

# Cross-border agglomeration benefits

Martijn J. Smit<sup>1,2</sup> 

Received: 14 September 2016 / Accepted: 5 July 2017 / Published online: 22 July 2017  
© The Author(s) 2017. This article is an open access publication

**Abstract** Spatial correlations exist for many economic phenomena. We also know such interactions are typically weaker across country borders than within countries, due to institutional, infrastructural or cultural factors. However, in spatial econometric analyses, the effect of borders is rarely taken into account, and all borders between regions are treated as equal. We distinguish between different types of borders by splitting our weight matrix in two, and creating separate lags for within-country and cross-border effects. We demonstrate the effectiveness of this method in an analysis of average productivity in a region for several manufacturing sectors. Without implying causality, we relate this productivity to average productivity in surrounding regions as well as to agglomeration externalities both in its own region and in surrounding regions. We find border effects are indeed quite present: spatial lags within the country are invariably different from those across borders.

**Keywords** Agglomeration externalities · Spatial econometrics · Border effects

**JEL Classification** C31 · R12

---

**Electronic supplementary material** The online version of this article (doi:[10.1007/s12076-017-0191-0](https://doi.org/10.1007/s12076-017-0191-0)) contains supplementary material, which is available to authorized users.

---

✉ Martijn J. Smit  
m.j.smit@uu.nl

<sup>1</sup> Department of Human Geography and Planning, Utrecht University, Heidelberglaan 2, Postbus 80115, 3508 TC Utrecht, The Netherlands

<sup>2</sup> Department of Spatial Economics, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands

## 1 Introduction

National borders are much stronger boundaries than regional borders within a country. The effects of this fact have been extensively studied in a wide array of fields, including political science, geography, and economics. However, among the ever growing number of studies that employ spatial econometrics to control for spatial effects, distinguishing between different types of borders, or taking borders into account at all, is rare. This goes both for spatial dependence, where the underlying relationship is one of direct influence, and for spatial heterogeneity, where the effects themselves vary over space (Getis 1994).

In this paper, we demonstrate with a simple model how this can be rectified. After a brief discussion of previous attempts in this direction, we discuss our data and the spatial setup of our analysis. Finally, we discuss the results, leading to the conclusion.

## 2 Border effects

Whole institutes devote themselves to border research, including the Department of Border Region Studies at the University of Southern Denmark; the Nijmegen Centre for Border Research; the Centre for International Borders Research in Belfast. Although these institutes focus on such widely diverging topics as cross-border shopping behaviour, ethno-national conflicts, and crime, even covering the economic stream of literature is an impossible task to adequately summarize. However, we will briefly touch upon a few papers which perform empirical economic analyses where border regions are taken into account in the spatial weights. None of these studies is concerned with agglomeration effects.

Niebuhr (2003) looks at regional unemployment levels in Western Europe, and constructs three weight matrices: one which includes all interregional effects, one which only includes intracountry effects (as we will show below) and one which gives cross-border effects a lower weight, using a border-specific factor. She also tests different values for a distance decay parameter in the range of spatial effects.

In the trade literature, where gravity model are the de-facto standard (Disdier and Head 2008), transnational borders are a given, but research has shown how some borders are stronger than others, for example due to tariffs, cultural distance, or institutional distance (Linders et al. 2008). Behrens et al. (2012) implement border impedance into such a model, using spatial econometrics.

Rietveld and Wintershoven (1998) separate cross-border from intracountry effects when looking at infrastructure provision, using intracountry and crossborder matrices like we do, but using a combined lag-error model instead of lagged variables. They find in one specification that for railway connections across the border, the density of the network across the border is a larger influence than the density in the home country.

### 3 Analysis

We construct our data using the Amadeus dataset.<sup>1</sup> Amadeus, coordinated by Bureau van Dijk, is a Europe-wide database of data gathered from the yearly statements of companies. Although it can be used to look up individual companies, it is also explicitly intended for sectoral analyses of trends and peer groups. It covers around 21 million companies all over Europe, and forms part of the global ORBIS database. Coverage by country varies from 30% in the Netherlands to over 90% in Romania, Slovakia and Estonia.<sup>2</sup>

We aggregate the microdata for 2011 available in that dataset to a cross-section<sup>3</sup> at the NUTS2 level, calculating the average added value per employee by 2-digit sector as well as four agglomeration variables by sectoral groupings discussed below. This strand of literature started with [Glaeser et al. \(1992\)](#), and has not been brought to a standstill yet, not even by the overview works of [Beaudry and Schiffauerova \(2009\)](#) or [de Groot et al. \(2015\)](#). In line with current developments in that field, we distinguish between three effects:

