Central European Journal of Economic Modelling and Econometrics

Regional Economic Impact Assessment with Missing Input-Output Data: A Spatial Econometrics Approach for Poland

Andrzej Torój*

Submitted: 26.02.2016, Accepted: 5.05.2016

Abstract

We propose a method of constructing multisector-multiregion input-output tables, based on the standard multisector tables and the tools of spatial econometrics. Voivodship-level (NUTS-2) and subregion-level data (NUTS-3) on sectoral value added is used to fit a spatial model, based on a modification of the Durbin model. The structural coefficients are calibrated, based on I-O multipliers, while the spatial weight matrices are estimated as parsimoniously parametrised functions of physical distance and limited supply in certain regions. We incorporate additional restrictions to derive proportions in which every cross-sectoral flow should be interpolated into cross-regional flow matrix. All calculations are based on publicly available data. The method is illustrated with an example of regional economic impact assessment for a generic construction company located in Eastern Poland.

Keywords: input-output model, spatial econometrics, Durbin model, multiregion analysis, regional economic impact assessment

JEL Classification: C21, C31, C67, D57, R12, R15

*Warsaw School of Economics; e-mail: andrzej.toroj@sgh.waw.pl

1 Introduction

Input-output (I-O) models, originated as early as in late 1930s and elaborated in the seminal works of Leontief (1936; 1941), are nowadays extensively used i.a. in corporate social responsibility reports to demonstrate the economic effects of a company's activity. The total effects consist not only of the direct effects, i.e. being a direct consequence of the expenditures borne by a company itself, but also of indirect effects, created along the company's supply chain, and induced effects, resulting from additional households' demand created via increased earnings (cf. Miller and Blair, 2009, ch. 6). Such effects can be of interest for country-level policymakers as far as decisions regarding a big enterprise, a group of enterprises or even a whole branch are concerned.

However, not all decision problems that require economic impact assessment as an information input are scaled for consideration on the central level of policymaking. Many economic decisions (including effective taxation or spatial development plans) are taken on various levels of local administration. Economic impact assessment at the country level is troublesome in such cases, for at least three reasons. Firstly, local policymakers might not be directly interested in country-level effects related e.g. to employment or taxation. National I-O tables for Poland (and many other countries) do not provide information necessary to decompose the computed effects into regions, and hence only rule-of-thumb estimates are available for local-level assessment. Secondly, the same absolute economic effects of an investment or an enterprise's activity may appear modest as compared to the national economy, but still be economically significant in relative terms for the local economy. Thirdly, country-level impact assessment is insufficient for regional development issues under consideration at the central level.

The list of regional economic impact assessment exercises conducted in the literature is long and includes the following examples: occurence of earthquakes (Rose *et al.*, 1997), hurricanes (Hallegatte, 2008), presence of recreational fisheries (Steinback, 1999) or even beetles (Patriquin *et al.*, 2007), ecological footprints (Turner *et al.*, 2007), tourism (Horváth and Frechtling, 1999), location of airports (Hakfoort *et al.*, 2001), FIFA World Cup stadiums (Baade and Matheson, 2004) or national defence installations (Atkinson, 1993). This list is by no means exhaustive.

Since I-O tables are normally available at national level (e.g. Central Statistical Office in Poland, 2014), the effects computed from I-O model are attributable to national economy, but without any additional regional breakdown. As a result, central policymakers may receive economically insignificant and perhaps irrelevant information on the country level, while local policymakers end up without information that, in turn, would be economically significant and relevant for them. Two extreme, naive solutions to the problem appear at first glance. The first one is to attribute the whole, computed economic effect to the local economy. This generates obvious upward bias, as it is extremely unlikely that there are no spillovers into other regions in the whole supply chain. The second solution is to spread the country-wide effects between

www.czasopisma.pan.pl PAN POLSKA AKADEMIA NAUK www.journals.pan.pl

regions proportionately to some measure of their economic size, e.g. (sectoral) value added. This, in turn, produces downward bias for the region of interest, as some goods or services may indeed be mainly delivered locally. The true regional impact can be located somewhere in this (very wide) interval.

Folmer and Nijkamp (1985) make a general recommendation to use models incorporating a number of cross-regional feedbacks to evaluate such effects. The leading example of such an approach is I-O analysis in a multi-sector, multi-region model (Miller and Blair, 2009, ch. 3, pp. 76–101). This requires the usage of I-O table with one dimension being a Carthesian product of the sets of sectors and regions (see Table 1 for a 2-sector, 2-region example). Contrary to this handbook requirement, available tables for many countries (including Poland) are constructed as in Table 2, i.e. without the regional sub-dimension. The problem of constructing table 1 has been labelled in the literature as the 'GRIT' problem (see West, 1990) – generation of regional input-output tables. The problem is highly context-dependent, whereby the availability of data is a key circumstance to be taken into consideration.

Table 1: Example multi-sector, multi-region input-output table

		Sector A		Sector B	
		Region 1	Region 2	Region 1	Region 2
Sector A	Region 1	$x^{A;A}_{1;1}$	$x^{A;A}_{1;2}$	$\cdots$	
	Region 2	$x^{A;A}_{2;1}$	$x^{A;A}_{2;2}$		
Sector B	Region 1	$\vdots$		$\ddots$	
	Region 2				

Table 2: Example multi-sector input-output table

		Sector A		Sector B	
		Region 1	Region 2	Region 1	Region 2
Sector A	Region 1 Region 2	$x^{A;A}$		$\cdots$	
Sector B	Region 1 Region 2	$\vdots$		$\ddots$	

The aim of this study is to find a practical approximation to the illustrative Table 1 for Poland, based on the available country-level I-O matrix as of 2010 (illustrated by Table 2) and available regional data that carry more information about the regional structure of value added creation. The data available via Local Data Bank from Central Statistical Office in Poland include:

1. voivodship-level data (16 units at NTS-2 level): value added, output and intermediate use (by companies from a region) in 20 NACE sections;
2. subregion-level data (66 units at NTS-3 level): value added in 7 groups of NACE sections.

Another, more general purpose is to lay ground for a universal approach to economic impact assessment at regional levels when the I-O tables in sufficient disaggregation do not exist. This is, in general, applicable to any country, even with available regional tables at, say, NTS-2 level, when effects for a smaller entity (NTS-3 or less) should be estimated.
The rest of the paper is organized as follows. In Section 2, we review the previous literature related to regional I-O matrices and analyses. In Section 3 we present assumptions and a formal derivation of the proposed method of interpolation of Table 2 into Table 1. Section 4 demonstrates how spatial econometric analysis is applied to solving the problem set in Section 3, while Section 5 discusses the quantitative results of this application. In Section 6, an example, subregion-level simulation is demonstrated. Section 7 concludes.

2 Regional input-output tables: literature review

Systematic, theoretical foundations for regional I-O analysis were laid by Leontief and Strout (1963), and an extensive review of later methodological developments and improvements can be found e.g. in Miller and Blair (2009, ch. 3 and 8).
The earliest regional applications from 1950s were intended for single-region problems (see Isard and Kuenne, 1953), especially for Washington, and the tables were based predominantly on surveys. The most widespread mathematical technique to develop a single-region (but sub-national) I-O tables is location quotient technique, allowing to approximate both flows within a region, as well as a region's 'exports' to other regions and countries when applied within a more complex model (see McCann, 2007, for an overview). It is until the very recent applications that some authors regionalize national I-O tables using the location quotient technique (Flegg and Tohmo, 2013). A construction of multi-region tables also requires the construction of a map of cross-regional flows.
Multi-region analyses from 1950s involved the cases of USA and Italy (Chenery, 1953; Moses, 1955). The US regions (in different configurations) were subject to the most extensive research on the topic, and hence the most popular contemporaneous application of regional I-O tables to the USA economy is the US MRIO model by Polenske (2004), consisting of 51 regions and 79 sectors. A number of multiregion models was created for Asian economies. Okamoto and Ihara (2005) elaborated a model for China with 30 sectors and 8 regions. Sub-national trade patterns across Japanese islands were analysed by Sonis *et al.* (2000). A 9-region, 25-sector table for Japan was created by Akita and Kataoka (2002). There are also multinational

applications of the multi-region framework: for the EU, the ASEAN-5 group, or the world's leading economies (WIOD database).
The baseline approach, proposed by Leontief and Strout (1963), was based on applying gravity models to data from individual regions. This method later evolved into hybrid methods, combining the gravity model approach with other data sources, including expert estimates (West, 1990). Hulu and Hewings (1993), in a model for Indonesia, impose further balancing restrictions. Some authors derive their tables from survey methods, e.g. commodity inflow survey in the USA (Liu and Vilain, 2004). The model MulitREG for Austria (Fritz *et al.*, 2001) is another example of survey approach in construction of cross-regional flows. Canning and Wang (2005) consider the construction of multiregion I-O tables as a constrained optimization problem. An extensive overview of recent applications and modelling directions is provided by Wiedmann *et al.* (2011).
The attempts to apply a regional I-O analysis for Poland are scarce. Welfe *et al.* (2008, chapter 1) apply the multiplier analysis to identify locally dominant branches in łódzkie voivodship. Chrzanowski (2013) constructed an I-O table for lubelskie voivodship based on location quotient technique. Currently, extensive efforts are taken to compute voivodship-level tables for Poland (Zawalińska *et al.*, 2014).
Our analysis can be located in the strand of literature based on gravity approach by Leontief and Strout (1963), further explored by Theil (1967) and generally positively validated by Polenske (1970). Later applications of this approach include Uribe *et al.* (1966), Gordon (1976) and Lindall *et al.* (2006). However, to the best of our knowledge, there have not been any previous attempts to formalize the problem of estimating multiregion I-O tables based on spatial econometric tools for Poland. According to Loveridge (2004), this could be a promising direction in the strand of so-called integrated econometric-input-output models. Another contribution of this paper is the estimation of such tables for Poland at sub-regional level, as we consider the voivodship level insufficient for many practical, small-size policy problems.

3 Feasible approximation to regional I-O matrix blocks: underlying assumptions

To focus our attention, let us consider a 4-region economy for which country-level, multisector I-O tables are available. Let $x^{s;v}$ denote the use of intermediate products from sector s in sector v as the s,v-th element of multisector I-O matrix $X^{S\times S}$ (S – number of sectors). We look for the following interpolation:

$$\mathbf{x^{s;v}} = \begin{bmatrix} x_{1;1}^{s;v} & x_{1;2}^{s;v} & x_{1;3}^{s;v} & x_{1;4}^{s;v} \\ x_{2;1}^{s;v} & x_{2;2}^{s;v} & x_{2;3}^{s;v} & x_{2;4}^{s;v} \\ x_{3;1}^{s;v} & x_{3;2}^{s;v} & x_{3;3}^{s;v} & x_{3;4}^{s;v} \\ x_{4;1}^{s;v} & x_{4;2}^{s;v} & x_{4;3}^{s;v} & x_{4;4}^{s;v} \end{bmatrix} \tag{1}$$

in which $x_{r;p}^{s;v}$ is the use of products from sector s, region r in sector v, region p. For a given pair of sectors (s, v) the sum of cross-regional, estimated flows must be equal to the known value from the I-O table without regional breakdown:

$$\sum_{r,p} x_{r;p}^{s;v} = x^{s;v} \tag{2}$$

Performing the interpolation, we set the following principles:

1. With the sectoral cost structure given, the recipient of a good is more likely to select a supplier in a less distant region, *ceteris paribus.* This reflects the basic intuition, and the quantification of this relationship is left for the econometric analysis.

