Rural development funding and agricultural labour productivity: A spatial analysis of the European Union at the NUTS2 level

Martijn J. Smit a,b,*, Eveline S. van Leeuwen a, Raymond J.G.M. Florax c,a,d, Henri L.F. de Groot a,d,e

a VU University Amsterdam, Netherlands
b Utrecht University, Netherlands
c Tinbergen Institute, Amsterdam/Rotterdam, Netherlands
d Ecorys, Rotterdam, Netherlands

1. Introduction

Traditionally, agriculture has formed the mainstay of rural economic activities, but this changed in the second half of the twentieth century in the direction of a much more balanced and diverse economy as more and more labour was driven out of the agricultural sector into employment in alternative sectors. This labour transition has, however, not taken place evenly across space, and continued innovation still drives labour productivity up in certain subsectors and geographical environments. Across the European Union, Regional Development Programmes (RDPs) have been put in place to enhance rural economies and societies.

In this paper we investigate the development of agricultural labour productivity across the European Union (EU), and focus specifically on how agricultural labour productivity is impacted by RDP spending. We will show how the development of both labour productivity and the effectiveness of EU funds vary geographically across regions.

A B S T R A C T

With more than 50 percent of the European population (EU-24) living in rural areas and a renewed focus on stimulating smart, sustainable and socially inclusive growth, Rural Development Programmes (RDPs) are an important instrument for economic, social and environmental policies. Evaluating the impact of rural development programmes is, however, complicated due to the widely varying policy targets of RDPs as well as their substantial heterogeneity across rural areas. In this paper we use spatial econometric techniques to evaluate RDPs in the European Union, at the NUTS2 level, and focus specifically on labour productivity in the agricultural sector. To address the clear spatial patterns in the distribution of agricultural labour productivity, we employ regression models in which spatial heterogeneity and spatial dependence are explicitly modelled to quantify the impact of RDPs.

There is an extensive literature on regional productivity and economic growth in the field of spatial economics. This literature includes a growing number of studies employing spatial econometrics (Abreu et al., 2005a; Gardiner et al., 2004; Bartelsman and Wolf, 2014; Van Oort et al., 2012). Although productivity can be measured at the level of the aggregate economy, a sector-specific approach is preferable in order to account for the well-known apples and oranges problem (Dollar and Wolff, 1993; Bernard and Jones, 1996).

Since we will be looking at relative changes in agricultural labour productivity, we effectively study technological change, its diffusion and adoption, and the resulting process of catching up, where less advanced regions copy techniques and routines from the technologically leading region. At the national or European level, the idea that spending money and improving conditions in one region may have positive benefits in other regions is very attractive for policy makers. If such regional spinoff effects are indeed real, econometric techniques need to explicitly allow for such effects in order to provide unbiased and efficient estimates of the impacts of RDPs. Concurrently, this may lead to changes in the perception of the effectiveness of rural development policy and/or to better targeting of scarce resources.
2. Labour productivity in agriculture

In agriculture, labour productivity depends on many factors, among which three main categories can be distinguished (Hayami and Ruttan, 1970):

- resource endowments (such as soil fertility and precipitation),
- technology (for instance, fertilizer and machinery), and
- human capital (education, physical strength, etc.).

These factors explain, for example, why labour-intensive wine growing in California or France yields much more value added per unit of labour than labour-intensive rice growing in western China.\(^1\) One should note that space (or travel time; Spiekermann and Wegener, 1994) should also enter into the equation. The famous model by Von Thünen (Forstner et al., 2009) predicts that even with the same soil type everywhere, areas nearer to the market will be able to specialize in different products due to their relatively low transportation costs. However, such models mainly explain the levels of labour productivity attained in a region. To investigate the impact of policies such as subsidies, we will need to look at the changes in labour productivity over time. The most important factor there is of course technological change and its diffusion.\(^2\) At a regional level, this could result in certain regions catching up (Abramovitz, 1986) when output or productivity levels converge between leading and lagging regions (Dollar and Wolff, 1993; Abreu et al., 2005a,b). Notably, farmers in less advanced regions have the advantage that they are able to copy techniques and routines from the region which is closest to the technological frontier (Dosi, 1982); such copying can be argued to be easier when the leading and lagging regions are geographically closer to each other.

3. European support fostering productivity

Productivity is a key factor on the Lisbon agenda, and so is territorial cohesion. European Union support for investments in agricultural holdings started already in the mid-1960s, and it has ever since been a permanent instrument of the Common Agricultural Policy (CAP). In the current implementation of the CAP, support for enhancing agricultural productivity is labelled “farm modernisation”, classified in the broader axis of “competitiveness” (axis 1 in Fig. 1). All regions are eligible for this type of support. By supporting individual holdings to innovate and increase their productivity in a sustainable way, region-wide economic growth and competitiveness are enhanced. Investing in agricultural productivity can have a positive effect on the economy as a whole (Gollin, 2010).

Within the abovementioned competitiveness axis, several policy measures have been identified. Measure 121 focuses on labour productivity (see Fig. 2), and it is one of the largest targets of RDP spending: it covers over one-tenth of the total budget across the European Union, ranging from 3 percent in Ireland to 51 percent in Belgium. The total amount of money spent under this measure over the entire programming period

---

1. Obviously, more factors than mentioned here play a role, such as the demand side of the market (market prices), competing sectors on the labour market, and management techniques.
2. At the micro level, the degree of flexibility of individual farmers has been studied in detail (Renner et al., 2014).
(2007–2013) will be over €15 billion. Appendix A shows a few examples of actions supported under this specific policy measure.