- Concentration of an industry: this type of agglomeration effects harks back to [Marshall \(1890\)](#), and covers, among others, shared inputs, dense labour markets, and knowledge spillovers. We measure it using a location quotient, as is quite standard in the literature ([Beaudry and Schiffauerova 2009](#)), attributing our sectors to the 51 industrial groupings of the European Cluster Observatory ([Sölvell et al. 2009](#)), and calculating the variable across the economy as a whole. This implies for example that sectors 2453 (casting of light metals) and 2910 (manufacture of motor vehicles) are grouped together as the automotive sector, which also includes sectors not analysed in our current dataset, such as 3040 (manufacture of military fighting vehicles).
- Population density as a measure for the concentration of general activity: non-sector-specific agglomeration also offers advantages, including a home market effect and possibly a general amount of creativity in the population. This effect is normally labelled “urbanisation”, and population density is the standard measure for it ([Melo et al. 2009](#)); however, it is often confused with the effects of relatedness discussed below.
- Related and unrelated variety: according to [Frenken et al. \(2007\)](#), the traditional [Jacobs \(1969\)](#) externalities of diversity or variety can be split in two aspects: on the one hand, an effect of relatedness within industrial groupings, fostering creativity and innovation, for example among high-tech firms ([Hartog et al. 2012](#));

---

<sup>1</sup> The author thanks Jakub Gloser of Ecorys and Henri de Groot of the Vrije Universiteit Amsterdam; together, we created the dataset on which this paper is based. He also thanks the anonymous reviewers at LSRS for their cooperative comments.

<sup>2</sup> These numbers relate to the share of gross output in the economy for 2011, and are found in Tables 6.1 and 6.2 of [Kalemli-Ozcan et al. \(2015\)](#).

<sup>3</sup> Many studies in this literature describe a dynamic process (as discussed by [de Groot et al. 2015](#)), but a number describe it at a static situation instead: among them [Condliffe et al. \(2008\)](#), [Drucker and Feser \(2007\)](#) and [van Oort \(2002\)](#). We have dynamic data for employment and turnover growth for the short period 2010–2012: these results are available online Electronic Supplementary Material or at [www.martijnjsmit.nl/appendix](http://www.martijnjsmit.nl/appendix).

**Table 1** Descriptives

Variable		Mean	SD	Min	Max
Productivity	In region	102,261	106,602	-365,150	1,021,862
Lagged productivity	Overall	392,617	401,180.9	0	3,229,691
Lagged productivity	Within-country	326,840	369,188.5	-113,545	2,860,440
Lagged productivity	Cross-border	65,777	159,379.6	0	1,837,066
Concentration (LQ)	In region	1.114776	0.742162	0.012257	5.714916
Concentration (LQ)	Overall	4.519649	3.30663	0	18.33563
Concentration (LQ)	Within-country	3.511927	2.697561	0	15.03329
Concentration (LQ)	Cross-border	1.007722	1.857123	0	14.29095
Unrelated variety	In region	2.142744	0.318501	1.283289	2.827869
Related variety	In region	0.24464	0.08988	0	0.367774
Average establishment size (employees)	In region	116.1606	283.0868	3.5	7642.333
Population density	In region	411.4069	1064.942	3.3	10,138.3
Population density	Overall	1094.754	1517.829	0	11,973.4
Population density	Within-country	981.8926	1544.267	0	11,973.4
Population density	Cross-border	112.8611	206.5473	0	1047.4

on the other hand, a portfolio effect of unrelated sectors across the rest of the economy. These two variables are measured using the entropy measures described by [Frenken et al. \(2007\)](#), using the 51 industries recently defined by the European Cluster Observatory ([Ketels and Protsiv 2014](#)) as the group, and the 3-digit NACE classification as the units.<sup>4</sup>

We hypothesize unrelated variety has no direct relationship to added value at the sectoral level in our cross-sectional analysis, but both related variety and concentration can have positive impacts. For urbanisation, the effect can depend on the life cycle of a (regional) industry, as was forcefully argued by [Duranton and Puga \(2001\)](#); in the absence of data, our hypothesis is that we will see mixed effects across the sectors in this analysis.

