A DIFFERENT LOOK TO THE AGGLOMERATION EFFECTS IN SPAIN

David Martín-Barroso
e-mail: dmartin@ccee.ucm.es

Juan A. Núñez-Serrano
e-mail: janunezs@ccee.ucm.es

Francisco J. Velázquez-Angona
e-mail: javel@ccee.ucm.es

Departamento de Economía Aplicada II
Universidad Complutense de Madrid
Campus de Somosaguas
28223, Pozuelo de Alarcón
Madrid
Tel.: +34 913 942 635
Fax: +34 913 942 457

Abstract

This paper explores the relationship between productivity and labour density at a regional NUTs-4 level for the Spanish economy and year 2001. Previous results on the mentioned relationship are confirmed. Whilst agglomeration effects at NUTs-3 level were important along the 1960s and 1970s, they seem to have disappeared along the second half of the 1980s. We show that agglomeration effects are still present, nonetheless when analysed at a higher degree of regional disaggregation. Recent amendments in regional governance and the creation of *Comunidades Autónomas* along the 1980s may have had something to do with this change in agglomeration patterns. We follow Ciccone (2002) methodology nonetheless extending the analysis to revision of some of the shortcomings pointed out in this seminal paper. In one hand we include additional instruments for labour density in instrumental variable estimations. In the other, we use three different databases, (i) the one used by Ciccone, which considers total area and non agricultural economic activities, (ii) a more appropriate adjustment where non agricultural economic activities are solely associated to non agricultural area, hence using non agricultural area when agriculture is excluded from the analysis and, (iii) total area and the whole of the economic activity.

JEL Classification: R10

Agglomeration, labour productivity, IV estimation.

1. Introduction

A common result in the new economic geography (NEG) literature is that the size of cities and regions is positively correlated with corresponding wages and productivity, amongst an important number of other economic variables. Rosenthal and Strange (2004) conclude from their literature survey that labour productivity elasticities with respect to size range between 4 and 8 per cent. Ciccone and Hall (1996) seminal paper enhances interest for this field of the NEG literature as they suggest that the effects of geographical externalities don't necessarily take place only within a given region, but also between the region and the regions of the neighbourhood. They confirm further that agglomeration effects are more robustly captured when labour density instead of absolute regional size is used as a measure of spatial concentration in the economic activity dimension. Subsequent applications of this theoretical framework to regions and contexts different to the originally tested scenario, United States counties, have confirmed the mentioned relationship between regional productivity and the *density* of the economic activity. In this sense, Ciccone (2002) finds that the previously observed elasticity of 6 per cent for US counties reduces to 4.6 per cent in the case of some European NUTs-3 regions. Many other papers as Dekle and Eaton (1999) for Japanese prefectures (observing elasticities between 1 and 2 per cent), Rice, Venables and Patacchini (2006) for British NUTs-3 regions (3.5 per cent), Ottaviano and Pinelli (2006) for Finish NUTs-3 (positive elasticity), Braunerhjelm and Borgman (2004) for Swedish labour market regions (positive elasticity), Cingano and Schivardy (2004) for Italian labour market regions (6.7 per cent), and Combes, Duraton, Gobillon and Roux (2008) for French labour market regions (4.8 per cent), they all confirm the positive relationship between regional productivity and the density of the economic activity.

Results in the Spanish case may turn out inconsistent for certain time periods and given levels of regional disaggregation. Ciccone (2002), using data on NUTs-3 regions of five European countries, including Spain for year 1986, obtained an elasticity of productivity on agglomeration of 5.1%. The analysis to a broader time perspective carried out in Martínez-Galarraga et al. (2007) using data only for Spanish NUTs-3 regions, shows that agglomeration effects on productivity were important along the 1860 to 1980 horizon, but have basically disappeared from the 1980s. In fact, using data of The Spanish Statistical Institute (INE) for years 1986 and 2001 and NUTs-3 Spanish regions only, we obtain an elasticity of 3.4[1] per cent for year 1986, but a statistically equal to zero elasticity for year 2001, independently of the number of sample replications considered when estimating by frequency weights.

Some of these results are very interesting precisely because of their theoretical implications. If larger agglomeration implies higher productivity levels there is an incentive for enterprises to locate in agglomerated cities. As these new enterprises install, the density of economic activity increases, consequently raising productivity and generating an apparently attractive virtuous circle. Nonetheless, between 1986 and 2001, the standard deviation of labour density for NUTs-3 Spanish regions has only increased by just over 6 per cent, whilst corresponding standard deviation of regional productivity has decreased by more than 44 per cent.

---

[1] This elasticity is statistically significant at the 99 per cent significance level when estimating by frequency weights and replicating the sample 5 times.

Of course, this virtual circle is broken when congestion problems emerge in an over-agglomerated scenario. Thus, one possible explanation behind the different results obtained in the Spanish economy could rely under the distinct stages in seeking or trying to achieve an optimal level of agglomeration. Spain is nevertheless a country with a low population density, 91.22 habitants per squared kilometre (INE data at 01/01/2008). It is the second European Union (EU-27) country by size, the fifth in terms of population, and in contrast, the 19th in terms of population density, with a value well below the EU-27 mean of 114.80 (Eurostat data for year 2006). Nonetheless, as shown in Viladecans (2004), agglomeration economies play an important role in the location processes of Spanish manufacturing firms.

Comparing the procedures in Viladecans (2004) and Martínez-Galarraga et al (2007), whilst the former uses data on large municipalities (NUTs-4), the later is carried out for Provinces (NUTs-3). Thus a reliable explanation could be related to the regional level at which agglomeration processes take place. Additionally, *Comunidades Autónomas* (NUTs-2) were created in the 1980s, implying a higher level of self-governance at this regional level. Along the 1980s, large amount of effort was dedicated to avoid those inter-regional migration movements that had been so important along the 1950s, the 1960s, and the first half of the 1970s. Thus these political measures could have succeeded in holding inter regional movements back, and hence blocking agglomeration processes at NUTs-3 level. The question we raise at this point is obvious; could agglomeration effects still be present, nonetheless at a lower geographical level? In fact, Audretsch and Feldman (1996) and Ciccone and Hall (1996) point out that the geographical level at which agglomeration phenomena is studied is relevant, suggesting the use of a fine level of geographical disaggregation. To this respect, this article is the first to use data at local NUTs-4 level covering the complete set of Spanish municipalities[2].

The aim of this paper is then to explore agglomeration effects at municipality level, considering the complete set of Spanish municipalities in year 2001, using basically the principles and methodology proposed by Ciccone. The paper is organised as follows. Next section summarises Ciccone's theoretical model. We then describe the municipal database used for the analysis. The empirical models and estimation procedures to capture (i) agglomeration effects on productivity, (ii) the possible influence of neighbours' agglomeration on own productivity, and (iii) neighbours' productivity spillovers are described in third place. Results are discussed just before finishing the article off with conclusions and final remarks.

