Regional Studies

ISSN: 0034-3404 (Print) 1360-0591 (Online) Journal homepage: https://rsa.tandfonline.com/loi/cres20

Regional input–output tables and trade flows: an integrated and interregional non-survey approach

Riccardo Boero, Brian K. Edwards & Michael K. Rivera

To cite this article: Riccardo Boero, Brian K. Edwards & Michael K. Rivera (2018) Regional input–output tables and trade flows: an integrated and interregional non-survey approach, Regional Studies, 52:2, 225-238, DOI: 10.1080/00343404.2017.1286009

To link to this article: https://doi.org/10.1080/00343404.2017.1286009

REGIONAL STUDIES, 2018
VOL. 52, NO. 2, 225–238
https://doi.org/10.1080/00343404.2017.1286009

Regional input–output tables and trade flows: an integrated and interregional non-survey approach

Riccardo Boero[a], Brian K. Edwards[b] and Michael K. Rivera[c]

ABSTRACT

Regional input–output tables and trade flows: an integrated and interregional non-survey approach. *Regional Studies*. Regional analyses require detailed and accurate information about dynamics happening within and between regional economies. However, regional input–output tables and trade flows are rarely observed and they must be estimated using up-to-date information. Common estimation approaches vary widely but consider tables and flows independently. By using commonly used economic assumptions and available economic information, this paper presents a method that integrates the estimation of regional input–output tables and trade flows across regions. Examples of the method implementation are presented and compared with other approaches, suggesting that the integrated approach provides advantages in terms of estimation accuracy and analytical capabilities.

KEYWORDS

input–output; regionalization; non-survey methods; multiregional models; domestic trade flows

摘要

区域投入—产出表和贸易流：一个整合性的跨区域非调查方法。*Regional Studies*. 区域分析需要关于区域经济之中和之间发生的动态的细緻且正确之信息。但区域投入—产出表与贸易流却鲜少受到观察，且它们必须使用最新的信息进行评估。一般的评估方法差异相当大，但却各自考量投入—产出表与贸易流。透过运用一般使用的经济预设与可取得的经济信息，本文呈现整合区域投入—产出表与区域间的贸易流的评估之方法。本文将呈现该方法的实际执行案例，并与其他方法进行比较，指出整合方法在评估准确度以及分析能力上具有优势。

关键词

投入—产出; 区域化; 非调查方法; 多重区域方法; 国内贸易流

RÉSUMÉ

Tableaux d’entrée/sortie régionaux et flux d’échanges: approche non observée intégrée et interrégionale. *Regional Studies*. Les analyses régionales nécessitent des informations détaillées et précises sur la dynamique survenant au sein d’économies régionales et entre celles-ci. Toutefois, on ne relève que rarement des tableaux d’entrée/sortie régionaux et des flux d’échanges, et on doit les évaluer en appliquant des informations à jour. Les approches à base d’estimations communes varient considérablement, mais tiennent compte de tableaux et flux indépendamment. En appliquant des hypothèses économiques répandues, et des informations économiques dont elle dispose, la présente communication présente une méthode intégrant l’estimation de tableaux d’entrée/sortie régionaux et des flux d’échanges dans les régions. Elle y présente des exemples de mise en œuvre de la méthode, que l’on compare avec d’autres approches, en soutenant qu’une approche intégrée présente des avantages en termes de précision des estimations et de capacités analytiques.

MOTS-CLÉS

entrées/sorties; regionalisation; méthodes non observées; modèles multi-régionaux; échanges intérieurs

CONTACT

[a] **(Corresponding author)** riccardo@lanl.gov
Los Alamos National Laboratory, Los Alamos, NM, USA.
[b] bke@lanl.gov
Los Alamos National Laboratory, Los Alamos, NM, USA.
[c] mkrivera@lanl.gov
Los Alamos National Laboratory, Los Alamos, NM, USA.

ZUSAMMENFASSUNG

Regionale Input-Output-Tabellen und Handelsströme: ein integrierter interregionaler, nicht erhebungsgestützter Ansatz. *Regional Studies*. Für regionale Analysen werden detaillierte und präzise Informationen über die Dynamiken innerhalb und zwischen regionalen Ökonomien benötigt. Allerdings werden regionale Input-Output-Tabellen und Handelsströme selten beobachtet, weshalb sie anhand aktueller Informationen geschätzt werden müssen. Die gängigen Schätzverfahren fallen höchst unterschiedlich aus, berücksichtigen die Tabellen und Ströme aber stets unabhängig voneinander. In diesem Beitrag stellen wir unter Nutzung von häufig verwendeten wirtschaftlichen Annahmen und verfügbaren wirtschaftlichen Informationen eine Methode vor, die die Schätzung von regionalen Input-Output-Tabellen und von Handelsströmen zwischen Regionen miteinander kombiniert. Ebenso stellen wir Beispiele für die Umsetzung der Methode vor und vergleichen sie mit anderen Ansätzen, wobei wir den Schluss ziehen, dass der integrierte Ansatz Vorteile hinsichtlich der Schätzgenauigkeit und analytischen Möglichkeiten bietet.

SCHLÜSSELWÖRTER

Input-Output; Regionalisierung; nicht erhebungsgestützte Methoden; multiregionale Modelle; Binnenhandelsströme

RESUMEN

Tablas regionales de entrada y salida y flujos comerciales: un enfoque integrado e interregional sin encuestas. *Regional Studies*. En los análisis regionales se requiere información detallada y exacta sobre las dinámicas que ocurren dentro y entre las economías regionales. Sin embargo, raras veces se observan las tablas de entrada y salida y los flujos comerciales regionales, y por eso deben calcularse utilizando información actualizada. Los métodos de estimación comunes varían en gran medida pero las tablas y los flujos se consideran de manera independiente. En este artículo tenemos en cuenta las suposiciones económicas habitualmente usadas y la información económica disponible, y presentamos un método que integra la estimación de las tablas de entrada y salida y los flujos comerciales regionales en todas las regiones. Mostramos ejemplos de la aplicación del método y lo comparamos con otros enfoques, lo que indica que el enfoque integrado ofrece ventajas en términos de la exactitud de la estimación y las capacidades analíticas.

PALABRAS CLAVES

entrada y salida; regionalización; métodos sin encuestas; modelos multirregionales; flujos comerciales nacionales

JEL C67, D57, O18, R15
HISTORY Received 6 January 2014; in revised form 17 January 2017

INTRODUCTION

Research and policy questions in regional economic analyses often call for improvements in how regional economic impacts are estimated. Most approaches for the modelling of economic impacts require regional input–output (IO) tables. When the focus is on economic dynamics both within and between regions, approaches such as interregional IO models and multipliers, and interregional computable general equilibrium models, also require a detailed description of linkages between regional economies. Besides regional IO tables, economic impacts assessment needs information about trade flows between regions.

Regional IO tables are often not immediately available, so analysts estimate using available data. At the same time, domestic trade flows are rarely and only partially observed and therefore also require estimation.

Regional economic analysts have developed several methods to estimate regional IO tables that do not rely on expensive and time-consuming surveys. These methods take advantage of the recent availability of frequently updated information about national-level IO accounts and regional labour markets. Other methods have been developed to estimate domestic trade flows that exploit similar information.

All methods available consider the estimation of regional IO tables and domestic trade flows as separate problems. Besides the different degrees of accuracy associated with each variant of those estimation methods, the resulting estimates are often inconsistent because the problems are treated independently.

This paper presents a novel approach for the estimation of the information needed for regional analyses. It is aimed at improving both analyses focused at the level of single regions and those with an interregional focus.

