

Alma Mater Studiorum Università di Bologna  
Archivio istituzionale della ricerca

A Simultaneous System of Dynamic Spatial Stochastic Frontier Models

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

*Published Version:*

Emili, S., Galli, F. (2025). A Simultaneous System of Dynamic Spatial Stochastic Frontier Models. Springer [10.1007/978-3-031-64350-7\_31].

*Availability:*

This version is available at: <https://hdl.handle.net/11585/1006211> since: 2025-03-04

*Published:*

DOI: [http://doi.org/10.1007/978-3-031-64350-7\\_31](http://doi.org/10.1007/978-3-031-64350-7_31)

*Terms of use:*

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).  
When citing, please refer to the published version.

(Article begins on next page)

# A simultaneous system of dynamic spatial stochastic frontier models

Silvia Emili and Federica Galli \*

University of Bologna, Bologna, Italy

\*: corresponding author

`federica.galli14@unibo.it`

**Abstract.** In this paper, we propose a simultaneous system of stochastic frontiers for dynamic panel data models with spatial dependence. Specifically, the model allows differentiating between productivity spillovers within a single production frontier and among different production processes. This model is particularly relevant in empirical studies considering multi-output and multi-input economic systems characterized by interrelated input-output linkages, spatial correlation, and temporal dynamics. As an example, an application to European OECD data for the agricultural and food and beverage industrial sectors is provided.

**Keywords:** Simultaneous equations; Stochastic frontiers; Multiple outputs; Spatial dynamic panel models.

## 1 Introduction

Stochastic frontier models (SF) were initially proposed by [1] and [12] in order to create a tool for efficiency analysis capable of distinguishing between two different error terms: the inefficiency component and random shocks. This paper extends the existing literature on SF models by introducing a multivariate dynamic spatial stochastic frontier model that incorporates simultaneous effects across different frontier functions, as well as time and spatial effects. The proposed model is particularly suited for empirical analyses of production processes characterized by (i) spatio-temporal dynamics and (ii) multi-output and multi-input factors as well as input-output linkages.

Economic systems are characterized by intensive relationships and dependencies, especially due to multisectoral and multiproduct production processes. Multivariate SF models have been developed to jointly model such complex production systems, taking into account distinct frontier functions and simultaneities in production processes [9, 4, 10]. However, to our knowledge, none of these models considers the possibility of accounting for spatial dependence within and between different production frontiers. Indeed, based on the economic geography theory and on empirical evidence on spatial phenomena like agglomeration and clustering [13, 2], in the last two decades, authors have strongly recognized the inadequacy of assuming cross-sectional independence in non-spatial SF models, leading to a growing body of literature addressing spatial dependence in the SF

framework [7, 8, 6, 5, 14]. However, there is a notable absence of works incorporating spatial correlation in a multivariate SF setting.

Therefore, this paper bridges the gap by merging techniques from spatial stochastic frontier literature with those from the multivariate framework, creating a simultaneous system of stochastic frontiers for dynamic panel data models with spatial dependence.

In the subsequent sections, we introduce the model specification and estimation in Section 2, we show an application to European data in Section 3, and we conclude and provide some future directions of research in Section 4.

## 2 Methodology

### 2.1 The model

The model specification is presented in Eq. (1)-(3) for  $N_m$  units observed over  $T$  time periods, where  $m$  denotes the number of equations identifying distinct interdependent frontier functions and  $N_m$  represents the number of units characterizing each production process. Throughout the remainder of this section, without loss of generality, we assume  $m = 2$  and a balanced panel structure with a constant number of units and temporal instants across equations.

$$Bz_t = \Psi W z_t + \Phi W z_{t-1} + P z_{t-1} + X_t \Pi + V_t - c U_t \quad (1)$$

$$V_t \sim N_m(0, \Sigma_v) \quad (2)$$

$$U_t \sim N_m^+(0, \Sigma_u) \quad (3)$$

In particular,  $z_t$  is a  $(Nm \times 1)$  vector of outcome variables representing production outputs and  $X_t = (I_m \otimes x_t)$  where  $x_t$  is a  $(N \times k)$  matrix of production inputs with associated  $(mk \times 1)$  vector of coefficients  $\Pi$ .  $\mathbf{B} = (B \otimes I_N)$  is an  $Nm \times Nm$  matrix capturing mutual impacts among the  $m$  output variables by the off-diagonal entries. For  $m = 2$ , this implies  $vec(B) = (1, b_2, b_1, 1)$ .

