSE
β_1	0.0006	0.0002
γ_0	-0.0622	0.0784
γ_1	0.0228	0.0155
ψ	-0.0031	0.0004
ρ	-0.0288	0.0417
E(u|ε)	-0.0035	0.0279
$\mathrm{r}_{u,\hat{u}}$	0.7480	
	(0.5911)	
E($\tilde{u}$|ε)	-0.0004	0.0200
$\mathrm{r}_{\tilde{u},\hat{\tilde{u}}}$	0.7472	
	(0.5714)	

$n = 256$ $T = 5$

	Bias	MSE
β_1	0.0000	0.0001
γ_0	-0.0363	0.0387
γ_1	0.0130	0.0083
ψ	-0.0024	0.0002
ρ	-0.0139	0.0228
E(u|ε)	-0.0013	0.0245
$\mathrm{r}_{u,\hat{u}}$	0.7581	
	(0.6027)	
E($\tilde{u}$|ε)	0.0004	0.0186
$\mathrm{r}_{\tilde{u},\hat{\tilde{u}}}$	0.7547	
	(0.5788)	

(b) $\rho = 0.7$

$n = 121$ $T = 5$

	Bias	MSE
β_1	0.0007	0.0003
γ_0	-0.0910	0.1265
γ_1	0.0373	0.0215
ψ	-0.0003	0.0007
ρ	-0.0154	0.0150
E(u|ε)	-0.0167	0.0725
$\mathrm{r}_{u,\hat{u}}$	0.7891	
	(0.7026)	
E($\tilde{u}$|ε)	0.0295	0.0210
$\mathrm{r}_{\tilde{u},\hat{\tilde{u}}}$	0.7618	
	(0.5855)	

$n = 256$ $T = 5$

	Bias	MSE
β_1	0.0000	0.0001
γ_0	-0.0421	0.0373
γ_1	0.0149	0.0084
ψ	-0.0027	0.0005
ρ	-0.0082	0.0079
E(u|ε)	0.0008	0.0630
$\mathrm{r}_{u,\hat{u}}$	0.8008	
	(0.7161)	
E($\tilde{u}$|ε)	0.0319	0.0193
$\mathrm{r}_{\tilde{u},\hat{\tilde{u}}}$	0.7679	
	(0.5915)	

Table 4: Simulation results for the SIM-TFE. The Spearman rank correlation coefficient is in parentheses ($\bar{\lambda}$=2).

(a) $\rho = 0.3$

$n = 121$ $T = 5$

	Bias	MSE
β_1	0.0008	0.0004
γ_0	-0.0570	0.0253
γ_1	0.0294	0.0079
ψ	0.0059	0.0008
ρ	-0.0213	0.0189
E(u|ε)	-0.0275	0.0593
$\mathrm{r}_{u,\hat{u}}$	0.8597	
	(0.7228)	
E($\tilde{u}$|ε)	-0.0074	0.0464
$\mathrm{r}_{\tilde{u},\hat{\tilde{u}}}$	0.8569	
	(0.6947)	

$n = 256$ $T = 5$

	Bias	MSE
β_1	0.0003	0.0002
γ_0	-0.0222	0.0078
γ_1	0.0099	0.0027
ψ	0.0020	0.0004
ρ	-0.0145	0.0103
E(u|ε)	-0.0163	0.0523
$\mathrm{r}_{u,\hat{u}}$	0.8641	
	(0.7308)	
E($\tilde{u}$|ε)	0.0001	0.0438
$\mathrm{r}_{\tilde{u},\hat{\tilde{u}}}$	0.8603	
	(0.7005)	

(b) $\rho = 0.7$

$n = 121$ $T = 5$

	Bias	MSE
β_1	0.0019	0.0006
γ_0	-0.1095	0.0734
γ_1	0.0601	0.0158
ψ	0.0353	0.0031
ρ	-0.0278	0.0127
E(u|ε)	-0.1398	0.1916
$\mathrm{r}_{u,\hat{u}}$	0.8559	
	(0.7825)	
E($\tilde{u}$|ε)	0.0442	0.0565
$\mathrm{r}_{\tilde{u},\hat{\tilde{u}}}$	0.8428	
	(0.6761)	

$n = 256$ $T = 5$

	Bias	MSE
β_1	0.0000	0.0002
γ_0	-0.0543	0.0183
γ_1	0.0287	0.0045
ψ	0.0208	0.0013
ρ	-0.0154	0.0050
E(u|ε)	-0.0924	0.1304
$\mathrm{r}_{u,\hat{u}}$	0.8642	
	(0.7939)	
E($\tilde{u}$|ε)	0.0540	0.0524
$\mathrm{r}_{\tilde{u},\hat{\tilde{u}}}$	0.8485	
	(0.6860)	