A number of studies has discussed the benefits of the farm modernization measure at the micro level (see Uthes et al., 2011, for an overview). Dwyer et al. (2008) show how restructuring and development of holdings leads to productivity gains and increased business turnover: studies in Brandenburg (Forstner et al., 2009) and Belgium (Beck and Dogot, 2006) found that labour productivity and the improvement of working conditions were the most important outcomes of the policy.

At the macro level, a large number of studies has looked at the impact of (EU) subsidies on regional growth and productivity. For example, Coppola and Destrateni (2011) look at the impact of structural funds on four broad sectors in Italy; they find a positive impact of structural funds outlays on agricultural productivity. Ball et al. (2001) found a large degree of catching up in agricultural productivity at the country level. Ezcurra et al. (2011) analyzed productivity levels in European regions in a model quite similar to our approach; they also use a regional typology, but they cover only Western Europe. Their results will be discussed in more detail below.

4. Theory and model

Although the EU expenditures on RDP amount to approximately 70 billion euro between 2007 and 2013 (European Communities, 2008), there are reasons why the effect of the spending itself might be small. Foremost is the displacement effect: if a subsidy to some farms makes these very competitive, they might actually push some of their competitors out of the market. To the extent that these competitors are in the same region, this will lead to a displacement effect. A second risk is a deadweight effect or a free-rider effect: if subsidies replace investments that would have taken place anyway, the amount of subsidies will not change the outcome (Meyer, 2006).

Another dimension of effectiveness of the spending relates to the spatial scope of the impact of the policy. Qualitative research, including interviews with experts, gives an expectation that spillover effects from RDP spending might be low at any spatial scale level (Uthes et al., 2011). Since we will be measuring spatial effects at the relatively high NUTS2 level for reasons of data availability, the probability of finding large spatial correlations is even lower, as these regions are large enough to incorporate most possible spillovers within their boundaries, and moreover, many NUTS2 areas coincide with planning regions for the RDP. However, in more general terms and for lower-level spatial units we can think of two main ways in which spillovers may occur:

- through copy-cating: EU-funded modernization measures on one farm may be copied by neighbouring farms, or by farmers within a local network. We note that the term ‘local’ is generic and does not imply a predefined area, certainly not the area to which an RDP applies. Moreover, it is well known (even in the spatial sciences) that proximity has other explanations than the obvious spatial one (Boschma, 2005; Torre and Rallet, 2005). Obviously, copy-cating can be actively promoted or actually deliberately pursued by local governments and farmers’ associations.

- through migration: some farmers may move their ‘business’ elsewhere, especially if they are not dependent on land or fixed assets. Many of the movers will move within a region, but some may cross regional boundaries. In some cases, a farmer who has scattered possessions may even receive money in one region but effectively spend part of the subsidy elsewhere.

Hence, the effects of RDP’s can easily cross regional boundaries, especially where these do not coincide with actual patterns of physical and/or cultural separation (Newman and Paasi, 1998). Some boundaries will be stronger than others. Physical separation, due to for example water or elevation, can hamper contact between otherwise similar regions. Cultural separation is especially severe across country borders (Hussler, 2004). Borders with a language barrier, both between countries and within countries (for instance, in Belgium), will provide an even more severe obstacle to spatial interaction.

Apart from the spatial configuration, the dimension of time is important as well. We would like to point out three aspects in particular:

- Investments in productivity increases take time to materialize. Previous research has shown that results are to be expected 2–3 years after the investments have taken place at the earliest (Forstner et al., 2009).

- Since we are interested in (long-term) structural developments of the economy, we need to focus on longer time periods. In doing so, we will avoid confounding long-term structural effects with temporary fluctuations caused by business cycles, of which different types, with different lengths occur in agriculture practices (Coase and Fowler, 1935; Da-Rocha and Restuccia, 2006).

- The current RDP (“RDP2”) started only in 2007. Hence, the amount of years for which data are available is still small. We have therefore matched these data to that of the previous RDP period (2000–2006), allowing us to analyze the period 2000–2010. To a certain extent this implies loss of detail, as some policy measures have changed between the two periods, and data for RDP1 are not well documented.1

4.1. Model approach

We base our model of labour productivity change on the basic Solow (1956)/Swan (1956) model which was first empirically tested by Mankiw et al. (1992; henceforth MRW). In the setup of MRW, total production \(Y\) is a Cobb–Douglas function of capital \(K\), labour \(L\), and a labour augmenting technological component \(A\). The factor \(A\) represents an amount of technology, endowments, institutions, etc. that is available locally: in short, \(A\) can be thought of as an estimated parameter capturing all local influences on production apart from physical capital and labour input. MRW take this factor to consist of a global constant plus a purely random component, but other explanatory variables can be easily included, although this will make interpretation of the core variables less straightforward. Differences in growth in this basic framework are driven by: (i) differences in population growth, (ii) differences in investments, and (iii) differences in the rate of return on capital accumulation. The latter are caused by differences in the distance to the steady state productivity level, and result in countries with relatively little capital to be characterized by relatively fast growth driven by what we may call a ‘convergence bonus’.