2. Model

2.1 Agglomeration effects on labour productivity

Three basic regression models constitute the basis of our whole analysis. Regression model 1 (R.1) estimates agglomeration effects θ, of employment density d_i, on labour productivity y_i, conditional on five different human capital levels.

$$\ln y_i = \alpha + \theta \ln d_i + \sum_{l=1}^{5} \delta_l \ln(1 + \frac{hk_{l,i}}{HK_i}) + \varepsilon_i \qquad \text{(R.1)}$$

[2] Viladecans (2004) uses this same level of geographical disaggregation nonetheless considering just municipalities of more than 15,000 inhabitants for only 14 of the 17 different Spanish NUTs-2 regions.

Human capital is expressed as the percentage of workers with education level l, thus $hk_{l,i}$ is the number of workers with education level l in municipality i, and HK_i the total number of workers in that same municipality.

2.2 Neighbouring agglomeration effects

Regression model (R.2) goes a step further and includes additional regressors capturing agglomeration effects across ten different neighbouring areas.

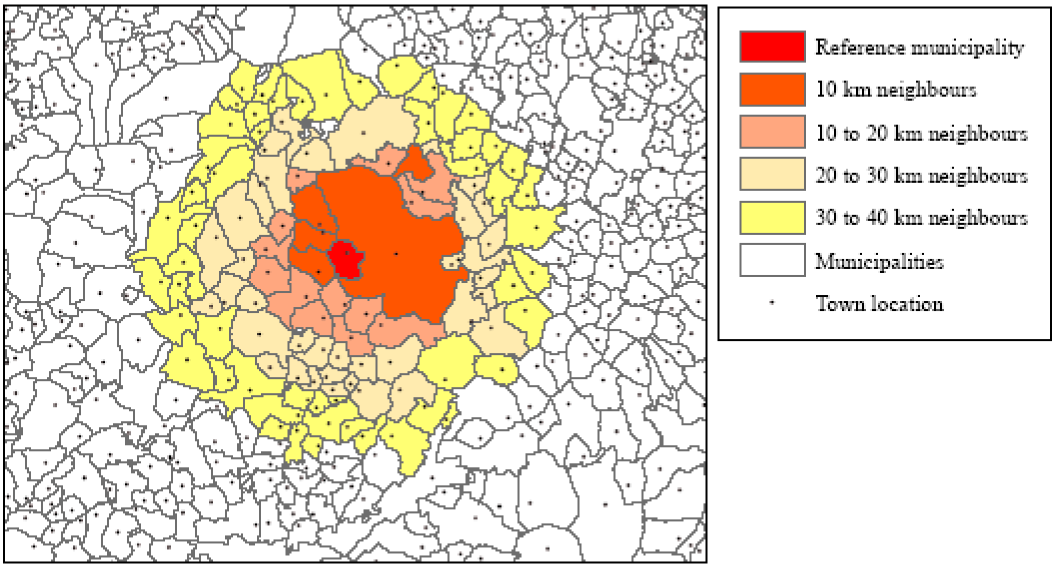
$$\ln y_i = \alpha + \sum_{j=1}^{10} \varpi_j \ln d_{j,i} + \theta \ln d_i + \sum_{l=1}^{5} \delta_l \ln(1 + \frac{hk_{l,i}}{HK_i}) + \varepsilon_i \qquad \text{(R.2)}$$

Thus $d_{j,i}$ is the average employment density of the different municipalities located around the neighbourhood of municipality i along area j. Neighbouring area $d_{1,i}$ includes all municipalities except municipality i, whose distances to i are at most 10 kilometres away, distances being calculated between town centres using the Great Circle Distance formula[3]. Neighbouring areas 2 to 10 are constructed in a slightly different manner, they include all municipalities whose distances to i are less or equal to $10j$ kilometres, and greater than $10(j - 1)$ kilometres, for $j = 2,\ldots,10$. Figure 1 illustrates the different irregular crowns that form areas $d_{1,i}$ to $d_{4,i}$ around the neighbourhood of municipality i.

Hence, regression model (R.2) estimates the elasticity of productivity with respect to employment density conditional on neighbouring agglomeration effects and human capital endowments.

[3] The Great Circle Distance formula gives the shortest distance between any two points on the surface of a sphere, measured along the closest path over the surface, as opposed to going through the sphere's interior. All distances in this paper are calculated this way.

Figure 1. Neighbouring areas of a given reference municipality. Illustration of irregular crown formation



Source: Own elaboration from Spanish INE and National Geographical Institute (IGN) geographical data.

2.3 Productivity spillovers across regions

Regression model (R.3) excludes labour density in neighbouring areas and instead conditions agglomeration elasticity to the existence of productivity spillovers across regions. To capture these productivity spillovers at the municipality level we use the maximum value for labour productivity, max $y_{d,i}$, observed in the area formed by the complete set of municipalities at a distance from municipality *i* no greater than *d*, determining one area for each different municipality, and distances *d* = 10,…,500, in 10 km intervals, hence resulting in 50 different regressions. The idea here is that the distance to locations with higher levels of productivity may have some influence over municipal labour productivity.

$$\ln y_i = \alpha + \varpi_d \ln(\max y_{d,i}) + \theta \ln d_i + \sum_{l=1}^{5} \delta_l \ln(1 + \frac{hk_{l,i}}{HK_i}) + \varepsilon_i \qquad \text{(R.3)}$$

Each of these models is extended to the inclusion of regional indicators at NUTs-2 level, *Comunidades Autónomas*, in one hand, and NUTs-3 level, *provincias*, in the other. By allowing the constant term α in Regressions (R.1) to (R.3) to vary across regions we try to capture differences in average total factor productivity between regions. These regional indicators could also be denoting differences in institutional settings due to the existence of a high degree of economic and political autonomy at the regional level, specially at NUTs-2 level. Furthermore, estimations of agglomeration effects take also into account the possibility of failure of regional fixed effects in accomplishing the requested task. If fixed effects do not entirely pick up exogenous differences in total factor productivity across regions, estimates may turn to be inconsistent due to endogeneity problems, i.e. regions with higher productivity levels will be attracting more labour and hence becoming more employment dense. We thus estimate each regression model first by ordinary least squares, OLS, and then using the 2-stage-least-squares, 2SLS, estimator. We try a complete set of instruments for employment density, d_i, and average neighbouring employment density, $d_{j,i}$.

3. Data

Spain has very rich statistical regional information. Main economic variables are available by *Comunidades Autónomas* (NUTs-2), and in some cases, the statistics are also published at *Provincias* level (NUTs-3). Unfortunately there is no such datasets at municipality level, there is only data for large cities (more than 15,000 inhabitants) and not all regions are complete, hence we estimate data for this level of regional disaggregation.