Table 5: Summary statistics (n=171, T=6)

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Total output (Kg)	1405.167	1921.757	42	20960
Seeds (Kg)	18.206	45.251	1	1250
Urea (Kg)	95.441	127.149	1	1250
Phosphate (Kg)	34.728	47.588	1	701
Total labor	388.447	484.204	17	4774
Cultivated area (Ha)	0.432	0.547	0.01	5.322
Pesticide	0.305	0.461	0	1
High yielding varieties	0.287	0.452	0	1
Mixed varieties	0.049	0.215	0	1
Wet season	0.5	0.5	0	1
Non sharecropper	0.717	0.451	0	1
Share of family labor	0.533	0.287	0	1
Wage above the average	0.516	0.5	0	1

Table 6: Estimation results on Indonesian rice farms data (n=171, T=6).

	PSF	NS-TRE	SIM-TRE	NS-TFE	SIM-TFE
			Frontier		
Seeds	0.158 ***	0.153 ***	0.134 ***	0.131 ***	0.126 **
	(0.025)	(0.025)	(0.034)	(0.028)	(0.038)
Urea	0.110 ***	0.100 ***	0.104 ***	0.087 ***	0.088 ***
	(0.016)	(0.017)	(0.021)	(0.020)	(0.025)
Phosphate	0.057 ***	0.060 ***	0.064 ***	0.064 ***	0.062 ***
	(0.010)	(0.010)	(0.011)	(0.012)	(0.015)
Total labor	0.214 ***	0.216 ***	0.230 ***	0.223 ***	0.233 ***
	(0.027)	(0.027)	(0.031)	(0.030)	(0.036)
Cultivated area	0.474 ***	0.473 ***	0.472 ***	0.447 ***	0.444 ***
	(0.029)	(0.029)	(0.044)	(0.033)	(0.054)
Pesticide	0.014	0.022	0.019	0.043	0.038
	(0.035)	(0.035)	(0.036)	(0.038)	(0.042)
High yielding varieties	0.156 ***	0.167 ***	0.130 ***	0.156 ***	0.149 ***
	(0.027)	(0.029)	(0.033)	(0.041)	(0.044)
Mixed varieties	0.124 *	0.136 **	0.129 **	0.147 **	0.155 **
	(0.049)	(0.049)	(0.043)	(0.053)	(0.052)
Wet season	0.070 **	0.073 **	0.074 **	0.085 ***	0.086 **
	(0.026)	(0.025)	(0.023)	(0.025)	(0.027)
Wet$\times$Pesticides	-0.017	-0.019	-0.016	-0.026	-0.016
	(0.045)	(0.043)	(0.042)	(0.042)	(0.047)
t	-0.263 ***	-0.257 ***	-0.251 ***	-0.243 ***	-0.236 ***
	(0.030)	(0.029)	(0.025)	(0.029)	(0.031)
t^2	0.039 ***	0.038 ***	0.039 ***	0.037 ***	0.038 ***
	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)
Constant	5.551 ***	5.618 ***	6.290 ***		
	(0.191)	(0.195)	(0.375)		
			Inefficiency		
Non sharecropper	0.848	0.610	0.444 *	0.259	0.387 *
	(0.574)	(0.396)	(0.226)	(0.132)	(0.181)
Share of family labor	0.350	0.326	-0.054	0.104	-0.022
	(0.644)	(0.488)	(0.298)	(0.175)	(0.232)
Wage above the average	-1.122	-0.856 *	-0.487 **	-0.317 ***	-0.334 ***
	(0.573)	(0.366)	(0.152)	(0.092)	(0.097)
Constant	-3.998 ***	-3.292 ***	-1.703 ***	-1.272 ***	-1.405 ***
	(1.133)	(0.866)	(0.409)	(0.223)	(0.323)
ρ			0.830 ***		0.866 ***
λ	0.518	0.791	0.729	1.284	1.187
$\bar{\sigma}_{\tilde{u}}$	0.160	0.217	0.199	0.309	0.277
σ_v	0.309 ***	0.275 ***	0.273 ***	0.241 ***	0.233 ***
τ		0.113 ***	0.111 ***		
			Estimated technical inefficiencies, $\hat{u}_{it}$		
Mean	0.128	0.173	0.922	0.243	1.623
SD	0.057	0.078	0.118	0.120	0.240
Min	0.047	0.054	0.667	0.046	1.141
Max	0.437	0.668	1.368	0.893	2.355
ℓ	-303.599	-289.838	-284.203^a	-347.690	-343.633^a

Notes: * 0.05, ** 0.01, *** 0.001.^{a} Simulated composite log-likelihood.