Following the literature on simultaneous equations for spatial dynamic panel data models (e.g. [15, 3]) we assume  $\Psi = \Psi \otimes I_N$  as an  $(Nm \times Nm)$  matrix capturing the impact of spatially lagged dependent variables at time  $t$  between and within frontiers, where  $\Psi$  is an  $(m \times m)$  matrix with  $vec(\Psi) = (\psi_{1,1}, \psi_{2,1}, \psi_{1,2}, \psi_{2,2})$  for  $m = 2$ . Similarly, the matrices  $\mathbf{P} = P \otimes I_N$  and  $\Phi = \Phi \otimes I_N$  are included to account for possible dynamic time and space-time effects respectively.

The vector  $U_t$  refers to the  $(Nm \times 1)$  i.i.d. half-normally distributed inefficiency error component, while  $V_t$  is a  $(Nm \times 1)$  vector of random error terms. Inefficiency is defined based on the internal transformation of resources into output, and its form depends on the type of frontier ( $c = 1$  for a production and  $c = -1$  for a cost frontier). The covariance matrices  $\Sigma_u$  and  $\Sigma_v$  capture the variances and covariances of these error components. The model allows for each equation to have its own inefficiency-generating process, considering inefficiency

as context-specific and thus,  $\Sigma_u = \Sigma_u \otimes I_N$ , where  $diag(\Sigma_u) = (\sigma_{u,1}, \sigma_{u,2})$ , for  $m = 2$  and zero off-diagonal entries. On the other hand, we assume that a random shock to the first production process of the system may also influence the second one due to productive interdependencies and thus, we define  $\Sigma_v = \Sigma_v \otimes I_N$ , where  $vech(\Sigma_v) = (\sigma_{v,1}, \sigma_{v,12}, \sigma_{v,2})$ , for  $m = 2$ . Finally, according to usual assumptions in SF literature, the random shocks  $U_t$  and  $V_t$  are assumed to be independently distributed.

## 2.2 Estimation

The loglikelihood function at time  $t$  associated with the model in Eq. (1)-(3) can be written as

$$\begin{aligned} \frac{1}{nT} \log L_t = & \frac{1}{n} \log |S| + m \log(2) - \frac{m}{2} \log(2\pi) - \frac{1}{2} \log |\Sigma_u| - \frac{1}{2} \log |\Sigma_v| \\ & + \frac{1}{2} \log |\Sigma_\mu| - \frac{1}{2nT} (\varepsilon_t' \Sigma_v^{-1} \varepsilon_t - \mu_t' \Sigma_\mu^{-1} \mu_t) + \frac{m}{nT} \log \Phi(\mu_t C^{-1}), \end{aligned} \quad (4)$$

where  $\mu_t = -c\varepsilon_t A^{-1}$ , with  $A = (\Sigma_v + \Sigma_u) \Sigma_u^{-1}$ ,  $\Sigma_\mu^{-1} = \Sigma_v^{-1} (\Sigma_v + \Sigma_u) \Sigma_u^{-1}$ , and  $\varepsilon_t = Bz_t - \Psi W z_t - \Phi W z_{t-1} - Pz_{t-1} - X_t \Pi$ . Moreover,  $C$  is obtained by Choleski decomposition of  $\Sigma_\mu$ , such that  $\Sigma_\mu = CC'$ ,  $\Phi$  represents the cumulative distribution function of the multivariate normal, and  $|S|$  is the determinant of the Jacobian deriving from the transformation from  $\varepsilon_t$  to  $z_t$  given by  $|B' \otimes I_N - (\Psi \otimes I_N)W|$ .

## 3 Application to European Data

In this section, we show, as an example, a brief empirical application to aggregate production in Europe by means of OECD data for 20 European countries in the time period 1995-2019. The system is given by two production frontiers, the first modelling the production process of the agricultural sector and the second referring to the food and beverage manufacturing sector. Specifically, we estimate the model in Eq.(1)-(3) using a Cobb-Douglas specification with value added as the outcome variable and fixed assets and annual employment as input variables (all in log-form). The spatial structure is defined as a combination of a first and a second order spatial weight matrix, the first capturing inter-sectoral spillovers (off-diagonal blocks of  $W$ ) and the second for intra-sectoral spatial effects (diagonal blocks of  $W$ ), ensuring identification of the model (see [3]). Indeed, as discussed by [11], spatial effects within a given sector are more likely to materialize and propagate across space when compared to inter-sectoral spillovers due to a shared technological and knowledge base. Therefore, we consider a larger number of neighbouring countries for intra-sectoral spatial effects by means of a second order contiguity matrix.