Finally, we control for average establishment size. Descriptive statistics can be found in [Table 1](#). Note that our productivity value, added value per employee, contains negative values: some industries incurred negative added value in 2011, probably due to the crisis.

Due to missing data for many regions, we restrict ourselves to a sample of several manufacturing industries that are well represented. These are the NACE2-sectors 20 and 22–28 ([Table 2](#)).

We aggregate data from these individual firms to the regional level in order to be able to create spatially lagged variables at that level. We drop the regions with no observations for at least one of the sectors, and this results in 175 NUTS2-regions,

<sup>4</sup> Note that spatial lags for related and unrelated variety do not make sense: see [Frenken et al. \(2007\)](#), endnote 7.

**Table 2** Sectors included in the analysis, and total number of firms in the database

NACE2	Description	#	%
20	Manufacture of chemicals and chemical products	8475	6.18
22	Manufacture of rubber and plastic products	13,600	9.92
23	Manufacture of other non-metallic mineral products	11,307	8.25
24	Manufacture of basic metals	5001	3.65
25	Manufacture of fabricated metal products, except machinery and equipment	56,329	41.1
26	Manufacture of computer, electronic and optical products	9379	6.84
27	Manufacture of electrical equipment	9565	6.98
28	Manufacture of machinery and equipment not elsewhere classified	23,381	17.06
Total		137,037	100

**Fig. 1** Regions included in the analysis**Table 3** Summary of the weight matrices

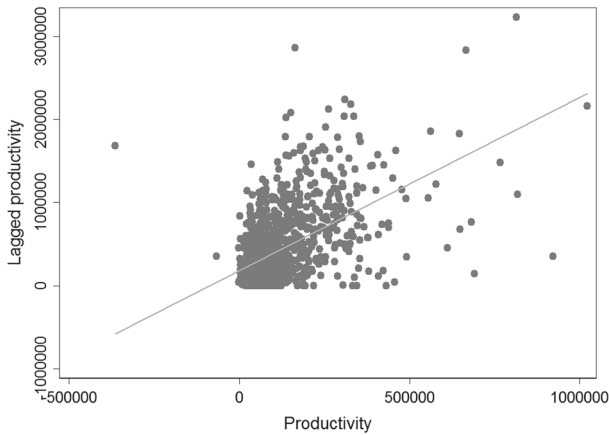
	Full matrix	Within-country matrix	Cross-border matrix
Dimensions	175 × 175	175 × 175	175 × 175
Total links	664	534	130
Min. links	0	0	0
Max. links	3.79	3.05	0.74
Mean links	11	9	7

**Table 4** Results

Dependent	Productivity (GVA/employee)			
	(1)	(2)	(3)	(4)
Lagged productivity	0.141*** (22.87)		0.175*** (24.32)	
Lagged productivity (within country)		0.146*** (21.54)		0.183*** (21.05)
Lagged productivity (across a border)		0.117*** (7.79)		0.168*** (8.41)
Concentration (LQ)	8104.2* (2.37)	7984.0* (2.33)	14,504.4*** (4.08)	14,260.7*** (4.02)
Lagged concentration			-6709.3*** (-7.12)	
Lagged concentration (within country)				-7507.1*** (-6.43)
Lagged concentration (across a border)				-512.1 (-0.23)
Related variety	-43,475.6 (-1.61)	-44,378.7 (-1.64)	-40,503.0 (-1.54)	-40,847.6 (-1.55)
Unrelated variety	-35,705.7*** (-4.60)	-31,814.3*** (-3.95)	-27,800.3*** (-3.62)	-26,263.3*** (-3.31)
Average establishment size (employees)	23.30** (2.60)	23.29** (2.60)	22.26* (2.55)	22.44* (2.58)
Population density	9.491*** (3.92)	9.476*** (3.92)	10.28*** (4.23)	9.953*** (4.10)
Lagged population density			-7.099*** (-4.37)	
Lagged pop. density (within country)				-7.159*** (-4.35)
Lagged pop. density (across a border)				-71.35*** (-3.30)
Constant	118,557.0*** (6.50)	110,542.1*** (5.88)	118,106.0*** (6.59)	117,003.4*** (6.35)
Observations	1400	1400	1400	1400
Adjusted $R^2$	0.341	0.342	0.374	0.379

displayed in Fig. 1. Unfortunately, we miss the Netherlands, Denmark, most of former East Germany, and a large central chunk of Italy.

