ASECU South-Eastern Europe Journal of Economics 2 (2013) 177-185 SEEJE

USING GRAVITY MODELS FOR THE EFFECTIVE DETERMINATION OF SOCIOECONOMIC LOCALITY: LOCAL LABOUR MARKETS IN CENTRAL MACEDONIA

ZACHARIAS SARTZETAKIS[a], GRIGORIS ZAROTIADIS[b*]

[a, b]Aristotle University of Thessaloniki, Greece

Abstract

The discussion about the effectiveness of active policy measures that consider spatial socioeconomic aspects has become especially important in light of recent socio-economic developments. The paper presents an alternative methodology for the determination of spatial boundaries for any socioeconomic locality, based on the gravity models tradition. We apply the described methodology to the Region of Central Macedonia and show that there are significant discontinuities between administrative spatial segregation on the one hand and socioeconomic on the other. The latter provides evidence for the inappropriateness of simply following the administrative boundaries and reveals the usefulness of our proposal.

JEL Classification: R12, R58, J68

Keywords: Spatial Socioeconomic Policy, Local Labour Market, Gravity Model

The paper was presented at the 3rd International Summer School of ASECU-Youth in Kotor, Montenegro, 2012.

**Corresponding Author:* Grigoris Zarotiadis, Department of Economics, Aristotle University of Thessaloniki, 541 24, Thessaloniki, Greece e-mail: gzarotia@econ.auth.gr

1. Introduction

The use of administrative areas as geographical units does not provide a valid insight into the functional reality of the territory. Administrative boundaries are not necessarily appropriate, either for scientific research or for policy making (Casado-Diaz, 2000). Commonly, employment and housing networking does not coincide with existing administrative units. Local district boundaries usually have a historical background that does not automatically correspond to current links and relations. The development of new transport or social infrastructure or new retail, commercial, industrial, residential and leisure facilities continuously redefines the current spatial socioeconomic structure (Duncan, 2010). Moreover, Ballas and Clarke (2000) pointed out that the use of specialized geographic tools increases the efficiency of applied socioeconomic policy. Given the recent socio-economic developments within the systemic crisis, the scientific interest in active policy measures with respect to local socioeconomic aspects has been amplified. The main purpose of a spatial planning framework is to ensure that policies cover local socioeconomic needs, counteract regional inequalities and favour a more balanced regional economic development (Duncan, 2010)[1]. In other words, it is of great importance to develop methodologies that ensure an efficient determination of each socioeconomic locality. This is especially the case in planning and applying locally specified, active employment policies. Spatial dimensions contribute a lot to the understanding of the labour market and thereby to the development of well adjusted measures.

The gravity model has been one of the most popular empirical methodologies for studying spatial interactions in economics over the last fifty years (Cieślik, 2009 and Rose, 2000). It imitates Newton's law of universal gravitation (1687): the socioeconomic mass of two agglomerations (measured for instance by the population or by locally produced income) along with their distance determine the magnitude of their relations in terms of goods, services and production factors. Gravity models "... *are simple in structure, fit the data well, and are in principle consistent with a wide range of theoretical underpinnings*" (Deardorff, 1998) and they provide "*some of the clearest and most robust findings in empirical economics*" (Leamer E. & Levinsohn J.,1995).

The main purpose of this paper is to present an innovative methodology for the determination of spatial boundaries for any socioeconomic locality, based on the gravity models tradition. Thereby, we provide an instrument for a more efficient preparation of spatially oriented measures (see also in Zarotiadis and Stamboulis, 2011). As an

1. *"In my mind, direct links between the European Union and regional and local authorities are more needed than ever"* (Martin Schulz - President of the European Parliament, www. 11/03/2013 - Reuters.com).

indicative example, we apply the presented methodology to a specific Greek Region and we show that there are significant discontinuities regarding administrative spatial segregation on the one hand and the actual socioeconomic segregation on the other.

The next section presents the proposed normative algorithm for empirical regionalization. The third section deals with the pilot application in the Region of Central Macedonia. We present the results on a map, contrasted with the picture of administrative segmentation and we go into the main differences. To conclude, the last section summarises the discussion and comments on policy implications.