2. With the sectoral cost structure given, the recipient is more likely to select a supplier in a region where the supply of certain type of goods is more abundant, *ceteris paribus.* This assumption assures that no artificial effects are created in the regions that are unable to supply, even to nearby locations (e.g. a greenfield investment in a charcoal power plant will not increase the coal mining in the nearby, if no coal mines exist).

3. The impact of the abovementioned two factors depends on the sort of good, i.e. the supplying sector, but not on the demanding sector. For example, an entrepreneur may prefer to buy agricultural products locally, but there may be no role of distance for manufacturing products – and it does not matter whether the recipient is an enterprise from a food industry or from the chemical industry.

4. There is no substitution between goods from different sectors and physical locations. We do not take into account the fact that missing local availability of some product (e.g. local repair services) may create incentives to switch to a different sector's goods (e.g. ordering a new device from a remote manufacturer). This is consistent with the assumption of constant sectorial cost structure (see assumptions 1 and 2).

5. The I-O proportions for Poland as a whole should still hold on the regional level. The sum of effects simulated via multi-region model should be equal to the effects simulated via country-level model.

6. Sectoral technology does not vary between regions. This implies the same sectoral cost structures, though they may vary as regards the geographical origins of the intermediate goods.

Interpolating (1), we follow a 2-step approach. In step 1, we split the inter-sectoral flow into column sums of the matrix (1):

$$x^{s;v} \rightarrow \begin{bmatrix} \sum_r x_{r;1}^{s;v} & \sum_r x_{r;2}^{s;v} & \sum_r x_{r;3}^{s;v} & \sum_r x_{r;4}^{s;v} \end{bmatrix} \tag{3}$$

Column sums in a block are allocated proportionately to the shares of individual regions in total value added of the demanding sector:

$$\sum_r x_{r;p}^{s;v} = \frac{va_p^v}{\sum_p va_p^v} x^{s;v} \tag{4}$$

where va_p^v denotes value added in sector v, region p. This reflects assumption (6), as identical cost structures within the same sector between regions imply identical cost-to-VA ratios. By using (4) to accomplish interpolation (3), we treat the sectoral technology as region-independent, in the sense that the proportions between inputs of intermediate goods do not vary from region to region, and also assume that labour-intensity and gross profitability of output is a sectoral, but not regional attribute.
In step 2, column sums have to be translated into individual column elements:

$$\begin{bmatrix} \sum_r x_{r;1}^{s;v} & \sum_r x_{r;2}^{s;v} & \sum_r x_{r;3}^{s;v} & \sum_r x_{r;4}^{s;v} \end{bmatrix} \rightarrow \begin{bmatrix} x_{1;1}^{s;v} & x_{1;2}^{s;v} & x_{1;3}^{s;v} & x_{1;4}^{s;v} \\ x_{2;1}^{s;v} & x_{2;2}^{s;v} & x_{2;3}^{s;v} & x_{2;4}^{s;v} \\ x_{3;1}^{s;v} & x_{3;2}^{s;v} & x_{3;3}^{s;v} & x_{3;4}^{s;v} \\ x_{4;1}^{s;v} & x_{4;2}^{s;v} & x_{4;3}^{s;v} & x_{4;4}^{s;v} \end{bmatrix} \tag{5}$$

This step is less obvious and the use of spatial econometric analysis in this paper is intended exactly for this point. The technicalities do not result here directly from assumptions 1-6. Given the quantities from sector s that sector v in each region wants to order, we need to allocate them between regions on the supply side. We describe this allocation by a supply-sector-specific matrix $\mathbf{W^s}$ (cf. assumption 3). This matrix should be composed of columns with non-negative, real elements. For a given demanding sector, a given region where demanding company is located and a given sector supplying intermediate goods, the elements in columns will indicate proportions between the regions of origin of demanded intermediate good. Every column describes one region where demanding companies from a given sector are located.
Two comments should be made as regards matrices $\mathbf{W^s}$. Firstly, the only information content of interest are the proportions between the elements in a column. Neither the proportions between column sums are relevant (as we define (3) before), nor the scaling of the entire matrix. However, additional macroeconomic assumption (5) allows to uniquely identify $\mathbf{W^s}$ matrices by imposing all column sums equal to unity, as will be proven in Section 4. Secondly, matrices $\mathbf{W^s}$ could in principle be derived for each sector pair individually as $\mathbf{W^{s;v}}$. However, assumption (3) allows to limit the number of $\mathbf{W^s}$ matrices to be found to S (from $S \times S$), as it implies that $\mathbf{W^s}$ should be a characteristic of the supplying sector (the sort of intermediate good), not the demanding sector. In practice, we assume here that e.g. coal-demanding companies represent the same geographical map of purchasing preferences, regardless of the sort of goods they produce.

We perform step 2 of the interpolation as in (5) by spreading the column sums into individual rows proportionately to the respective elements in the respective columns of $\mathbf{W^s}$:

$$x^{s;v}_{r;p} = \left(\sum_r x^{s;v}_{r;p}\right) \mathbf{W^s}(r,p) \tag{6}$$

with $\mathbf{W^s}(r,p)$ indicating r,p-th element of matrix $\mathbf{W^s}$. Equation (6) may be viewed as a modification of Leontief's and Strout's (1963) gravity formula that allows for an empirical investigation in the spatial econometrics framework.

Our principles of interpolation (1) are summarized in Table 3, for an example, 2-sector 2-region economy with sectors A and B and regions 1 and 2. Note that $\mathbf{va^v} = \begin{bmatrix} va^v_1 & va^v_2 & \dots & va^v_R \end{bmatrix}^T$ denotes the vector of value added in sector v, in various regions (R – number of regions). While all the elements $x^{s;v}$ and vectors $\mathbf{va^v}$ are known, the matrices $\mathbf{W^s}$ remain unobservable. The next section explains the approach adopted for their estimation, for each s.

Table 3: Information inputs for constructing multisector-multiregion table (illustration)

		Sector A		Sector B	
		Region 1	Region 2	Region 1	Region 2
Sector A	Region 1	block based on:		block based on:	
	Region 2	$x^{A,A}$; $\left(\mathbf{va^A}\right)^T$;$\mathbf{W^A}$		$x^{A,B}$; $\left(\mathbf{va^B}\right)^T$;$\mathbf{W^A}$	
Sector B	Region 1	block based on:		block based on:	
	Region 2	$x^{B,A}$; $\left(\mathbf{va^A}\right)^T$;$\mathbf{W^B}$		$x^{B,B}$; $\left(\mathbf{va^B}\right)^T$;$\mathbf{W^B}$	

4 Spatial econometric analysis: methodology

In this Section, we derive the equations that serve the purpose of estimating $\mathbf{W^s}$ matrices for all s, along with the accompanying restrictions. A number of authors, including Fritz *et al.* (2001), acknowledge the advantages from combining the econometric approach and the input-output models. However, to the best of our knowledge, in the previous literature these efforts were related to the estimation of parameters in private consumption block (Fritz *et al.* 2001) or to dynamisation of input-output coefficients (Kratena and Zakarias, 2004) rather than as a cost-efficient and internationally replicable method of estimating cross-regional flow map in input-output tables, especially at a high level of disaggregation.

The demand responsible for the creation of value added in sector s (va^s) is either intermediate demand from other sectors (proportional to value added in these sectors) or final demand (y^s):

$$va^s = \sum_{v=1}^{S} \beta_v^s va^v + \beta_0^s y^s \tag{7}$$

It is econometrically infeasible (and economically senseless) to estimate the parameters β_0^s and β_v^s in a linear, cross-sector regression equation. Econometrically, there are too many parameters to estimate and, additionally, va^s would be regressed on itself, not to mention endogeneity problem arising when equations for all sectors s are taken into account. More importantly, in economic terms, β_0^s and β_v^s are not free parameters as long as the validity of constant, countrywide I-O ratios is assumed (see assumption 5 in Section 3).

The specification of equation (7) reflects the fact that there are $S + 1$ sources of demand for global output in sector s: intermediate demand from S sectors plus the final demand. Let's focus on v-th sector's impact on va^s to calibrate the coefficient β_v^s:

Δva^v generates global output in sector v equal to $\Delta x^v = \frac{x^v}{va^v}\Delta va^v$ (whereby x^v denotes global output in sector v);

additional global output in sector v translates into intermediate demand for goods produced by sector s, equal to $a^{s,v}\Delta x^v$ (whereby $a^{s,v}$ denotes s,v-th element of cost structure matrix $\mathbf{A^{S\times S}}$);

this additional intermediate demand becomes part of global output in sector s, x^s, hence $\Delta x^s = a^{s,v}\Delta x^v$;

value added in sector s grows in respective proportion to the global output in the same sector: $\Delta va^s = \frac{va^s}{x^s}\Delta x^s$;

collecting terms, we obtain: $\Delta va^s = \underbrace{\frac{va^s}{x^s} a^{s;v} \frac{x^v}{va^v}}_{\beta_v^s} \Delta va^v$ and for each s and v, β_v^s can be computed as:

$$\beta_v^s = \frac{va^s}{x^s} a^{s,v} \frac{x^v}{va^v} \tag{8}$$

Computation of β_0^s is straightforward as y^s directly becomes part of global output in s, hence $\Delta va^s = \frac{va^s}{x^s}\Delta x^s = \frac{va^s}{x^s}\Delta y^s$ and:

$$\beta_0^s = \frac{va^s}{x^s} \tag{9}$$

Under calibration (8) and (9), equations (7) – for every sector s – become identities for the period and country, for which the I-O ratios $\frac{va^s}{x^s}$, $a^{s,v}$ and $\frac{x^v}{va^v}$ were derived. However, any disaggregation or extrapolation, either in space or time, render this

equation stochastic.
Let us consider such a regional disaggregation of (7) in the case of a 3-sector, 4-region economy. Let us focus on sector s, and let va_r^v denote value added in sector v (including s) in region $r = 1, 2, 3, 4$:

$$\begin{bmatrix} va_1^s \\ va_2^s \\ va_3^s \\ va_4^s \end{bmatrix} = \begin{bmatrix} va_1^1 & va_1^2 & va_1^3 \\ va_2^1 & va_2^2 & va_2^3 \\ va_3^1 & va_3^2 & va_3^3 \\ va_4^1 & va_4^2 & va_4^3 \end{bmatrix} \begin{bmatrix} \beta_1^s \\ \beta_2^s \\ \beta_3^s \end{bmatrix} + \begin{bmatrix} y_1^s \\ y_2^s \\ y_3^s \\ y_4^s \end{bmatrix} \beta_0^s + \begin{bmatrix} \varepsilon_1^s \\ \varepsilon_2^s \\ \varepsilon_3^s \\ \varepsilon_4^s \end{bmatrix} \tag{10}$$