The approach is based on the recognition that regional IO tables and domestic trade flows are two observations of the same phenomenon that represent the regional organization of economic activities. It is an integrated approach because it considers at the same time what happens within a region (i.e., regional IO tables) and the links with other regions (i.e., domestic trade flows). The integration improves the precision of the estimation because it extends the set of constraints to be considered.

As other approaches in the literature, it is a non-survey approach because it is based on commonly available datasets limiting the need for expensive ad hoc surveys. In particular, it relies on national-level IO accounts, on regional labour market data and on measures of transport costs between regions.

Similar to other non-survey approaches, it relies on two main economic assumptions on the homogeneity of adopted technology and preferences. Further, it considers transport cost minimization as the main dynamics determining domestic trade flows.

In order to fully present the method and the differences with other approaches, the paper is organized as follows. The next section summarizes the common features of other non-survey and hybrid approaches to the regionalization of IO tables and of gravity models for the estimation of domestic trade flows. The third section presents the integrated approach in detail. The fourth section tackles the difficult task of comparing the integrated approach with others by presenting results obtained in two examples of implementation of the method. First, results are presented about observed and estimated domestic trade flows of commodities at the state level in the US economy in 2002 and 2007. Second, the IO table estimated with the integrated approach for the state of Washington in 2007 is compared with the similar one estimated with a hybrid method. The final section concludes, discussing the results and pointing out the advantages and limitations of the proposed method.

ESTABLISHED METHODS

Most of the methods used to regionalize IO tables, whether non-survey methods (Round, 1983), hybrid methods (Lahr, 1993), or ready-made or short-cut techniques (Jensen, 1990), have some common features. First, they rely on national IO tables and regional-level labour market data. Second, they are based on two main theoretical assumptions: (1) that regional and national technologies are very much similar (Flegg, Webber, & Elliott, 1997); and (2) that customers have very similar preferences within a nation (Isard et al., 1998; Miller & Blair, 2009). Third, they are implemented using a similar procedure.

In particular, non-survey methods usually work as follows. First, direct coefficients of the most recent national IO table at disposal are computed. Second, using one of the many variants of the approach of location quotients (LQs) for each industry an LQ-based measurement is computed (Bonfiglio & Chelli, 2008; Flegg et al., 1997; Flegg & Tohmo, 2013; Flegg & Webber, 2000; Flegg, Webber, & Elliott, 1995; Frechtling & Horvath, 1999; Kowalewksi, 2015; Lehtonen & Tykkyläinen, 2014; Tohmo, 2004). Such measurements compare regional and national outputs (usually approximated using labour market data, e.g., number of employees in industries), because of reasons of data availability (Isserman, 1977), to quantify the capability of regional industries to satisfy regional demands for their product. Third, if the measurement points out that the industry is relatively smaller at the regional level than at the national one, the national coefficients are decreased accordingly along the row that represents the industry of focus.

By means of these modifications to national coefficients, new coefficients should represent the use of intermediates produced locally within the region. At the same time, the implicit assumption is that technology is homogenous in regions comprising the nation. If there is significant variation between the composition of intermediate production between the national and regional economies, then this difference is attributable to the fact that intermediates considered in regional tables are only the ones produced locally.

Non-survey methods then require two more steps to finally obtain a regional IO table. First, vectors of final demand have to be estimated in order to transform regional coefficients to a full transaction table. The estimation is usually done assuming homogenous preferences within the nation and thus regional final demand is estimated using an approach similar to LQ. Second, the resulting regional transaction matrix has to be balanced with methods (Dewhurst, 1992; Robinson, Cattaneo, & El-Said, 2001) such as iterative proportional fitting (i.e., the RAS method).

Methods that include the use of regional information collected with ad hoc surveys, usually called hybrid methods or partial-survey methods (Harris & Liu, 1998; West, Morison, & Jensen, 1984), follow a procedure very similar to the one depicted above. In particular, they only substitute the modification of national coefficients on LQ measurements with modifications based on information collected through the survey.

In summary, non-survey methods are based upon the availability of two data sources: national IO transaction tables and regional data about economic output. Further, non-survey methods are based on two main economic assumptions of similarity between the regional economy and the national one. In fact, production technology is assumed to be the same across regions composing the nation and equal to the average one observed at the national level. Secondly, consumers, either private or public, are assumed to have the same preferences in all regions and equal to the average ones observed at the national level.

A large literature has pointed out the biases and large estimation errors associated with non-survey methods when estimating regional IO tables, focusing on the many existent variants of LQ methods (e.g., Flegg et al., 1995; Lehtonen & Tykkyläinen, 2014; Round, 1983). However, two more sources of error and inconsistency are worth considering. First, because the estimation procedure does not consider the balancing of the transaction table as a constraint, at the end of the procedure non-survey methods use balancing methods that might compound existing estimation errors. Unfortunately, adopters of non-survey methods very seldom evaluate the impact of such methods on regional tables. The evaluation of the error introduced is easily done by computing statistics, such as the mean absolute percentage error (Miller & Blair, 2009), on the regional direct input coefficients derived from the national ones (either as resulting from the LQ method or from the information collected through the survey) and on the same coefficients after the balancing procedure.

Further, non-survey methods present specific criticalities for interregional analyses. In particular, they lack both consistency in terms of aggregation and the capability to support the estimation of interregional trade flows. Non-survey methods do not allow the aggregation of regional tables since they are focused on locally produced and consumed intermediates, and thus they lack the information to effectively integrate tables for larger areas. Further, because transaction tables include biases and errors introduced by balancing techniques, aggregation can mean propagating and enlarging such biases.

Most of these critiques can be extended to hybrid methods and to other less common non-survey approaches that do not use LQ measurements to modify national coefficients, such as, for instance, supply–demand pool approaches, and the ones based on fabrication effects and regional purchase coefficients (Round, 1972, 1983; Schaffer & Chu, 1969; Stevens & Trainer, 1976, 1980).

Non-survey and hybrid methods are aimed at estimating transactions within a single region and ignore the information that good estimates of interregional trade flows can provide (Harris & Liu, 1998). Trade flows, however, have been the focus of methods introduced in the literature in parallel to non-survey methods.

Although trade flows have emerged early on as an important topic in the field (Isard, 1951), the attempts to estimate interregional trade flows are few (e.g., Boomsma & Oosterhaven, 1992; Canning & Wang, 2005; Escobedo & Ureña, 2008; Miller & Blair, 1983; Riddington, Gibson, & Anderson, 2006; Shao & Miller, 1990). Most of them are variants of the seminal method for estimating interregional trade flows presented by Leontief and Strout (1963) that is based on a gravity model. Gravity models are based on three main components: demand, supply and distance. For instance, if the trade flow from region i to region j of commodities produced by sector s is the one in focus, model components are the supply of s in region i, the demand of s in region j, and a measure of the distance between i and j. The model specification is usually non-linear (hence, the resemblance with gravity) and it is motivated, from the viewpoint of economic theory, in terms of differentiation of products, assuming that largest economies supply and demand more differentiated products than small ones. Considering the distance between regions means to consider the role of transport costs.

As a simple example of a model, define:

$$t_{s,ij} = Su^{\alpha}_{s,i} De^{\beta}_{s,j} Tr^{\gamma}_{ij},$$

where $t_{s,ij}$ is the trade flow of products of sector s from region i to j; $Su_{s,i}$ is the supply of s in region i; $De_{s,j}$ is the demand of s in region j; and Tr_{ij} is the transport cost from i to j. The parameters of the model are α, β and γ.