In an extension to this basic model, MRW include human capital \(H\) as a production factor. This results in the following production function:

\[
Y_t = K_t^\alpha H_t^\beta (A_t L_t)^{1-\alpha-\beta}
\]  

1 Our data on RDP expenditure were provided by the European Commission through the CATS database. Budget codes for 2000–2003 were grouped into broad classes, and then classified as agriculture-related measures (agri-)environmental measures, or other measures. These were then taken to be the equivalents of axis 1, axis 2 and the total of the other axes of the RDP2 period.

2 This is a convenient model, but this is not the only way to model productivity. For example, Hayami and Ruttan (1970) (in their appendix; see discussion in Huffman and Evenson, 1992, p. 366) follow the approach where wages in agriculture is the key variable of the analysis.
in which \( \alpha \) and \( \beta \) are the production elasticities of physical and human capital, respectively. The sum of these two elasticities is assumed to be smaller than unity \((\alpha + \beta < 1)\). Under the assumption of a region-specific rate of population growth \((n_i)\), an equal initial level of technology \((A_0)\), a constant global rate of technological progress \((g)\), a constant global depreciation rate of physical capital \((\delta)\) and a region-specific rate of investments in physical capital \((s_k)\), one can derive a steady state labour productivity level by log-linearizing the steady state labour productivity level to arrive at (MRW, p. 418):

\[
\ln \frac{Y_{it}}{L_{it}} = \ln A_0 + gt - \frac{\alpha}{1 - \alpha} \ln(n_i + g + \delta) + \frac{\alpha}{1 - \alpha} \ln(s_k) - \frac{\beta}{1 - \alpha} \ln(n_i^*)
\]

Islam (1995, p. 1136) subsequently reformulated this model to take advantage of existing variation across time as well. Spatial panel modelling, incorporating both space and time, is feasible nowadays based on substantial progress that has been made in incorporating spatial autoregressive processes in the analysis of panel data (e.g., Anselin, 2006; Elhorst, 2003, 2010; Millo and Piras, 2012). Since we are interested in productivity growth as a long-run phenomenon one needs to be cautious not to over-value temporal fluctuations in productivity data, especially over short time periods. The reliability of the available annual data series is limited; negative figures for some region-year combinations are indicative of substantial complications in accounting procedures. We therefore choose not to focus on a full-blown panel version in the spirit of the Islam model, but rather to provide two “snapshots” of a spatial cross-sectional model for two different time periods.\(^6\)

4.2. Data

We use data from a Cambridge Econometrics database, called the Regional Economic Model. This database provides comprehensive data at the NUTS2 level, as well as a rather restricted set of data at the NUTS3 level. Cambridge Econometrics has employed “deflation, interpolation, and summation constraints” to clean and verify the data, mainly based on information taken from the Eurostat REGIO database.

For our analysis we connect this database with two other data sources. First, we use data on RDP spending by NUTS2 region, gathered by the European Commission in the so-called CATS database. Secondly, we make use of Eurostat data, which have been conveniently bundled in the so-called MetaBase, developed by LEI (Dol and Godeschalck, 2011). From the enormous amount of available data there, we have selected some relevant proxies available at the NUTS2 level, including the size of agricultural holdings, the number of holdings with livestock, and the amounts of land (as a percentage of the total area) used for pasture, woodland and vineyards, as well as the share of land in Less Favoured Areas.

The Cambridge Econometrics dataset and Eurostat both comprise 283 NUTS2 regions, covering the entire European Union (including overseas regions) as well as Norway and Switzerland. Basic economic data are available for 1980–2014 (extrapolations after 2010); data on CAP spending only for 1999–2010. Therefore, we base our analysis on the 11-year period between 2000 and 2010. For econometric reasons, we restrict ourselves to regions within the European part of the EU, with a few additional outliers removed (see below), yielding a sample of 261 regions. All regions are eligible to receive (at least some) RDP subsidies.

The variables we include in the model, besides labour productivity, the \( n + g + \delta \) term from the MRW model (where we set \( g + \delta \) equal to 0.05 as is done is most applied studies; see, e.g., MRW and Ederveen et al., 2006), investments and RDP spending, are:

- population and motorway density intended as a proxy for access to consumers;
- the share of agricultural land in mountainous and less favoured areas to account for less favourable production circumstances;
- the share of large farms (larger than 50 ha) in a region to capture economies of scale or the use of less intensive forms of agriculture;
- climate data, taken from the EDENext data portal (Hijmans et al., 2005) at high resolution, and joined to the NUTS2 regions using ArcMap;
- and measures for some specific types of activities, namely utilized agricultural land, woodlands, vineyards, flowers and livestock, which all have their specific technological and climatic differences.

The number of variables is large, compared to the number of regions. In the steady-state regressions reported in Section 5.1 we include all the variables described above; but the models in Section 5.2 also include a series of variables for RDP spending, and we chose to include only the first three variables from the list above, which we consider to be the most important controls. Based on previous literature, we gathered some other variables,\(^6\) but since these never rendered statistically significant results, we omitted them from the final regressions. Appendix B provides descriptive statistics of all variables used.

We construct our dependent variable “labour productivity in agriculture” as gross value added (GVA) per employee. Fig. 3 shows how agricultural labour productivity varied across Europe in 2010. It is easy to notice clear spatial patterns of high productivity in Nordic Europe, North-Eastern France, and the Netherlands; very low productivity in Poland, Romania and Bulgaria, and also in Alpine Austria, Greece, Slovenia and southern Italy. When we look at the map of spending in Rural Development (Fig. 4), we see that the Czech Republic has spent large amounts of money (measured per holding). Labour productivity has gone up markedly in this country in the period 2000–2010. Apart from this region, Sweden and Finland as well as the Anglo-Scottish border regions have received the largest amounts of rural development money.

