Spatial nonstationarity in the stochastic frontier model: an application to the Italian wine industry

F. Vidoli | University of Roma Tre, Dept. of Political Science | Rome | Italy
C. Cardillo | CREA, Consiglio per la ricerca in agricoltura e l’analisi dell’economia agraria | Rome | Italy
E. Fusco | University of Rome, La Sapienza, Dept. of Statistical Sciences | Rome | Italy
J. Canello | University of Queensland, School of Economics | Brisbane | Australia
DOI: 10.1481/icasVII.2016.b08b

ABSTRACT
This research estimates the efficiency of a representative sample of Italian wine producers from the Italian FADN survey following a recent spatial stochastic frontier framework that allows to isolate the spatial dependence among units and to evaluate the role of intangible local factors in influencing the economic performance of firms. The empirical exercise shows that the specific territorial patterns in the data cannot be merely explained using a standard set of contextual factors. This intangible component can be interpreted as the role of the local business climate: in most localities, the presence of an embedded community stimulates a process of local learning that generates the diffusion of tacit knowledge through continuous interaction among the local actors. This effect is found to be different across firm size, with a larger impact on small firms.

Keywords: Spatial nonstationarity, Spatial Stochastic Frontier Model, Wine, Efficiency

1. Introduction
The globalization of productive processes and liberalization of trade activities have generated a strong competition between regional economic systems: paradoxically, rather than drastically reducing the role of spatial proximity, this new open scenario has shed new light on the key relevance of local and agglomeration externalities in the generation of competitive advantage (Porter, 2000). The relevance of these aspects is particularly evident in certain sectors, such as wine production, where the rapid transformations occurred in the last few decades have fostered a rapid process of technological change in which firms are constantly required to be at the forefront of the productive process in order to survive in the competitive arena (Cusmano et al., 2010). In this context, the role of intangible factors associated with the ‘business climate’ is crucial in stimulating the process of knowledge accumulation and learning through continuous interaction with peers located in close proximity: in several circumstances, the presence of these mechanisms ensures the diffusion of new productive practices that prevents local firms from increasing their gap with the technological frontier.

The classical stream of the literature linking productive efficiency to the territorial determinants assumes that the dynamic process leading firms to concentrate in specific subregions is only associated to specific tangible aspects: this assumption leads to neglect the role of spatial non-stationarity, intended as “a condition in which a simple global model cannot explain the relationships between some sets of variables” (Brunsdon et al., 1996). This problem is particularly evident in the parametric frontier framework, where it is essential to specify a priori an explicit functional form of the boundary of the production set: however, in the early contributions the spatial dependence among productive units has often been ignored and associated to the stochastic error. A number of recent works have attempted to address this issue by specifically including a set of contextual factors in the model (see e.g. Hughes et al., 2011, Brehm, 2013): however, such a strategy is not always effective as it ignores the fact that the relationship between the dependent variable and the covariates (a) tends to vary in a continuous rather than a discrete manner among spatial units and (b) may not be necessarily related to measurable local factors. This problem is particularly evident in specific spatial contexts, such as industrial districts, characterized by the presence of global intangible factors that cannot be measured empirically (Vidoli and Canello, 2016).

Ignoring spatial autocorrelation among residuals limits the validity of the empirical investigation for several reasons. First, it causes serious consequences to statistical inference, reducing both the efficiency and consistency of the estimations and generating a negative impact on the validity of testing procedures and on the predicting capability of the model. This drawback generates significant distortions in the interpretation of the stochastic frontier model, as higher values of the inefficiency term
may be associated with a territorial effect rather than the ability of a productive unit to generate more output with the same amount of inputs. In this respect, the inclusion of spatial autocorrelation into the stochastic frontier production framework has been the subject of a lively debate in the econometrics literature in the recent past, generating a multitude of approaches aimed to address this issue: in this context, the specification proposed by Fusco and Vidoli (2013) appears to be particularly suitable in that the spatial autoregressive specification is modelled in the error term, generating results that can be directly compared with those of the classic stochastic frontier approach.

In this paper, the above mentioned spatial stochastic frontier approach is implemented on a sample of Italian firms specialized in wine production, using data extracted from the 2013 FADN Survey. This archive is particularly suitable for the scope of the analysis, as it allows to account for a wide variety of structural and economic factors that are believed to influence the territorial effects: moreover, the presence of specific reference to a wide variety of inputs allows to build a solid production function with several benefits for the estimation process. The aim of the empirical exercise proposed in this paper is to evaluate the contribution of both tangible and intangible factors in influencing the performance of these firms, discussing how the space can play a different role for the different members of a local network. In this respect, the specification proposed is of particular use as it allows to isolate the local intangible factors, often statistically and economically difficult to capture through specific proxies, that nonetheless are determinant in influencing the firms productivity. The role of tangible factors is nonetheless evaluated through a second stage estimation.

The remainder of the paper is structured as follows: section 2 presents the main features of the Italian wine sector, highlighting how the recent technological and structural changes have increased the importance of local learning and continuous access to new knowledge. In section 3, a brief review of the parametric frontier models is introduced, stressing on the recent debate on spatial approaches and presenting the proposed method and the benefits deriving for its application in the present context. In section 4, the focus is moved to the empirical application: after the main features of the database used for the analysis are introduced and discussed, the results of the spatial stochastic frontier model are introduced, comparing them with those of the traditional approach and highlighting the effectiveness of the Spatial Stochastic Frontier (SSFA) specification (see section 3) in isolating the spatial dependence that are present in the data. Finally, the structural features of the spatial effect are inspected and variables is not sufficient to explain the variability of local effect: this result provides evidence of the existence of an intangible local effect in the wine industry that influence firm performances in different ways. Section 5 summarizes the main finding of the paper and proposes some concluding remarks and possible directions for future research.