The method for estimating interregional trade flows with gravity models is as follows. First, a reference period is chosen for which observations on trade flows are available, and is the most recent available. Second, after the log-linearization of the original model specification, a linear regression is conducted on observed data in order to calibrate the parameters that allow explaining trade flows of commodity s from region i to j in terms of supply of s in i, corresponding demand in j, and distance between the regions. The estimation is done by industry s using all the possible trade flows as regression sample and thus leading to a single value for each parameter (α, β and γ in the example) in each sector over all pairs of regions.

Finally, using the parameters calibrated in the period of reference and the values of supply and demand in the period of interest, regional flows are estimated.

Gravity models for the estimation of interregional trade flows thus present some practical challenges, in particular regarding the availability of data. In fact, gravity models need to be empirically calibrated and even if it is plausible to assume that the effects that are captured by the model parameters are not particularly time variant (and thus even data about periods much in the past can be used as reference), the problem is to have complete data for all industries. In fact, surveys about commodity flows are rare, often incomplete and not very spatially detailed, and empirical information about trade flows in service sectors is almost completely non-existent.

Taking the broader perspective of interregional analyses, the established methods present several criticalities that go beyond the problems of each individual variant. In particular, the methods for the regionalization of IO tables and the ones for estimating trade flows are independent and separated, with the possibility to obtain inconsistent estimates. For instance, total regional demand is poorly estimated in IO tables obtained with LQ-based regionalization since the focus is only on demand for locally produced intermediates. This means that demand estimated in this way cannot be used in gravity models, and that, from the opposite perspective, estimates of domestic trade flows can be inconsistent with the supply and demand data described in regional tables. Further, when aggregating information from regionalized tables and trade flows, it is very likely that resulting tables are unbalanced and that they require another balancing procedure, possibly adding additional errors.

AN INTEGRATED APPROACH

It is possible to estimate regional tables and interregional trade flows simultaneously, exploiting the advantages provided by treating the problem as one and not as two independent ones. The integrated method is based on the same economic assumptions used in other non-survey methods and uses the same data sources.

The method has two components, namely, estimating regional unbalanced transaction tables and estimating trade flows. The estimation problem is unique. The trade flow model is constrained by information in unbalanced tables. The aggregation of unbalanced tables and trade flows provides accurate symmetrical regional tables and information about interregional linkages.

The starting point of the method is an accurate estimation of total supply and demand at the regional level by industry. This is accomplished by preparing a regional transaction table that considers the entire regional demand of intermediates. The regional transaction table is calculated as follows. First, under the assumption of homogeneous production technology and by proceeding column by column, the output of each regional industry is computed by multiplying each value in the corresponding industry column of the national IO table by the ratio of the product in that sector in the region and the same product at the national level. Regional and national products can be approximated by measures of employment, consistently with the hypothesis of homogenous technology and thus homogenous contribution of labour to output.

By means of this first step, the regional transactions table is filled with intermediate demand for each industry (and in each industry), and with value added, determining the total regional supply.

Secondly, final demand can be estimated under the assumption of homogeneity of consumer preferences. In particular, all the national values of final demand are transformed into regional values by multiplying the former by the ratio of income measured at the regional level and income measured at the national level. It is possible to use as approximated measure of income the total number of employees, population, or value added.

The allocation of regional imports and exports to foreign countries can be also estimated using these two approaches. In fact, assuming that regional exports by industry are proportional to the concentration observed at the regional level relative to the national level, regional exports by industry can be estimated using the first step as described above. Regional imports from foreign countries, in contrast, should be proportional to income and thus computed as in the second step following the assumption on homogeneity of preferences.

Executing the two steps just described allows obtaining regional unbalanced transaction tables that describe regional supply (at the first step) and net regional demand (intermediate demand computed at the first step, net final demand at the second step). The unbalance of the table at this stage of the estimation process is not a problem but an opportunity: it is in fact an important information source for the computation of interregional trade flows. In fact unbalanced regional tables are missing only the information about domestic trade flows, which is the element that balances them.

The second component of the integrated approach is a trade flow model used to estimate interregional trade flows by relying on regional supply and demand, which are estimated by the first component, and transport costs between regions.

From the theoretical viewpoint, it is possible to rely on economic dynamics such as utility maximization on transport costs without referring to other supplemental economic dynamics such as product differentiation. Consumers, when choosing their suppliers, try to spend less in order to maximize their utility and thus chose suppliers that minimize transport costs.

In particular, three different models of cost minimization can be considered. The first two consider minimization of transport costs alone, but they will differ since the first considers simultaneous decision-making, while the second employs sequential decision-making. Like the second method, the third method incorporates sequential decision-making but relies on minimizing costs associated with competing for access to scarce resources instead.

The first trade flow model, if seen from the demand side, can be formalized as follows for every region i of the N regions composing the country, and for each sector in the economy:

$$\begin{aligned} \min_{d} \quad & \sum_{j=0}^{N} t_{ji} d_{ji} \\ & \text{s.t.} \sum_{j=0}^{N} d_{ij} = S_i, \qquad \forall\, i \\ & \sum_{j=0}^{N} d_{ji} = D_i, \qquad \forall\, i \\ & d_{ji} \geq 0, \qquad \forall\, i, j \end{aligned}$$

Transport costs from j to i are represented by t_{ji}; d_{ij} is the trade flow of the good or service from region j to region i; and S_i and D_i are, respectively, the total supply and demand for that product in region i. Trade flows cannot be negative, and because trade flows within regions are estimated too, their aggregation must correspond to the amount of supply and demand.

The model is based on dispersed decisions, optimization does not work on system-level variables, and consumers in all regions minimize transport cost simultaneously (i.e., there are N regions concurrently optimizing). Because transport costs are lower within a region than across regions, the model does not allow cross-hauling to emerge and consumers in regions first search for local products and then, if the region has more demand than supply, they search for suppliers in other regions.

The second trade flow model has the same formal specification as the first one but decision-making differs. Consumers search for suppliers sequentially in a random order and without coordination (i.e., there are N regions optimizing in a random sequence), allowing the emergence of cross-hauling, while in the first model they search for suppliers simultaneously. The sequential access to the market is what allows the emergence of cross-hauling. For example, there are two near regions A and B in the same country. A is a region with an excess of demand for a particular product and B is a region with a perfect balance between supply and demand for that same product. By chance, the demand of consumers in A emerges first during the period and thus they first search for suppliers in their own region and then later search in B. At this point B, which was a region with a balance between supply and demand, shows an excess of demand since part of the supply has been sold to consumers in A. According to this example, when B consumers will have the chance to search for suppliers in other regions, resulting trade flows will show cross-hauling because B will be both a domestic importer and an exporter of the same product.

The second trade flow model therefore reveals the problematic role of the timing of decision-making in regions and the probable competition dynamics among potential customers from different regions.

By way of contrast, the third trade model incorporates competitive dynamics in supplying across regions and requires modifying the model specification as presented below. It introduces a measure of competition in accessing products produced in other regions, p, that is the average transport cost that consumers are paying in those regions:

$$
\begin{aligned}
\min_{d} \quad & \sum_{j=0}^{N} (p_j + t_{ji}) d_{ji} \\
\text{s.t. } & p_i = \frac{\sum_{j \neq i}^{N} t_{ji} d_{ji}}{\sum_{j \neq i}^{N} d_{ji}}, \quad \forall\, i \\
& \sum_{j=0}^{N} d_{ij} = S_i, \quad \forall\, i \\
& \sum_{j=0}^{N} d_{ji} = D_i, \quad \forall\, i \\
& d_{ji} \geq 0, \quad \forall\, i, j
\end{aligned}
$$

Customers minimize their costs as before, but costs now comprise two elements, namely, transport costs and a measure of the difficulty to gaining access to products. The difficulty of getting products from a region is approximated by the average transport cost paid by customers in the region considered to be a potential supplier (source) of that product. For example, a region A has two potential suppliers with identical transport costs, regions B and C. Region B is already importing some products from other regions and thus is paying positive transport costs, while C is not importing anything. In this example, A chooses C as supplier because the costs to obtain such products from C are lower not having to compete with customers in that region and from others.