SABI database is used for these purposes. This dataset is the Spanish branch of AMADEUS family of databases and is generated by the private firms INFORMA and Bureau Van Dyck. This database contains accounts and useful information for enterprises. The main problem is the lack of sample representativeness in both, the sector and region dimensions. In fact, there are no rules for the inclusion of enterprises in the dataset. For this reason we calculate expansion coefficients for each enterprise considering its headquarters regional location as well as the type of performed economic activity, i.e. the industry or sector of economic activity in which the firm operates.

The National Institute of Statistics (INE) provides Spanish Regional Accounts with data on value added and employment at NUTs-2 level and industry classification NACE A-31 classification, let us refer to it as INE-2, as well as at NUTs-3 level and NACE A-6, let us call this data INE-3. Expansion coefficients are thus built following a two-stage mechanism.

We exclude enterprises with negative or null value added and also those firms with no information on the employment variable. Value added is calculated for each enterprise as the difference between operating revenues and intermediate consumption plus other operating expenses, excluding labour costs.

We then obtain value added and employment for the sample of valid SABI enterprises at NUTs-2 level and A-31 industry classification, let us call this data SABI-2. The initial expansion coefficient ($e1$) is thus calculated as the ratio between the value given by the universe i.e. that given by Regional Accounts (INE-2), and the sample value calculated from aggregation (SABI-2). Thus this expansion coefficient is calculated for a firm i belonging to A-31 sector s, and located in region NUTs-2 R, following expression (1).

$$e_{1,s,R} = \frac{\text{INE-2}_{s,R}}{\sum_{i \in s,R} SABI_{i,s,R}} \tag{1}$$

Multiplying original SABI data by this expansion coefficient and aggregating resulting information to NUTs-3 and A-6 levels (let's refer to this as SABI-3), allows calculation of a second coefficient (*e*2) by simply dividing Regional Accounts INE-3 data by expanded SABI-3 data. This is, for all the firms operating in A-6 sector *S*, note that sector *s* belongs to sector *S*, that are located in NUTs-3 region *r*, where *r* is located within NUTs-2 region *R*, the expansion coefficient is thus calculated by expression (2).

$$e_{2,S,r} = \frac{\text{INE-3}_{S,r}}{\sum_{i \in S} \sum_{i \in r} e_{1,s,R} \cdot \text{SABI}_{i,S,r}}, s \in S \text{ and } r \in R \tag{2}$$

The final expansion coefficient (*e*) is obtained by multiplying *e*1 by *e*2. Thus the municipal dataset is consequently built by expanding original and valid values of SABI microdata with expansion coefficients e. The nature of original microdata obliges to assume that firms are solely located on headquarters and produce in the declared main sector of economic activity. In the case of Spain, multiplant firms are only a small proportion of total firms and hence this assumption is not especially restrictive. We end up having two different datasets on value added and employment observed at NUTs-4, one including the agricultural sector and the other excluding it.

Data on human capital comes from 2001 Spanish Population Census. These statistics have information at NUTs-4 level and are available for five different education levels, nonetheless they are based on resident population and not on workers.

Area is obtained from the Spanish National Institute of Statistics. To obtain the non-agricultural surface, we use data from the 1999 Agricultural Census. Nevertheless, this information is obtained from a survey to owners and agricultural entrepreneurs. For this reason, agricultural surface is assigned to the municipality where farmer lives, inducing an important bias in measuring agricultural area. We consequently introduce a procedure based on calculation of a given radius of influence around each municipality enough to correct biasness. A coefficient is calculated by dividing declared agricultural area of each municipality and its surrounding neighbours located at a maximum distance of 50 km, by the corresponding total area. Agricultural area is subsequently generated by multiplying this resulting coefficient by declared agricultural area.

Table A1 in the Appendix presents summary statistics for the complete set of main variables used in the estimations of proposed empirical models described in next section. Statistics are calculated for the three data sets and the sample which turns to be valid when running regressions. Municipalities are removed from the sample when they have no operating firms as recorded by SABI. There are 8,110 municipalities in Spain, from which 2,043 have no private economic activity and 228 have only agricultural firms. We only consider private non-proprietary firms as in Ciccone and Hall (1996). Estimation results are presented and discussed along next section.

4. Agglomeration Effects on Productivity: Some Results on Municipal Data

4.1 Ciccone's results at municipality level

The results from estimation of model (R.1) over the dataset that includes total area and excludes agriculture and forestry sectors are registered in Table 1. The first column presents OLS results. The next five correspond to 2SLS regression results where employment density has been instrumented by (a) area, (b) average area of neighbour municiplatites in a 5 km radius, (c) average neighbouring area of neighbours (again, neighbours correspond to municipalities whose distances to reference municipality are less or equal to 5 km), (d) elevation, and (e) 2 period lagged employment density. Moomaw (1981) is the first one to document the simultaneity problems associated to the estimation of agglomeration economies. Ciccone and Hall (1996) introduce the idea of using as instruments for labour/population densities their past values, based on the strong persistence of population's spatial distribution. Additionally, Combes, Duraton, Gobillon and Roux (2008) assert that geological aspects are important determinants of settlement patterns, and therefore they recommend the nature of soils as a relevant variable to explain actual labour distributions.

Table 1. Agglomeration effects with human capital controls and regional indicators. Without agriculture and total area dataset

	Regional Indicator	OLS	(a) 2SLS	(b) 2SLS	(c) 2SLS	(d) 2SLS	(e) 2SLS
Parameter θ (%)		5.04***	-3.82**	6.13**	-.83	4.43***	5.53***
Standard error of θ (%)		.46	1.17	2.20	1.29	.65	.46
R^2 (%)		8.91	1.09	7.96	5.20	8.87	9.53
ΔR^2 (%)	*No regional indicators*	-	4.45	4.50	6.44	13.15	69.18
Parameter θ (%)		5.12***	-5.5***	5.73*	-3.21†	3.06***	5.67***
Standard error of θ (%)		.49	1.40	2.86	1.67	.81	.49
R^2 (%)		10.28	.42	10.00	4.65	9.91	10.87
Wald test (*ccaa = 0*)		7.07***	9.33***	6.80***	7.16***	7.44***	6.72***
Wald test (*ccaa ≠ pro = 0*)	*NUTs-2*	6.62***	10.64***	6.41***	7.14***	7.14***	6.34***
Parameter θ (%)		5.90***	-4.07**	9.62**	-2.39	5.13***	6.56***
Standard error of θ (%)		.52	1.57	3.61	1.98	.86	.52
R^2 (%)		12.63	4.54	11.28	7.96	12.58	13.60
Wald test (*pro = 0*)	*NUTs-3*	85.1***	69.59***	4.55***	4.41***	81.55***	89.08***

*** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$.

Instruments: (a) municipality area, (b) average area of neighbours, (c) average neighbouring area of neighbours, (d) elevation, and (e) lagged density.