Figure 1: Kernel densities of technical inefficiency scores.

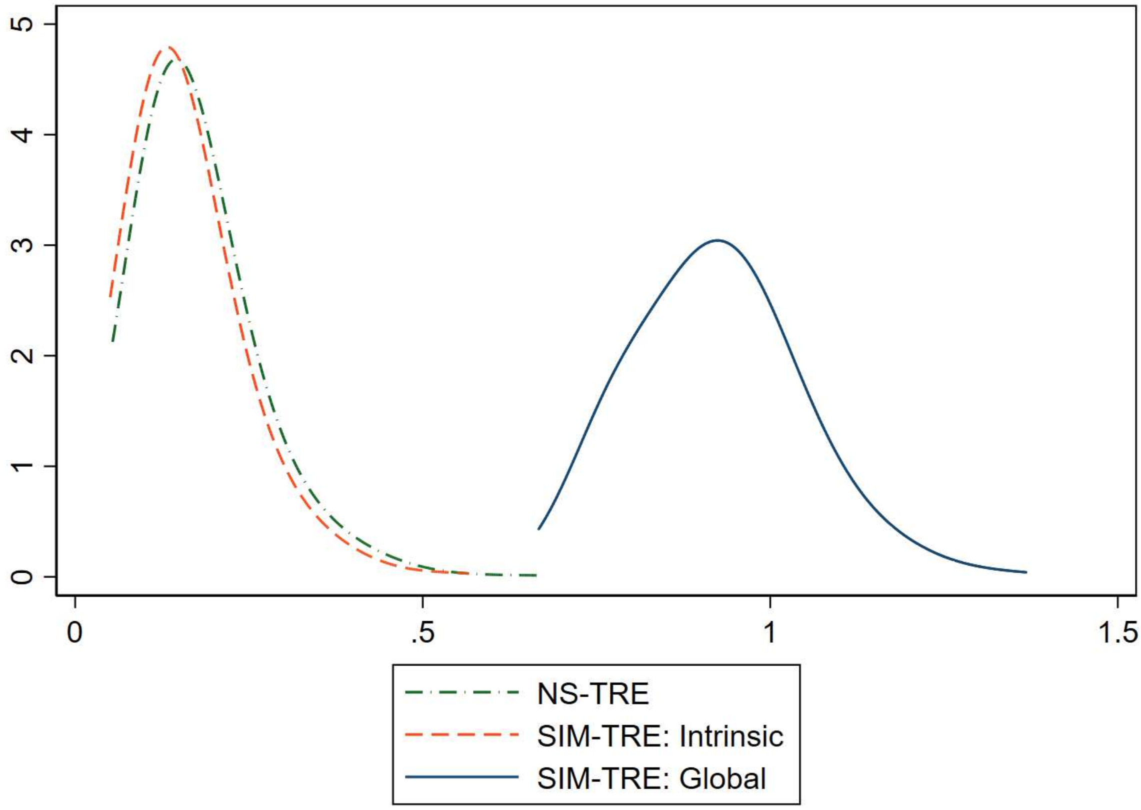


Table 7: Spearman rank correlations for inefficiencies.

	PSF	NS-TRE	NS-TFE	SIM-TRE*	SIM-TFE*
PSF	1.000	0.934	0.738	0.625	0.413
NS-TRE	0.934	1.000	0.913	0.686	0.519
NS-TFE	0.738	0.913	1.000	0.662	0.576
SIM-TRE*	0.625	0.686	0.662	1.000	0.948
SIM-TFE*	0.413	0.519	0.576	0.948	1.000

Figure 2: Dynamic of direct, indirect and global technical inefficiency for a selected farm (SIM-TRE).