In summary, the third trade flow model has customers searching for suppliers and considering both transport costs and competition in acquiring products. In this framework, cross-hauling can emerge because of the sequential access of customers to the market.

The three models utilize different assumptions surrounding the role played by economic dynamics for the estimation of trade flows. They all rely on similar information and are based on cost minimization. Moreover, they imply that for each sector the estimation is done for all regions at the same time. Consequently, the entire approach has to be applied including all regions composing a country.

As a final remark, in order to estimate trade flows, transport costs information is needed. A first approximation of transport costs could be the simple computation of the geodesic distance between the centroids of regions, as often done in the literature using gravity models, but that does not consider the crucial role of infrastructure.

Today several research centres make available more sophisticated transport models mainly developed for the analysis of policies aimed at infrastructure improvement. It would probably be preferable to select two transport models, one for goods and one for services because the infrastructure for moving goods and people is only partially overlapped and transport costs may be different (for an example related to Europe, see Thissen, Diodato, & Van Oort, 2013).

Moreover, since transport costs are determined by transport infrastructure, which takes long times to significantly change, it is plausible to assume that they are relatively stable over time and that exogenous shocks, such as those on oil prices, affect homogeneously transport costs and markets, and thus relative distances between regions are rather stable. In fact, it is not needed a measure of transport costs in specific prices (e.g., current, real, total, etc.) but just a measure of the distance or impedance between regions based on any kind of unit of measurement, even without a direct economic measurement (see the example in the following section). Further, more than transport models what is needed by the procedure is the model result, which is a matrix of distances between all regions composing the nation.

In conclusion, similarly to hybrid and partial-survey methods, the approach presented here can easily gain advantage from the availability of further data sources (Dewhurst, 1992). For example, if a survey provides details about the technology adopted in the economy of a region, this information can be easily inserted in tables specification. Similarly, if specific patterns of consumption are known for a region, they can be considered in estimating regional demand (Trigg & Madden, 1994).

Further, the fact that with this method all regions composing a country have to be estimated together allows extending the informative value of supplemental regional specific information. In fact, when further information is known for a single region within the country, the estimation of other regions is improved as well because region specific information is a constraint added to the estimation. In other words, the availability of detailed supplementary information for a region allows improving the estimation for that region but also for others.

TEST OF THE METHODOLOGY

In order to start evaluating the proposed method, this section presents two applications of the model that can be compared with observed data and with results from other methods.

The area of interest is the United States and the initial focus is on interstate trade flows of commodities. The analysis of commodity flows is aimed at comparing results obtained with the different alternative trade flow models introduced above, and with a gravity model.

Second, the procedure is used to estimate the regional IO table of Washington State in 2007, which is compared with a table of the same state and period developed with a hybrid method.

The empirical data that are used are introduced before presenting the examples of application.

Data sources

Being focused on the US economy, the procedure uses counties as the spatial scale of estimation. Tables are then aggregated to the state scale for the sake of the analysis. In order to regionalize IO tables according to the method presented in the preceding section of this work, three data sources are needed: national IO accounts, measures of regional (i.e., county level) output and transport costs between counties.

As already mentioned, since the starting point of any attempt to estimate regional IO tables is the national IO table for the period of interest, including the necessary degree of industrial specification (Comer & Jackson, 1997), the first data needed to build unbalanced county-level IO tables are national IO accounts.

National supply and make tables for the United States are published by the US Bureau of Economic Analysis (BEA). The BEA provides several kinds of IO economic accounts. For this work, annual tables are used at the summary level of aggregation (65 sectors), which roughly corresponds to the three-digit level in the North American Industrial Classification System (NAICS). The BEA tables considered here include current producers' prices and follow standard commodity and industry definitions (i.e., columns of make tables represent commodities and rows represent industries). Make-and-use tables are aggregated to a symmetrical, industry-by-industry IO table, based on the assumption of a fixed-product sales structure according to the EUROSTAT procedure 'model D' (EUROSTAT, 2008, p. 349), where each product has its own specific sales structure, regardless of the industry where it is produced.

To estimate county outputs, it is common to rely on labour market data due to matters of reliability and availability. For national and local labour markets, annual data of employment and average weekly salaries are taken from the US Quarterly Census of Employment and Wages (QCEW) published by the US Bureau of Labor Statistics (BLS). These data are published at the national, state and county levels, and here the three-digit NAICS level is used. Data are published for establishments of any size and for all kinds of ownership. QCEW data are reliable because they are collected by states in order to manage unemployment insurance (further details about possible biases associated with this data source are presented in Appendix A in the supplemental data online).

QCEW records any job and wage observed in the period, regardless of the duration of the job. The number of employees recorded there is thus inflated by part-time jobs and average salaries are underestimated because they do not refer to full-time employees only. Considering these peculiarities of QCEW, the total payment to the labour factor, computed by multiplying the total number of employees and their average weekly salary, is used as the approximate measure of national and regional output for each industry.

Final demand is estimated using the assumption on homogeneous preferences and distributing national consumption according to the ratio between regional total gross domestic product (GDP) and the national one. As measures of transport costs between US counties, results from the 2007 intermodal transportation network, managed by the Center for Transportation Analysis of Oak Ridge National Laboratory, are used. The transportation model covers the entire highway, railroad and water transportation infrastructure in the United States. In particular, the networks covered include private and for-hire trucks, railroads, inland waters and Great Lakes, ocean vessels and intermodal terminals, including all their characteristics, such as number of lanes, tolls, congestions etc. (Peterson, 2000).

A county-to-county distance matrix is estimated by running an intermodal network algorithm on the transportation model that minimizes the impedance between every two county centroids. Distances use 1 highway-mile (0.9 for 1 mile on rural interstates) as the impedance unit and assigns impedances for modal transitions according to the characteristics of modes and intermodal terminals used. Impedance is used as a non-monetary measure of transport costs.

Data on the movement of goods in 2002 and 2007 contained in the Commodity Flow Survey (CFS) administered in partnership by the Census Bureau and the Bureau of Transportation Statistics are used as a data source for validation of trade flows estimation. The CFS provides information on domestic trade flows of commodities from manufacturing, mining, wholesale, and some retail and services establishments. In particular, data on the economic value of transported commodities by origin and destination are used at the spatial scale of states.

CFS data show origin and destination of commodities in terms of transportation, but they do not equal the origin and destination in terms of regional IO tables where they mean, respectively, the place of production and the place of consumption. To make the point clearer, CFS observes a commodity being transported from A to B, but that commodity could have been produced in C and it could be consumed in D. That is the reason why state-level data are used instead of data with a higher level of spatial detail (up to metropolitan areas). It is in fact plausible to assume that much of the distribution networks and transport hubs are enclosed by state boundaries, but this does not preclude the possibility of errors due to the fact that the CFS observes segments of the economic trade flow between production and consumption places. Looking at state-level data reduces the error, but it still can be relevant in particular for smaller states and for industries concentrated in a few areas.

The Office of Financial Management (OFM) of the state of Washington publishes an IO table for the economy in the state in 2007. The table is estimated using a hybrid or partial-survey procedure (Beyers & Lin, 2014).