In order to evaluate the quality of the different instruments we run the OLS regression municipality labour density as a function of NUTs-3 indicators, and then, one additional regression for each instrument which simply adds to the former set of regressors the log of the variable used as an instrument for labour density. We register the gain in R^2, ΔR^2 row in Table 1, associated to this extended regression. In this respect, municipality area as instrument of labour density presents the minimum gain, 4.45 per cent, and the 2 period lagged employment density, the maximum, with a 69.18 per cent gain. Area is the instrument chosen by Ciccone in his NUTs-3 analysis, resulting in a positive agglomeration effect on labour productivity. The analysis at NUTs-4 level shows that municipality area cannot be a good instrument for labour density as it predicts a negative relationship between productivity and agglomeration, with elasticities going from -5.5 to -3.8 per cent. Another instrument that should be discarded is average neighbouring area of neighbours, which only gives a statistically significant elasticity of -3.2 at the 90 per cent significance level, when NUTs-2 indicators are included. Average neighbouring area provides statistically significant positive values for the elasticity of labour density on labour productivity, and in principle, these values are higher than those reported by OLS regressions. The explanatory power of this instrument is nonetheless low, with an R^2 gain of just 4.5 per cent. Moreover, the standard errors of estimated θ parameters are high, oscillating between 2.2 to 3.6 per cent. More convincing results are found with the elevation instrument. This variable is correlated to labour density —a 13.15 per cent gain in explanatory power is achieved when regressing productivity on elevation in addition to NUTs-3 indicators— and it should not be related to exogenous total factor productivity. The elasticities are always statistically significant at the 99.9 per cent significance level, with values in the range

3.1 —when NUTs-2 indicators are included in regression—to 5.1 per cent —when regional differences in total factor productivity are captured by NUTs-3 indicators—, and associated standard errors are in all cases well below 1 per cent. The remaining instrument, the 2 period lagged employment density, offers most promising results. These results should nonetheless be taken with caution. Endogeneity problems may have not been removed by just considering employment of year 1999.

Focusing now on OLS results, the elasticities of labour productivity with respect to employment density are always statistically significant at the 99.9 per cent significance level, and go from 5.04 per cent with a robust standard error of .46 per cent, when agglomeration effects are estimated conditional on human capital levels, to 5.90 per cent and a robust standard error of .52 per cent, when conditioning is augmented to the inclusion of NUTs-3 fixed effects. These values are very similar to those obtained by Ciccone (2002) for Spain (5.1 per cent) in year 1986 for NUTs-3 regions. The elasticity of productivity with respect to agglomeration increases when introducing regional fixed effects, attaining higher values when NUTs-3 indicators are included. Thus agglomeration effects are slightly higher when controlling for exogenous differences in total factor productivity across NUTs-3 regions, indicating that regional idiosyncratic factors may to some minor extent limit the agglomeration effects on labour productivity. Goodness of fit oscillates between 8.9 per cent and 12.6, and regional indicators are jointly significant at the 99.9 per cent significance level, even those representing multi-provincial NUTs-2 regions.

Thus using the Ciccone equivalent dataset, nonetheless for year 2001, NUTs-4 level, and using elevation as a valid instrument for labour density, the elasticity of agglomeration on labour productivity is along the range 3.06 (.81) to 5.90 (.52) at the 99.9 per cent significance level, standard errors in parenthesis, depending on the estimation method and the inclusion of regional fixed effects.

Next we turn to measure agglomeration effects nonetheless considering non agricultural land instead of total land, and only keeping 2SLS results for the elevation instrument.

4.2 Agglomeration excluding non agricultural land

Agglomeration effects are as expected, slightly higher when considering only non agricultural land when calculating area and hence labour density. Agriculture and forestry are much more land use intense than manufacturing and services, and their weight in total economic activity is limited. Main results are shown in Table 2.

Table 2. Agglomeration effects with human capital controls and regional indicators. Without agriculture and non agricultural land dataset

	Regional Indicator	**OLS**	**Elevation 2SLS**
Parameter θ (%)		5.55***	5.65***
Standard error of θ (%)		.49	.83
R^2 (%)		9.20	9.20
ΔR^2 (%)	*No regional indicators*	-	12.41
Parameter θ (%)		5.35***	3.46***
Standard error of θ (%)		.50	.92
R^2 (%)		10.41	10.10
Wald test (*ccaa = 0*)		6.89***	7.33***
Wald test (*ccaa ≠ pro = 0*)	*NUTs-2*	7.01***	7.64***
Parameter θ (%)		5.89***	5.66***
Standard error of θ (%)		.53	.95
R^2 (%)		12.56	12.56
Wald test (*pro = 0*)	*NUTs-3*	93.34***	93.27***

*** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$.

Independently of the estimation method, elasticities of productivity with respect to labour density attain maximum levels when controlling for provincial fixed effects, whilst minimum values are observed when considering NUTs-2 indicators. Parameter values oscillate between 3.46 (.92) to 5.89 (.53), with associated standard errors in parenthesis, at the 99.9 per cent significance level. Goodness of fit slightly increases and ranges from 9.2 to 12.6 per cent. Regional indicators are as usual statistically significant at the 99.9 per cent significance level.

The lowest agglomeration effects are observed when replicating estimations considering total economic activity and total land (see Table A.2 in the Appendix). We now turn to extensions (R.2) and (R.3), focusing on estimation results over the data set that excludes non agricultural land. The results for the remaining 2 data sets can be consulted in the Appendix through Tables A.3 to A.4

5. Agglomeration Effects across Neighbouring Municipalities

In this section we estimate θ conditioned on the possible presence of neighbours' agglomeration effects. Further, the inclusion of neighbours' labour densities allows quantification of the geographical magnitude of agglomeration economies. Instrumental variable estimations are tried for a large number of instrument combinations. In one side we consider elevation and lagged employment as instruments for labour density, and in the other, neighbouring labour densities are instrumented by its 2 period lagged values and the average area of neighbouring areas $d_{i,j}$, Estimation results of regression model (R.2) are thus presented in Table 3.