2. Algorithm for empirical regionalization

The algorithm we present provides the optimal spatial segregation of socioeconomic locality (for instance, as we will see in the pilot application, of local labour markets), based on the quantitative consideration of the socioeconomic mass of various agglomerations, the distance, administrative constraints, infrastructure as well as other socioeconomic parameters. It proceeds through the following steps:

i) Determine an elementary or augmented log-linear formulation of a gravity model:

$$ln(Mob_{i,j}) = ln(b) + a_1 ln(p_i) + a_2 ln(p_j) + a_3 ln(distance_{i,j}) + a_4 lnF + \varepsilon_{i,j} \qquad (1)$$

where $Mob_{i,j}$ is a measure for the magnitude of the relation between any pair of considered agglomerations ($i,j = 1,2,...,N,\ i \neq j$ and N is the number of considered agglomerations); p_i and p_j is the significance (the socioeconomic mass) of any agglomeration; $distance_{i,j}$ shows the distance between them; F is a vector of various dummy variables representing the different physical, technical and administrative boundaries that exist between i and j. (Note that a_4 is the vector of relevant coefficients.)

ii) Based on the estimated coefficients b, a_1, a_2, a_3 and a_4, calculate the estimated relation for each pair of considered agglomerations in the region of interest, $eMob_{i,j}$[2]. Construct the double entry matrix A:

$$A = \begin{array}{cccccc}
0 & eMob_{1,2} & .... & eMob_{1,v} & .... & eMob_{1,N} \\
eMob_{2,1} & 0 & .... & eMob_{2,v} & .... & eMob_{2,N} \\
eMob_{3,1} & eMob_{3,2} & 0. & eMob_{3,v} & .... & eMob_{3,N} \\
..... & ..... & .... & 0 & .... & ..... \\
eMob_{\kappa,1} & eMob_{\kappa,2} & .... & eMob_{\kappa,v} & 0 & eMob_{\kappa,N} \\
.... & .... & .... & .... & .... & 0 \\
eMob_{N,1} & eMob_{N,2} & .... & eMob_{N,v} & .... & eMob_{N,N}
\end{array}$$

2. There is a precondition for this step: the estimated equation has to be significant enough and unbiased. Otherwise, the usefulness of $eMob_{i,j}$ can be doubtful. There are no generally valid rules for having a significant, unbiased, estimated equation, depending on the type of data and the applied econometric model and method.

iii) Calculate the relative significance, $R_{i,j} \in (0,1)$, of each $eMob_{i,j}$ by dividing it by the sum of all $eMob_{i,j}$. For instance, if we measure $Mob_{i,j}$ by the number of employees that move from one agglomeration to the other (as we do in the pilot application we present below), $R_{i,j}$ represents the relative significance of the estimated mobility between i and j, related to the overall mobility in the wider area.
Construct the double entry matrix B:

$$B = \left[\begin{array}{ccccccc|c} R_{1,1} & R_{1,2} & \cdots & R_{1,v} & \cdots & R_{1,N} & & R_{1,\cdot} \\ R_{2,1} & R_{2,2} & \cdots & R_{2,v} & \cdots & R_{2,N} & & R_{2,\cdot} \\ R_{3,1} & R_{3,2} & \cdots & R_{3,v} & \cdots & R_{3,N} & & R_{3,\cdot} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & & \cdots \\ R_{\kappa,1} & R_{\kappa,2} & \cdots & R_{\kappa,v} & \cdots & R_{\kappa,N} & & R_{\kappa,\cdot} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & & \cdots \\ R_{N,1} & R_{N,2} & \cdots & R_{N,v} & \cdots & R_{N,N} & & R_{N,\cdot} \\ \hline R_{\cdot,1} & R_{\cdot,2} & \cdots & R_{\cdot,v} & \cdots & R_{\cdot,N} & & \mathbf{1} \end{array}\right]$$