A simple transformation of (7) to (10) assumes an extremely unlikely case of regional autarky. For example, it assumes that value added in sector 1, region 1 only depends on intermediate and final demand in the same region, and *vice versa* – intermediate and final demand in region 1 only affects value added in region 1. Note that an error term appears in (10), as it is no more an identity at the local level.
To remove the assumption of autarky, let us introduce a weighting scheme

$$\mathbf{W^s} = \begin{bmatrix} w_{11}^s & w_{12}^s & w_{13}^s & w_{14}^s \\ w_{21}^s & w_{22}^s & w_{23}^s & w_{24}^s \\ w_{31}^s & w_{32}^s & w_{33}^s & w_{34}^s \\ w_{41}^s & w_{42}^s & w_{43}^s & w_{44}^s \end{bmatrix},$$

sized 4×4 in accordance with the number of regions:

$$\begin{aligned} \begin{bmatrix} va_1^s \\ va_2^s \\ va_3^s \\ va_4^s \end{bmatrix} &= \begin{bmatrix} w_{11}^s & w_{12}^s & w_{13}^s & w_{14}^s \\ w_{21}^s & w_{22}^s & w_{23}^s & w_{24}^s \\ w_{31}^s & w_{32}^s & w_{33}^s & w_{34}^s \\ w_{41}^s & w_{42}^s & w_{43}^s & w_{44}^s \end{bmatrix} \begin{bmatrix} va_1^1 & va_1^2 & va_1^3 \\ va_2^1 & va_2^2 & va_2^3 \\ va_3^1 & va_3^2 & va_3^3 \\ va_4^1 & va_4^2 & va_4^3 \end{bmatrix} \begin{bmatrix} \beta_1^s \\ \beta_2^s \\ \beta_3^s \end{bmatrix} \\ &+ \begin{bmatrix} w_{11}^s & w_{12}^s & w_{13}^s & w_{14}^s \\ w_{21}^s & w_{22}^s & w_{23}^s & w_{24}^s \\ w_{31}^s & w_{32}^s & w_{33}^s & w_{34}^s \\ w_{41}^s & w_{42}^s & w_{43}^s & w_{44}^s \end{bmatrix} \begin{bmatrix} y_1^s \\ y_2^s \\ y_3^s \\ y_4^s \end{bmatrix} \beta_0^1 + \begin{bmatrix} \varepsilon_1^s \\ \varepsilon_2^s \\ \varepsilon_3^s \\ \varepsilon_4^s \end{bmatrix} \end{aligned} \tag{11}$$

Now sector 2 impacts sector 1, region 1 not just via va_1^2, but via $w_{11}^1 va_1^2 + w_{12}^1 va_2^2 + \dots$. Note that the same matrix is applied to all regressors in va^1 equation. This is a direct

www.czasopisma.pan.pl PAN www.journals.pan.pl
POLSKA AKADEMIA NAUK

consequence of assumption (3) in Section 3. By rearranging terms, we obtain:

$$\begin{bmatrix} va_1^s \\ va_2^s \\ va_3^s \\ va_4^s \end{bmatrix} = \begin{bmatrix} w_{11}^s & w_{12}^s & w_{13}^s & w_{14}^s \\ w_{21}^s & w_{22}^s & w_{23}^s & w_{24}^s \\ w_{31}^s & w_{32}^s & w_{33}^s & w_{34}^s \\ w_{41}^s & w_{42}^s & w_{43}^s & w_{44}^s \end{bmatrix} \begin{bmatrix} va_1^s \\ va_2^s \\ va_3^s \\ va_4^s \end{bmatrix} \beta_s^s + \begin{bmatrix} w_{11}^s & w_{12}^s & w_{13}^s & w_{14}^s \\ w_{21}^s & w_{22}^s & w_{23}^s & w_{24}^s \\ w_{31}^s & w_{32}^s & w_{33}^s & w_{34}^s \\ w_{41}^s & w_{42}^s & w_{43}^s & w_{44}^s \end{bmatrix} \begin{bmatrix} va_1^{v_1} & va_1^{v_2} & y_1^s \\ va_2^{v_1} & va_2^{v_2} & y_2^s \\ va_3^{v_1} & va_3^{v_2} & y_3^s \\ va_4^{v_1} & va_4^{v_2} & y_4^s \end{bmatrix} \begin{bmatrix} \beta_{v_1}^s \\ \beta_{v_2}^s \\ \beta_0^s \end{bmatrix} + \begin{bmatrix} \varepsilon_1^s \\ \varepsilon_2^s \\ \varepsilon_3^s \\ \varepsilon_4^s \end{bmatrix} \tag{12}$$

whereby $v_1, v_2 \neq s$. The above formulation is a special case of the Durbin model in spatial econometrics (cf. LeSage and Pace, 2009, p. 82):

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\alpha}_\mathbf{1} + \mathbf{W}\mathbf{X}\boldsymbol{\alpha}_\mathbf{2} + \boldsymbol{\varepsilon} \tag{13}$$

with

$$\mathbf{y} = \mathbf{va^s}, \quad \boldsymbol{\alpha}_\mathbf{1} = \mathbf{0}, \quad \rho = \beta_s^s, \quad \mathbf{W_s} = \mathbf{W}$$

$$\boldsymbol{\alpha}_\mathbf{2} = \begin{bmatrix} \beta_{v_1}^s \\ \beta_{v_2}^s \\ \beta_0^s \end{bmatrix}$$

and

$$\mathbf{X} = \begin{bmatrix} \mathbf{va^{v_1}} & \mathbf{va^{v_2}} & \mathbf{y^s} \end{bmatrix}.$$

Note that

$$\mathbf{y^s} = \begin{bmatrix} y_1^s \\ y_2^s \\ y_3^s \\ y_4^s \end{bmatrix}.$$

Note the following differences between (12) and the standard Durbin model:

1. There are no local regressors without spatial interactions, $\boldsymbol{\alpha}_\mathbf{1} = \mathbf{0}$.
2. This is a multi-equation model (there is one block of equations for each sector, equation (12) is for sector s only).
3. It is not ρ and $\boldsymbol{\alpha}_\mathbf{2}$ that will be estimated, but $\mathbf{W^s}$. In the standard Durbin model, the opposite is the case: $\mathbf{W}$ is given and the rest is estimated.
4. There is no reason to set $\mathbf{W^s}$ as row-stochastic matrices, as we do not need an interpretation of ρ as the spatial autocorrelation coefficient.

In line with assumption (5) from Section 3, we want to comply with the I-O ratios and multipliers at the country level, regardless of regional disaggregation. In other words, a unit increase in va^v should lead to β_v^s increase in va^s at the country level, for any geographical distribution of the impulse in va^v and response in va^s. One can prove that this is the case when the columns of $\mathbf{W^s}$ sum to unity, for all s. To see this, sum the impacts of intermediate demand changes on $\mathbf{va^s}$ over regions:

$$
\begin{aligned}
& \Delta va_1^s + \Delta va_2^s + ... = \\
= \; & \left(\beta_1^s \sum_r w_{1r}^s \Delta va_r^1 + \beta_2^s \sum_r w_{1r}^s \Delta va_r^2 + \beta_3^s \sum_r w_{1r}^s \Delta va_r^3\right) + \\
& + \left(\beta_1^s \sum_r w_{2r}^s \Delta va_r^1 + \beta_2^s \sum_r w_{2r}^s \Delta va_r^2 + \beta_3^s \sum_r w_{2r}^s \Delta va_r^3\right) + ... = \\
= \; & \beta_1^s \sum_r \left(w_{1r}^s \Delta va_r^1 + w_{2r}^s \Delta va_r^1 + ...\right) + \beta_2^s \sum_r \left(w_{1r}^s \Delta va_r^2 + w_{2r}^s \Delta va_r^2 + ...\right) + \\
& + \beta_3^s \sum_r \left(w_{1r}^s \Delta va_r^3 + w_{2r}^s \Delta va_r^3 + ...\right) = \\
= \; & \beta_1^s \sum_r \left[\Delta va_r^1 \left(\sum_p w_{pr}^s\right)\right] + \beta_2^s \sum_r \left[\Delta va_r^2 \left(\sum_p w_{pr}^s\right)\right] + \beta_3^s \sum_r \left[\Delta va_r^3 \left(\sum_p w_{pr}^s\right)\right]
\end{aligned}
\tag{14}
$$

The three components in (14) are associated with three different sectors in which an impulse in intermediate demand may arise. Consider a unit change in any of them, $\Delta va^v = 1$. To comply with country-wide I-O ratios, va^s must change by β_v^s, and consequently:

$$
\underset{v}{\forall} \quad \sum_r \left[\Delta va_r^v \left(\sum_p w_{pr}^s\right)\right] = 1 \tag{15}
$$

Consider the extreme case when the unit growth of va^v is concentrated in a single region r, i.e. $\Delta va_r^v = 1$. Then:

$$
\underset{r,s}{\forall} \quad \sum_p w_{pr}^s = 1 \tag{16}
$$

i.e. the sum of every column r in every matrix $\mathbf{W^s}$ must be equal to 1. If these equalities hold, the condition to preserve the country-level I-O ratios also holds in the case when growth of va^v is distributed over multiple regions.

Still, there are far too many elements of matrices $\mathbf{W^s}$ to estimate them freely. We propose a parsimonious parametrisation based on assumptions (1) and (2) from Section 3, i.e. the principle that the weights for supplying regions are a function of (i) distance to demanding region, (ii) sector-specific supply constraints:

$$
\mathbf{W^s} = f\left(\mathbf{W^s_{distance}}, \mathbf{W^s_{supply}}\right) \tag{17}
$$

The distance-based argument is related to the known symmetric distance matrix $\mathbf{W}^*$, sized $R \times R$. The elements of this matrix are physical distances between centroids of regions at NTS-2 level (voivodships) and NTS-4 (counties). For subregions (NTS-3), centroid for its most populated NTS-4 has been selected as a representative. In our simple example:

$$\mathbf{W}^* = \begin{bmatrix} 0 & w_{12}^* & w_{13}^* & w_{14}^* \\ . & 0 & w_{23}^* & w_{24}^* \\ . & . & 0 & w_{34}^* \\ . & . & . & 0 \end{bmatrix} \tag{18}$$

Individual elements (r, p) of $\mathbf{W^s_{distance}}$ were computed using the gamma cumulative distrubution function $\Gamma(.)$ as:

$$w^s_{distance}(r,p) = \frac{1 - \Gamma\left(\theta_1^s, \theta_2^s, w^*_{r,p}\right)}{\sum_p \left[1 - \Gamma\left(\theta_1^s, \theta_2^s, w^*_{r,p}\right)\right]} \tag{19}$$

θ_1^s and θ_2^s denote sector-specific distance and shape parameters. The choice of the gamma-function-based approach (cf. Jackson *et al.*, 2006, for a similar problem) was motivated by the fact that this functional form transforms the distance very flexibly, while its parametrisation is highly parsimonious. For example, it can be fitted to three different situations: (i) when local suppliers are strongly preferred, and the demanding company is relatively indifferent between supplier located 50 km and 1000 km away, (ii) local suppliers are preferred, but not strongly, to distant suppliers, and the utility from distance supplies is decreasing very gradually, (iii) the demanding company is relatively indifferent between supplies up to some threshold, e.g. from 0 to 100 km, above which the preference for supplies is decreasing sharply.
The supply-based argument is linked with the vector value added across regions, $\mathbf{va^s} = \begin{bmatrix} va_1^s \\ va_2^s \\ va_3^s \\ va_4^s \end{bmatrix}$, in the following way:

$$w^s_{supply}(r,p) = \left(\frac{va_r^s}{\sum_i va_i^s}\right)^{\gamma^s} \tag{20}$$

with $\gamma^s \in [0;1]$. For the extreme case of $\gamma^s = 1$, the probability of selecting a producer from a given location as a supplier of sector s good is proportional to the mass of supply offered in this region, relative to other regions, other thing being equal (i.e. distance). In the other extreme case, $\gamma^s = 0$, the supply factor does not matter and other regions are equally probable in this respect, so as to let other factors decide (i.e. distance).