The estimation is conducted starting from the direct coefficients derived from the benchmark table of the BEA for the US economy in 2002. The coefficients are then updated and regionalized at the same time using the information collected through a survey. The sample includes 2531 establishments, covering about 50% of the gross business income recorded in the state five years previously, i.e., in 2002. The survey sample is not stratified and information is not corrected through weights to ensure representativeness at the population level. Direct coefficients are then used with sale information collected in the survey and the table is balanced with the RAS procedure. No information is provided on the degree of error introduced by the balancing procedure.

Interstate trade flows

The comparison is limited by the number of commodities surveyed by the CFS, and thus only 17 industries producing manufactured goods can be compared. For the sake of comparison and validation, trade flows are estimated with four methods: one utilizes the approach taken by gravity models and three utilize the trade flow models described above. In all trade flow models two main sources of information are used. First, regional supply and demand are estimated as described above by means of unbalanced county-level IO tables. Second, the impedance between centroids of counties described before is used as measure of transport costs.

For the first method, a gravity model is estimated over the years of interest, 2002 and 2007, and at the state level in order to use the CFS data as a dependent variable. Its specification starts from the standard one:

$$t_{s,ij} = Su_{s,i}^{\alpha} De_{s,j}^{\beta} Tr_{ij}^{\gamma},$$

where $t_{s,ij}$ is the trade flow of the products of sector s from state i to state j; $Su_{s,i}$ is the supply of s in i; $De_{s,j}$ is the demand of s in j; and Tr_{ij} is the transport cost from i to j. The parameters α, β and γ are the ones to be estimated; and γ is expected to be negative so that increasing transport cost means decreasing trade flows. The model is then log-linearized and dummy variables are added for sectors (δ) and years (θ), transforming the model in:

$$\begin{aligned}\log(t_{s,ij}) = &\alpha \log(Su_{s,i}) + \beta \log(De_{s,j}) + \gamma \log(Tr_{ij}) \\ &+ \delta + \theta + \varepsilon,\end{aligned}$$

where ε is the error term.

Supply and demand at the state level are computed by summing the supply and demand of the counties that comprise each state. Transport costs within a state are computed as the average impedance between counties in each state. Transport costs between two states are computed as the average impedance between counties in one state and counties in the other state.

The model is estimated using ordinary least squares (OLS) and coefficients are significant ($p = 0.000$) and $\alpha = 0.689$, $\beta = 0.603$ and $\gamma = -1.554$. Dummy variables are significant as well. Trade flows are then predicted with the calibrated model, and results are presented in Table 1 (in the column labelled *gm*).

Having at one's disposal county-level measures of supply and demand and transport costs within and between counties, three micro-simulation models are developed to implement at the county level the three trade flow models described in the previous section. Models are implemented as micro-simulations, which are models that operate at the high level of detail of county-level supply and demand for each good and service, because they are based on autonomous and decentralized decision-making happening in sequential order (Savard, 2003; van Wissen, 2000).

The first micro-simulation (called *t*1 from now on) implements the first trade flow model that considers trade flows determined only by excesses in supply and demand and without cross-hauling. The micro-simulation proceeds as follows: for each industry, a county with unbalance (i.e., with local supply and demand differing) is randomly selected and its trade deficit or surplus is balanced with nearest neighbours presenting a complementary situation. For example, if the selected county presents a trade surplus, the procedure creates a trade flow with the nearest county with a trade deficit. If the values of surplus and deficit in the two counties are the same both counties are balanced. Otherwise, one of the two remains unbalanced and the procedure restarts from the beginning, searching for the nearest county that can balance the selected one. The procedure continues until all counties are balanced. Balance for all counties is ensured by the way regional supply and demand are estimated, and by using a starting national IO table that is balanced.

The second micro-simulation model (*t*2) is similar and it implements the second trade flow model. It starts by randomly picking an unbalanced county. Then a trade partner is selected according to transport costs and the availability of goods for domestic trade, a situation which does not imply having trade imbalance. In fact, while in *t*1 trade partners are selected according to distances and to complementary imbalances (i.e., a county in surplus with a county in deficit), under *t*2 the procedure searches for nearest counties with the capacity of providing or absorbing goods, even if that means for them to become imbalanced or to worsen a pre-existing imbalance.

The capacity of providing and absorbing goods for domestic trade is defined, respectively, as the difference between local output and total export (either abroad or domestic) and as the difference between local consumption and total imports (either international or domestic). Supply and demand for domestic products are thus measures of the local capacity to provide goods for the national market and to absorb national products. For example, if a county is already exporting abroad and to regions all of its output, it cannot provide more goods for exchange in the domestic market. Similarly, if a county is already importing from other regions and international markets all of the goods it consumes, it cannot import more products from other counties.

As the procedure iterates, randomly selecting unbalanced counties and establishing trade flows with nearest regions with the capacity to provide or absorb domestic goods, there is no guarantee of convergence. The number of unbalanced counties may increase. In fact the balancing of a county can happen at the expenses of the balance of one

Table 1. Domestic trade flows comparison: flows between and within 50 US states, 2002 and 2007.

	Average state-level trade flows (US$, millions)					R^2			
	CFS	*gm*	*t*1	*t*2	*t*3	*gm*	*t*1	*t*2	*t*3
Agriculture–Food–Forestry	394.0	116.0	265.0	605.0	280.0	.582	.860	.892	.868
Chemical Products	300.0	115.0	116.0	238.0	123.0	.160	.581	.654	.573
Electronic and Electrical Products	298.0	193.0	103.0	174.0	110.0	.529	.755	.713	.738
Fabricated Metal Products	105.0	63.7	58.7	180.0	62.3	.378	.632	.830	.643
Furniture and Related Products	43.0	33.1	25.3	52.7	27.5	.464	.417	.645	.502
Machinery	184.0	127.0	57.3	142.0	61.3	.476	.535	.733	.638
Mining – Except Oil and Gas	13.7	7.2	34.0	46.7	35.0	.152	.272	.294	.302
Motor Vehicles – Bodies and Trailers – and Parts	249.0	156.0	112.0	235.0	118.0	.488	.468	.639	.441
Nonmetallic Mineral Products	54.2	26.2	25.7	68.5	27.2	.353	.737	.826	.743
Oil–Gas–Coal – and Products	321.0	84.5	325.0	374.0	340.0	.570	.762	.835	.743
Other Transportation Equipment	20.3	28.5	43.6	91.8	45.5	.141	.110	.368	.125
Paper Products	72.5	46.6	45.0	85.3	48.7	.432	.593	.803	.696
Plastics and Rubber Products	141.0	95.6	55.4	112.0	58.0	.292	.546	.686	.623
Primary Metals	126.0	92.7	60.7	127.0	65.9	.502	.770	.844	.812
Printed Products	48.0	33.9	24.6	67.6	26.2	.508	.522	.702	.497
Textile and Leather Products	133.0	107.0	34.1	57.2	35.6	.377	.530	.627	.547
Wood Products	56.9	31.8	33.7	65.4	35.3	.466	.687	.843	.701
All industries	151.0	80.0	83.5	160.0	88.2	.319	.680	.735	.680

or more other counties. Because the procedure can take a very long time to converge (if it can converge at all), a closure condition is added to the micro-simulation: if the number of balanced counties has not increased in the last 100 repetitions of the procedure, remaining unbalanced counties are balanced according to *t*1.

The third micro-simulation model, *t*3, is identical to *t*2 except for the choice of nearest suppliers by customers. In fact customers search for counties with the capacity of providing or absorbing goods and select suppliers that minimize the total cost that results from summing transport costs and those related to local competition on goods, as in the third trade flow model described before. Costs of competition in a county are equal to the average transport cost paid by customers in the county.