iv) Calculate $eR_{i,j} = R_{.,j}\ R_{i,}.$
$eR_{i,j}$ represents the indicative mobility between the two agglomerations *i* and *j*. In a way, $eR_{i,j}$ stands for the mobility that would arise only due to the total relative repulsion of each departure city ($R_{i,}.$) and the total relative attraction of each arrival city ($R._{,j}$), without having any additional reasons that give rise to an intensified dependence of *i* and *j*. In other words, we treat the two variables (city of departure, city of arrival) as being independent.
Construct the double entry matrix C:

$$C = \begin{array}{cccccc} eR_{1,1} & eR_{1,2} & \cdots & eR_{1,v} & \cdots & eR_{1,N} \\ eR_{2,1} & eR_{2,2} & \cdots & eR_{2,v} & \cdots & eR_{2,N} \\ eR_{3,1} & eR_{3,2} & \cdots & eR_{3,v} & \cdots & eR_{3,N} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ eR_{\kappa,1} & eR_{\kappa,2} & \cdots & eR_{\kappa,v} & \cdots & eR_{\kappa,N} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ eR_{N,1} & eR_{N,2} & \cdots & eR_{N,v} & \cdots & eR_{N,N} \end{array}$$

v) Construct matrix $D = B - C$
If an element ($R_{i,j}$ - $eR_{i,j}$) of matrix *D* is positive (negative), the interaction from departure agglomeration *i* and arrival agglomeration *j* is (not) gravitational. (The closer to one an element is, the more gravitational is the relation.)

From this point on, the following steps are provisional and may be adjusted according to the characteristics of the specific study. In general, we should proceed with the most appropriate form of clustering analysis. For instance, in the pilot application presented in the following section, we first picked out the most *intense* pairs of agglomerations (the highest $R_{i,j}$ - $eR_{i,j}$). Note that at this point of the clustering procedure, we did not consider the biggest city of the specific Region (in terms of population). Thereafter, pairs were combined into groups by considering the strongest mutual gravitational interactions[3]. Next, we can also proceed with consolidating groups that were characterized by the higher number of mutual gravitational interactions. Finally, we include the biggest agglomeration (city) of the Region to the group that has the mightiest mean mutual gravitational interaction with this city.

3. A pilot application to the Region of Central Macedonia

The study we present in this part of the paper provides a good example of how we can use the described methodology. Moreover it shows what could be the resulting differentiations compared to the actual administrative segregation of the region, even when we do not use a very detailed analysis of agglomerations, as is the case, for reasons of simplicity, in the following.

We used data from the Greek Social Data Bank (GSDB) in the National Centre of Social Research (EKKE) on the mobility of employees between the municipalities (133) of the Region of Central Macedonia. As we mentioned before, for reasons of simplicity we restricted our sample to the agglomerations whose population is over 20,000 inhabitants: Alexandreia Imathias (AI), Aridaia (A), Edessa (E), Giannitsa (G), Katerini (Ka), Kilkis (Ki), Naousa (N), Serres (S), Thessaloniki (T) and Veroia (V). Additionally, we included also Irakleia Serron (IS), Litohoro (L), Poligiros (P) and Nea Moudania (NM) in order to have a more effective representation of the whole geographical area.

The Pearson-test showed that none of the various conceivable dummy variables (if the city is an administrative centre, existence of administrative border, if there is a main road connection or not) should be included in the model. Therefore, we formulated the following equation:

$$ln(Mob_{i,j}) = ln(b) + a_1 ln(p_i) + a_2 ln(p_j) + a_3 ln(distance_{i,j}) + \varepsilon_{i,j} \qquad (2)$$

OLS regression gave the results presented in the following table[4]:

Estimated coefficients are significant and confirm the theoretical assumptions: as expected pair-wise labour mobility is being positively affected by the size of the two cities and negatively by their distance. R^2- and F-statistic, as well as the results of the regression specification error test, indicate that the model is well specified and provides a satisfactory explanatory ability.