As $\mathbf{W^s_{distance}}$ and $\mathbf{W^s_{supply}}$ shall be aggregated into a single matrix $\mathbf{W^s}$, a functional form for (17) needs to be specified. Our preferred aggregation method is multiplicative, because we intend to allow the two determinants interact. Some regions do not offer some sort of goods (e.g. charcoal), and these goods cannot be ordered by other sectors (e.g. energy sector) in the nearby even if their instances (e.g. power plants) are located in the proximity. The aggregate matrix in equation (17) is additionally rescaled so as to fulfil the condition (16) and consists of the following elements:

$$w^s_{rp} = \frac{w^s_{distance}(r,p) \cdot w^s_{supply}(r,p)}{\sum_r w^s_{distance}(r,p) \cdot w^s_{supply}(r,p)} \tag{21}$$

We perform a joint estimation of equations for all sectors, written in the block form as:

$$\begin{aligned} \begin{bmatrix} \mathbf{I_4} & & \\ & \mathbf{I_4} & \\ & & \mathbf{I_4} \end{bmatrix} \begin{bmatrix} \mathbf{va^1} \\ \mathbf{va^2} \\ \mathbf{va^3} \end{bmatrix} &= \begin{bmatrix} \beta^1_1\mathbf{W^1} & \beta^1_2\mathbf{W^1} & \beta^1_3\mathbf{W^1} \\ \beta^2_1\mathbf{W^2} & \beta^2_2\mathbf{W^2} & \beta^2_3\mathbf{W^2} \\ \beta^3_1\mathbf{W^3} & \beta^3_2\mathbf{W^3} & \beta^3_3\mathbf{W^3} \end{bmatrix} \begin{bmatrix} \mathbf{va^1} \\ \mathbf{va^2} \\ \mathbf{va^3} \end{bmatrix} \\ &+ \begin{bmatrix} \beta_{0,1}\mathbf{W}^1 \\ \beta_{0,2}\mathbf{W}^2 \\ \beta_{0,3}\mathbf{W}^3 \end{bmatrix} \begin{bmatrix} y^s \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon^1} \\ \boldsymbol{\varepsilon^2} \\ \boldsymbol{\varepsilon^3} \end{bmatrix} \end{aligned} \tag{22}$$

with $\mathbf{y} = \begin{bmatrix} y^1 \\ y^2 \\ y^3 \end{bmatrix}$. Based on (22), the following $\mathbf{A}$ and $\mathbf{B}$ matrices can be defined:

$$\underbrace{\left(\mathbf{I_{12}} - \begin{bmatrix} \beta^1_1\mathbf{W^1} & \beta^1_2\mathbf{W^1} & \beta^1_3\mathbf{W^1} \\ \beta^2_1\mathbf{W^2} & \beta^2_2\mathbf{W^2} & \beta^2_3\mathbf{W^2} \\ \beta^3_1\mathbf{W^3} & \beta^3_2\mathbf{W^3} & \beta^3_3\mathbf{W^3} \end{bmatrix}\right)}_{\mathbf{A}} \begin{bmatrix} \mathbf{va^1} \\ \mathbf{va^2} \\ \mathbf{va^3} \end{bmatrix} = \underbrace{\begin{bmatrix} \beta_{0,1}\mathbf{W}^1 \\ \beta_{0,2}\mathbf{W}^2 \\ \beta_{0,3}\mathbf{W}^3 \end{bmatrix}}_{\mathbf{B}} \begin{bmatrix} y \end{bmatrix} + \underbrace{\begin{bmatrix} \boldsymbol{\varepsilon^1} \\ \boldsymbol{\varepsilon^2} \\ \boldsymbol{\varepsilon^3} \end{bmatrix}}_{\boldsymbol{\varepsilon}} \tag{23}$$

Data about final output, $\mathbf{y}$, is only available from the I-O tables in sectoral breakdown, but no territorial breakdown is available. For the purpose of estimating system (23), we spread y^s for each sector s into y^s_r, whereby $y^s = \sum_r y^s_r$, in proportion to local populations. We assume thereby that the sectoral composition of final demand is the same across regions. This should not be controversial as regards household consumption, and possibly government consumption, but is clearly a strong assumption for investment expenditures and exports.

An ideal, but unfeasible indicator variable should be correlated to final output and be available at the regional and sectoral level (jointly). This excludes GDP (estimates exist at voivodship levels, but no sectoral breakdown on the distribution side) and most measures of 'economic activity' (that include intermediate demand, not only the final demand – and this differentiation is the clue for this regression being successful). The adopted approach is based on the following consideration: the final demand can be decomposed between consumption (including government consumption), capital formation and exports that, as of 2010, accounted for approximately 56%, 13% and 31% of the final demand, respectively. If we assume that consumption volumes and tastes per head do not differ significantly region by region, interpolation by the population allows to correctly approximate 56% of the target variable. However, the two remaining categories are non-negligible and this issue should definitely be addressed in future research.

Equation (23) is estimated in two versions: for voivodship- and subregion-level data. We do not opt in advance for a single model because there is a trade-off between sectoral and regional disaggregation of value added data. For 16 voivodships, a breakdown into as many as 20 NACE sections is available. For 66 subregions, there are only 7 groups of sections. This trade-off is non-negligible when we take into account the information content of the parameters to be estimated: the distance parameters (which are clearly more precise for subregion-level data, though only available for groups of sections at this level) and the supply-constraint parameters (which differ substantially between broad and narrow categories). Two examples can be evoked. Firstly, suppose that office cleaning services are ordered from firms located up to 30 km away. Voivodship-level data is far too imprecise to capture this. Secondly, imagine tertiary education as a sector: it appears to be quite supply-constrained when we look at subregions (universities are usually located in big, voivodship cities), but will not seem to be constrained in voivodship-level data, nor in subregion-level data when aggregated with primary and secondary education institutions.

For both variants of the model, we use annual panel data for the period 2008-2012. This period has been selected as centred around 2010, the year for which last available I-O tables for Poland are available. Recall that this table underlies the derivation of coefficients β_v^s and the approximation of vector $\mathbf{y}$, so that further expansion of the sample may be a source of bias. In this panel, vector $\mathbf{y}$ is time-invariant.

Adjusting the likelihood function in the Durbin model appropriately, we obtain:

$$lnL = -\frac{n}{2}\ln(2\pi) + n\ln|\mathbf{A}| - \frac{n}{2}\ln|\mathbf{D}| + \boldsymbol{\varepsilon}^T\mathbf{D}^{-1}\boldsymbol{\varepsilon} \tag{24}$$

We assume spherical $\boldsymbol{\varepsilon^1}$, $\boldsymbol{\varepsilon^2}$, $\boldsymbol{\varepsilon^3}$ and hence $E\left(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T\right) = \mathbf{D}$ being diagonal (but not spherical). The maximum is found using Matlab via simulated annealing as a global optimisation method. The standard errors are evaluated by the standard formula for nonlinear models (cf. Cameron and Trivedi, 2005, p. 156):

$$E\left(\hat{\boldsymbol{\alpha}}\hat{\boldsymbol{\alpha}}^T\right) = \mathbf{G}^T\mathbf{D}^{-1}\mathbf{G} \tag{25}$$

with $G = \frac{\partial \mathbf{g}(\boldsymbol{\alpha})}{\partial(\boldsymbol{\alpha})}|_{\hat{\boldsymbol{\alpha}}}$, $\mathbf{g}$ being the right-hand side of (22).

5 Spatial econometric analysis: results

5.1 Voivodship-level estimates

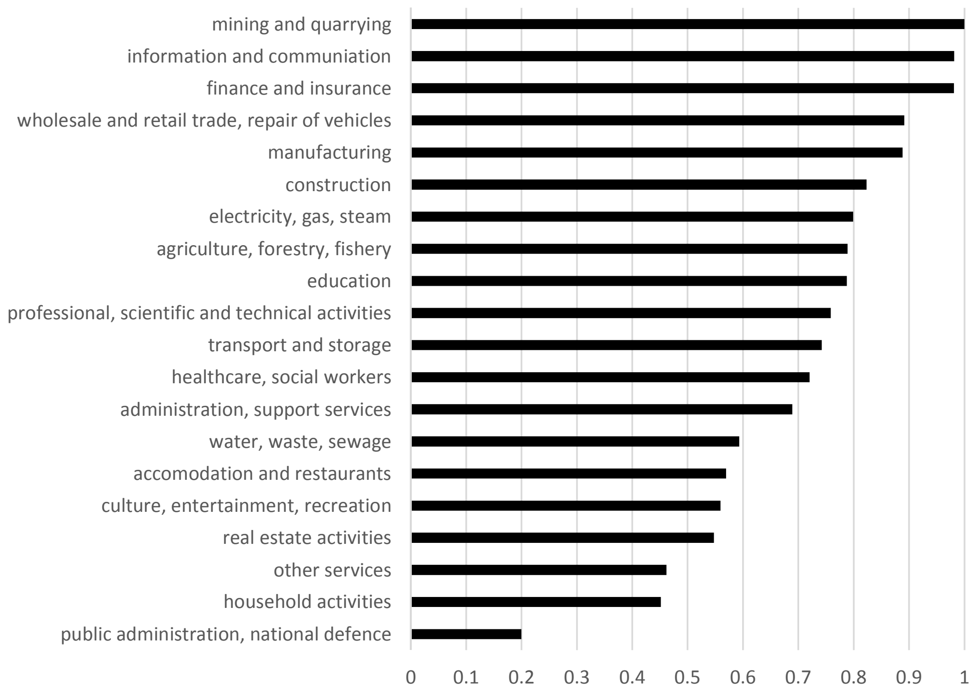
The estimates on the voivodship level seem to be precise enough to draw cross-sectoral, comparative conclusions (see Table 4). In line with expectations, the highest parameter value γ^s has been obtained for mining and quarrying (1.000), which implies the most restrictive supply constraints – mines are only located in some voivodships. Relatively high values can also be attributed to service sectors with a high share of Warsaw-based companies, i.e. information and communication, as well as finance and insurance. On the other hand, supply of public administration services seem to be the least supply-constrained (0.200), followed by two specific sectors: household activities (0.451), based on imputed data rather than any official reporting, and the residual, heterogeneous group of 'other services' (0.462).