We note that subsidies can be evenly distributed across all farms in the region, but it can also be targeted at a small number of holdings. Although different spatial distributions of spending are likely to affect effectiveness of the subsidies (e.g., concentration might improve the effectiveness if positive economies of scale are present) we lack information to empirically address this issue. At the planning level, this is basically an issue of spatial scale and of the so-called modifiable areal unit problem (Briant et al., 2010; Burger et al., 2010). Some of the case studies within the SPARD project (e.g. Desjeux et al., in this issue) delve deeper into these issues.

4.3. Space

Spatial clustering can exist due to two processes. On the one hand, there can be unique spatial characteristics that create a spatial pattern. This pattern is due to spatial heterogeneity: the mean

---

\(^5\) We have also split the sample into two periods, viz. 2000–2005 and 2005–2010, to check the robustness of the cross-sectional results and to see whether parameters evolve over time; most coefficients and significance levels proved quite stable. Detailed results are available upon request.

\(^6\) This included data on smaller farm types, and the share of family labour in total labour, which has been widely discussed in development economics (Bardhan, 1973; Gershon, 1985).
and variance of a random variable are not the same (‘stationary’) across space. If this influence is due to characteristics such as physical geography, cultural differences, institutions, climate, or any other spatially varying characteristic, then allowing parameters to vary across these characteristics provides an optimal way to deal with spatial heterogeneity. For the effects of RDP spending, we address this problem by classifying the European NUTS-2 regions into six types. These six types are organized along two axes: on
the one hand, southern versus northern regions and on the other, rural vs. intermediate vs. urban regions. We hypothesize that RDP spending has different effects according to the population density of an area: agriculture in urban areas faces different challenges and opportunities (Dekkers, 2010; Von Thünen, 1826) as compared to peripheral, low-density areas. In addition, there are clear institutional differences within Europe, and in considering the importance of institutions for the regional economy (e.g., Rodrik et al., 2004; Ederveen et al., 2006), we choose to also distinguish between northern and southern countries. The resulting classification into six large regions is shown in Fig. 5.

The other cause that potentially leads to a noticeable pattern of similar values across space is spatial dependence. Here, a process in one region is related to what happens in neighbouring regions. It may be that labour productivity in one region influences labour productivity of the neighbours in a contagious fashion, as for a disease. Alternatively, it may be that an increase in labour productivity in one region increases welfare in surrounding regions directly, without affecting labour productivity there. It is obvious that the actual causal chain of effects cannot be inferred from the mere use of spatial econometric techniques, so caution is needed in interpreting the results.

We hypothesize that productivity in one region can be correlated with spending in neighbouring regions. Therefore, we include spatially lagged RDP spending variables. We also allow for spatial patterns that we have not been able to account for explicitly through the incorporation of spatial heterogeneity and a spatially lagged RDP variables. These patterns can be operationalized as direct relationships between the outcome variable (labour productivity) in neighbouring regions (a so-called spatial lag model), or as associations between unknown processes in neighbouring regions (a so-called spatial error model). To construct the spatially lagged variables, we need a definition of neighbours, or, in other words, a spatial weight matrix. We choose a Gabriel matrix for this;

more details are given below. A series of Lagrange Multiplier tests (Anselin et al., 1996) is used to investigate whether the abovementioned spatial process models are warranted from a statistical point of view.

A final issue in spatial analyses is the possibility of spatial crowding out. This can occur for example in regional employment under the influence of localized capital investments, as De Castri and Pellegrini (2012) show for southern Italy. A spatial lag model, in which per capita agricultural productivity levels are explicitly linked across space, does take account of such crowding out and spatial spillover effects. In agriculture, however, the phenomenon of crowding out may be substantially less prevalent, especially in the short run, since agricultural production is spatially tied to productive soils and production is therefore generally not considered footloose.

4.4. Gabriel matrix

Gabriel plots are well-known in graph theory and other areas of mathematics, and sometimes used in spatial econometrics (Bivand and Portnov, 2004). They have the advantage of providing a neighbour to every observation in the dataset – so that, unlike with the more customary contiguity matrices, also islands and remote regions will be included in the set of neighbours. Moreover, Gabriel-type relationships compare well with distance-based neighbour definitions in areas with heterogeneous region sizes. In a European context, distance-based weights have the disadvantage that the cut-off distance that determines whether two regions are neighbours has to be quite large in order to provide neighbours for a few rather distant regions in Scandinavia. A side effect of a large cut-off distance is, however, that regions in denser areas of the Union have many neighbours. The Gabriel graph does not have this disadvantage. The principle behind the Gabriel neighbour definition is to include as neighbours only those regions where there is no
intervening geographical centroid of a third region that falls within an imaginary circle around two geographical midpoints. Fig. 6a–c illustrates how the Gabriel neighbour criterion works as well as the non-zero links present in our spatial weights matrix.

There are 262 NUTS2 regions in our dataset. The Gabriel plot has 1080 links in total, which amounts to an average of 4.1 links per region. Only three regions have just one link; these are the islands of Cyprus, Malta and the Canary Islands. There are four regions that have the largest number of links totalling seven.