Table 3. Agglomeration effects with human capital controls, neighbouring agglomeration and regional indicators

	Regional Indicator	OLS	(a) 2SLS	(b) 2SLS	(c) 2SLS
Parameter (%)		5.28***	6.14***	8.50*	5.33**
Standard error (%)		.59	.59	3.42	1.70
Parameter 70 (%)		1.72†	1.71†	-	1.73†
Standard error (%)	*No regional*	.98	.97	-	.99
R^2 (%)	*indicators*	9.60	10.21	-	9.63
Parameter (%)		5.27***	6.13***	6.94*	-
Standard error (%)		.60	.60	2.72	-
R^2 (%)		10.73	11.35	-	-
Wald test (*ccaa = 0*)		6.92***	6.21***	1.21	-
Wald test (*ccaa ≠ pro = 0*)	*NUTs-2*	7.32***	6.84***	1.32	-
Parameter (%)		5.46***	6.33***	-	3.99**
Standard error (%)		.61	.60	-	1.55
R^2 (%)		12.95	13.99	-	12.88
Wald test (*pro = 0*)	*NUTs-3*	5.50***	5.83***	-	5.43***

*** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$.
(a) Labour density and average labour density across different neighbouring areas are instrumented with the 2 period lagged values, (b) labour density is instrumented with its 2 period lagged values, and density across neighbours with the average area observed along each considered irregular crown, (c) labour density is instrumented with the elevation variable, and density across neighbours with its 2 period lagged values.

The consideration of average labour density for different neighbouring areas does not practically affect municipal agglomeration economies. OLS results are in fact very similar to those presented in Table 2. The elasticity of productivity with respect to labour density oscillates between 5.27 (.60) to 5.46 (.61) per cent, and the explanatory power of estimated regressions ranges from 9.6 to 13.0 per cent. Only in the absence of regional fixed effects, the average labour density of the area formed by municipalities situated more than 60 and at most 70 km far apart has a statistically significant elasticity at the 90 per cent significance level of 1.72 per cent with associated robust standard error of .98 per cent. This may be capturing the effect of the main cities and their close neighbourhood over municipal labour productivity. The result tells that if average productivity of neighbouring area within the 60 to 70 km distance doubles, municipal productivity increases by 1.7 per cent. The average radius of Spanish provinces (NUTs-

3) excluding Ceuta and Melilla is around 60 km. This fact together with the disappearance of neighbouring agglomeration effects when introducing regional fixed effects reinforces the idea that neighbouring agglomeration effects as here defined are somehow capturing the positive correlation between the largest municipalities of each province, e.g. the provincial capital effect.

In terms of instrumental variable estimation results, things are not as straightforward as in previous section. It is definitely harder to find a right combination of instruments for labour density and average neighbouring labour densities. When the 2 period lagged values are used for both variables, elasticities turn out just slightly higher and the same conclusions as reported for OLS follow here. Average area of neighbours may not be a good instrument for average neighbouring labour densities. Elasticities increase substantially as well as their corresponding standard errors. Furthermore, regional indicators can only be included at NUTs-2 level, and they end up being equal to zero. Elevation again offers some neater results when average neighbouring labour densities is instrumented by their lagged values. The elasticity of labour productivity with respect to the agglomeration variable is of 5.33 per cent, with standard error of 1.70 per cent. Neighbouring area between 60 and 70 km again shows a statistically significant elasticity of 1.73 (.99) per cent at the 90 per cent significance level, which vanishes off when statistically significant NUTs-3 level indicator are introduced to capture regional differences in total factor productivity. In this last case, θ elasticity is just below 4 per cent with a standard error of 1.55 per cent.

The results for the remaining data sets are registered in Table A.3 of the Appendix. Without agriculture and total land dataset offers elasticities ranging from 3.24 (1.44) to

6.32 (.60) per cent, and neighbouring agglomeration has no effect over municipality productivity. Consideration of all sectors of economic activity and total land brings in contrast statistically significant positive effects of agglomeration across neighbours within a 10 km radius, over municipality productivity. These elasticities range from 1.49 (.68) to 1.74 (.68) per cent and vanish off when estimating by 2SLS. Another positive externality emerges at neighbouring area d_{10}, municipalities at more than 90 and at most 100 km away, with values ranging from 1.73 (1.04) (OLS with NUTs-3 indicators) to 2.23 (.94) per cent, when no regional indicators are included and labour density is instrumented with elevation variable and neighbouring agglomeration with its lagged values. A negative externality across neighbours in d_5 of -1.77 (1.03) per cent appears just in the basic OLS estimation with no regional indicators. The results associated to this dataset reflect the peculiarities of the agrarian sector, often located along rural areas formed by small municipalities with low agglomeration levels.

To have a better idea of the extent and importance of agglomeration effects and their impact on labour productivity, Table 4 provides average values for the appropriate θ elasticities estimated up to now in one hand, and in the other, the median of labour densities along the four different quartiles and their proportional change from one quartile to the next one. The numbers in bold represent the expected gain in labour productivity associated to the registered increase in densities once the corresponding elasticities are applied. The productivity gains range from 13 per cent —when shifting from the median of the first quartile to that of the second one, in the with agriculture dataset—, to more than 52 per cent, corresponding to the change in labour density from the third to the forth quartile median in the without agriculture and total area dataset.

Table 4. Productivity gains associated to density increases

Average		Percentile			
θ	Variable	12.5	37.5	62.5	87.5
	Without agriculture and non agricultural land dataset				
	Labour density[4]	1.90	9.70	41.52	340.21
	Proportional change in labour density	-	4.11	3.28	7.19
5.26 (a)	Increase in labour productivity (%)	-	**21.59**	**17.25**	**37.84**
5.39 (b)	Increase in labour productivity (%)	-	**22.14**	**17.69**	**38.79**
	With agriculture dataset				
	Labour density	.33	1.78	7.84	75.05
	Proportional change in labour density	-	4.39	3.40	8.57
3.99 (c)	Increase in labour productivity (%)	-	**17.54**	**13.59**	**34.22**
3.93 (d)	Increase in labour productivity (%)	-	**17.25**	**13.37**	**33.66**
	Without agriculture and total area dataset				
	Labour density	.28	1.42	6.50	71.57
	Proportional change in labour density	-	4.07	3.58	10.01
5.16 (e)	Increase in labour productivity (%)	-	**21.01**	**18.46**	**51.66**
5.21 (f)	Increase in labour productivity (%)	-	**21.20**	**18.62**	**52.11**

Elasticities are calculated taking the mean of the θ values in: (a) Table 2; (b) Table 2 and Table 3: columns OLS, a and c; (c) Table A.2; (d) Table A.2 and Table A.3; (e) Table 1: columns OLS, d and e; (f) Table 1: columns OLS, d and e, and Table A.3.

We now turn to analyse if labour productivity is affected by the presence of efficient municipalities in some given neighbourhood.

6. Productivity spillovers across regions

Results to estimations of (R.3) set of regressions are presented in Table 5, for the without agriculture dataset, and Table A4 in the Appendix, for remaining two datasets. Figure 2 illustrates the complete group of statistically significant values at the 95 per cent significance level of θ and ω_d elasticities, as well as corresponding R^2 coefficients of determination, along the *y*-axis. The radius in km of influential neighbourhood area is registered along the *x*-axis.