For these regions, we create three weight matrices based on contiguity, using Stata. The within-country matrix is the hadamard product of the full matrix with a matrix of ones and zeroes indicating all regions that are within the same country; this means all pairs in different countries have a value of zero. The cross-border matrix covers those,



**Fig. 2** Scatter plot of productivity and its spatial lag

with zeroes within the country. Statistics for all three weight matrices can be found in Table 3.

## 4 Results

The cross-sectional character of our data implies we have to be careful with statements on causality; in the absence of acceptable exogenous proxies for sectoral agglomeration externalities,<sup>5</sup> we will show correlations and provide arguments on the supposed direction of causality. Previous literature has shown, however, that the relationship between agglomeration externalities and productivity is often bidirectional (Graham et al. 2010).

Table 4 shows our results, in four models of increasing complexity. We find that the spatial lag of productivity has a positive relationship with productivity in a region—that is, when your neighbouring regions have a high productivity, you are more likely to have the same. A bivariate Moran scatterplot of productivity against its spatial lags shows the same effect, even with no controls present (Fig. 2). When we distinguish this effect between within-country and cross-border effects, we see both are statistically significant, but, as expected, the cross-border effect is smaller; at the 10% level, the effects are different, with an F-test resulting in 3.06.

Concentration is slightly statistically significant, with a 1-point increase in the location quotient leading to an €8000 increase in added value per employee in the non-spatial models; however, adding the spatial lag shows a concentration (cluster) in the region itself has a much higher positive effect, whereas concentrations in nearby regions diminish this effect considerably, as long as they are within the country itself. This suggests agglomeration benefits of the Marshallian type do not carry across any country borders, and that positive effects are to be expected only at the NUTS2 level

<sup>5</sup> For urbanisation, or non-sectoral agglomeration, proxies are often used (e.g. Ades and Glaeser 1994), in particular in the literature on infrastructure and urban growth (García-López et al. 2015).

or below that. This is confirmed in some strands of literature; some studies suggest the appropriate scale is much, much smaller than NUTS2 regions, from cities (Viladecans-Marsal 2004) to even buildings (Kabo et al. 2014).

Unrelated variety also has a negative and statistically significant relationship with productivity in a region; for related variety, the effect is statistically insignificant. Population density has a consistently positive effect, but again the spatial lag is negative, both within a country and across a border. However, the cross-border effect is by far the strongest, pointing to a lower productivity in regions that have a foreign city close by.

## 5 Conclusions

Border effects are strong, and should be included in any spatial econometric analysis that covers multiple countries. Doing so by splitting the weight matrix renders more precise results, which are moreover more valuable for national policies aiming to maximize the intracountry effect and cross-border policies aiming to promote interactions between countries.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Ades, A.F., Glaeser, E.L.: Trade and circuses: explaining urban giants. Working Paper No. 4715, National Bureau of Economic Research (1994)
- Beaudry, C., Schiffauerova, A.: Who's right, Marshall or Jacobs? The localization versus urbanization debate. *Res. Policy* **38**, 318–337 (2009). doi:[10.1016/j.respol.2008.11.010](https://doi.org/10.1016/j.respol.2008.11.010)
- Behrens, K., Ertur, C., Koch, W.: “Dual” gravity: using spatial econometrics to control for multilateral resistance. *J. Appl. Econ.* **27**, 773–794 (2012). doi:[10.1002/jae.1231](https://doi.org/10.1002/jae.1231)
- Condliffe, S., Latham, W., Le Bas, C., Miribel, F.: Agglomeration Economies within IT-producing and IT-consuming Industries in U.S. Regions (2008). doi:[10.2139/ssrn.1162373](https://doi.org/10.2139/ssrn.1162373)
- de Groot, H.L.F., Poot, J., Smit, M.J.: Which agglomeration externalities matter most and why? *J. Econ. Surv.* **30**, 756–782 (2015). doi:[10.1111/joes.12112](https://doi.org/10.1111/joes.12112)
- Disdier, A.C., Head, K.: The puzzling persistence of the distance effect on bilateral trade. *Rev. Econ. Stat.* **90**, 37–48 (2008)
- Drucker, J.M., Feser, E.: Regional industrial dominance, agglomeration economies, and manufacturing plant productivity (2007). US Census Bureau Center for Economic Studies Paper No. CES-07-31. doi:[10.2139/ssrn.1103067](https://doi.org/10.2139/ssrn.1103067)
- Durantón, G., Puga, D.: Nursery cities: urban diversity, process innovation, and the life cycle of products. *Am. Econ. Rev.* **91**, 1454–1477 (2001)
- Frenken, K., van Oort, F.G., Verburg, T.: Related variety, unrelated variety and regional economic growth. *Reg. Stud.* **41**, 685–697 (2007)
- García-López, M.-Á., Holl, A., Viladecans-Marsal, E.: Suburbanization and highways in Spain when the Romans and the Bourbons still shape its cities. *J. Urban Econ.* **85**, 52–67 (2015). doi:[10.1016/j.jue.2014.11.002](https://doi.org/10.1016/j.jue.2014.11.002)
- Getis, A.: Spatial dependence and heterogeneity and proximal databases. In: Fotheringham, S., Rogerson, P. (eds.) *Spatial Analysis and GIS*, pp. 105–120. CRC Press, Boca Raton (1994)