Utilizing the spatial weights matrix, a LISA (local indicator of spatial association) cluster map provides a first indication of the local spatial clustering of the intensity of spending on Rural Development Programmes. Fig. 7 shows the cluster maps for RDP spending and for labour productivity. The extent of spatial correlation, measured by means of the (global) spatial correlation coefficient Moran’s I, is 0.30 for RDP spending and 0.52 for labour productivity, both of which are statistically significantly different from zero. The labour productivity LISA map matches the map of Fig. 3 quite closely. The map of RDP spending in Fig. 7a, on the other hand, only partially reproduces the spatial patterns of the map in Fig. 4. For example, the LISA map shows significant clustering of high spending in and around the Czech Republic, but it does not indicate any significant clusters in Scandinavia or in the Anglo-Scottish border region, even although those appear quite prominently in Fig. 4. Fig. 7a also indicates spatial clusters of low values across Italy, in Romania and Bulgaria, and in parts of Spain and Poland, but these do not match areas of low agricultural labour productivity. Relatively low labour productivity is mainly found in Poland, Romania and Bulgaria.

4.5. Caveats

The RDP spending data we use are organized by year, but these years are not regular calendar years; instead, they start with three months in the previous calendar year, and then contain the first nine months of the given year. In other words, spending for 2010 refers to the period from October 2009 until September 2010. However, in some years the figures have been corrected for what apparently were mistakes or perhaps ex-post changes to previously allocated funding. In a handful of cases, this results in negative RDP spending for a particular year. In the analyses presented here, this does not pose an immediate problem, since we consider the total RDP spending over 11 years, but it will lead to some imprecision, as the data are apparently organized by accounting
Moreover, famous analysis and 5.

However, we did choose to concentrate on the European part of the EU, removing information on the Spanish exclaves of Ceuta and Melilla (in Africa) as well as on the Portuguese Azores, but retaining the Canary Islands. The dropped regions are quite special and, in the case of the Azores, so far away that interaction is severely limited. We also dropped the city region of Brussels, as improbably high amounts of RDP spending are reported there, and the Beleařes, where total value added in agriculture was negative and hence implausible.

5. Results

MRW models typically concentrate on growth towards an equilibrium steady-state outcome, and therefore occupies a rather prominent position in the literature on convergence (Abreu et al., 2005a,b). However, we can also assume the status in a given year to be the steady state – i.e., all regions are in their steady state equilibrium – or that all regions are at the same distance from their respective steady states. This is a somewhat heroic assumption (surely there are still technological improvements which for sound economic reasons will be implemented in rural Bulgaria in the near future), but it gives us an interesting background with which to compare our results of a growth analysis, which we present below. When aiming to explain the productivity in the steady state, we would expect aspects such as the quality of the soil, hours of sunshine, level of technology and human capital to affect the type and efficiency of activities, and thus labour productivity (Section 5.1). However, when we move towards a growth model, the picture is fundamentally different, since we explain the dynamics (Section 5.2). In the case of growth, spatial effects such as knowledge spillovers can play an important role.

5.1. Steady-state model

Table 1 presents the results of our basic MRW model in which we assume that labour productivity reflects a steady state level determined by structural factors. We have not included RDP spending. The rather high goodness of fit is especially due to the country fixed effects, which are presented in a map in Fig. 8. Our base model, estimated by ordinary least squares (OLS), is given in the first column with numerical results.

The Breusch–Pagan test rejects the null hypothesis of homoskedastic error terms, with varying coefficients as the alternative hypothesis. We therefore re-estimate the base model with robust standard errors. Moreover, the base model estimated by OLS exhibits some form of spatial dependence as indicated by the LM tests (at the bottom of the table). The diagnostic LM tests (Anselin et al., 1996) are performed on the residuals of the base model to test if the errors are spatially correlated and/or if the spatial lag of the dependent variable has erroneously been excluded from the specification. The LM tests, both the unidirectional as well as the robust versions, indicate that the model with the spatial lag of the dependent variable is a suitable alternative specification. These results are presented in the third model. We therefore present the total impact of our model in an additional column, since the spatial lag structure implies a chain of influences from neighbour to neighbour and back. As a result, the OLS estimator we originally use is both biased and inconsistent.

Our original model also suffered from heteroskedasticity, as indicated by the Breusch–Pagan test; since the spatial version is

---

7 As a robustness check, we looked at shorter periods of time (see note 5). Farm structure surveys are performed irregularly, which limits the possibilities for an analysis of very short periods. Moreover, agricultural business cycles, such as the famous pig cycle, will be more visible in the data over very short periods.

8 In theory, we can obviously imagine a steady-state model that includes the total sum of all subsidies ever received, but this model cannot be operationalized because of missing subsidy data.

9 We use the R package ‘sandwich’ to estimate the robust OLS model (vcovHC). For the spatial models, we use the ‘spdep’ package (Bivand and Piras, 2015).
Table 1
Steady-state models.