3. As a rule of thumb, at least half of the existing mutual relations had to be gravitational in order to allow the combination of two groups. The specific restriction was applied in order to ensure a minimum internal consistency.
4. We applied White-correction in order to remove heteroskedasticity. Estimation is free of multi-collinearity.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
ln ($distance_{i,j}$)	-2,71	0,19	-13,95	0,00
ln (p_i)	1,15	0,06	20,27	0,00
ln (p_j)	1,12	0,04	25,78	0,00
Constant	-10,35	1,22	-8,47	0,00
R-squared:	0,90	Mean dependent var:		3,01
Adjusted R-squared:	0,89	S.D. dependent var:		2,06
S.E. of regression:	0,68	Akaike info criterion:		2,13
Sum squared resid:	27,85	Schwarz criterion:		2,26
Log Likelihood:	-64,19	Hannan-Quinn criterion:		2,18
F-statistic:	173,04	Burbin-Watson statistic:		1,69
Prob (F-statistc):	0,00			

Based on the above estimations, we constructed the four matrices of our methodology, following the procedures mentioned above:

Table A.

	AI	A	V	G	E	IS	T	Ka	Ki	L	N	NM	P	S	Sum
AI	0	2	42	76	4	0	260	10	3	0	7	0	0	1	406
A	2	0	7	8	36	0	49	1	1	0	5	0	0	0	109
V	43	7	0	35	20	0	296	13	2	1	144	1	0	2	563
G	77	8	34	0	21	0	457	8	4	0	10	1	0	2	622
E	4	36	20	21	0	0	91	2	1	0	23	0	0	1	200
IS	0	0	0	0	0	0	48	0	1	0	0	0	0	47	96
T	291	55	323	503	101	55	0	439	452	22	75	159	103	277	2853
Ka	11	1	13	8	2	0	405	0	2	30	3	1	0	2	478
Ki	3	1	2	4	1	1	407	2	0	0	0	1	0	6	428
L	0	0	1	0	0	0	19	28	0	0	0	0	0	0	48
N	7	5	141	10	23	0	67	3	0	0	0	0	0	1	257
NM	0	0	1	1	0	0	141	1	1	0	0	0	9	1	155
P	0	0	0	0	0	0	91	0	0	0	0	9	0	0	100
S	1	1	2	2	1	49	255	2	7	0	1	1	0	0	322
Sum	439	116	585	667	210	106	2587	509	474	54	269	172	113	340	6639

Table B.

	AI	A	V	G	E	IS	T	Ka	Ki	L	N	NM	P	S	Sum
AI	0,00	0,00	0,01	0,01	0,00	0,00	0,04	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,06
A	0,00	0,00	0,00	0,00	0,01	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02
V	0,01	0,00	0,00	0,01	0,00	0,00	0,04	0,00	0,00	0,00	0,02	0,00	0,00	0,00	0,08
G	0,01	0,00	0,01	0,00	0,00	0,00	0,07	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,09
E	0,00	0,01	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,03
IS	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,01
T	0,04	0,01	0,05	0,08	0,02	0,01	0,00	0,07	0,07	0,00	0,01	0,02	0,02	0,04	0,43
Ka	0,00	0,00	0,00	0,00	0,00	0,00	0,06	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,07
Ki	0,00	0,00	0,00	0,00	0,00	0,00	0,06	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,06
L	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01
N	0,00	0,00	0,02	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,04
NM	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02
P	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02
S	0,00	0,00	0,00	0,00	0,00	0,01	0,04	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,05
Sum	0,07	0,02	0,09	0,10	0,03	0,02	0,39	0,08	0,07	0,01	0,04	0,03	0,02	0,05	1,00

Table C.