The estimation results presented in Table 1 suggest that the outcome of the

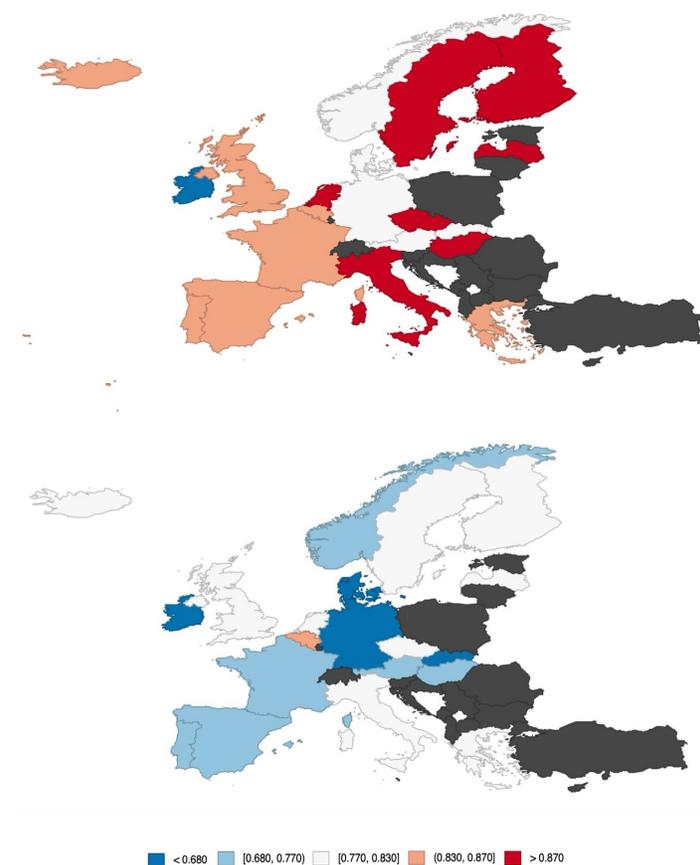
agricultural sector is positively associated with the one of the food and beverage industry, while indicating a negative opposite relationship. Accounting for spatial effects, we observe positive and statistically significant coefficients at time  $t$  both within and across sectors. Additionally, temporal effects are statistically significant, with spatial effects at time  $t - 1$  showing negative values and lagged outputs positive values. Moreover,  $diag(\hat{\Sigma}_u) = (0.03, 0.07)$  and  $vech(\hat{\Sigma}_v) = (0.20, 0.41, 0.85)$ . Overall, these estimates corroborate the idea that industrial sectors composing the agricultural fili re are highly linked and shaped by intense simultaneous relationships driven by cross-sectoral effects, spatial correlation and time dynamics. However, ad hoc evaluation of long and short run impacts should be carried out through the computation of marginal effects (as further development of this study).

**Table 1.** Estimation results

	Agriculture		Food & Beverage	
	Coeff.	SD	Coeff.	SD
$z_{At}$	-	-	-0.29 ***	0.02
$z_{Mt}$	0.42 ***	0.01	-	-
$Wz_{At}$	0.12 ***	0.01	0.27 **	0.03
$Wz_{Mt}$	0.21 ***	0.01	0.47 ***	0.01
$Wz_{At-1}$	-0.13 **	0.01	-0.38 ***	0.03
$Wz_{Mt-1}$	-0.26 ***	0.01	-0.45 ***	0.02
$z_{At-1}$	0.84 ***	0.01	0.13	0.04
$z_{Mt-1}$	-0.40 ***	0.01	0.87 ***	0.04
$L_{At}$	-0.02	0.01	-0.05	0.01
$L_{Mt}$	0.07	0.02	0.23	0.04
$K_{At}$	0.02	0.01	0.03	0.01
$K_{Mt}$	0.20 ***	0.01	0.34 ***	0.02