Table 4: Estimation results for voivodships

sector (s)	γ^s	$SE(\gamma^s)$	θ_1^s	$SE\left(\theta_1^s\right)$	θ_s^2	$SE\left(\theta_2^s\right)$	var(ε)
agriculture, forestry, fishery	**0.789**	0.00027	**12.766**	0.038	**9.891**	0.027	**10.310**
mining and quarrying	**1.000**	0.00046	**12.080**	0.161	**19.197**	0.206	**11.786**
manufacturing	**0.888**	0.00005	**12.435**	0.019	**26.987**	0.038	**21.729**
electricity, gas, steam	**0.799**	0.00037	**8.939**	0.038	**18.846**	0.080	**7.311**
water, waste, sewage	**0.593**	0.00060	**11.182**	0.172	**23.981**	0.314	**6.734**
construction	**0.823**	0.00007	**21.995**	0.126	**22.652**	0.134	**8.251**
wholesale and retail trade, repair of vehicles	**0.892**	0.00004	**16.389**	0.031	**24.339**	0.046	**16.237**
transport and storage	**0.742**	0.00031	**3.490**	0.019	**19.867**	0.076	**11.252**
accomodation and restaurants	**0.569**	0.00061	**6.301**	0.067	**19.438**	0.187	**8.422**
information and communication	**0.981**	0.00022	**14.181**	0.048	**15.900**	0.050	**7.828**
finance and insurance	**0.981**	0.00022	**19.171**	0.083	**13.338**	0.053	**5.351**
real estate activities	**0.547**	0.00021	**8.450**	0.021	**13.546**	0.030	**13.108**
professional, scientific and technical activities	**0.759**	0.18397	**6.241**	3.151	**6.271**	1.688	**16.109**
administration, support services	**0.689**	0.00027	**13.783**	0.065	**15.826**	0.080	**8.768**
public administration, national defence	**0.200**	0.04396	**4.135**	0.283	**10.704**	0.301	**8.841**
education	**0.787**	0.00018	**8.122**	0.023	**25.146**	0.074	**9.278**
healthcare, social workers	**0.720**	0.00026	**6.728**	0.026	**17.439**	0.061	**7.071**
culture, entertainment, recreation	**0.559**	0.00066	**15.081**	0.255	**18.413**	0.279	**5.970**
other services	**0.462**	0.18225	**7.233**	47.345	**5.910**	25.374	**6.143**
household activities	**0.451**	0.00188	**12.727**	0.544	**25.422**	0.976	**6.775**

The distance parameters, θ_1^s and θ_2^s, should be interpreted jointly. For this purpose, Figure 2 presents the distance functions for products supplied by different sectors, computed according to (19). The demanding companies appear to be the most

Figure 1: Supply parameter: voivodship-level estimates

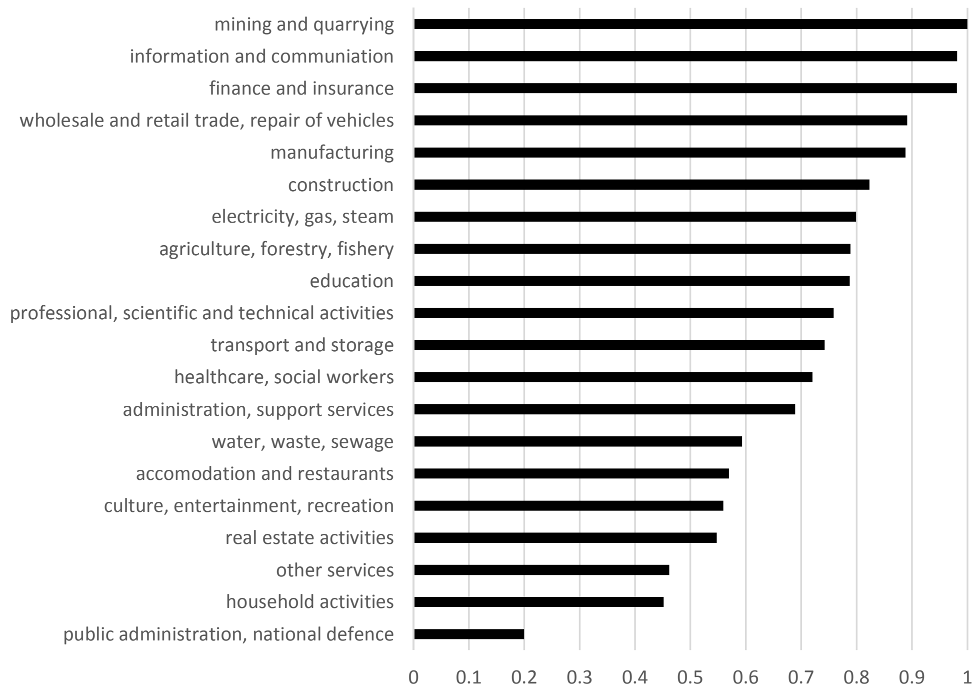


indifferent between local and remote supply of construction products – the function is almost flat from 0 to ca. 350 km. It is followed by wholesale and retail trade and manufacturing products (flat up to 200-250 km). On the other hand, the most 'local' branches are professional, scientific and technical services, public administration, transport and storage (likely driven by the latter) and 'other services'.

5.2 Subregion-level estimates

Less intuitive conclusions emerge when we analyse subregion-level data for 7 groups of NACE sections. The most supply-constrained category seems to be: trade, repair, transport, storage, accommodation, restaurants and communication (sections G, H, I, J). Note that none of these categories reached the value of $\gamma^s = 1.000$ with voivodship data. After the jump to finer territorial breakdown, much more spatial variation has emerged in these service categories and they seem to be more supply-constrained than on higher aggregation level. Sections G-J are followed by manufacturing and construction. Since both of these categories are represented individually in Table 4, we can compare the isolated impact of territorial aggregation. Neither manufacturing nor construction belongs to top-ranked sectors as regards the supply constraints on

Figure 2: Distance functions – voivodship-level estimates



the voivodship level. However, the breakdown of voivodships into subregions reveals uneven distribution of value added in these sections within voivodships, and hence boosts the role of constraints.

On the other side of the ranking, industry (except manufacturing) belongs to the least supply-constrained sectors. This is in spite of the inclusion of mining and quarrying sector, which seemed to be the most spatially supply-constrained when analysed in isolation. Interestingly, section A (agriculture, forestry and fishery) seems now to be less supply-constrained than on the voivodship level, and the least supply-constrained group of sections in Table 5. This counterintuitive result clearly deserves further investigation, and possible reasons may include at least the multiplicative specification and the resulting interaction with the estimate for distance (which has considerably changed), as well as inclusion of low variability within voivodships.

The most interesting subregion-level results are related to the distance functions (see Figure 3), which can be estimated much more precisely for subregions. A dichotomy can be observed. Some groups of sections are relatively indifferent between distant and local supplying up to 200-250 km. Unsurprisingly, the most distance-independent type of supply originates in the finance, insurance and real estate activity group (sections K and L). However, it is directly followed by agriculture, forestry and fishery (section

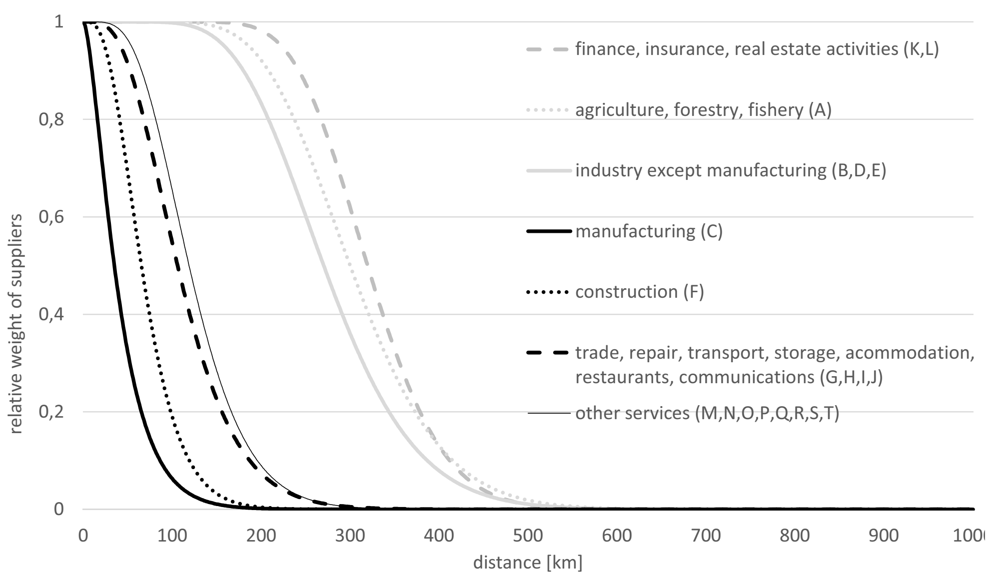
www.czasopisma.pan.pl PAN POLSKA AKADEMIA NAUK www.journals.pan.pl

Table 5: Estimation results for subregions

sector (s)	γ^s	$SE(\gamma^s)$	θ_1^s	$SE(\theta_1^s)$	θ_2^s	$SE(\theta_2^s)$	var(ε)
agriculture, forestry, fishery (A)	**0.552**	0.00019	**14.137**	0.074	**21.725**	0.103	1.323
industry except manufacturing (B,D,E)	**0.785**	0.00006	**11.588**	0.042	**24.007**	0.076	6.467
manufacturing (C)	**0.996**	0.00009	**1.723**	0.001	**24.888**	0.010	8.932
construction (F)	**0.974**	0.00012	**3.972**	0.004	**17.950**	0.015	4.687
trade, repair, transport, storage, acommodation, restaurants, communications (G,H,I,J)	**1.000**	0.00002	**4.450**	0.002	**25.715**	0.008	8.705
finance, insurance, real estate activities (K,L)	**0.871**	0.00003	**23.667**	0.079	**13.695**	0.043	3.040
other services (M,N,O,P,Q,R,S,T)	**0.939**	0.00002	**5.846**	0.003	**21.556**	0.010	4.940

A), which seems counterintuitive, at least in the case of agriculture (but not any more in the case of fishery). Also, industry except manufacturing exhibits indifference between short- and long-distance supplying, which is probably related to the presence of mining and quarrying in the aggregate. On the other hand, short-distance supplies are strongly preferred for goods of manufacturing (C), construction (F), and all sorts of services (sections G-J do not differ much from sections M-T).

Figure 3: Distance functions – subregion-level estimates



5.3 Voivodship versus subregion level results: is there room for coherent conclusions?

In two previous Subsections, the estimates for the voivodship and subregion level were discussed. To answer policy questions, however, one needs to provide one, consistent model, and a need for evaluation or unification of both approaches arises. As discussed before, there is a trade-off between sectorial and regional disaggregation, and the nature of estimated parameters clearly indicates that they may depend on the territorial and sectorial aggregation level. One may think of various estimation strategies incorporating both datasets. To understand their implications, one should systematically compare both result sets obtained before. To this aim, we compute weighted averages of 20 sectorial parameter vectors on the voivodship level so as to obtain 7 vectors of parameters corresponding to the groups of sectors from the subregion level, using the shares in global output from the I-O table as weights.

Figure 4 compares seven distance functions obtained in this way. The results are only consistent for 'other services' (sections M-T) and, to a lesser extent, industry ex. manufacturing (sections B, D, E). All other groups exhibit a qualitative difference. The following groups of sections appear to exhibit more tolerance for distance when analysed on subregion level: agriculture, forestry, fishery (A) and finance, insurance, real estate activities (K, L). On the contrary, manufacturing, construction and service sections G-J seem to be more indifferent to distance when analysed on voivodship level.