2. The recent trends in the Italian wine industry and the role of agglomeration effects

Italy has played, together with France, a key role in the wine industry for several decades, dominating the international scenario in terms of both exported volumes and values. This established pattern has radically changed since the 1990s, when the entrance of the New World producers (United States, Australia, Chile, Argentina, South Africa) in the global market has fostered a radical transformation of the existing competitive arena (Cusmano et al., 2010). The increased complexity of the new global environment has further been influenced by several exogenous risk factors, such as the increased climate variability and the radical changes in wine consumption habits, with a shift in preferences towards high quality wines (Bardaji and Iraizoz, 2015). In this context, the sector has experienced a process of rapid modernization and technological change, identified by Crowley, 2000 as a “wine revolution”. This radical transformation process is pushing wine producers to adopt improvement strategies in the quality and production process and acquire new knowledge in order to effectively respond to the volatile needs of the global markets. The potential gains from selecting an effective strategy are especially important in the wine sector where, despite the existence of a moderate correlation between price and quality, several price setting possibilities are available for wine producers given the incomplete quality information held by consumers (Oczkowski and Doucouliagos, 2015).

The current scenario generates several challenges for the Italian wine sector, which has recently faced a significant downturn in domestic demand and is characterized by a higher degree of fragmentation relative to other countries, such as Australia or Chile (Cusmano et al., 2010). Italian wine producers are often small and medium businesses that lack the financial and managerial resources to handle the increased complexity of the surrounding environment. This limit is especially problematic in the new global context, characterized by the constant need to update productive knowledge and acquire new skills and competences. In fact, small businesses are not generally equipped to gather relevant information outside the locality in which they are embedded; moreover, they cannot rely on the same formal channels used by leader firms, such as formal collaborations with research institutions (Giuliani et al., 2010) and interaction with foreign competitors, often through the presence of foreign subsidiaries (Felzensstein and Deans, 2013) or simply through the creation of relational networks with producers that are at the forefront of the industry (Turner, 2010). Given these opportunities are not generally accessible to small producers, the main source of learning and developing new competencies is the community in
which these entrepreneurs are embedded. The role of this factor is especially relevant in Italy, where contrary to other countries, such as Australia, the wine sector cannot count on institutional assets and top-down measures to stimulate the above mentioned upgrading process.

Although several contributions have stressed the increasing importance of codified knowledge in the wine industry (e.g. Giuliani, 2007), the sector is still dependent on context-specific and localized informal practices of learning, that are crucial to take advantage of the specificities of each terroir (Turner, 2010). The effect generated by the local business climate and the informal interaction among local actors can be explained through the concept of "industrial atmosphere", that has generally been used in the industrial districts literature (Marshall, 1920); other than generating tangible benefits such as reputation, greater international demand and access to skilled labour pool, spatial proximity among wine producers stimulates everyday interaction, facilitating the opportunities for face-to-face contacts that are crucial to generate tacit knowledge flows and incremental learning. In this respect, wine clusters can be seen as communitarian networks, characterized by resource sharing and continuous informal interaction (Turner, 2010). The presence of interpersonal networks can be beneficial in many respects: producers can be rapidly informed of the presence of new business opportunities, but also of new sellers or providers that can form new partnerships and generate further spillovers. Inter-firms market cooperation can also foster marketing collaboration strategies, facilitating development of joint sales in foreign markets and allowing to overcome the limited exporting capabilities of several small and medium firms (Felzensztein and Deans, 2013). More importantly, the presence of a collaborative environment can allow small producers to fill the technological gap with competitors, as collaboration can foster the shared use of new technology, exchange of technical advices and information on the effective use of machinery and inputs (Morrison and Rabellotti, 2009).

A certain number of agglomeration externalities generates spontaneously as a consequence of spatial proximity between wine producers. For example, the successful performance of neighbouring wineries stimulates the development of positive marketing-related externalities for the whole area (Giuliani and Bell, 2005): these positive spillovers in terms of reputation for the neighbouring producers have been classified by Beebe et al. (2013) as "halo" effect. However, spatial proximity itself is not sufficient to guarantee the diffusion of agglomeration externalities among all the members of a local community. Indeed, two elements are required to enhance this process, i.e. the willingness of givers to share their knowledge and the absorptive capacity of the receiver: these conditions are generally meet when the cognitive proximity among the members of a network is present (Boschmaa, 2005). The presence of diversified abilities/attitudes to access to local informal knowledge has been documented in different regional contexts in the wine industry (Giuliani and Bell, 2005; Morrison and Rabellotti, 2009); according to Giuliani and Bell (2005), the presence of barriers to knowledge exchange is testified by the presence of different production methods within the same wine cluster.

What is the profile of those firms which are more often engaged in networking activities? The core of these local networks is generally represented by small firms, which are generally more inclined to cooperate and share information in order to overcome their structural limits: the lack of competencies among small firms act as stimulus to share different experiences and spread knowledge among the community. On the other hand, large firms tend to be located at the periphery of the local network and provide a limited contribution to the local learning system: these actors generally have stronger connections with external sources of knowledge and often prefer to share the acquired competencies with a restricted number of partners that are directly involved in their production process (Morrison and Rabellotti, 2009). This trend is confirmed by the empirical investigation of Turner (2010), who has shown that small wine producers are more interested in marketing practices associated with the territory while large firms are more interested in developing their own brand.