The results using *t*2 and *t*3 show similar dynamics in each different sector. After an initial increase in the number of counties with imbalance, such a number significantly decreases towards a positive small value, which is then balanced by the closure condition similar to *t*1.

All trade flow models and procedures do not guarantee the uniqueness of solutions. In each sector there may be multiple different configurations of trade flows that equally minimize objective functions respecting constraints, in particular if economic agents make decision sequentially in a random order. However, any solution provided is valid from the perspective of the information at disposal and of the integrated approach because it allows perfectly balancing regional IO tables, minimizing transportation and access costs, and respecting considered constraints. Studying the variation of results in repetitions of micro-simulations can support the identification of sectors for which more information should be collected. Preliminary analyses of multiple solutions, however, point out the high degree of similarity between solutions. In fact, a randomly selected sample of 10 different solutions (i.e., county-to-county trade flows) in the 'farms' sector in 2010 obtained with the *t*2 micro-simulation model shows an average 1.04% variation of the mean of domestic trade flows and very small differences in their distribution (i.e., the average Bhattacharyya coefficient measured on all the possible couples of solutions in the sample is equal to 0.99998).

Table 1 presents averages by sector of state-level trade flows and coefficients of determination between CFS data and values estimated according to the gravity model (*gm*) and micro-simulation models (*t*1, *t*2, *t*3). The last row of the table presents total values across sectors.

On aggregate, *t*2 provides the best estimates of trade flows, *t*3 estimates are not much different from those obtained with *t*1, and the gravity model *gm* provides the worst estimates. Over the 17 industries considered, *t*1 performs better than *t*2 in 'electronic and electrical products' only, with an R^2 that is about 6% better than *t*1. In 'mining except oil and gas' *t*3 outperforms *t*2 with an R^2 about 3% better. In all other industries *t*2 is more accurate than *t*1 and *t*3 and on average (over all industries) the improvement is about 8%, and about 130% in comparison with the gravity model. Although such improvements in estimation accuracy, *t*2 provides inaccurate estimations for the industries 'mining except oil and gas' and 'other transportation equipment' where the R^2 is lower than 0.5.

Table 2. Comparison of type I multipliers in the state of Washington, 2007.

Office of Financial Management (OFM) sectors	US Bureau of Economic Analysis (BEA) sectors (summary level)	OFM multiplier	Integrated approach multiplier
Crop Production; Animal Production	Farms	1.391	1.486
Forestry and Logging; Fishing, Hunting, and Trapping	Forestry – Fishing – and Related Activities	1.516	1.344
Mining	Mining	1.456	1.050
Electric Utilities; Gas Utilities; Other Utilities	Utilities	1.363	1.287
Highway and Street Construction; Other Construction	Construction	1.464	1.404
Food, Beverage and Tobacco Manufacturing	Food and Beverage and Tobacco Products	1.470	1.485
Textiles and Apparel	Textile Mills and Textile Product Mills; Apparel and Leather and Allied Products	1.384	1.142
Wood Product Manufacturing	Wood Products	1.734	1.486
Paper Manufacturing	Paper Products	1.439	1.497
Printing	Printing and Related Support Activities	1.361	1.165
Petroleum and Coal Products	Petroleum and Coal Products	1.238	1.408
Chemical Manufacturing	Chemical Products	1.236	1.117
Nonmetallic Mineral Products Manufacturing	Nonmetallic Mineral Products	1.276	1.239
Primary Metals	Primary Metals	1.469	1.260
Fabricated Metals	Fabricated Metal Products	1.361	1.124
Machinery Manufacturing	Machinery	1.454	1.138
Computer and Electronic Product	Computer and Electronic Products	1.479	1.334
Electrical Equipment	Electrical Equipment – Appliances – and Components	1.317	1.181
Aircraft and Parts; Ship and Boat Building; Other Transportation	Motor Vehicles – Bodies and Trailers – and Parts; Other Transportation Equipment	1.136	1.778
Furniture	Furniture and Related Products	1.342	1.266
Other Manufacturing	Miscellaneous Manufacturing; Plastics and Rubber Products	1.399	1.189
Wholesale	Wholesale Trade	1.204	1.154
Non-Store Retail; Other Retail	Retail Trade	1.194	1.214
Air Transportation	Air Transportation	1.367	1.367
Water Transportation	Water Transportation	1.470	1.573
Truck Transportation	Truck Transportation	1.483	1.286
Other Transportation/Postal Offices	Rail Transportation; Transit and Ground Passenger Transportation; Pipeline Transportation	1.465	1.272
Support Activities for Transportation, Warehousing and Storage	Other Transportation and Support Activities; Warehousing and Storage	1.527	1.206
Software Publishers and Internet Service Providers	Publishing Industries (includes Software)	1.208	1.482
Telecommunications	Broadcasting and Telecommunications	1.442	1.459
Other Information	Motion Picture and Sound Recording Industries; Information and Data Processing Services	1.126	1.206
Credit Intermediation and Related Activities	Federal Reserve Banks – Credit Intermediation – and Related Activities	1.605	1.233

(Continued)

Table 2. Continued.

Office of Financial Management (OFM) sectors	US Bureau of Economic Analysis (BEA) sectors (summary level)	OFM multiplier	Integrated approach multiplier
Other Finance and Insurance	Securities – Commodity Contracts – and Investments; Insurance Carriers and Related Activities; Funds – Trusts – and Other Financial Vehicles	1.718	1.217
Real Estate and Rental and Leasing	Real Estate; Rental and Leasing Services and Lessors of Intangible Assets	1.153	1.257
Legal/Accounting and Bookkeeping/ Management Services	Legal Services; Management of Companies and Enterprises	1.108	1.185
Architectural and Engineering/ Computer Systems Design and Related Services	Computer Systems Design and Related Services; Miscellaneous Professional – Scientific – and Technical Services	1.191	1.198
Educational Services	Educational Services	1.614	1.324
Ambulatory Health Care Services	Ambulatory Health Care Services	1.491	1.250
Hospitals; Nursing and Residential Care Facilities, Social Assistance	Hospitals and Nursing and Residential Care Facilities; Social Assistance	1.395	1.282
Arts, Recreation, and Accommodation; Food Services and Drinking Places	Performing Arts – Spectator Sports – Museums – and Related Activities; Amusements – Gambling – and Recreation Industries; Accommodation	1.456	1.259
Food Services and Drinking Places	Food Services and Drinking Places	1.473	1.339
Administrative/Employment Support Services	Administrative and Support Services	1.116	1.187
Waste Management/Other, and Agriculture Services	Waste Management and Remediation Services; Other Services – except Government	1.503	1.349
	Federal General Government; Federal Government Enterprises; State and Local General Government; State and Local Government Enterprises	n.a.	1.267

Note: n.a., Not available.

Intraregional impacts

The state level IO table of Washington is obtained aggregating county-level IO tables estimated using the *t*2 micro-simulation model for the US economy using 2007 data. The table is at the 65-sector BEA summary level. For purposes of comparing it with the OFM table produced with the hybrid method, sectors are reduced to 44 as presented in the second column of Table 2. Surprisingly, the OFM table does not consider the government as an industrial sector that acquires intermediates to supply other sectors with its services.

For the sake of the analysis, type I multipliers are computed both on the OFM table and on the one estimated with the integrated approach (Frechtling & Horvath, 1999). Results are presented in Table 2.