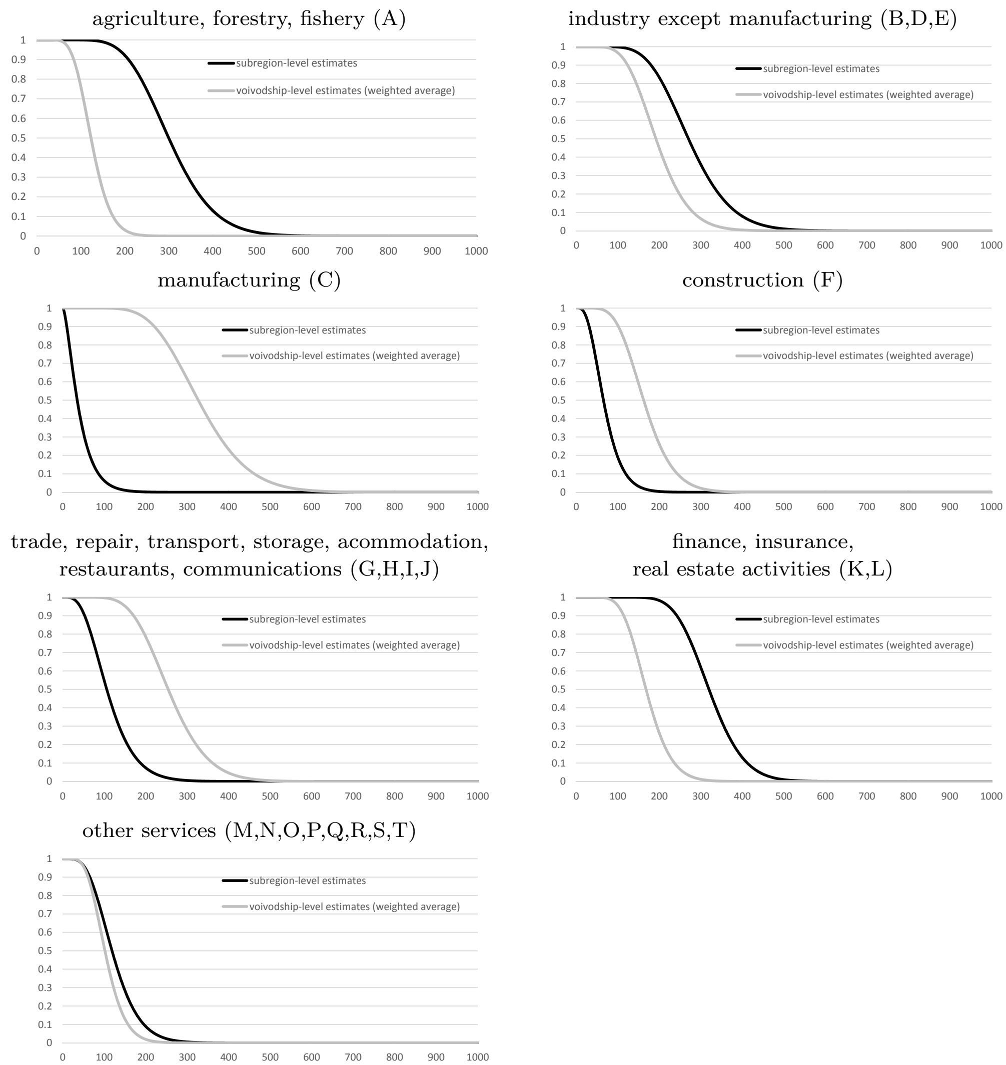
Importantly, the same line of division can be applied to the discrepancies in estimated parameters γ^s (Figure 5). The former groups of sections appear to be more constrained on the supply-side when analysed from the perspective of subregions, while the latter – from the perspective of voivodships. This clearly indicates that, for the same groups of sections, there is some interplay between distance and supply-constraint parameters depending on the levels of aggregation in the data.

The differences discussed above should not be seen as surprising, for the reasons mentioned in Section 4. In further analysis, we can adopt the following strategies to handle them:

1. Select one, more viable set of estimates, based on economic criteria. This solution has been adopted in simulation analysis in Section 6, where we opt for the results obtained with voivodship-level data (see Subsection 5.1). This is because sectors are ranked more in line with economic intuition, and because there is less bias in supply constraint parameters resulting from sectoral aggregation. Still, we disregard the fact that distance functions could be estimated more imprecisely than with subregion data.

2. Investigate the influence of spatial disaggregation in isolation from the sectorial disaggregation. This could be done by either (i) aggregating 20 sectors into 7 (for 16 voivodships) and comparing the estimation results with Subsection 5.1

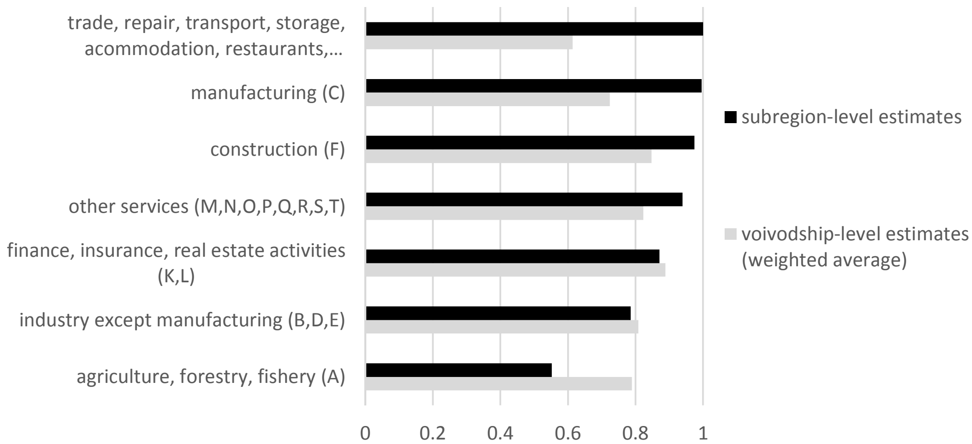
Figure 4: Subregion vs voivodship level estimates: gamma functions



or (ii) aggregating subregions into voivodships (for 7 sectors) and comparing the estimation results with Subsection 5.2.

3. Combine both datasets into a single estimation framework, stepwise or jointly. One could e.g. first estimate the distance functions for groups of sections with subregional data, and then attribute them to every respective section for

Figure 5: Subregion level estimates (versus weighted average of voivodship-level estimates)



voivodship data. Note that this would not solve the problem of structurally different γ^s values at different levels of aggretation, and hence additional information would be required.

4. Incorporate prior knowledge. Using Bayesian methods, we could additionally supplement one of the approaches with prior knowledge from another, or with prior knowledge from outside the models, e.g. by differentiating between goods and services, more or less sophisticated kinds of services, or between goods with different average price-to-mass ratios.

Solutions 2-4 open a number of issues, and hence we leave them for future research.

6 Example simulation

Having estimated γ^s, θ_1^s and θ_2^s for every sector (see Table 4), we can proceed with the construction of a multisector-multiregion I-O matrix $\mathbf{X}^{\mathbf{RS}\times\mathbf{RS}}$ by using formulae (4) and (6). Using additionally data on global output (by subregions and sectors), we can construct the corresponding cost structure matrix $\mathbf{A}^{\mathbf{RS}\times\mathbf{RS}}$. Before proceeding to simulations, two extensions to this matrix need to be applied. First, we extend it by additional R columns and rows to incorporate the sector of households (necessary to compute the induced effects). Secondly, we include one additional row and column to account for the enterprise in question. The row only contains zeros, and the column describes the cost structure of the company, by products and regions jointly. The final cost structure matrix will be denoted as $\mathbf{A}^{[\mathbf{R}(\mathbf{S}+\mathbf{1})+\mathbf{1}]\times[\mathbf{R}(\mathbf{S}+\mathbf{1})+\mathbf{1}]}$, and the corresponding Leontief matrix $\mathbf{L}^{[\mathbf{R}(\mathbf{S}+\mathbf{1})+\mathbf{1}]\times[\mathbf{R}(\mathbf{S}+\mathbf{1})+\mathbf{1}]} = \mathbf{I}_{\mathbf{R}(\mathbf{S}+\mathbf{1})+\mathbf{1}} +$

$-\mathbf{A}^{[\mathbf{R(S+1)+1}]\times[\mathbf{R(S+1)+1}]}$. The simulation is performed according to the standard formula:

$$\triangle\mathbf{x}^{\mathbf{R(S+1)+1}} = \left\{\mathbf{L}^{[\mathbf{R(S+1)+1}]\times[\mathbf{R(S+1)+1}]}\right\}^{-1} \cdot \triangle\mathbf{y}^{\mathbf{R(S+1)+1}} \tag{26}$$

whereby $\triangle\mathbf{y}^{\mathbf{R(S+1)+1}}$ is the vertical vector of final output and $\triangle\mathbf{x}^{\mathbf{R(S+1)+1}}$ – the resulting vertical vector of global output across all sectors (including households) and regions. Note that $\triangle\mathbf{y}^{\mathbf{R(S+1)+1}}$ exhibits a specific structure: it is a zero vector except the last element, where the value of the analysed company's output is contained (last element, i.e. the row corresponding to the analysed company in matrices $\mathbf{A}^{[\mathbf{R(S+1)+1}]\times[\mathbf{R(S+1)+1}]}$ and $\mathbf{L}^{[\mathbf{R(S+1)+1}]\times[\mathbf{R(S+1)+1}]}$).

In practice, it may be challenging to construct the last column of matrix $\mathbf{A}^{[\mathbf{R(S+1)+1}]\times[\mathbf{R(S+1)+1}]}$. While it is not unusual to decompose costs according to sectors from which materials were purchased (or, more precisely, the sectors where the purchased intermediate goods were produced), the region of their origin may often remain unknown, even to the purchasing company or household. Note that it is not the place of purchase, but the location of the production, that allocates individual sectoral cost across regions. With no information available, it might be helpful to use a respective column of the matrix $\mathbf{A}^{\mathbf{RS}\times\mathbf{RS}}$ as a proxy.

In our illustrative generic simulation, we use a breakdown into 77 sectors from Polish I-O tables and seek to evaluate the economic impact for any of 66 subregions in Poland. We look at an example, hypothetical construction company (sector 34 from Polish I-O table) from lubelskie subregion (subregion 11 according to Polish CSO nomenclature). This company is assumed to sell final output worth 1m PLN per period (at constant prices of 2010). The following assumptions about its cost structure are made:

1. sectoral cost structure is representative for the entire construction branch, i.e. the split across sectors is identical to column 34;