The brief review presented in this section has shown that the tangible and intangible local effects play a key role in determining the performance of wine producers. Against this background, the aim of the following section is to propose an empirical framework that can be effectively used to account for both effects in the estimation of productive efficiency, allowing evaluate the role of the different spatial factors in a consistent manner.

3 Disentangling the role of spatial effects on firm efficiency: the spatial stochastic frontier models

Efficiency estimation has been subject to considerable research during the last decades, generating important contributions both in the econometrics (see e.g. Aigner et al., 1977; Meesussen and van den Broeck, 1977; Aigner and Chu, 1968) and operational research literature (see e.g. Farrell, 1957; Charnes et al., 1978; Deprins et al., 1984; Grosskopf, 1996; Daraio and Simar, 2007). In this context, one of the most widely used parametric models is the Stochastic Frontier Analysis (SFA), which was originally proposed in two different contexts by Aigner et al. (1977) and Meesussen and van den Broeck (1977): under specific assumptions, this specification allows to estimate the parameters of the frontier production function and to perform hypothesis testing procedures to validate the model (see e.g. Kumbhakar and Lovell, 2000 for a detailed introduction to frontier analysis).
In a cross-sectional framework, the productive process of the \( p \)th firm using \( N \) inputs \( x = (x_1, \ldots, x_N) \) to produce \( M \) outputs \( y = (y_1, \ldots, y_M) \) can be specified as follows:

\[
\log(y) = \log(f(x, \beta)) + v - u
\]

(1)

where:

- \( v \sim \text{iid } N(0, \sigma_v^2 I) \) is the random term;
- \( u \sim \text{iid } N^+(0, \sigma_u^2 I) \) is the inefficiency term;
- \( v \) and \( u \) are assumed to be independently and identically distributed.

The traditional SFA specification estimates firm-level efficiency from the residuals, assuming that all producers in the sample are independent; however, this assumption rules out the possibility to account for spatial effects in the theoretical model. The limitations of this approach were already known in early contributions, considering Farrell (1957) highlighted the importance of incorporating the correlation between technical efficiency and variables representing location, temperature and rainfall. In his analysis, focused on the efficiency patterns in US agricultural firms, he argued that "the apparent differences in efficiency [...] reflect factors like climate, location and fertility that have not been included in the analysis, as well as genuine differences in efficiency".

When spatial effects are significant, the traditional techniques used to estimate the SFA (MLE or its variants) generate biased results: indeed, if the disturbances are spatially correlated, the assumption of a spherical error covariance matrix is violated, leading to biased and inconsistent estimators (LeSage, 1997). In order to overcome these issues, a number of recent contributions in frontier analysis have proposed alternative specifications aiming to incorporate spatial effects in the baseline models. This literature stream follows the approach used by spatial econometrics, a specific branch of econometrics that deals with spatial interaction (spatial autocorrelation) and spatial structure (spatial heterogeneity) in both cross-sectional and panel data (Paelinck and Klaassen, 1979; Anselin, 1988).

Subsequently, given the need to consider for spatial dependence also in frontier analysis, some models have been developed; they can be divided in two major fields distinguishing those that explain inefficiency/efficiency in terms of exogenous determinants analysing the heterogeneity from those that consider the spatial dependence by including in the model a spatial autoregressive specification.

As far as the first stream is concerned, some authors have proposed to analyse heterogeneity by including contextual factors as regressors or to modelling the inefficiency term. In particular, Lavado and Barrios (2010) used contextual factors to modelling the inefficiency part of a stochastic frontier model embedding a sparse spatial autoregression (SAR) in the deterministic part and a general linear mixed model into the efficiency equation; Hughes et al. (2011) considered specific spatial effects in the stochastic production frontier by adding climate effects as dependent variables; Jeleskovic and Schwanebeck (2012) proposed a two step deterministic estimation model to differentiate heterogeneity and inefficiency in world healthcare systems: (i) in the first step different fixed effects panel spatial models have been estimated; (ii) in the second step the obtained inefficiency has been regressed (also with various fixed effects panel spatial models) as dependent variable onto country specific variables that identify the heterogeneity; finally, Brehm (2013) proposed a correction of the SFA error term for panel data by introducing spatially correlated factors variables that affect the production process.

In the second set of analysis, others proposals consider spatial dependence by including a spatial lag into the dependent variable or into the covariates. More specifically, Affuso (2010) included spatial lag on the dependent variable reformulating the stochastic frontier density function; Glass et al. (2013), Glass et al. (2014) and Glass et al. (2016) introduced the concept of efficiency spillover, extending the non-spatial Cornwell et al. (1990) model to the case of spatial autoregressive dependence; Adetutu et al. (2015) proposed a local spatial stochastic frontier model that accounts for spatial interaction by allowing spatial lags on the inputs and on the exogenous variables to shift the production frontier technology; Han et al. (2013) proposed a method for investigating spillovers effects in panel data by maintaining the Schmidt and Sickles (1984) hypothesis of time-invariant inefficiency, but allowing global spatial dependence through the introduction of a spatial lag on the dependent variable.