[4] Labour density is measured in workers per squared kilometre. We only consider valid data for regressions, i.e. those with strictly positive employment.

The elasticity of labour productivity with respect to labour density seems not to be affected when conditioned to the presence of productivity spillovers across neighbours.

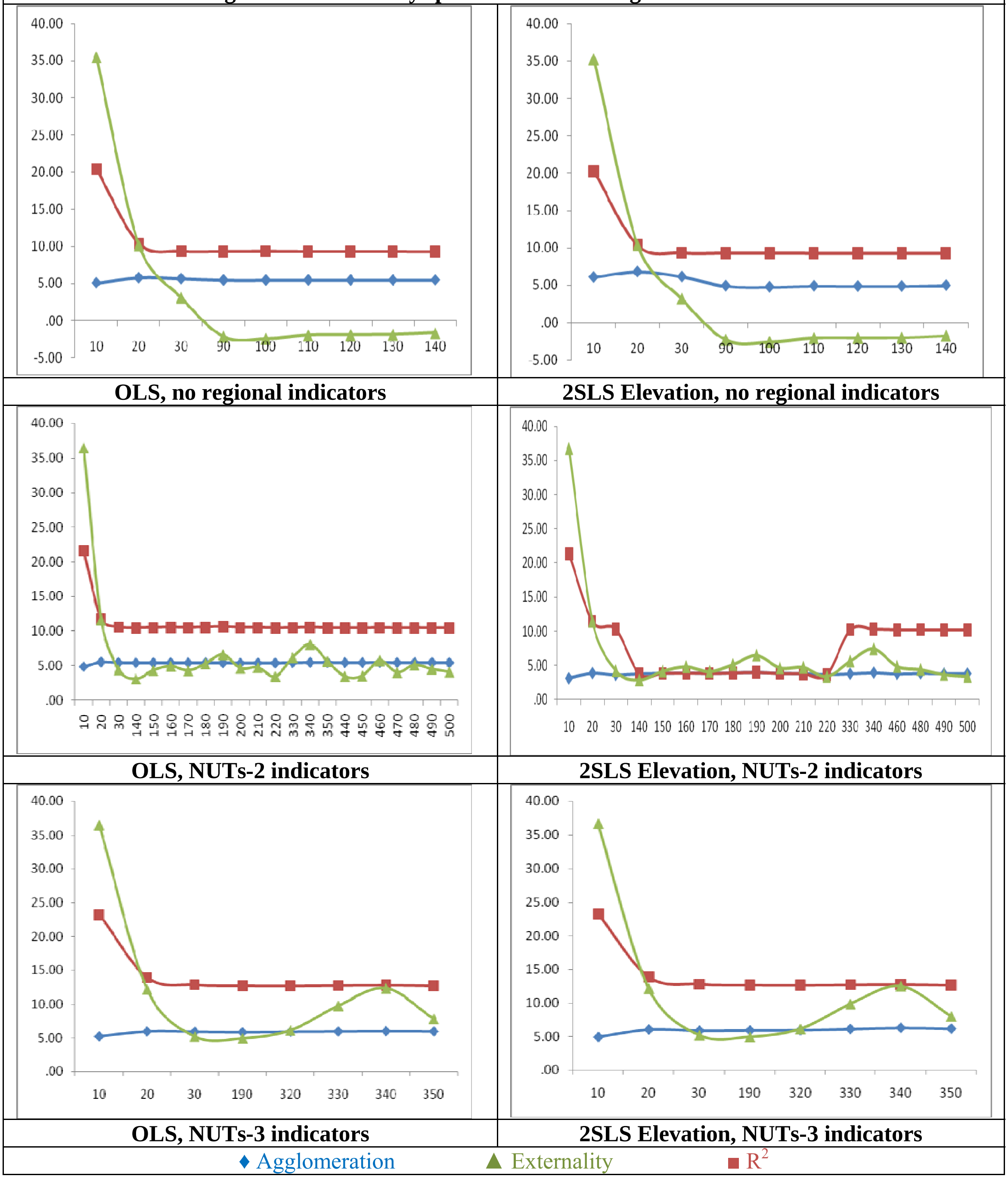
Table 5. Spillover effects of neighbours. Without agriculture dataset

	Regional Indicator	OLS	Elevation 2SLS
Parameter (%)	*NO*	5.06***-5.78***	4.74***-6.81***
Standard error (%)		.45-.48	.85-.82
+ve externalities		3.08*-35.45***	3.19*-35.30***
-ve externalities		-2.46**--1.66†	-2.59**--1.77*
R^2 (%)		9.25-20.37	9.23-20.27
Parameter (%)	*NUTs-2*	4.84***-5.50***	3.11***-4.03***
Standard error (%)		.46-.50	.87-.92
+ve externalities		3.08†-36.53***	2.84†-36.77***
-ve externalities			
R^2 (%)		10.44-21.64	10.22-21.38
Wald test	*ccaa = 0*	7.02***-9.22***	7.4***-9.53***
Wald test	*ccaa != pro = 0*	6.55***-10.78***	7.08***-11.44***
Parameter (%)	*NUTs-3*	5.25***-6.03***	4.92***-6.27***
Standard error (%)		.48-.53	.90-.99
+ve externalities		4.93*-36.54***	4.93*-36.60***
-ve externalities			
R^2 (%)		12.62-23.32	12.62-23.31
Wald test	*pro = 0*	21.87***-92.15***	19.93***-92.33***

*** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$.

Values are very close to those obtained for (R.1) and (R.2) family of regressions. Extremes are both found in 2SLS estimations, elevation being the instrument for the agglomeration variable, and range from 3.11 (.87) to 6.81 (.82) per cent. OLS results are more centred on the 5 per cent level, oscillating between 4.84 (.46) and 6.03 (.53). Strong positive externalities occur along a close neighbourhood of at most 10 km radius, with elasticities that range from 35.30 to 36.77 per cent. Thus municipal productivity benefits substantially from high level productivity close neighbours. Here we are probably capturing the influence of metropolitan areas. Some negative externalities appear along the different neighbourhood areas of radius 90 to 140 km, with elasticities ranging from -2.59 to -1.66. These negative externalities disappear nonetheless as soon as regional indicators are included in regressions. Positive externalities emerge when total factor productivity regional differences are captured with regional indicators, for distances as far as 350 km when NUTs-3 indicators, and 500 km if NUTs-2, with elasticities that wonder around the neighbourhood of corresponding θ values, see Figure 2. In fact, these distances are somehow representative of the distances between main metropolitan areas is Spain[5].