	AI	A	V	G	E	IS	T	Ka	Ki	L	N	NM	P	S
AI	0,000	0,001	0,005	0,006	0,002	0,001	0,024	0,005	0,004	0,000	0,002	0,002	0,001	0,003
A	0,001	0,000	0,001	0,002	0,001	0,000	0,006	0,001	0,001	0,000	0,001	0,000	0,000	0,001
V	0,006	0,001	0,000	0,009	0,003	0,001	0,033	0,006	0,006	0,001	0,003	0,002	0,001	0,004
G	0,006	0,002	0,008	0,000	0,003	0,002	0,037	0,007	0,007	0,001	0,004	0,002	0,002	0,005
E	0,002	0,001	0,003	0,003	0,000	0,000	0,012	0,002	0,002	0,000	0,001	0,001	0,001	0,002
IS	0,001	0,000	0,001	0,001	0,000	0,000	0,006	0,001	0,001	0,000	0,001	0,000	0,000	0,001
T	0,028	0,007	0,038	0,043	0,014	0,007	0,000	0,033	0,031	0,003	0,017	0,011	0,007	0,022
Ka	0,005	0,001	0,006	0,007	0,002	0,001	0,028	0,000	0,005	0,001	0,003	0,002	0,001	0,004
Ki	0,004	0,001	0,006	0,006	0,002	0,001	0,025	0,005	0,000	0,001	0,003	0,002	0,001	0,003
L	0,000	0,000	0,001	0,001	0,000	0,000	0,003	0,001	0,001	0,000	0,000	0,000	0,000	0,000
N	0,003	0,001	0,003	0,004	0,001	0,001	0,015	0,003	0,003	0,000	0,000	0,001	0,001	0,002
NM	0,002	0,000	0,002	0,002	0,001	0,000	0,009	0,002	0,002	0,000	0,001	0,000	0,000	0,001
P	0,001	0,000	0,001	0,002	0,000	0,000	0,006	0,001	0,001	0,000	0,001	0,000	0,000	0,001
S	0,003	0,001	0,004	0,005	0,002	0,001	0,019	0,004	0,003	0,000	0,002	0,001	0,001	0,000

Based on the results of table D, we may proceed with the appropriate form of clustering analysis. In order to provide a complete pilot application of the proposed methodology, we first picked out the most *intense* pairs of agglomerations (the highest $R_{i,j}$ - $eR_{i,j}$) and we combined pairs into groups according to the mutual gravitational interaction. We also proceeded with consolidating groups, getting finally a picture which, despite the fact that that we restricted our sample to relatively big agglomerations, reveals significant differences compared to the current administrative structure. According to our estimates the prefecture of Thessaloniki and that of Kilkis should be treated as a single labour market. The same is true for Pella and Imathia –in fact they could also be included in the socioeconomic area formed around Thessaloniki. The prefectures of Serres, Pieria and Chalkidiki can be treated as separate markets.

Table D.

	AI	A	V	G	E	IS	T	Ka	Ki	L	N	NM	P	S
AI	0,000	-0,001	0,001	0,005	-0,001	-0,001	0,015	-0,003	-0,004	-0,000	-0,001	-0,002	-0,001	-0,003
A	-0,001	0,000	-0,000	-0,001	0,005	-0,000	0,001	-0,001	-0,001	-0,000	0,000	-0,000	-0,000	-0,001
V	0,001	-0,000	0,000	-0,003	0,000	-0,001	0,012	-0,005	-0,006	-0,001	0,018	-0,002	-0,001	-0,004
G	0,005	-0,000	-0,003	0,000	0,000	-0,001	0,032	-0,006	-0,006	-0,001	-0,002	-0,002	-0,002	-0,005
E	-0,001	0,005	0,000	0,000	0,000	-0,000	0,002	-0,002	-0,002	-0,000	0,002	-0,001	-0,000	-0,001
IS	-0,001	-0,000	-0,001	-0,001	-0,000	0,000	0,002	-0,001	-0,001	-0,000	-0,001	-0,000	-0,000	0,006
T	0,015	0,001	0,011	0,033	0,002	0,001	0,000	0,033	0,037	-0,000	-0,006	0,013	0,008	0,020
Ka	-0,003	-0,001	-0,004	-0,006	-0,002	-0,001	0,033	0,000	-0,005	0,004	-0,002	-0,002	-0,001	-0,003
Ki	-0,004	-0,001	-0,005	-0,006	-0,002	-0,001	0,036	-0,005	0,000	-0,001	-0,003	-0,002	-0,001	-0,002
L	-0,000	-0,000	-0,000	-0,001	-0,000	-0,000	0,000	0,004	-0,001	0,000	-0,000	-0,000	-0,000	-0,000
N	-0,002	0,000	0,018	-0,002	0,002	-0,001	-0,005	-0,003	-0,003	-0,000	0,000	-0,001	-0,001	-0,002
NM	-0,002	-0,000	-0,002	-0,002	-0,001	-0,000	0,012	-0,002	-0,002	-0,000	-0,001	0,000	0,001	-0,001
P	-0,001	-0,000	-0,001	-0,002	-0,000	-0,000	0,008	-0,001	-0,001	-0,000	-0,001	0,001	0,000	-0,001
S	-0,003	-0,001	-0,004	-0,005	-0,001	0,007	0,020	-0,003	-0,002	-0,000	-0,002	-0,001	-0,001	0,000