Finally, others papers proposed to consider spatial dependence by including a spatial lag on the inefficiency term. Druska and Horrace (2004) extended the Kelejian and Prucha (1999) specification for cross-sectional data based on a standard fixed effects model by assuming an autoregressive specification of the error term and estimating inefficiency with the Generalized Moments Method; Schmidt et al. (2009) used a Bayesian approach to include latent spatial effects, that explain geographical variation of firms’ outputs and inefficiency, dependent on a parameter that captures the unobserved spatial characteristics; e.g. Areal et al. (2010) suggested, with the aim of measuring the overall effect of spatial factors that affect the production, to include a spatial lag directly into inefficiency allowing the splitting
of the inefficiency into a spatial component and into a specific term for every firm through a Bayesian procedure. Instead, Pavlyuk (2010) proposed to include spatial lags on the overall standard SFA model (see also recent enhancements, Pavlyuk, 2012, 2013). Following the approach implemented by Areal et al. (2010), Fusco and Vidoli (2013) have proposed to measure the global effect of spatial factors by including a spatial lag only in the inefficiency term of a stochastic frontier (SFA), not using a Bayesian procedure but by reformulating the SFA density function with a spatial error autoregressive specification (SEM). The Spatial Stochastic Frontier model (SSFA) is defined as:

\[
\begin{align*}
\log(y) &= \log(f(x; \beta)) + v - u \\
&= \log(f(x; \beta)) + v - (I - \rho W)^{-1}\tilde{u}
\end{align*}
\] (2)

where:

- \( v \sim iid \ N(0, \sigma_v^2 I) \);
- \( u \sim N^+(0, \begin{bmatrix} (I - \rho W)^{-1} & -1 \end{bmatrix} \begin{bmatrix} \sigma_u^2 & \sigma_v^2 \\
\sigma_v^2 & \sigma_u^2 \end{bmatrix} ) \);
- \( v \) and \( u \) are independent of each other and of the regressors;
- \( \tilde{u} \sim iid \ N(0, \sigma_{u|\tilde{u}}^2 I) \).

In this paper, an empirical application of the SSFA has been proposed in order to avoid the subjective choice of the exogenous determinants and to focus the analysis of the spatial dependence only on the inefficiency term: this approach reduces the complexity of the model and makes the estimation easier computationally1 as well as to be immediately comparable with the classical SFA model when no spatial autocorrelation is assumed on the error term.

It is important to note, moreover, that the formulation of equation (2) hide an interesting property: the identification of the error part imputable to the spatial proximities different for each unit allows to change (positively or negatively) the intercept of the model; therefore, in the SSFA framework the intercept is to be understood only as the “medium level” cleansed by the individual spatial effect.

Figure 1 clarifies the issue: departing from high spatial correlated simulated data (one input vs one output), OLS, SFA and SSFA has been estimated; in particular it can be noted that the SSFA specification impact both on the specific estimation of each unit (red continue line) both on the average slope (orange dashed line).

Figure 1 - OLS, SFA, and SSFA fitted values, simulated data

\[ \text{Figure 1 - OLS, SFA, and SSFA fitted values, simulated data} \]

\[ \text{Thanks to the SSFA R package (Fusco and Vidoli, 2015b, Fusco and Vidoli, 2015a).} \]
4. Application

In this section, the spatial stochastic frontier specification introduced in section 3 is used to evaluate the role of local effects in the Italian wine industry; as highlighted in section 2, the recent technological advances in the sector has increased the importance of both the tangible and intangible factors associated with the specific territorial effects that cannot always be captured by the inclusion of contextual variables. In this respect, the spatial technique proposed in the previous section appears to be especially effective to account for these factors and evaluate their role in influencing firm-level efficiency. The empirical investigation is focused on year 2013 and is implemented on a detailed database that includes a wide variety of economic and structural variables: the main features of this database are presented in the following subsection.

4.1 Italian FADN survey: the data source used for the empirical analysis

The Farm Accountancy Data Network (FADN) is a yearly survey carried out by the Member States of the European Union and established in 1965 by the Council Regulation No 79/65/EEC: this measure was aimed to establish a network for the systematic collection of accountancy data on incomes and business operations of agricultural holdings in the European Economic Community. According to Regulation n. 1859/82, this database includes all the agricultural holdings having an economic size equal to or greater than a minimum threshold, i.e. that identified to be considered commercial: in Italy, this threshold is set on 4,000 e/standard output. The selection of the holdings that take part to the survey is carried out according to sampling plans defined at the national level, following the guidelines and recommendations provided by the European Commission. The sampling procedure must ensure the representativeness of the identified subset and defines the number of farms to be selected, specifying the approach followed to select the productive units. According to FADN methodology, stratification variables are territorial location, economic size and type of farming.

The Italian section of the survey is based on the Agricultural Census, updated on a two-year basis by the Farm Structure Survey (FSS) carried out by ISTAT: this main data source is complemented with further sources of agricultural statistics. The Italian FADN sample is selected using the stratified random sampling technique described above: in particular, the territorial location corresponds to the 21 administrative regions; the economic size is expressed in terms of Standard Output and defined through several classes, the lowest of which starts from 4,000 e/standard output. The selection of the farms that take part to the survey is carried out according to sampling plans defined at the national level, following the guidelines and recommendations provided by the European Commission. The sampling procedure must ensure the representativeness of the identified subset and defines the number of farms to be selected, specifying the approach followed to select the productive units. According to FADN methodology, stratification variables are territorial location, economic size and type of farming.