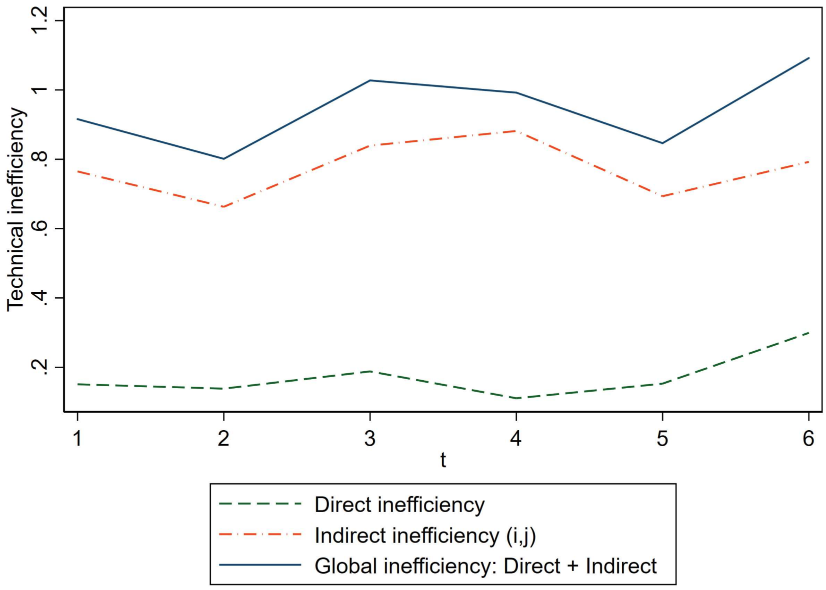
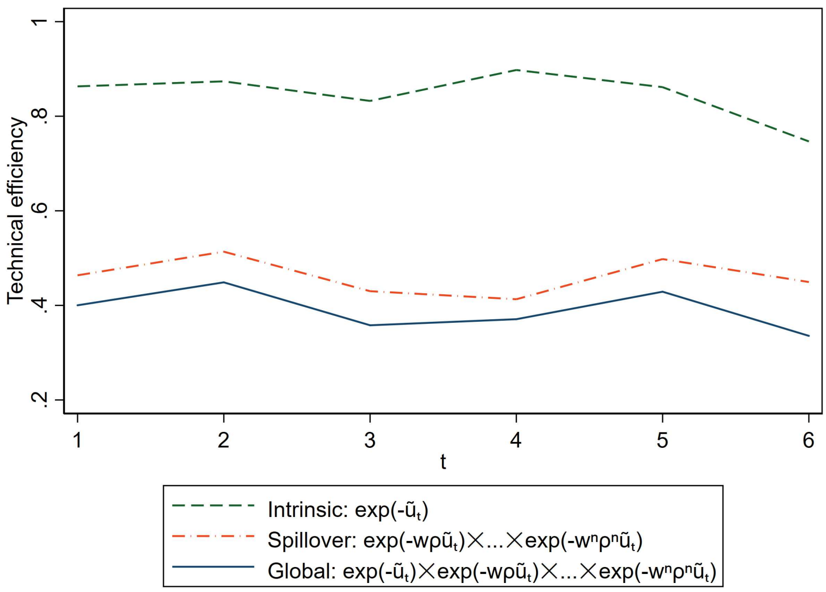


Figure 3: Dynamic of intrinsic, spillover and global technical efficiency for a selected farm (SIM-TRE).



RECENT PUBLICATIONS BY *CEIS Tor Vergata*

Hurricanes, Climate Change Policies and Electoral Accountability
Stefano Gagliarducci, M. Daniele Paserman and Eleonora Patacchini
CEIS Research Paper, 458, May 2019

Demography and Productivity in the Italian Manufacturing Industry: Yesterday and Today
Carlo Ciccarelli, Matteo Gomellini and Paolo Sestito
CEIS Research Paper, 457, May 2019

Real Effective Exchange Rates Determinants and Growth: Lessons from Italian Regions
Silvia Calò and Mariarosaria Comunale
CEIS Research Paper, 456, March 2019

Predictability, Real Time Estimation, and the Formulation of Unobserved Components Models
Tommaso Proietti
CEIS Research Paper, 455, March 2019

Sovereign Spread Volatility and Banking Sector
Vivek Sharma and Edgar Silgado-Gómez
CEIS Research Paper, 454, March 2019

The Societal Benefits of Money and Interest Bearing Debt
Luis Araujo and Leo Ferraris
CEIS Research Paper, 453, February 2019

A Horse Race in High Dimensional Space
Paolo Andreini and Donato Ceci
CEIS Research Paper, 452, February 2019

Robust Tests for Convergence Clubs
Luisa Corrado, Thanasis Stengos, Melvyn Weeks and M. Ege Yazgan
CEIS Research Paper, 451, February 2019

Forecasting Volatility with Time-Varying Leverage and Volatility of Volatility Effects
Leopoldo Catania and Tommaso Proietti
CEIS Research Paper, 450, February 2019

The Financial Decisions of Immigrant and Native Households: Evidence from Italy
Graziella Bertocchi, Marianna Brunetti and Anzelika Zaiceva
CEIS Research Paper, 449, January 2019