<table>
<thead>
<tr>
<th>Labour productivity in agriculture in 2010 (log)</th>
<th>OLS model</th>
<th>Spatial lag model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff.</td>
<td>Rob.</td>
</tr>
<tr>
<td>Population density (log)</td>
<td>−0.048</td>
<td>0.22</td>
</tr>
<tr>
<td>Motorway density (log)</td>
<td>0.043</td>
<td>0.16</td>
</tr>
<tr>
<td>% of land that is utilized agricultural land</td>
<td>0.004</td>
<td>0.04</td>
</tr>
<tr>
<td>% of agricultural land in less favoured areas</td>
<td>−0.004</td>
<td>0.12</td>
</tr>
<tr>
<td>% of surface that is woodlands</td>
<td>−0.003</td>
<td>0.54</td>
</tr>
<tr>
<td>% of surface that is vineyards</td>
<td>0.011</td>
<td>0.63</td>
</tr>
<tr>
<td>% of surface that is pastures</td>
<td>0.013</td>
<td>0.64</td>
</tr>
<tr>
<td>% of surface that is flowers</td>
<td>0.138</td>
<td>0.17</td>
</tr>
<tr>
<td>% of farms with livestock</td>
<td>−0.294</td>
<td>0.04</td>
</tr>
<tr>
<td>Climate: mean minimum temp. in January</td>
<td>−0.016</td>
<td>0.15</td>
</tr>
<tr>
<td>Share of large farms (&gt;50 ha)</td>
<td>0.686</td>
<td>0.02</td>
</tr>
<tr>
<td>Constant</td>
<td>2.480</td>
<td>0.00</td>
</tr>
<tr>
<td>Country fixed effects</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>261</td>
<td>261</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>Breusch–Pagan test</td>
<td>92.2</td>
<td>0.00</td>
</tr>
<tr>
<td>Rho</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LM test</th>
<th>Chi-square</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error model</td>
<td>12.10</td>
<td>0.00</td>
</tr>
<tr>
<td>Lag model</td>
<td>13.67</td>
<td>0.00</td>
</tr>
<tr>
<td>Robust error model</td>
<td>0.88</td>
<td>0.35</td>
</tr>
<tr>
<td>Robust lag model</td>
<td>2.45</td>
<td>0.12</td>
</tr>
<tr>
<td>SARMA</td>
<td>14.55</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Note: Significant results (at 5%) are shown in bold. Columns marked “Rob.” report p-values based on robust standard errors. For the impacts, p-values are based on 5000 simulations.

We also present in the fourth model a variant of the spatial model with a correction for heteroskedasticity.

Regarding the density variables, we see that population density has a negative but statistically insignificant effect on labour productivity in agriculture. When looking at environmental variables and the variables that indicate the use of land, we see that productivity is higher in areas where there are more pastures. We would expect this variable to cover areas where soil or climate do not permit intensive agriculture, where one might suspect low value added; but apparently, value added is rather high in areas with a lot of pasture, so extensive agriculture can apparently prove profitable, at least if measured in GVA per employee. Note that if we would have had access to micro-data at the farm level, we could compare farms in northern Scotland that attempt to grow vines with those that have sheep, and likewise on the Côte d’Azur, and we would probably find large differences. As the data stands, however, the choices as to what activities are deployed in a region are limited by soil conditions and climate. Moreover, the inclusion of country fixed effects at least partly corrects for variation in climate and soil conditions. However, we did construct one climate variable: from the daily minimum temperatures recorded (or reconstructed) per square kilometre across Europe, we took the monthly average for January, and then the regional average by NUTS2 region. This variable had no statistically significant effect. In an analysis on the production of specific crops with their own specific sensitivities (e.g., a minimum amount of sun in the growing season, no rain during the harvest, no frost in winter), more climate variables could and should probably be constructed, and might prove more influential.

The country fixed effects of the OLS model are presented in Fig. 8 (without taking their statistical significance into account). Among the countries with a high level of labour productivity in agriculture, we see Hungary, the Czech Republic, Slovakia, Denmark, the Netherlands, Belgium and France. Much lower levels can be discerned in other parts of Eastern Europe. These differences can be due to any variable we did not include; this can range from exchange rates through specific weather conditions to institutional factors. As for the East–West difference, differences in capital availability, entrepreneurial spirit, and easy access to the latest technology can also play a role. De Wit et al. (2011) note that catching-up between East and West has been less than expected, and claim that the high dependency of the Eastern European rural population on agriculture plays an important role.

5.2. Growth models

We now estimate the dynamic model, relating the change in labour productivity between 2000 and 2010 to initial levels of labour productivity in 2000 and a series of variables influencing change. These include the \( n + g + \beta \) term from the standard MRW model, investments, and the controls from the regressions presented above. We also add RDP spending. Results are presented in Table 2. We will first discuss the basic model, presented in the first column. We will then proceed to discuss our main model, which uses some region-specific effects, and finally the more elaborate model with correction for spatial correlation.

First, we look at the basic variables. As expected, we find that labour productivity in 2000 has no statistically significant influence on the change in productivity from 2000 to 2010: we discern no catching up of regions with a lower labour productivity.

---

10 A spatial version of the Breusch–Pagan test is unfortunately not implemented in the ‘spdep’ package, nor elsewhere, for the type of 2SLS models we use here. We therefore perform the correction based upon the indication for the model estimated by OLS.

11 The model originally included a constant, and one country served as the omitted category for the fixed effects; these results have been recalculated to absolute levels by adding the value of the constant.
Furthermore, the technical term \((n + g + f)\) has a negative effect and investments have a positive effect; these are both as expected. However, the effect of investments is not statistically significant.

Regarding the density variables, we see that population density now has a negative and statistically significant effect on labour productivity in agriculture. In other words, productivity growth seems to be lower in (urban) areas, with high population densities. This might be somewhat mitigated for regions that are easy to access by car/truck (motorway density), but that effect is not statistically significant.