The methodology used to allocate the sample among the strata is a combination of Neyman and Bethel methods (Bethel, 1989). The main benefits associated with the use of this database can be summarized by the following two aspects:

- **harmonization**: FADN is the only source of micro-economic data that is harmonised at European level, i.e. the book-keeping principles are the same in all countries, and it represents an important tool for the evaluation of the income of agricultural holdings and the impacts of the Common Agricultural Policy. In Italy the FADN survey is carried out by the Center for Policy and Bio-economy of the Council for Research in Agriculture and the Agricultural Economics Analysis - CREA3, as liaison agency between EU and Member State.

- **information assets**: The FADN survey collects more than 1,000 variables that refer to physical and structural data, such as location, crop areas, livestock units, labour force. It also contain economic and financial data, such as the value of production of the different crops, stocks, sales and purchases, production costs, assets, liabilities, production quotas and subsidies, including those connected with the application of CAP measures and recently were added also information linked to environmental aspects. These variables are extremely convenient for the purpose of this paper, as they allow to create both a solid production function and include a wide variety of local effects that are associative with the performance of these firms.

---

2 The coefficient of variation of a variable is the ratio between the standard deviation of the variable layer and the estimate of the total layer of variable.

3 Previously National Institute of Agricultural Economics - INEA
In this paper a sub sample of 853 wineries has been extracted from the Italian FADN database, which includes a total number of 11,319 farms in year 2013. Using this data, the application presented in the following sections compares the results of the traditional specifications of the production function with those of the SSFA model, showing the benefits associated with the use of the latter approach. An important caveat relates to the variable used to evaluate output, i.e. the litres of wine produced by each unit: given the information at disposal does not allow to evaluate the qualitative aspects of production (which are nonetheless relevant in the sector), the concept of efficiency should be interpreted from a technical point of view, avoiding any considerations on the quality of output produced.

4.2 Production function

The production function of the Italian wine firms has been initially estimated using a simple OLS approach, choosing a Cobb-Douglas log-log functional form and relating the produced quantity of output with labour, machinery, water-energy-fuel and land capital inputs. The results of the estimation (Table 1) appear to confirm the validity of the specification, given the significance of all covariates and the high $R^2 = 0.593$; it is also worth noting that the intercept is negative and statistically significant.

Table 1 - Wine production function - OLS estimators

<table>
<thead>
<tr>
<th>Dependent variable: production output (log)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
</tr>
<tr>
<td>Labour input</td>
</tr>
<tr>
<td>Machinery capital input</td>
</tr>
<tr>
<td>Water, energy and fuel input</td>
</tr>
<tr>
<td>Land capital input</td>
</tr>
<tr>
<td>Observations</td>
</tr>
<tr>
<td>$R^2$</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
</tr>
</tbody>
</table>

|               | Estimate | Std. Error | $z$ value | $Pr(>|z|)$ |
|---------------|----------|------------|-----------|------------|
| Intercept     | -1.411159| 0.322292   | -4.3785   | 1.195e-05  *** |
| Labour input  | 0.637796 | 0.055230   | 11.5480   | < 2.2e-16  *** |
| Machinery capital input | 0.314146 | 0.044574   | 7.0477    | 1.819e-12  *** |
| Water, energy and fuel input | 0.071348 | 0.013375   | 5.3346    | 9.537e-08  *** |
| Land capital input | 0.233026 | 0.037678   | 6.1847    | 6.224e-10  *** |
| $\sigma^2$   | 1.003306 | 0.116481   | 8.6135    | < 2.2e-16  *** |
| $\gamma$     | 0.581710 | 0.097161   | 5.9871    | 2.136e-09  *** |

**Note:** $^*p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Using OLS as baseline for the analysis, the stochastic frontier model has been estimated (Table 2) and the results of the two specifications compared: the analysis confirms the stability of the latter model, since the values of the coefficients are similar in the two cases, except for the intercept that decreases in absolute value. This trend is expected and can be explained by the fact that the production function has been shifted from the average values to efficient ones without affecting the relationship between output and inputs. The specific parameters of SFA ($\sigma^2$, $\gamma$ and the average efficiency equal to 0.59) confirm the validity of the proposed model. In particular, $\gamma = \sigma^2 u / \sigma^2 v$ depends on two relevant parameters, $\sigma^2 u$ and $\sigma^2 v$, that are the variances of the noise and inefficiency effects. Note that $\gamma$ varies from 0 to 1; when the value is close to zero deviations from the frontier are attributed to noise, while in the opposite case the deviations are entirely explained by the technical inefficiency of the firm. As discussed in Section 3, the stochastic frontier model is based on the hypothesis of mutual independence among the productive units: therefore, this specification ignores the role of any spatial effects that may be present in the data. However, the evolutionary trends emerged from the brief overview of the Italian wine industry presented in this paper suggests that efficiency in this sector could be influenced by a multiplicity of tangible and intangible local factors.

Table 2 - Wine production function - SFA estimator

<table>
<thead>
<tr>
<th>Dependent variable: production output (log)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
</tr>
<tr>
<td>Labour input</td>
</tr>
<tr>
<td>Machinery capital input</td>
</tr>
<tr>
<td>Water, energy and fuel input</td>
</tr>
<tr>
<td>Land capital input</td>
</tr>
<tr>
<td>$\sigma^2$</td>
</tr>
<tr>
<td>$\gamma$</td>
</tr>
</tbody>
</table>

|               | Estimate | Std. Error | $z$ value | $Pr(>|z|)$ |
|---------------|----------|------------|-----------|------------|
| Intercept     | -1.411159| 0.322292   | -4.3785   | 1.195e-05  *** |
| Labour input  | 0.637796 | 0.055230   | 11.5480   | < 2.2e-16  *** |
| Machinery capital input | 0.314146 | 0.044574   | 7.0477    | 1.819e-12  *** |
| Water, energy and fuel input | 0.071348 | 0.013375   | 5.3346    | 9.537e-08  *** |
| Land capital input | 0.233026 | 0.037678   | 6.1847    | 6.224e-10  *** |
| $\sigma^2$   | 1.003306 | 0.116481   | 8.6135    | < 2.2e-16  *** |
| $\gamma$     | 0.581710 | 0.097161   | 5.9871    | 2.136e-09  *** |

**Note:** $^*p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

4 This model has been also estimated using a Translog specification for the production function. However, given the lack of significance of composite terms, a simpler model has been chosen for this part of the analysis.