The comparison of multipliers points out two important results. First, on average multipliers derived from the table estimated with the integrated approach are about 7% smaller than the ones derived from the OFM table. This result is valid across categories of industries, being integrated approach multipliers on average 8.6% smaller than OFM ones in 'agriculture, mining and construction', 5.8% in 'manufacturing' and 6.6% in 'services'. Integrated approach multipliers are smaller than OFM ones in 29 industries of the 43 considered (excluding from the comparison 'government' where the OFM estimation is not available). The industries in which this negative difference is larger are 'mining' and the ones related to credit and finance (all multipliers in these industries are at least 30% smaller than OFM ones).

Second, at the same time sectors of regional specialization and concentration in Washington have very small multipliers in the OFM estimation and very large multipliers in the other case. For example, the manufacturing of transportation equipment, which includes the aerospace industry, presents the lowest multiplier of all manufacturing sectors in the OFM table. On the contrary, with the integrated approach that sector presents the largest multiplier in manufacturing, which is 56.5% larger than OFM one. Similar arguments can be done about the software industry, where the integrated approach multiplier is 22.7% larger than the respective OFM one and the second largest one in 'services' (after 'water transportation').

DISCUSSION AND CONCLUSIONS

The comparative analysis of estimated interstate trade flows points out three main results. First, most accurate estimations are those obtained with the trade flow model that considers minimization of transport costs only and that allows the emergence of cross-hauling through the sequential access of customers to the market.

Second, even if the third trade flow model allows for cross-hauling through sequential decision-making, it generates results very much similar to the ones obtained with the first trade flow model, where cross-hauling is impossible and trade flows are determined by regional trade deficits only. Incorporating competition costs does not improve trade flows estimation and, on the contrary, increases costs of domestic trade so much that the only flows estimated are those strictly necessary to balance trade deficits.

Third, the gravity model performs very poorly, even if it is calibrated with the data that it is supposed to predict, and thus it should effectively minimize prediction errors. It is surprising that an empirical model based on the minimization of prediction errors performs worse than other models, in particular when considering that is evaluated over the same dataset that is used for its parameters calibration.

One possible explanation for this is that the gravity model specification might be inappropriate for this approach and for the data used. In fact, trade flows estimation is here conducted on accurately estimated values of total regional supply and demand. The unknown is the trade flow, which is how supply is distributed to demand, and transport cost minimization seems to have most of the explanatory power in this context. Product differentiation, and other dynamics such as localization economies that can be modelled by means of gravity models, are probably already captured by values of regional supply and demand. If product differentiation gives rise to the concentration of supply and demand in some areas, the data used by the model already capture these effects and make the gravity model specification not particularly effective. If other data about regional size are used or if product differentiation can be observed, results may be different (Alonso-Villar, 2007; Brander, 1981; Kronenberg, 2009).

From the perspective of accuracy in regional IO tables, the analysis is more problematic. The comparison presented above concerns a table estimated with a hybrid method. Without sounder empirical reference at disposal, an analysis of estimation accuracy as the one conducted for trade flows is impossible. Results about multipliers in sectors of local specialization however are particularly interesting.

Washington multipliers estimated with the hybrid method show that sectors of regional specialization rely less than other sectors on local supply. This result is inconsistent with the evidence (Bai, Du, Tao, & Tong, 2004; Belderbos, Capannelli, & Fukao, 2001; Krugman & Venables, 1993; Turok, 1993) that largest sectors have a strong capability to purchase from local suppliers, given their local market power, and that local specialization requires local external economies. The hybrid method multiplier for the manufacturing of transportation equipment is the smallest multiplier of all manufacturing sectors and, from the technical viewpoint, this is probably due to not accounting for local purchasing power.

The non-consideration of the relative size of purchasing sectors is typically observable also in non-survey methods based on the simple LQ, and it has been the main reason for the development of other LQ-based approaches (Flegg & Webber, 2000; Morrison & Smith, 1974; Schaffer & Chu, 1969).

From the viewpoint of the evaluation of the integrated approach presented here, results show that it is not affected by this problem, even if this problem has not been explicitly considered in its design.

In conclusion, the integrated approach seems to allow a reasonably good degree of accuracy both in regional IO tables and in domestic trade flows, which in this case are mutually dependent. Further, by definition it allows obtaining scalability. If all regions are aggregated, the national table will be obtained back. But scalability concerns also the important issue (Andrew & Peters, 2013; Andrew, Peters, & Lennox, 2009; Harris & Liu, 1998) of propagation of effects across different spatial and analytical scales. Precisely modelling the propagation of effects means to compute more precisely total effects at the national level.

Disadvantages and limits of the proposed approach are mainly two. First, it implies a computational burden. If, for example, the focus is on a single county in the United States, the integrated approach requires computing IO tables in all counties in the country, which are more than 3000, and estimating a number of unknown trade flows that is equal to the product between the number of sectors and the square of the number of counties. Second, the approach does not allow multiregional analyses because it does not differentiate trade flows between intermediate and final consumption.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors.

FUNDING

The authors acknowledge the financial support of the National Infrastructure Simulation and Analysis Center (NISAC) at Los Alamos National Laboratory (LANL).

SUPPLEMENTAL DATA

Supplemental data for this article can be accessed at http://dx.doi.org/10.1080/00343404.2017.1286009

ORCiD

Riccardo Boero http://orcid.org/0000-0002-7468-9096

REFERENCES

Alonso-Villar, O. (2007). A reflection on the effects of transport costs within the new economic geography. *Review of Urban and Regional Development Studies*, *19*, 49–65. doi:10.1111/j.1467-940X.2007.00117.x

Andrew, R. M., & Peters, G. P. (2013). A multi-region input–output table based on the global trade analysis project database (GTAP-MRIO). *Economic Systems Research*, *25*, 99–121. doi:10.1080/09535314.2012.761953

Andrew, R. M., Peters, G. P., & Lennox, J. (2009). Approximation and regional aggregation in multi-regional input–output analysis for national carbon footprint accounting. *Economic Systems Research*, *21*, 311–335. doi:10.1080/09535310903541751

Bai, C.-E., Du, Y., Tao, Z., & Tong, S. Y. (2004). Local protectionism and regional specialization: Evidence from China's industries. *Journal of International Economics*, *63*(2), 397–417. doi:10.1016/S0022-1996(03)00070-9

Belderbos, R., Capannelli, G., & Fukao, K. (2001). Backward vertical linkages of foreign manufacturing affiliates: Evidence from Japanese multinationals. *World Development*, *29*(1), 189–208. doi:10.1016/S0305-750X(00)00086-3

Beyers, W. B., & Lin, T.-W. (2014). *The 2007 Washington input–output study, full report, September 2014 revision*. Olympia: Washington State Office of Financial Management.

Bonfiglio, A., & Chelli, F. (2008). Assessing the behaviour of non-survey methods for constructing regional input–output tables through a Monte Carlo simulation. *Economic Systems Research*, *20*, 243–258. doi:10.1080/09535310802344315

Boomsma, P., & Oosterhaven, J. (1992). A double-entry method for the construction of bi-regional input–output tables. *Journal of Regional Science*, *32*, 269–284. doi:10.1111/j.1467-9787.1992.tb00186.x

Brander, J. A. (1981). Intra-industry trade in identical commodities. *Journal of International Economics*, *11*, 1–14. doi:10.1016/0022-1996(81)90041-6

Canning, P., & Wang, Z. (2005). A flexible mathematical programming model to estimate interregional input–output accounts. *Journal of Regional Science*, *45*, 539–563. doi:10.1111/j.0022-4146.2005.00383.x

Comer, J. C., & Jackson, R. W. (1997). A note on adjusting national input–output data for regional table construction. *Journal of Regional Science*, *37*, 145–153. doi:10.1111/0022-4146.00047

Dewhurst, J. H. L. (1992). Using the RAS technique as a test of hybrid methods of regional input–output table updating. *Regional Studies*, *26*, 81–91. doi:10.1080/00343409212331346791

Escobedo, F., & Ureña, J. M. (2008). The influence of transport links on disaggregation and regionalization methods in interregional input–output models between metropolitan and remote areas. In R. Minguez, J. Sarabia, N. Balakrishnan, & B. C. Arnold (Eds.), *Advances in mathematical and statistical modeling* (pp. 231–241). Boston: Birkhäuser.