- Glaeser, E.L., Kallal, H.D., Scheinkman, J.A., Shleifer, A.: Growth in cities. *J. Polit. Econ.* **100**, 1126–1152 (1992)
- Graham, D.J., Melo, P.S., Jiwattanakulpaisarn, P., Noland, R.B.: Testing for causality between productivity and agglomeration economies. *J. Reg. Sci.* **50**, 935–951 (2010). doi:[10.1111/j.1467-9787.2010.00676.x](https://doi.org/10.1111/j.1467-9787.2010.00676.x)
- Hartog, M., Boschma, R., Sotarauta, M.: The impact of related variety on regional employment growth in Finland 1993–2006: high-tech versus medium/low-tech. *Ind. Innov.* **19**, 459–476 (2012). doi:[10.1080/13662716.2012.718874](https://doi.org/10.1080/13662716.2012.718874)
- Jacobs, J.: *The Economy of Cities*. Random House, New York (1969)
- Kabo, F.W., Cotton-Nessler, N., Hwang, Y., Levenstein, M.C., Owen-Smith, J.: Proximity effects on the dynamics and outcomes of scientific collaborations. *Res. Policy* **43**, 1469–1485 (2014). doi:[10.1016/j.respol.2014.04.007](https://doi.org/10.1016/j.respol.2014.04.007)
- Kalemli-Ozcan, S., Sorensen, B., Villegas-Sanchez, C., Volosovych, V., Yesiltas, S.: How to construct nationally representative firm level data from the ORBIS global database. Working Paper No. 21558, National Bureau of Economic Research (2015)
- Ketels, C., Protsiv, S.: Methodology and Findings Report for a Cluster Mapping of Related Sectors, European Cluster Observatory. European Cluster Observatory, European Commission, Brussels (2014)
- Linders, G.-J.M., Burger, M.J., van Oort, F.G.: A rather empty world: the many faces of distance and the persistent resistance to international trade. *Camb. J. Reg. Econ. Soc.* **1**, 439–458 (2008). doi:[10.1093/cjres/rsn019](https://doi.org/10.1093/cjres/rsn019)
- Marshall, A.: *Principles of Economics: An Introductory Volume*. Macmillan, London (1890)
- Melo, P.S.C., Graham, D.J., Noland, R.B.: A meta-analysis of estimates of urban agglomeration externalities. *Reg. Sci. Urban Econ.* **39**, 332–342 (2009)
- Niebuhr, A.: Spatial interaction and regional unemployment in Europe. *Eur. J. Spat. Dev.* **5**, 1–26 (2003)
- Rietveld, P., Wintershoven, P.: Border effects and spatial autocorrelation in the supply of network infrastructure. *Pap. Reg. Sci.* **77**, 265–276 (1998). doi:[10.1111/j.1435-5597.1998.tb00717.x](https://doi.org/10.1111/j.1435-5597.1998.tb00717.x)
- Sölvell, O., Ketels, C., Lindqvist, G.: *The European Cluster Observatory: EU Cluster Mapping and Strengthening Clusters in Europe* (2009). <https://www.scribd.com/document/29279536/The-European-Cluster-Observatory-EU-Cluster-Mapping-and-Strengthening-Clusters-in-Europe>
- van Oort, F.G.: Innovation and agglomeration economies in the Netherlands. *TESG* **93**, 344–360 (2002)
- Viladecans-Marsal, E.: Agglomeration economies and industrial location: city-level evidence. *J. Econ. Geogr.* **4**(5), 565–582 (2004)