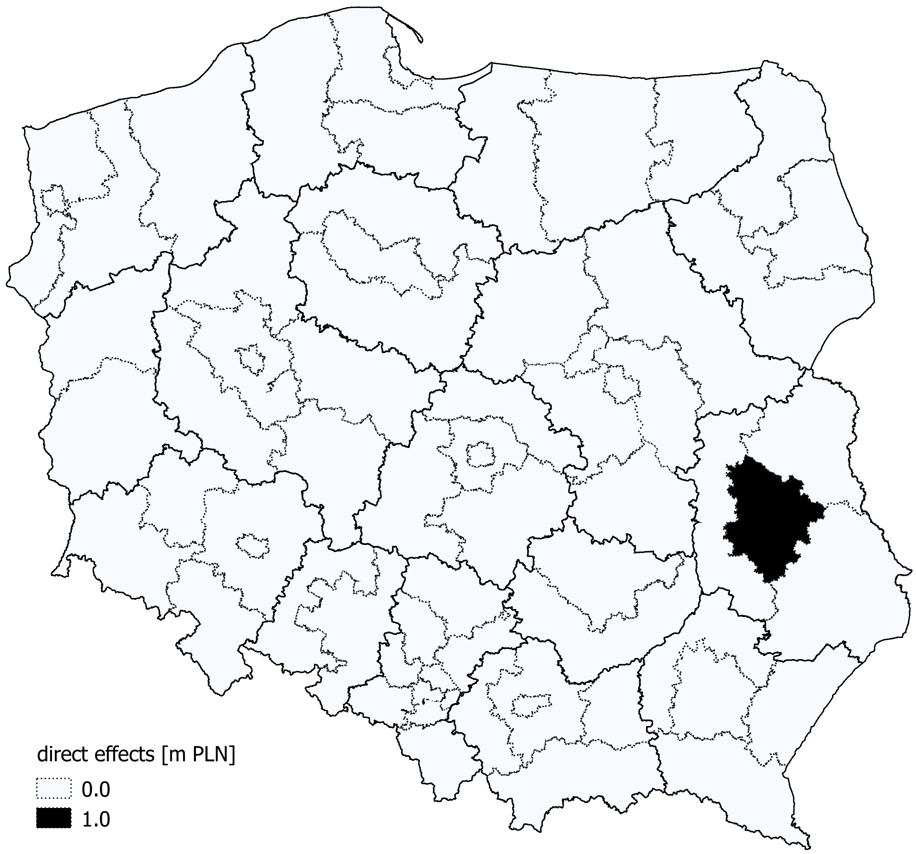
2. for 4 highest cost components, i.e. intermediate demand for construction (34), retail trade (36), wholesale trade (37) and transport (38) – 60% of the cost is allocated to lubelskie subregion, 30% is allocated to other subregions in lubelskie voivodship (chełmsko-zamojski, bialski, puławski) and equally split and 10% is allocated to all other 62 subregions in Poland (and equally split between them);

3. for all branches except 34, 36, 37 and 38, the cost structures in the respective column in matrix $\mathbf{A}^{\mathbf{RS}\times\mathbf{RS}}$ (i.e. column for sector 'construction', subregion 'lubelskie') were replicated.

By definition, all direct effects are allocated to lubelskie subregion (see Figure 6).

Indirect effects (see Figure 7) are highly concentrated in lubelskie subregion, but we can also see some interesting spatial patterns. Firstly, four other subregions generate output higher than 0.12 m PLN: 3 of them belong to lubelskie voivodship (9, 10 and 12) and profit both from geographical proximity and our assumptions as regards

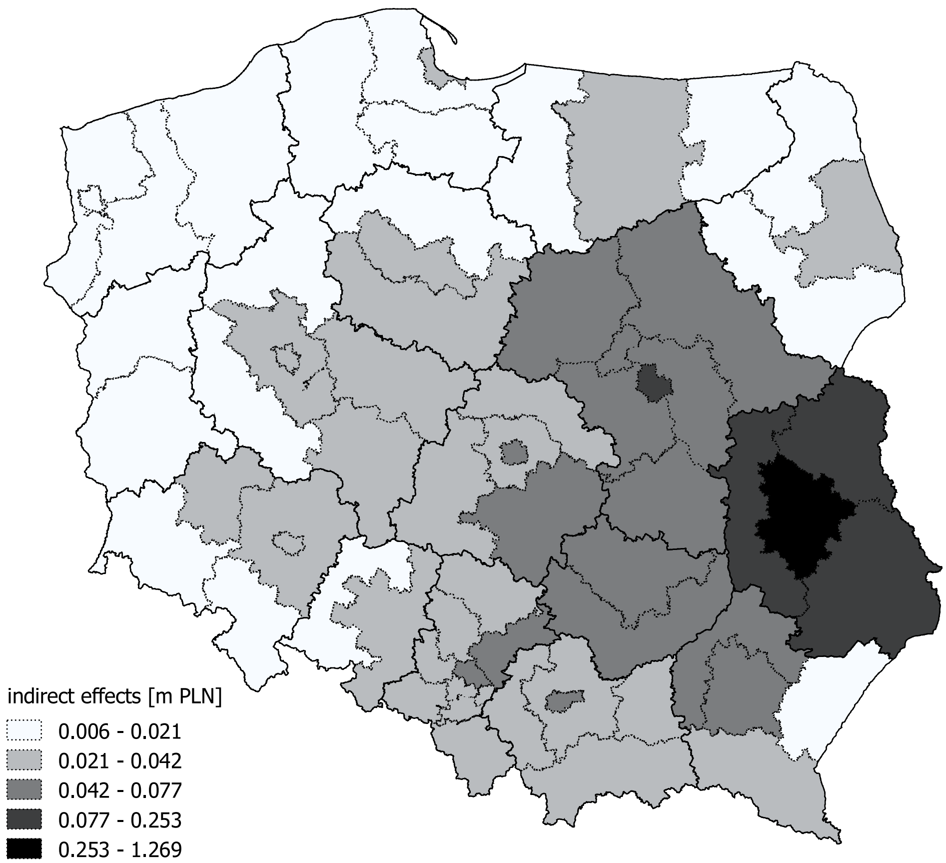
Figure 6: Example simulation: direct effects on global output



the cost structure. The fourth one is the Warsaw subregion (28), which is relatively close to lubelskie and has a very high contribution to Poland's value added, i.e. is unlikely to be constrained on the supply side in many branches. This is why it is estimated to profit more than e.g. Warsaw-East subregion (29), located closer to lubelski subregion. Still, North-Western part of mazowieckie voivodship still seems to be an area where economic effects of the company can be relatively visible, with more than 0.06 m PLN. The same applies to tarnobrzeski subregion in the North of podkarpackie voivodship (geographical proximity) and Łódź (high value added). Note that effects are slowly dying out towards Western Poland, being the lowest in the voivodships along the Western and Northern border.
A similar picture emerges in the case of induced effects (see Figure 8). Note that

www.czasopisma.pan.pl PAN www.journals.pan.pl
POLSKA AKADEMIA NAUK

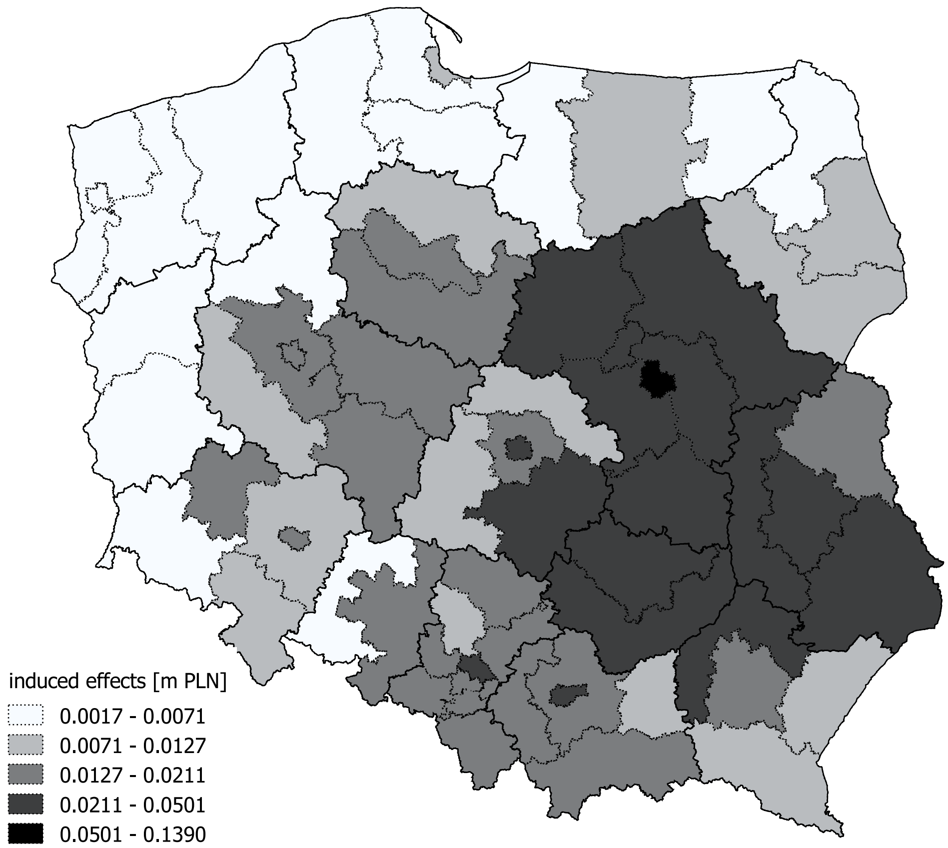
Figure 7: Example simulation: indirect effects on global output



the difference between Figures 7 and 8 is mainly due to spatial distribution of labour-intensive branches. Note that, in the I-O framework, a branch appears as more labour-intensive when i.e. it is mainly located in the areas where labour costs are higher than elsewhere. This is why the induced effects are higher in Warsaw subregion than in the satellite subregions of lubelskie voivodship.

Total effects are presented in Figure 9. They suggest, in line with intuition, that lubelskie subregion generates the most output, but that additional effects are also located in neighbourhood voivodships: mazowieckie and świętokrzyskie. There are also certain subareas of łódzkie, śląskie and podkarpackie that take profits – whereby śląskie is a major a supplier of steel for construction, and podkarpackie is another neighbouring voivodship. The map of total spillovers is understandably not indicating

Figure 8: Example simulation: induced effects on global output



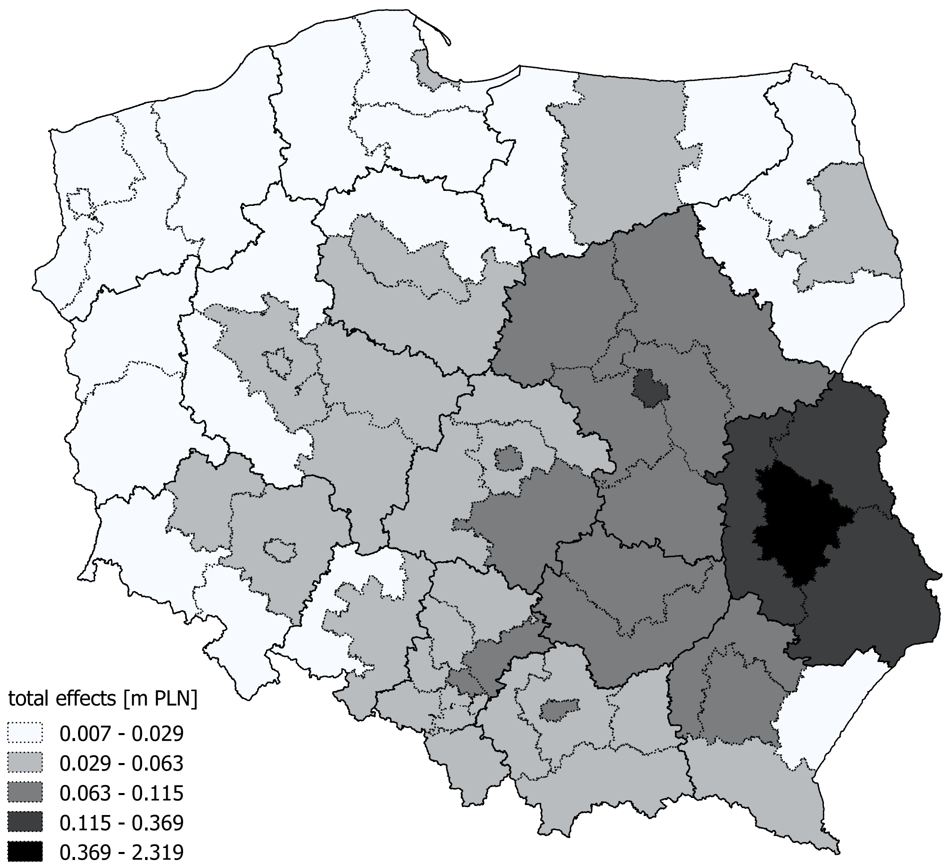
substantial effects in zachodniopomorskie and lubuskie – distant voivodships, where no unique production plants for construction are located.