EUROSTAT. (2008). *EUROSTAT manual of supply, use and input–output tables*. Luxembourg: EUROSTAT.

Flegg, A. T., & Tohmo, T. (2013). Regional input–output tables and the FLQ formula: A case study of Finland. *Regional Studies*, *47*, 703–721. doi:10.1080/00343404.2011.592138

Flegg, A. T., & Webber, C. D. (2000). Regional size, regional specialization and the FLQ formula. *Regional Studies*, *34*(6), 563–569. doi:10.1080/00343400050085675

Flegg, A. T., Webber, C. D., & Elliott, M. V. (1995). On the appropriate use of location quotients in generating regional input–output tables. *Regional Studies*, *29*, 547–561. doi:10.1080/00343409512331349173

Flegg, A. T., Webber, C. D., & Elliott, M. V. (1997). On the appropriate use of location quotients in generating regional input–output tables: Reply. *Regional Studies*, *31*, 795–805. doi:10.1080/713693401

Frechtling, D. C., & Horvath, E. (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. doi:10.1177/004728759903700402

Harris, R. I., & Liu, A. (1998). Input-output modelling of the urban and regional economy: The importance of external trade. *Regional Studies*, *32*, 851–862. doi:10.1080/00343409850118004

Isard, W. (1951). Interregional and regional input–output analysis: A model of a space-economy. *Review of Economics and Statistics*, *33*, 318–328. doi:10.2307/1926459

Isard, W., Azis, I. J., Drennan, M. P., Miller, R. E., Saltzman, S., & Thorbecke, E. (1998). *Methods of interregional and regional analysis*. Aldershot: Ashgate.

Isserman, A. M. (1977). The location quotient approach to estimating regional economic impacts. *Journal of the American Institute of Planners*, *43*, 33–41. doi:10.1080/01944367708977758

Jensen, R. C. (1990). Construction and use of regional input–output models: Progress and prospects. *International Regional Science Review*, *13*, 9–25. doi:10.1177/016001769001300102

Kowalewksi, J. (2015). Regionalization of national input–output tables: Empirical evidence on the use of the FLQ formula. *Regional Studies*, *49*(2), 240–250. doi:10.1080/00343404.2013.766318

Kronenberg, T. (2009). Construction of regional input–output tables using nonsurvey methods: The role of cross-hauling. *International Regional Science Review*, *32*, 40–64. doi:10.1177/0160017608322555

Krugman, P., & Venables, A. (1993). *Integration, specialization, and the adjustment* (Working Paper No. w4559). Cambridge, MA: National Bureau of Economic Research (NBER). doi:10.3386/w4559

Lahr, M. L. (1993). A review of the literature supporting the hybrid approach to constructing regional input–output models. *Economic Systems Research*, *5*, 277–293. doi:10.1080/09535319300000023

Lehtonen, O., & Tykkyläinen, M. (2014). Estimating regional input coefficients and multipliers: Is the choice of a non-survey technique a gamble? *Regional Studies*, *48*(2), 382–399. doi:10.1080/00343404.2012.657619

Leontief, W., & Strout, A. (1963). Multiregional input–output analysis. In T. Barna (Ed.), *Structural interdependence and economic development* (pp. 119–149). London: Macmillan.

Miller, R. E., & Blair, P. (1983). Estimating state-level input–output relationships from U.S. multiregional data. *International Regional Science Review*, *8*, 233–254. doi:10.1177/016001768300800304

Miller, R. E., & Blair, P. (2009). *Input–output analysis, foundations and extensions* (2nd ed.). Cambridge: Cambridge University Press.

Morrison, W. I., & Smith, P. (1974). Nonsurvey input–output techniques at the small area level: An evaluation. *Journal of Regional Science*, *14*(1), 1–14. doi:10.1111/j.1467-9787.1974.tb00425.x

Peterson, B. (2000). *CFS multi-modal network documentation*. Oak Ridge: Center for Transportation Analysis.

Riddington, G., Gibson, H., & Anderson, J. (2006). Comparison of gravity model, survey and location quotient-based local area tables and multipliers. *Regional Studies*, *40*, 1069–1081. doi:10.1080/00343400601047374

Robinson, S., Cattaneo, A., & El-Said, M. (2001). Updating and estimating a social accounting matrix using cross entropy methods. *Economic Systems Research*, *13*, 47–64. doi:10.1080/09535310120026247

Round, J. I. (1972). Regional input–output models in the U.K.: A re-appraisal of some techniques. *Regional Studies*, *6*, 1–9. doi:10.1080/09595237200185011

Round, J. I. (1983). Nonsurvey techniques: A critical review of the theory and the evidence. *International Regional Science Review*, *8*, 189–212. doi:10.1177/016001768300800302

Savard, L. (2003). *Poverty and income distribution in a CGE-household micro-simulation model: Top-down/bottom-up approach* (CIRPEE Working Paper No. 03-43). Montreal: CIRPEE.

Schaffer, W. A., & Chu, K. (1969). Nonsurvey techniques for constructing regional interindustry models. *Papers of the Regional Science Association*, *23*(1), 83–101. doi:10.1007/BF01941876

Shao, G., & Miller, R. E. (1990). Demand-side and supply-side commodity-industry multiregional input–output models and spatial linkages in the US regional economy. *Economic Systems Research*, *2*, 385–406. doi:10.1080/09535319000000029

Stevens, B. H., & Trainer, G. A. (1976). *The generation of error in regional input–output impact models* (Working Paper No. A1-76). Amherst: Regional Science Research Institute.

Stevens, B. H., & Trainer, G. A. (1980). Error generation in regional input–output analysis and its implications for nonsurvey methods. In S. Pleeter (Ed.), *Economic impact analysis: Methodology and applications* (pp. 68–84). Boston: Martinus Nijhoff.

Thissen, M., Diodato, D., & Van Oort, F. G. (2013). *Integrated regional Europe: European regional trade flows in 2000* (Publication No. 1035). The Hague: PBL Netherlands Environmental Assessment Agency.

Tohmo, T. (2004). New developments in the use of location quotients to estimate regional input–output coefficients and multipliers. *Regional Studies*, *38*, 43–54. doi:10.1080/00343400310001632262

Trigg, A. B., & Madden, M. (1994). Using a demand system to estimate extended input–output multipliers. *Economic Systems Research*, *6*, 385–395. doi:10.1080/09535319400000031

Turok, I. (1993). Inward investment and local linkages: How deeply embedded is 'Silicon Glen'? *Regional Studies*, *27*(5), 401–417. doi:10.1080/00343409312331347655

Van Wissen, L. (2000). A micro-simulation model of firms: Applications of concepts of the demography of the firm. *Papers in Regional Science*, *79*(2), 111–134. doi:10.1007/s101100050039

West, G. R., Morison, J. B., & Jensen, R. C. (1984). A method for the estimation of hybrid interregional input–output tables. *Regional Studies*, *18*(5), 413–421. doi:10.1080/09595238400185411