Figure 2. Productivity spillovers. Without agriculture dataset



---

[5] Some examples of these distances in kilometres, calculated by the Great Circle Distance formula, are: Madrid-Barcelona: 504, Madrid-Málaga: 419, Barcelona-Alicante: 408, Madrid-Sevilla: 394, Madrid-Alicante: 360, Madrid-Bilbao: 321, Barcelona-Valencia: 304, Madrid-Valencia: 304.

Regional indicators are always and jointly statistically significant at the 99.9 per cent significance level and R^2 coefficients range from 9.25 to 23.32 per cent.

The same general patterns are observed for remaining two datasets, where the conditional elasticities of labour productivity with respect to agglomeration are slightly lower in both cases as expected, and in the same order as commented in section 4, being the agriculture dataset the one with lowest elasticity values.

7. Conclusions

The analysis confirms that agglomeration processes seem to no longer have any effect on labour productivity from the second half of the 1980s, when studied at NUTs-3 level. This change may possibly be due to the conformation of *Comunidades Autónomas*, Spanish NUTs-2 regions. Nonetheless, agglomeration processes respond mainly to economic factors and hence they must be still taking place at a lower level of geographical disaggregation. Results corroborate the existence of agglomeration effects —agglomeration being measured by labour density— on labour productivity at the local NUTs-4 level, with elasticities slightly over 5 per cent, in consonance with the results obtained in Ciccone (2002) at NUTs-3 level and year 1986. A positive effect of neighbouring agglomeration is also captured, in particular within the influential area of provinces, most probably signalling the provincial capital effect and the interrelations of those large towns leading the agglomeration processes of the economic activity. Productivity spillovers occur at different geographical dimensions, in one hand, along a very close neighbourhood of each municipality —most probably indicating the strong economic links amongst those local towns that integrate large metropolitan areas— and in the other, some positive externalities arise across large distances, suggesting that agglomeration processes are at least, national wide phenomena. To conclude, obtained results reveal the importance of working at the appropriate level of geographical disaggregation, which turns out to be crucial to properly identify actual agglomeration effects in the Spanish economy.

References

Audretsch, D.B and Feldman, M.P (1996): "R&D Spillovers and the Geography of Innovation and Production," American Economic Review, American Economic Association, 86 (3), 630-40, June.

Braunerhjelm, P. and Borgman, B. (2004): "Geographical Concentration, Entrepreneurship and Regional Growth: Evidence from Regional Data in Sweden, 1975-99," Regional Studies, Taylor and Francis Journals, 38 (8), 929-947, January.

Ciccone, A. (2002): “Agglomeration effects in Europe”, European Economic Review, 46, 213-227.

Ciccone, A. and Hall, R.E. (1996): “Productivity and the density of economic activity”, American Economic Review, 86, 54-70.

Cingano, F. and Schivardi, F. (2004): "Identifying the Sources of Local Productivity Growth," Journal of the European Economic Association, MIT Press, 2 (4), 720-742, 06.

Combes, P.P., Duranton, G., Gobillon, L. and Roux, S. (2008): "Estimating Agglomeration Economies with History, Geology, and Worker Effects," CEPR Discussion Papers 6728, C.E.P.R. Discussion Papers.

Dekle, R. and Eaton, J. (1999):"Agglomeration and Land Rents: Evidence from the Prefectures," Journal of Urban Economics, 46 (2), 200-214, September.

Moomaw, R.L (1981): "Productivity and City Size? A Critique of the Evidence", The Quarterly Journal of Economics, MIT Press, 96 (4), 675-88, November.

Ottaviano, G.I.P. and Pinelli, D. (2006): "Market potential and productivity: Evidence from Finnish regions," Regional Science and Urban Economics, Elsevier, 36 (5), 636-657, September.

Rice, P., Venables, A. J. and Patacchini, E. (2006): "Spatial determinants of productivity: Analysis for the regions of Great Britain," Regional Science and Urban Economics, 36 (6), 727-752, November.

Rosenthal, S. S. and Strange, W. C. (2004): “Evidence on the Nature and Sources of Agglomeration Economies”, in Henderson, J.V. and J.-F. Thisse, eds., *Handbook of Urban and Regional Economics*, Volume 4. Amsterdam: Elsevier, 2119-2171.

Viladecans-Marsal, E. (2004): “Agglomeration Economies and Industrial Location: City-Level Evidence”, Journal of Economic Geography, 4(5), 565-582.

Martínez-Galarraga, Paluzie, Pons and Tirado (2007): “Agglomeration and labour productivity in Spanish industry: a long-term analysis”, Working Papers in Economics, 175, Universitat de Barcelona.

APPENDIX

Table A1. Summary statistics

Variable	Data	Obs	Mean	Std. Dev.	Min	Max
Municipalities		8,110	-	-	1	8,110
Comunidad Autónoma		8,110	-	-	1	18
Provincia		8,110	-	-	1	52
Value added[6]	Without agriculture and non agricultural area	5,839	82,402.09	1,078,722	.00	67,700,000
Employment[7]		5,839	2,192.99	23,796.48	.32	1,468,000
Area[8]		5,839	12.15	17.90	.01	376.11
Labour productivity[9]		5,839	29,560.43	48,053.70	.00	2,206,340
Labour density[10]		5,839	345.59	3,861.25	.02	252,177.40
Elevation[11]		5,839	529.86	334.88	2.00	1,692.00
Average neighbouring area		5,839	4.33	6.82	.00	144.39
Lagged labour density		5,839	307.67	3,128.06	.00	198,815.30
Average neighbouring area of neigh.		5,839	4.60	6.90	.00	144.39
Illiterates (%)[12]		5,839	.03	.03	.00	.24
No studies (%)		5,839	.16	.13	.00	.89
Primary education (up to 16) (%)		5,839	.30	.12	.00	.78
Secondary education (up to 18) (%)		5,839	.42	.12	.06	1.00
University degree (%)		5,839	.08	.05	.00	.50
Value added	With agriculture	6,067	83,228.07	1,063,211	.00	67,900,000
Employment		6,067	2,295.84	23,746.50	.32	1,479,800
Area		6,067	71.91	103.77	.08	1,750.30
Labour productivity		6,067	29,019.65	39,577.58	.00	2,206,416
Labour density		6,067	71.54	855.32	.01	62,993.25
Elevation		6,067	538.78	336.16	2.00	1,692.00
Average neighbouring area		6,067	23.91	35.41	.00	682.84
Lagged labour density		6,067	65.19	689.01	.00	49,666.21
Average neighbouring area of neigh.		6,067	25.73	36.72	.00	682.84
Illiterates (%)		6,067	.03	.03	.00	.24
No studies (%)		6,067	.16	.13	.00	.89
Primary education (up to 16) (%)		6,067	.31	.13	.00	.86
Secondary education (up to 18) (%)		6,067	.42	.12	.06	1.00
University degree (%)		6,067	.08	.05	.00	.50
Value added	Without agriculture and total area	5,839	82,402.09	1,078,722	.00	67,700,000
Employment		5,839	2,192.99	23,796.48	.32	1,468,000
Area		5,839	72.89	105.33	.08	1,750.30
Labour productivity		5,839	29,560.43	48,053.70	.00	2,206,340
Labour density		5,839	70.39	869.81	.01	62,993.25
Elevation		5,839	529.86	334.88	2.00	1,692.00
Average neighbouring area		5,839	23.96	35.68	.00	682.84
Lagged labour density		5,839	63.29	699.47	.00	49,663.51
Average neighbouring area of neigh.		5,839	25.82	37.03	.00	682.84

[6] Thousands (Euros).
[7] Number of workers.
[8] Squared kilometres.
[9] Euros per worker.
[10] Workers per squared kilometre.
[11] Meters.
[12] Levels of human capital expressed as a proportion of total workers.