Interestingly, the variable for RDP spending shows no statistically significant effect. To be precise, we observe that at the NUTS2 level, over a period of 11 years, no effect can be measured of the total sum spent on competitiveness (axis 1) in rural development programmes on labour productivity. This does not mean that rural development programmes have no impact on labour productivity at all; especially at the micro-level, i.e. at the level of individual farmers, we do expect an impact; for an example of such impacts, see Allaire et al. (in this issue).

We also include spending on axis 2 (environmental measures) and on the other axes (socioeconomic and other measures); these might have a counterbalancing effect. In particular, encouraging farmers to take environmental measures (e.g., by reducing fertilizers or farming less intensively) labour productivity is likely to fall. Other measures, among which we find socioeconomic projects, can also have a positive influence, as it is possible that they manage to revitalize rural areas, attracting new (young, sophisticated) farmers and boosting investments. However, the coefficients are statistically insignificant for these variables as well.

Finally, we include the spatial lags of these three RDP variables, in order to evaluate whether there are any correlations with RDP spending in neighbouring regions. Only the last of these variables, for the axes other than 1 and 2, is statistically significant in the first model. The effect of such spending in the region itself is significant at the 10% level, but with a lower coefficient as compared to the spatially lagged version. It is difficult to explain how such spending may be more influential when coming from neighbouring areas.

5.2.1. Region-specific effects

We have reason to believe that there is spatial heterogeneity in the effects of RDP spending, both for climatic and for institutional reasons, as outlined above. Therefore, we employ interaction effects with six regions (also discussed above) to allow for regionally varying effectiveness of RDP spending. Besides RDP spending, we also use a region-specific constant. These results are presented in the second column of Table 2.

In the first model, RDP spending on axis 1 over all regions has no statistically significant effect. Now that we allow for regional variation in the effects of RDP spending on this axis, we find that RDP spending is statistically significant in southern Europe, but not in northern Europe. Spending appears most effective in regions with intermediate population density. Also, the spatial lag of spending on axis 1 is statistically significant; spending on this axis in neighbouring regions also has a positive effect on labour productivity in a region. Spending on axis 2 is also statistically significant, but with the opposite sign, suggesting agro-environmental measures in neighbouring regions are negatively related to labour productivity in a region. We have no intuitive explanation for why this

---

12 This cannot be caused by a correlation with RDP spending; the correlation between investment on the one hand and spending on the first, second and other axes is \(-0.06, 0.17\) and \(-0.17\) respectively.

13 Since we also have a region-specific constant, our estimation is a small step towards so-called regimes, which allow for regionally specific coefficients for all variables (Anselin, 2007).

14 The constants are not reported, but they are available upon request.
relationship occurs only in the spatial lag; the effect of spending on axis 2 within the region itself is not statistically significant.

The specific agricultural variable in this model shows that large farms have a negative impact on labour productivity in these estimations. As a robustness check, we tested whether the use of the number of holdings as the denominator in the spending variable influenced the results, by using RDP spending over GVA in agriculture instead of RDP spending over the number of agricultural holdings (Carillo and Maietta, 2014). The results are, however robust to this change in specification.

5.2.2. Spatial dependence

For both models estimated by OLS we perform diagnostic tests for spatial dependence, again using the Gabriel weight matrix presented above. Results are reported at the bottom of Table 2. These LM tests indicate that there are spatial effects unaccounted for in these models. The LM tests on both models suggest a spatial lag model as the preferred specification. We estimate such a model (using the ‘lagsarlm’ maximum likelihood estimator) and report the results in the final column. The results change rather drastically. RDP spending is no longer statistically significant in the southern regions. For intermediate regions, the p-value is now 0.06; at the 10% level of statistical significance, we accept this as statistically significantly different from zero. The spatially lagged RDP spending on axis 2 and on the other axes persists, however. The coefficients for the other variables vary slightly, with none of the significant variables changing sign. The final column presents impacts, as in Table 1.

6. Discussion and conclusions

Spending within the regional development programmes on the competitiveness programme (axis 1) seems to have a statistically significant positive relationship with the increase of agricultural labour productivity in southern Europe, but this effect disappears when we properly control for spatial effects, especially for rural and urban areas. This shows how not taking spatial econometrics into account can lead to erroneous (policy) conclusions.

A negative relationship with spending on agro-environmental measures (axis 2) was expected, but it was not statistically significant in our models – except when we looked at the effect of agro-environmental regions in neighbouring regions, which was indeed negative and statistically significant.

We measured the effects over an 11-year period, at the NUTS2 level for all of Europe, with RDP spending aggregated within each axis, and looking at the agricultural sector as a whole. It may very well be that spending has local effects for individual farms, and that it can be traced at very low levels of spatial aggregation (e.g., municipalities). It can also be the case that the effects are different in the long term, or work only in tandem with national and regional spending – for both of these possibilities, we do not have the proper data available to provide statistical tests.

As for the role of spatial data analysis, we have shown by means of exploratory spatial data analysis that agricultural labour productivity has a clear spatial pattern, that some of these patterns can be explained by the use of spatial variables, but that LM tests still
indicate that results obtained with a spatial econometric model are more appropriate.