7 Conclusions

Regional economic impact assessment for a single company – especially at a level lower than NUTS-2 (voivodships) – is not straightforward in the absence of regional I-O tables. Intuitively, indirect and induced effects should be concentrated overproportionately in the area of impulse. In this paper, a novel method to quantify this intuition is proposed. It is based on publicly available data, replicable across EU countries, and scalable for different levels of regional and sectoral aggregation.

www.czasopisma.pan.pl PAN POLSKA AKADEMIA NAUK www.journals.pan.pl

Figure 9: Example simulation: total effects (direct, indirect, induced) on global output



Our approach is based on a spatial econometric analysis of regional value added data. The specification is based on the spatial Durbin model, but we modify it by calibrating the structural parameters and estimating the spatial weight matrix, being a function of physical distance between regions and region-specific supply constraints. In the paper, we derive the relationship between the resulting spatial weight matrices and the interpolation of cross-sectoral flows into cross-regional flow matrices. A list of assumptions underlying this interpolation is explicitly provided.

The estimation results – on the voivodship and subregion level – generally confirm a significant role of supply constraints and distance in spatial distribution of indirect and induced effects. However, the estimates on both levels are partly incompatible, mainly because supply constraints for aggregate goods and territories are a function of the level of aggregation. Additional research is needed to reconcile the information derived from both sources, and suggested research avenues were described in Subsection 5.3. Example simulation results confirm the intuitive pattern of spillover into subregions (i) with a high share in country-wide value added and (ii) at geographical proximity. Future research can deal with a number of extensions to the proposed framework that were not covered here. Firstly, the vector of final output – not available at subregional level – was derived from the I-O tables and the population shares of subregion. This method can certainly be refined so as to better reflect exports and investment as components of the final demand. Secondly, although the functional form of the distance function and aggregating function for $\mathbf{W^s}$ matrices was subject to extensive sensitivity testing, there may still be room for improvement or inclusion of further factors. Thirdly, the assumption about identical technology within sectors and between regions may not be plausible for subregions located near the borders, especially borders with other EU countries. In such cases, the share of imports can be higher than the country-average. This deviation would have to be levelled out by an appropriate downward correction of import intensity for interior subregions, and various scale of this correction should be expected in different sectors. Fourthly, the nexus between induced effects and regional wage differentiation also deserves some attention.

Acknowledgements

The author is grateful to an anonymous referee, Michał Kowalczuk, Aleksander Olechnowicz, Paweł Opala and Marek Rozkrut for valuable comments and discussions, and to Piotr Pękała – additionally for substantial help with data visualisation. This paper also benefited from discussions with participants of 42^{nd} Macromodels International Conference 2015 in Trzebieszowice, especially Henryk Gurgul. All errors and omissions are mine.

References

[1] Akita T., Kataoka M. (2002), Interregional Interdependence and Regional Economic Growth: An Interregional Input-Output Analysis of the Kyushu Region, *Review of Urban and Regional Development Studies*, 14, 18–40.

[2] Atkinson R.D. (1993), Defense Spending Cuts and Regional Economic Impact: An Overview, *Economic Geography*, 69(2), 107–122.

[3] Baade R.A., Matheson V.A. (2004), The Quest for the Cup: Assessing the Economic Impact of the World Cup, *Regional Studies*, 38(4), 343–354.

[4] Cameron A.C., Trivedi P.K. (2005), *Microeconometrics. Methods and Applications*, Cambridge University Press, Cambridge.

[5] Canning P., Wang Z. (2005), A Flexible Mathematical Programming Model to Estimate Interregional Input-Output Accounts, *Journal of Regional Science*, 45, 539–563.

[6] Central Statistical Office in Poland (2014), Input-Output Table at Basic Price in 2010, Statistical Information and Elaborations.

[7] Chenery H.B. (1953), The Structure and Growth of the Italian Economy, chapter *Regional Analysis*, 97–129, Rome: US Mutual Security Agency.

[8] Chrzanowski M. (2013), Wykrycie lokalnych obszarów kooperacji w województwie lubelskim (analiza output-input), Kapitał Intelektualny Lubelszczyzny 2010-2013, URL `http://www.kil.lubelskie.pl/wp-content/uploads/2013/08/Wykrycie-lokalnych-obszar%C3%B3w-kooperacji-w-wojew%C3%B3dztwie-lubelskim-analiza-output-input-dr-Marek-Chrzanowski.pdf`.

[9] Flegg A., Tohmo T. (2013), Regional Input-Output Table and the FLQ Formula: ACase Study for Finland, *Regional Studies*, 47(5), 703–721.

[10] Folmer H., Nijkamp P. (1985), Methodological aspects of impact analysis of regional economic policy, *Papers of the Regional Science Association*, 57(1), 165–181.

[11] Fritz O., Kurzmann R., Pointner W., Streicher G., Zakarias G. (2001), *Modeling and Control of Economic Systems*, chapter Modeling the Regional Economy: A Regional Econometric Input-Output Approach, Elsevier, Oxford.

[12] Gordon I.R. (1976), Gravity Demand Functions, Accessibility and Regional Trade, *Regional Studies*, 10, 25–37.

[13] Hakfoort J., Poot T., Rietveld P. (2001), The Regional Economic Impact of an Airport: The Case of Amsterdam Schiphol Airport, *Regional Studies*, 35(7), 595–604.

[14] Hallegatte S. (2008), An Adaptive Regional Input-Output Model and its Application to the Assessment of the Economic Cost of Katrina, *Risk Analysis*, 28(3), 779–799.

[15] Horváth E., Frechtling D.C. (1999), Estimating the Multiplier Effects of Tourism Expenditures on a Local Economy through a Regional Input-Output Model, *Journal of Travel Research*, 37(4), 324–332.

[16] Hulu E., Hewings G.J.D. (1993), The Development and Use of Interregional Input-Output Models for Indonesia Under Conditions of Limited Information, *Review of Urban and Regional Development Studies*, 5, 135–153.

[17] Isard W., Kuenne R.E. (1953), The Impact of Steel upon the Greater New York-Philadelphia Industrial Region, *Review of Economics and Statistics*, 35, 289–301.

[18] Jackson R.W., Schwarm W.R., Okuyama Y., Islam S. (2006), A Method for Constructing Commodity by Industry Flow Matrices, *Annals of Regional Science*, 40, 909–920.

[19] Kratena K., Zakarias G. (2004), Input Coefficient Change using Bi-proportional Econometric Adjustment Functions, *Economic Systems Research*, 16, 191–203.

[20] Leontief W. (1936), Quantitative Input-Output Relations in the Economic System of the United States, *Review of Economics and Statistics*, 18, 105–125.

[21] Leontief W. (1941), *The Structure of American Economy 1919-1939*, New York: Oxford University Press.

[22] Leontief W., Strout A. (1963), *Structural Interdependence and Economic Development*, chapter Multiregional Input-Output Analysis, 119–149, London: Macmillan (St. Martin's Press).

[23] LeSage J.P., Pace R.K. (2009), *Introduction to Spatial Econometrics*, Chapman & Hall.

[24] Lindall S., Olson D., Alward G. (2006), Deriving Multi-Regional Models Using the IMPLAN National Trade Flows Model, *Journal of Regional Analysis and Policy*, 36, 76–83.

[25] Liu L.N., Vilain P. (2004): Estimating Commodity Inflows to a Substate Region Using Input-Output Data: Commodity Flow Survey Accuracy Tests, *Journal of Transportation and Statistics*, 7, 23–37.

[26] Loveridge S. (2004), A Typology and Assessment of Multi-sector Regional Economic Impact Models, *Regional Studies*, 38(3), 305–317.

[27] McCann P. (2007), *Urban and Regional Economics*, Oxford University Press.

[28] Miller R.E., Blair P.D. (2009), *Input-Output Analysis*. Foundations and Extensions, Cambridge University Press, Cambridge.

[29] Moses L.N. (1955), The Stability of Interregional Trading Patterns and Input-Output Analysis, *American Economic Review*, 45, 803–832.

[30] Okamoto N., Ihara T. [eds.] (2005), *Spatial Structure and Regional Development in China. An Interregional Input-Output Approach*, UK: Palgrave Macmillan.

[31] Patriquin M.N., Wellstead A.M., White W.A. (2007), Beetles, trees, and people: Regional economic impact sensitivity and policy considerations related to the mountain pine beetle infestation in British Columbia, Canada, *Forest Policy and Economics*, 9, 938–946.

[32] Polenske K.R. (1970), An Empirical Test of Interregional Input-Output Models: Estimation of 1963 Japanese Production, *American Economic Review*, 60, 76–82.

[33] Polenske K.R. (2004), *Wassily Leontief and Input-Output Economics*, chapter Leontief's 'Magnificent Machine' and Other Contributions to Applied Economics, 269–289, Cambridge University Press, Cambridge.

[34] Rose A., Benavides J., Chang S.E., Szczesniak P., Lim D. (1997), The regional economic impact of an earthquake: direct and indirect effects of electricity lifeline disruptions, *Journal of Regional Science*, 37(3), 437–458.

[35] Sonis M., Hewings G.J., Okuyama Y. (2000), Feedback loops analysis of Japanese interregional trade, 1980-85-90, *Journal of Economic Geography*, 1.

[36] Steinback S.R. (1999), Regional Economic Impact Assessments of Recreational Fisheries: An Application of the IMPLAN Modeling System to Marine Party and Charter Boat Fishing in Maine, *North American Journal of Fisheries Management*, 19, 724–736.

[37] Theil H. (1967), *Economics and Information Theory*, New York: American Elsevier.

[38] Turner K., Lenzen M., Wiedmann T., Barrett J. (2007), Examining the Global Environmental Impact of Regional Consumption Activities – Part 1: A Technical Note on Combining Input-Output and Ecological Footprint Analysis, *Ecological Economics*, 62(1), 37–44.

[39] Uribe P., de Leeuw C., Theil H. (1966), The Information Approach to the Prediction of Interregional Trade Flows, *Review of Economic Studies*, (33), 209–220.

[40] Welfe W., Świeczewska I., Florczak W., Karp P. (2008), Rozwój regionalny - ujęcie ekonometryczne, Społeczna Wyższa Szkoła Przedsiębiorczości i Zarządzania w Łodzi, URL `http://www.eedri.pl/pdf/116.pdf`.

[41] West G.R. (1990), Regional Trade Estimation: A Hybrid Approach, International Regional Science Review, 13, 103–118.

[42] Wiedmann T., Wilting H.C., Lenzen M., Lutter S., Palm V. (2011), Quo Vadis MRIO? Methodological, data and institutional requirements for multi-region input-output analysis, *Ecological Economics*, 70, 1937–1945.

[43] Zawalińska K., Płaszaj A., Celińska-Janowicz D., Rok J. (2014), Regionalne przepływy międzygałęziowe - możliwości analityczne i wyzwania metodologiczne, URL `http://www.euroreg.uw.edu.pl/dane/web_euroreg_seminary_files/665/dr_hab._zawaliska_regionalne_przepywy_midzygaziowe_22.05.14.pdf`.