Table A2. Agglomeration effects with human capital controls and regional indicators. With agriculture and total land dataset

	Regional Indicator	OLS	Elevation 2SLS
Parameter θ (%)		4.14***	4.80***
Standard error of θ (%)		.46	.68
R^2 (%)	*No regional indicators*	7.23	7.19
ΔR^2 (%)		-	14.09
Parameter θ (%)		3.92***	2.77***
Standard error of θ (%)		.49	.80
R^2 (%)		8.34	8.23
Wald test (*ccaa = 0*)		6.33***	6.61***
Wald test (*ccaa ≠ pro = 0*)	*NUTs-2*	5.65***	5.81***
Parameter θ (%)		4.38***	3.94***
Standard error of θ (%)		.52	.83
R^2 (%)		10.18	10.16
Wald test (*pro = 0*)	*NUTs-3*	69.92***	66.39***

*** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$.

Table A.3. Agglomeration effects with human capital controls, neighbouring agglomeration and regional indicators

	Regional Indicator	OLS	(a) 2SLS	(b) 2SLS
Without agriculture and total land dataset				
Parameter (%)		5.14***	5.99***	4.68**
Standard error (%)	*No regional*	.59	.59	1.42
R^2 (%)	*indicators*	9.28	9.88	9.3
Parameter (%)		5.19***	6.03***	-
Standard error (%)		.6	.6	-
R^2 (%)		10.62	11.23	-
Wald test (*ccaa = 0*)		7.36***	6.85***	-
Wald test (*ccaa ≠ pro = 0*)	*NUTs-2*	7.9***	7.65***	-
Parameter (%)		5.47***	6.32***	3.24*
Standard error (%)		.61	.6	1.44
R^2 (%)		12.98	13.98	12.72
Wald test (*pro = 0*)	*NUTs-3*	5.86***	6.24***	5.85***
With agriculture and total land dataset				
Parameter (%)		3.43***	4.17***	3.94**
Standard error (%)		.6	.62	1.46
Parameter 10 (%)		1.66*	-	-
Standard error (%)		.67	-	-
Parameter 50 (%)		-1.77†	-	-
Standard error (%)		1.03	-	-
Parameter 100 (%)		2.13*	2.21*	2.26*
Standard error (%)	*No regional*	.94	.93	.94
R^2 (%)	*indicators*	7.79	8.15	7.75
Parameter (%)		3.44***	4.16***	-
Standard error (%)		.6	.63	-
Parameter 10 (%)		1.49*	-	-
Standard error (%)		.68	-	-
R^2 (%)		8.77	9.11	-
Wald test (*ccaa = 0*)		5.63***	5.06***	-
Wald test (*ccaa ≠ pro = 0*)	*NUTs-2*	5.5***	4.8***	-
Parameter (%)		3.6***	4.35***	-
Standard error (%)		.61	.63	-
Parameter 10 (%)		1.74*	-	-
Standard error (%)		.68		-
Parameter 100 (%)		1.73†	-	-
Standard error (%)		1.04	-	-
R^2 (%)		10.66	-	-
Wald test (*pro = 0*)	*NUTs-3*	4.5***	4.55***	-

*** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$.

(a) Labour density and average labour density across different neighbouring areas are instrumented with the 2 period lagged values, (b) labour density is instrumented with the elevation variable, and density across neighbours with its 2 period lagged values.

Table A4. Spillover effects of neighbours

	Regional Indicator	OLS	Elevation 2SLS
Without agriculture and total area dataset			
Parameter (%)	*NO*	4.65***-5.26***	3.73***-5.32***
Standard error (%)		.42-.45	.67-.65
+ve externalities		1.6†-35.6***	1.67†-35.59***
-ve externalities		-2.25*--1.52†	-2.54**--1.58†
R^2 (%)		8.94-20.19	8.84-20.19
Parameter (%)	*NUTs-2*	4.57***-5.3***	2.75***-3.56***
Standard error (%)		.45-.50	.77-.81
+ve externalities		3.23*-36.49***	2.89†-36.78***
-ve externalities			
R^2 (%)		10.32-21.49	10.02-21.2
Wald test	*ccaa = 0*	7.01***-8.09***	7.44***-8.65***
Wald test	*ccaa != pro = 0*	5.14***-7.91***	5.79***-8.78***
Parameter (%)	*NUTs-3*	5.22***-6.09***	4.46***-5.7***
Standard error (%)		.47-.52	.81-.90
+ve externalities		3.54†-36.45***	3.57†-36.6***
-ve externalities			
R^2 (%)		12.68-23.33	12.65-23.28
Wald test	*pro = 0*	10.86***-85.06***	10.12***-81.81***
With agriculture and total area dataset			
Parameter (%)	*NO*	3.79***-4.33***	4.31***-5.45***
Standard error (%)		.43-.46	.72-.67
+ve externalities		3.91**-33.92***	4.08**-33.86***
-ve externalities		-1.96†--1.96†	-1.87†--1.87†
R^2 (%)		7.29-17.72	7.28-17.69
Parameter (%)	*NUTs-2*	3.47***-4.00***	1.61*-3.24***
Standard error (%)		.45-.49	.77-.82
+ve externalities		2.53†-34.84***	2.52†-35.08***
-ve externalities			
R^2 (%)		8.37-18.67	8.29-18.38
Wald test	*ccaa = 0*	6.47***-7.97***	6.7***-7.94***
Wald test	*ccaa != pro = 0*	5.09***-7.94***	5.27***-7.81***
Parameter (%)	*NUTs-3*	3.86***-4.44***	2.92***-4.16***
Standard error (%)		.48-.52	.80-.84
+ve externalities		3.07†-35.39***	3.05†-35.52***
-ve externalities			
R^2 (%)		10.23-20.26	10.22-20.19
Wald test	*pro = 0*	8.19***-69.51***	7.4***-66.75***

*** *p* < .001, ** *p* < .01, * *p* < .05, † *p* < .10.