Acknowledgements

This research was carried out as part of a research project entitled “SPARD: Spatial Analysis of Rural Development Measures”, funded by the European Community’s Seventh Framework Programme (FP7/2007-2013) under Grant Agreement No. 244944. Parts of this article, including most of the figures and a previous version of the regression analysis, are also available in the SPARD report Reinhard et al. (2013). The authors wish to thank the anonymous reviewers and project partners for their helpful comments.

Appendix A.

See Table A.1.

Table A.1 Examples of investments supported under the measure “farm modernization”.

<table>
<thead>
<tr>
<th>Thematic area</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction of new technologies and innovation</td>
<td>• Automated animal identification system</td>
</tr>
<tr>
<td></td>
<td>• Milk meter</td>
</tr>
<tr>
<td></td>
<td>• Farm business management/recording software</td>
</tr>
<tr>
<td></td>
<td>• Global Positioning System</td>
</tr>
<tr>
<td></td>
<td>• Electronic tag reader</td>
</tr>
<tr>
<td>Improved animal welfare and health</td>
<td>• Automated/robotic slurry scraping system</td>
</tr>
<tr>
<td></td>
<td>• Cow cubicle mats</td>
</tr>
<tr>
<td></td>
<td>• Rotary livestock scratching brush</td>
</tr>
<tr>
<td></td>
<td>• Mobile sheep shower</td>
</tr>
<tr>
<td>Increased hygiene control and product storage</td>
<td>• Vermicom post bulk feed bin</td>
</tr>
<tr>
<td></td>
<td>• Potato store ambient cooling ventilation system</td>
</tr>
<tr>
<td>Enhanced Occupational Safety and Business Efficiency</td>
<td>• Calving gate incorporating dead lock gate</td>
</tr>
<tr>
<td></td>
<td>• Weighing platform or load bars for cattle crush</td>
</tr>
<tr>
<td>Increased energy efficiency</td>
<td>• Electric/water heat pads for farrowing and weaner accommodation</td>
</tr>
<tr>
<td></td>
<td>• Solar panel water heating system</td>
</tr>
<tr>
<td></td>
<td>• Rainwater harvesting pre-fabricated covered tank with filter and pump</td>
</tr>
<tr>
<td>Enhanced environmental status</td>
<td>• Weather station for crop pest/disease monitoring</td>
</tr>
<tr>
<td></td>
<td>• Steam boiler for soil/compost sterilization</td>
</tr>
<tr>
<td></td>
<td>• Quad/ATV fertiliser sower</td>
</tr>
<tr>
<td>Source: Department of Agriculture and Rural Development (s.d.)</td>
<td></td>
</tr>
</tbody>
</table>

Appendix B.

See Table B.1.

Table B.1 (Continued)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of land that is utilized</td>
<td>21.0</td>
<td>17.8</td>
<td>0.0</td>
<td>76.3</td>
</tr>
<tr>
<td>agricultural land</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% of agricultural land in less favoured areas</td>
<td>14.2</td>
<td>16.4</td>
<td>0.0</td>
<td>63.5</td>
</tr>
<tr>
<td>% of surface that is pastures</td>
<td>13.4</td>
<td>13.9</td>
<td>0.0</td>
<td>59.7</td>
</tr>
<tr>
<td>% of surface that is vineyards</td>
<td>0.6</td>
<td>1.4</td>
<td>0.0</td>
<td>10.9</td>
</tr>
<tr>
<td>% of surface that is woodlands</td>
<td>5.7</td>
<td>9.0</td>
<td>0.0</td>
<td>69.9</td>
</tr>
<tr>
<td>% of surface that is flowers</td>
<td>0.1</td>
<td>0.3</td>
<td>0.0</td>
<td>3.6</td>
</tr>
<tr>
<td>Climate: mean minimum temp. in January</td>
<td>−2.3</td>
<td>4.2</td>
<td>−18.0</td>
<td>10.5</td>
</tr>
<tr>
<td>RDP spending per holding (axis 1)</td>
<td>3.2</td>
<td>6.1</td>
<td>0.0</td>
<td>53.9</td>
</tr>
<tr>
<td>RDP spending per holding (axis 2)</td>
<td>9.5</td>
<td>19.1</td>
<td>0.0</td>
<td>205.0</td>
</tr>
<tr>
<td>RDP spending per holding (axis 1) (1000 euros)</td>
<td>3.0</td>
<td>7.0</td>
<td>0.0</td>
<td>76.5</td>
</tr>
<tr>
<td>RDP spending per holding (other axes)</td>
<td>0.2</td>
<td>0.8</td>
<td>0.0</td>
<td>12.8</td>
</tr>
<tr>
<td>RDP spending per € GVA (axis 1)</td>
<td>0.3</td>
<td>0.4</td>
<td>0.0</td>
<td>32.6</td>
</tr>
<tr>
<td>RDP spending per € GVA (other axes)</td>
<td>0.1</td>
<td>0.7</td>
<td>0.0</td>
<td>10.3</td>
</tr>
</tbody>
</table>

Note: RDP data are totals over 2000–2010, in thousands of euros PPP. Farm size data are for 2005, due to missing values for 2000. Other data are for 2000, except where otherwise indicated. Note that the maximum found for investments is 3.8 times the total GVA in that region; this outlier does not significantly influence results.

References


Gershon, F., 1985. The relation between farm size and farm productivity: the role of family labor, supervision and credit constraints. J. Dev. Econ. 18 (2–3), 297–313.