5 The basic statistics of the variables used in the analysis and the relative units of measurement are given in Annex, Table 5.
To evaluate the role of these effects, a formal statistical test has been implemented to verify the presence of spatial correlation among residuals: specifically, the global and local indicators proposed by Geary (1954) have been used, previously specifying a distance matrix to map the neighbourhood of each production unit. The correct definition of the matrix is crucial to ensure the consistency of the spatial analysis: in this respect, the identification of a correct unit of distance must be driven by economic considerations associated with the peculiarities of the sector under investigation. In this specific case, characterized by the presence of productive units often concentrated in narrow geographical areas, a particularly close neighbourhood (nearest neighbour, n = 10) has been chosen in order to account for the specificities of the wine industry: the contiguity matrix resulting from the application of this criterion is graphically represented in Figure 2.

Using this distance matrix, the presence of spatial autocorrelation among residuals for the SFA model has been formally tested using the Geary C statistic\(^6\): the estimated value of this variable (0.733) leads to reject the null hypothesis of mutual independence among firms, confirming the presence of a positive neighbourhood effect among the Italian wineries that cannot be isolated and estimated through the traditional stochastic frontier model. This scenario motivates the need to use a spatial stochastic frontier approach with the data at disposal: in this respect, the SSFA model proposed in equation (2) seems a particularly effective tool to isolate and evaluate the territorial component separately from the individual performance of the productive units. The results of the estimation are reported in Table 3: in all cases, the value of the coefficients for the inputs are consistent with those obtained from the SFA specification, with the expectation of the intercept that becomes not significant; however, this result is expected as the spatial specification generates a further shift in the production curve with respect to the SFA as a consequence of the isolation of the spatial effect, transforming the average value of \(b_0\) into a multiplicity of individual effects. This pattern as already been shown in section 3 (Figure 1). Interestingly, the value of the \(g\) parameter (0.433) is lower than the one estimated with the SFA model; this evidence supports the hypothesis that part of the technical inefficiency was mistakenly attributed to the production process rather than to neighbourhood effects.

Table 3 - Wine production function - SFA estimator

| Dependent variable: production output (log) | Estimate | Std. Error | z value | Pr(>|z|) |
|-------------------------------------------|----------|------------|---------|----------|
| Intercept                                 | -0.0707772 | 0.7387170  | -0.09581 | 0.923671 |
| Labour input                              | 0.6081372  | 0.0525574  | 11.57091 | <2e-16 *** |
| Machinery capital input                   | 0.3328368  | 0.0436632  | 7.62282  | <2e-16 *** |
| Water, energy and fuel input              | 0.0500784  | 0.0132213  | 3.78770  | 0.000152 *** |
| Land capital input                        | 0.325007   | 0.0401612  | 8.09240  | <2e-16 *** |
| \(\sigma^2_{\text{error}}\)              | 0.3253159  | 0.1250316  | 2.60187  | 0.009272 ** |
| \(\sigma^2_{\text{nu}}\)                 | 0.3591916  | 0.0456068  | 7.87584  | <2e-16 *** |

Note: *p<0.1; **p<0.05; ***p<0.01

\(^6\) The value of Geary C lies between 0 and 2. Values lower than 1 demonstrate increasing positive spatial autocorrelation, whereas values higher than 1 indicate increasing negative spatial autocorrelation. C=1 is consistent with no spatial autocorrelation in the data.
The benefits of the proposed framework are also evident in terms of global and local spatial autocorrelation: indeed, the global Geary C statistic equal to 1.048 suggests that global autocorrelation has been effectively removed from the residuals, while local ci (for each unit i) associated with the SSFA estimates is significantly lower with respect to the unconditional SFA scenario (Figure 3).

4.2.1 Explaining the spatial effect through the analysis of territorial imbalances

The analysis presented in the previous section has confirmed the presence of a spatial effect in the data that has been successfully isolated using the SSFA specification proposed in section 3. In this part of the empirical investigation, the focus is moved to the spatial effect itself, in an attempt to explain its structural characteristics and interpret its nature in light of the considerations emerged in the brief review of the wine sector presented in section 2. In order to do so, the analysis is focused on the territorial imbalances, defined as the difference between the efficiency term estimated in the SFA specification and that identified with the SSFA approach7: in general, higher values of territorial imbalance suggest the presence of a stronger territorial component.

A geographical representation of the territorial imbalances is presented in Figure 4: the map shows the presence of a heterogeneous distribution of the spatial effect, with areas characterized by a strong territorial factors while in other case the role of the locality appears to be negligible in determining the performance of the productive units.