Note: \*:p-value $\leq$ 0.10,\*\*:p-value $\leq$ 0.05;\*\*\*:p-value $\leq$ 0.01

Based on the estimates shown in Table 1, efficiency scores equalling 1 for fully efficient countries and 0 for fully inefficient ones can be computed. Figure 1 displays the map of mean efficiency scores throughout the entire period (agricultural sector in the upper panel and the food and beverage manufacturing industry in the lower panel). Overall, technical efficiency scores related to the agricultural sector tend to be higher than those of the manufacturing industry. However, countries demonstrating high efficiency levels in one sector also tend to exhibit satisfactory efficiency scores in the other. Analyzing cross-country variations, the Figure indicates that Italy, Sweden, Finland, the Netherlands, Latvia, the Czech Republic, and the UK consistently outperform other countries. In contrast, Ireland, Germany, Austria, Slovakia, Denmark, and Norway achieve lower efficiency levels.



**Fig. 1.** Map of the mean efficiency scores by sector  
 Notes: Black areas indicate countries not in the estimation sample

## 4 Conclusions and Further Developments

This paper develops a simultaneous system of stochastic frontiers for dynamic panel data models with spatial dependence. The applied relevance of the proposed model is demonstrated through an empirical application to aggregate production in Europe for agriculture and the manufacturing sector using OECD data. The estimation results indicate the significant influence of cross-sectoral connections and spatio-temporal effects at the European level between the agricultural sector and the food and beverage manufacturing industry. Consequently, policymakers should consider these spatial and cross-sectoral feedbacks when formulating international development plans and economic programs to effectively support the growth of the whole agri-food economy.

Future developments of this paper should consider computing short-run and long-run marginal effects of input factors as well as of an innovation in a given

sector, and the inclusion of a partially time-varying spatial structure, allowing each production outcome to be influenced by a specific combination of spatial weight matrices that can vary over time. Moreover, robustness checks for alternative spatial weight matrices and controlling for country-specific characteristics should be provided. Finally, alternative estimation methods using copula-based techniques may be introduced and tested.

## References

1. Aigner, D., Lovell, C.K., Schmidt, P. (1977). *Formulation and estimation of stochastic frontier production function models*. *Journal of Econometrics*, 6(1), 21-37.
2. Delgado, M., Porter, M.E., Stern, S. (2014). *Clusters, convergence, and economic performance*. *Research Policy*, 43, 1785–1799.
3. Elhorst, J. P., Emili, S. (2022). *A spatial econometric multivariate model of Okun's law*. *Regional Science and Urban Economics*, 93, 103756.
4. Fernández, C., Koop, G., Steel, M. (2000). *A Bayesian analysis of multiple-output production frontiers*. *Journal of Econometrics*, 98, 47-79.
5. Galli, F. (2023b). *A spatial stochastic frontier model introducing inefficiency spillovers*. *Journal of the Royal Statistical Society: Series C*, 72(2), 346-367.
6. Galli, F. (2023a). *A spatial stochastic frontier model including both frontier and error based spatial cross-sectional dependence*. *Spatial Economic Analysis*, 18(2), 239–258.
7. Glass, A.J., Kenjegalieva, K., Sickles, R.C. (2016). *A spatial autoregressive stochastic frontier model for panel data with asymmetric efficiency spillovers*. *Journal of Econometrics*, 190(2), 289-300.
8. 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.
9. Kumbhakar, S.C., Lai, H. (2021). *A multi-output multi-input stochastic frontier system with input- and output-specific inefficiency*. *Economics Letters*, 201, 109807.
10. Liu, J., Wang, M., Ma, J., Rahman, S., Sriboonchitta, S. (2020). *A simultaneous stochastic frontier model with dependent error components and dependent composite errors: An application to Chinese banking industry*. *Mathematics*, 8(238), 1-23.
11. Malerba, F., Mancusi, M.L., Montobbio, F. (2014). *Innovation, international R&D spillovers and the sectoral heterogeneity of knowledge flows*. *Review of World Economics*, 149(4), 697-722.
12. Meeusen, W., van Den Broeck, J. (1977). *Efficiency estimation from Cobb-Douglas production functions with composed error*. *International Economic Review*, 18(2), 435-444.
13. Porter, M. (1998). *Clusters and the New Economics of Competition*. *Harvard Business Review*, 77-90.
14. 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.
15. Yang, K., Lee, L.-F. (2019). *Identification and estimation of spatial dynamic panel simultaneous equations models*. *Regional Science and Urban Economics*, 76, 32-46.