4. Conclusions

Administrative boundaries usually have a historical background that does not automatically correspond to current links and relations. Therefore, it is not necessarily the appropriate way to determine socioeconomic localities.

At the same time, the recent systemic crisis and the resulting deficiencies amplify the importance of spatially specified, active policies. Locally accustomed planning and applications are the most secure way to counteract regional inequalities and to favour a more balanced regional economic development. This is especially the case when we deal with active employment policies. Spatial dimensions contribute greatly to the understanding of the labour market and thereby to the development of well adjusted measures.

Consequently, it is of great importance to develop methodologies that ensure an efficient determination of each socioeconomic locality. This is exactly the contribution of the present paper: based on the tradition of the gravity model, we develop an innovative algorithm that considers current demographic data, various socioeconomic relations, measurable or not, spatial dimension and transport infrastructure, in order to provide an objective segregation of local economies and societies.

In the second part of the paper we proceed with a pilot application of the developed methodology using labour mobility data from the main cities in the region of Central Macedonia. Despite the fact that we use a less detailed analysis of agglomerations, excluding from our sample those that have less than 20,000 inhabitants, and although we lack similar applications that would provide us with a set of comparable estimates simplifying the decision upon the clustering analysis, the results of our methodology prove the existence of significant discontinuities between the socioeconomic and the administrative segregation of the area. This verifies the inappropriateness of simply following the administrative boundaries and reveals the usefulness of our proposal. As far as optimal determination of socioeconomic locality is a prerequisite for effective planning, the developed methodology could be an important contribution for the enhancement of applied policies in the future.

References

Ballas D., Clarke G.P. (2000): GIS and microsimulation for local labour market policy analysis, Computers, Environment and Urban Systems, 24, pp. 305-330.

Casado-Diaz J.M. (2000): Local Labour Markets in Spain: A case study. Taylor and Francis Journals, Vol. 34 (9), pp. 843-856.

Cieślik Andrzej (2009): Bilateral trade volumes, the gravity equation and factor proportions, Journal of International Trade & Economic Development, Taylor and Francis Journals, 18(1), pp. 37-59.

Converse, P. D. (1949): New Laws of Retail Gravitation. Journal of Marketing, Volume 14, January, pp. 379-384.

Deardorff A.V. (1998): Determinants of bilateral trade: Does gravity work in a neoclassical world? in: Frankel J.A. (eds), The regionalization of the world economy. Chicago: The University of Chicago Press.

Duncan Bowie (2010): People, Planning and Homes in a World. Routledge

Leamer Edward E., Levinsohn James (1995): International Trade Theory, the Evidence. In Grossman and Rogoff (eds), Handbook of International Economics 3. North-Holland.

Rose A. (2000): One money, one market: the effect of common currencies on trade. Economic Policy, 30, pp. 9-45.

Zarotiadis G., Stamboulis M. (2011): Using Gravity Models in Regional Socioeconomic Policy: pilot applications in active labour market policies, Portuguese Journal of Quantitative Methods, Vol 1-II, pp. 63-73.