Figure 3 - Local ci kernel density of the SFA and SSFA efficiency

Figure 4 - Differences between SFA and SSFA efficiencies per quantile, q = 4

7 this term is generally positive, given in the SFA the spatial effect is mistakenly incorporated into the error term, generating higher values relative to those estimated with the SSFA
Having established the presence of a relevant and heterogeneous spatial effect in the Italian wine industry, the immediate question is whether this effect can be satisfactorily explained by a plurality of tangible local factors. To address this issue, a second stage analysis has been implemented regressing the territorial imbalances on a plurality of contextual variables, following previous findings in earlier studies (see e.g. Hani et al., 2003, Reig-Martinez et al., 2011 and Bardaji and Iraizoz, 2015) and incorporating other determinants that are believed to have an impact on firm-level productivity in agriculture. In this respect, the wide variety of variables available in the FADN database can be effectively used to incorporate a number of relevant factors that are believed to explain the spatial effect in the wine industry: specifically, the set of covariates include (i) endogenous factors linked to the productive process or to the corporate characteristics, (ii) exogenous physical factors and (iii) exogenous economical indicators related to the local supply factors.

The results of the estimation are reported in Table 4: in most cases, the coefficients are significant and the sign is that expected. As far as the endogenous factors are concerned, the first interesting result is determined by the key role played by quality and reputation: in fact, firms producing a better wine quality (OTE = 3510, wineries specialized in the production of quality wine) appear to benefit from a larger territorial effect. Moreover, being located in an area characterized by a higher reputation of the production process (i.e. DOCG production area) is also associated with an increase in territorial imbalances, irrespective of the quality of wine produced by the firm. The data also reveals the importance of the territorial effect for family-owned firms (Family owned) and those productive units characterized by an higher degree of product differentiation.

Regarding the exogenous physical factors it is worth noting that the coefficients of both Physical disadvantage (climate) and the Biophysical disadvantage are negative and statistically significant: this result is not surprising considering the beneficial effects associated with a temperate climate, advantageous slope inclination and slope exposure on wine production. Finally, a set of exogenous economical factors have been included in the model, in order to account for the role of elements external to firms and internal to the region that generate a competitive advantage among economic agents. The results are consistent with the expectations, showing that a higher level of the surrounding economic and network infrastructure has an indirect beneficial effects for the productive units specialized in wine production; on the other hand, the presence of a lower level of human capital in the region (higher Scholastic drop-out indicator) has a negative impact on the territorial imbalances. Despite the results of the estimation confirm the important role played by the above mentioned variables, it is worth noting that the presence of these tangible factors is not sufficient itself to explain the variance of the territorial imbalances (R² =0.214). This evidence indirectly confirms the presence of intangible factors associated with context specific and informal practices of learning that cannot be evaluated through the mere inclusion of specific contextual variables in the model: in this respect, the implementation of a SSFA approach can effectively address this issue, allowing to isolate the intangible effects associated with tacit knowledge flows and incremental learning that are peculiar of the wine industry and cannot be merely proxied through the inclusion of specific contextual variables.

4.2.2 Do spatial effects vary with size? The different role played by the local network in small and large firms’ efficiency

The analysis implemented in the previous section has highlighted that a combination of tangible and intangible factors explain the presence of a spatial effect in the Italian wine industry. In this scenario, the role of intangible effects appears to be particularly significant and possibly associable with the local business climate and informal interaction among local actors that has been discussed in Section 2: in this respect, the results support the empirical evidence emerged in previous contributions focused on case studies in specific wine regions, suggesting the presence of communitarian networks, characterized by resource sharing and informal interaction.
However, as several recent contributions have shown, access to these networks appears not to be uniform among local actors: not infrequently, small firms tend to be more inclined to engage in informal interaction practices with local peers, given the higher expected benefits, while large firms tend to acquire new knowledge from external sources, using the wide variety of formal channels at their disposal.

The possible presence of this differential has been evaluated by plotting the difference between the SFA and the SSFA efficiencies against the size of the firm. Figure 5 confirms that an inverse relationship exists between the two variables, with a higher territorial effect for small firms that tends to decrease when firm size become larger: however, the presence of a negligible spatial effect among large wine producers seems not to be associated with lower levels of technical efficiency: in fact, Figure 5 shows that these firms are more efficient than small producers.

This apparently contrasting trend can be easily interpreted in light of the patterns already identified in Section 2: the presence of a negligible spatial effect among large firms should not be motivated by their difficulty to access to local networks, but rather by a voluntary choice aimed at focusing on alternative sources of knowledge, such as internal learning, interaction with producers located outside the neighbourhoods and formal collaborations with institutional actors. The choice of these alternative forms of learning enables these firms to stay at the forefront of technical development, focusing on the most efficient technologies and maintaining high levels of technical efficiency. On the other hand, small producers who cannot access to external knowledge are required to rely on the informal learning practices associated with continuous interaction with the local community, generating the spatial effect identified for this subset of the firm population in the model. Although these intangible factors allow to reduce the gap with the leaders, small firms still display a lower level of efficiency relative to the larger ones.

The main finding introduced in this section confirms that spatial proximity does not necessarily generate knowledge spillovers. The main reasons explaining this pattern are probably two: on the one hand, large producers generally do not need to link to local networks to access informal sources of learning; on the other hand, leader firms may not be willing to share the knowledge acquired through access to external formal sources, generating positive externalities for small firms located in the close neighbourhood. In this respect, the behaviour of large firms could be interpreted as a rational strategy aimed at retaining a competitive advantage in the production process, ensuring higher levels of technical efficiency.

**Figure 5 - SSFA efficiency and difference with SFA per produced output (log) and ownership.**

5. Final remarks

The empirical exercise, implemented on a sample of wine producers extracted from the Italian FADN survey, shows that the spatial specification proposed by Fusco and Vidoli (2013) is extremely effective in disentangling the spatial effect that is present in the data, isolating a specific component that is erroneously attributed to the error term in the standard SFA approach.
The main features of the spatial effect are evaluated through a second stage analysis, in which the territorial imbalances (i.e. the difference between the inefficiency term calculated with the SFA specification and that identified with the SSFA approach) are regressed against a set of contextual variables that are generally associated with the presence of a stronger effect: although the role of these factors is confirmed by the results of the estimation, a relevant share of the variance in the model remains unexplained, suggesting that a key role is played by intangible factors that cannot be formally included in the model. Following the recent findings in the literature, it is reasonable to assume that this intangible component is the consequence of a network effect associated with the local business climate: in most localities, the presence of an embedded community stimulates a process of local learning that generates the diffusion of tacit knowledge through continuous interaction among the local actors. Although the investigation does not allow to evaluate whether this flow relates more to business information or technical knowledge, the identification of such aspect is by itself a key finding of the contribution.

The analysis of the degree to which the spatial effect varies with firm size provides evidence of a clear tendency of this effect to be significantly lower among large firms. This finding is in line with previous research, confirming that firms interacting in economic networks are not an homogeneous entity, but play different roles in the local scenario: although it is not a direct consequence of the results of the paper, it can be speculated that the different size of the territorial effect found in small and large wine producers is associated with different abilities and willingness to interact and share knowledge with neighbours located in close vicinity: such a scenario would confirm the trend already identified in case studies on wine sector (Giuliani, 2007; Morrison and Rabellotti, 2009), showing that large firms have a strong tendency to access to external and formal sources of knowledge, sharing the information acquired from outside with a small number of firms with which they collaborate on a regular basis. The regular interaction with external sources of knowledge enable these firms to stay at the forefront of the technological frontier, enabling them to face the challenges required by the rapid technological change: such a trend would be consistent with the higher levels of technical efficiency found among large firms in the empirical analysis.

The results of this investigation open some interesting avenues for further research. The SSFA specification can be extended to accommodate the use of panel data: the implementation of such an approach would be particularly convenient to control for seasonal or other unobserved factors that can influence harvesting in a particular year, such as the presence of parasites or other transient factors.

### Annex

#### Table 5 - Production function variables - main statistics

<table>
<thead>
<tr>
<th>Statistic</th>
<th>N</th>
<th>Mean</th>
<th>St. Dev</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Production output (Physical units) (log)</td>
<td>853</td>
<td>6.599</td>
<td>1.236</td>
<td>2.890</td>
<td>10.953</td>
</tr>
<tr>
<td>Labour input (Hours) (log)</td>
<td>853</td>
<td>8.071</td>
<td>0.746</td>
<td>6.125</td>
<td>11.613</td>
</tr>
<tr>
<td>Machinery capital input (Kw) (log)</td>
<td>853</td>
<td>6.672</td>
<td>0.786</td>
<td>1.609</td>
<td>7.553</td>
</tr>
<tr>
<td>Water, energy and fuel input (£) (log)</td>
<td>853</td>
<td>5.857</td>
<td>2.331</td>
<td>0.000</td>
<td>12.468</td>
</tr>
<tr>
<td>Land capital input (Ha) (log)</td>
<td>853</td>
<td>6.806</td>
<td>1.097</td>
<td>3.434</td>
<td>10.787</td>
</tr>
<tr>
<td>OTH = 3510 (yes)</td>
<td>853</td>
<td>0.715</td>
<td>0.452</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Physical disadvantage (slope) (yes=1)</td>
<td>853</td>
<td>0.216</td>
<td>0.412</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Physical disadvantage (climate) (yes=1)</td>
<td>853</td>
<td>0.284</td>
<td>0.451</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Biophysical disadvantage (yes=1)</td>
<td>853</td>
<td>0.699</td>
<td>0.459</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Economic infrastructure indicator (index)</td>
<td>853</td>
<td>96.038</td>
<td>57.679</td>
<td>26.668</td>
<td>397.647</td>
</tr>
<tr>
<td>Network infrastructure Indicator (index)</td>
<td>853</td>
<td>86.356</td>
<td>30.926</td>
<td>18.400</td>
<td>187.976</td>
</tr>
<tr>
<td>Scholastic drop-out indicator (index)</td>
<td>853</td>
<td>13.499</td>
<td>4.842</td>
<td>8.394</td>
<td>24.026</td>
</tr>
<tr>
<td>DOCO production area (yes=1)</td>
<td>853</td>
<td>0.356</td>
<td>0.479</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>EU subsidies (1,000 €)</td>
<td>853</td>
<td>2.339</td>
<td>8.093</td>
<td>0.000</td>
<td>87.108</td>
</tr>
<tr>
<td>Financial charges (1,000 €)</td>
<td>853</td>
<td>0.313</td>
<td>2.771</td>
<td>-51.183</td>
<td>0.000</td>
</tr>
<tr>
<td>Family owned (Direct family members=1)</td>
<td>853</td>
<td>0.431</td>
<td>0.496</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Gender (M=1)</td>
<td>853</td>
<td>0.779</td>
<td>0.415</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Young owner (yes=1)</td>
<td>853</td>
<td>0.117</td>
<td>0.322</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Diversified production (yes=1)</td>
<td>853</td>
<td>0.109</td>
<td>0.312</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Organic production (yes=1)</td>
<td>853</td>
<td>0.036</td>
<td>0.187</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

### References


