Survey based versus non survey based Multi-Regional Input-Output Tables – the case of Austria

First draft – please do not cite

Bartlomiej Rokicki
University of Warsaw, Faculty of Economic Sciences; and
Regional Economics Applications Laboratory, University of Illinois Urbana/Champaign.
E-mail: brokicki@wne.uw.edu.pl

Oliver Fritz
Austrian Institute of Economic Research (WIFO)
E-mail: Oliver.Fritz@wifo.ac.at

Jonathan M. Horridge
Centre of Policy Studies, Victoria University, Melbourne, VIC, Australia.
E-mail: Mark.Horridge@vu.edu.au

Geoffrey J.D. Hewings
Regional Economic Applications Laboratory, University of Illinois, Urbana-Champaign, USA
hewings@illinois.edu

ABSTRACT:

Spatial CGE models rely on detailed information from multiregional input-output systems. Multiregional input-output tables (IOT) are usually not available and have to be compiled. This paper compares two different approaches to compile regional IOT – algorithm based approach that mechanically regionalizes national IOT using a predefined set of regional variables and hybrid approach that uses as much regional data as possible. We aim at verifying whether a use of a given approach has a significant impact on CGE simulation results. In our case, we compile regional IOT for Austria applying ready-made Horridge algorithm and a hybrid approach. We find that aggregate simulation results are surprisingly similar. As a result, we could claim that algorithm based approach is in fact an effective way of regionalizing national IOT. However, once we compare the results at the sectoral level they start to differ significantly. This may raise serious concerns about the adequacy of certain results based on ready-made disaggregation models.

Keywords: input-output tables, CGE modelling, regionalization.

JEL codes: C67, O18, R15.

Acknowledgments

B. Rokicki gratefully acknowledges financial support from the Polish National Science Centre under the grant 2017/26/M/HS4/00732.

1. Introduction

Input-output and computable general equilibrium (CGE) modelling have become a common tools in macroeconomic modelling in the last few decades. While differing in certain methodological assumptions (e.g. West, 1995), both approaches rely on the same input-output tables. The latter are usually compiled by national statistical offices that use survey data from many different sources. The problem, however, is that in most of the cases such tables are not provided at regional level. Hence, the researchers that want to construct regional models almost always need to rely on information derived from the corresponding national input-output tables. Data required for such input-output tables, e.g. on the use of intermediate goods by firms in different sectors or on goods consumed by households, is not readily available at regional level. This is true in particular for data on regional exports and imports, which usually are neither collected by national statistical offices[1] nor easily retrievable even by regional producers themselves. Moreover, surveys to fill the existing secondary data gaps are very costly. Therefore regional table construction activities have shifted away from survey-based tables to tables based on so-called non-survey or hybrid (partial survey) methods (e.g. Greenstreet, 1989).

The methods used to regionalize national tables differ with respect to the amount of regional data used. The simplest non-survey methods merely rely on information concerning the basic sectoral structure of the region, e.g. measured by the sectoral employment distribution, and adjust national input-output tables by the means of regional location coefficients. Other methods (partial survey methods or hybrid methods) attempt to use more extensive regional data and often focus efforts to integrate this regional information on those sections of the table that are deemed most important in terms of regional multipliers or linkages (e.g. Greenstreet, 1989; Lahr, 1993)[2]. Nowadays, some software providers offer ready-made models that can estimate regional IOT, multiplicators, linkages etc. The most well known examples include IMPLAN, REMI or RIMS.

There exists a growing body of literature that compares different non-survey regionalization methods (e.g. Kowalewski, 2015), using different statistical criteria. However, to the best of our knowledge, there are no papers that actually compare the accuracy of simulation results

---

[1] Data recently published by Statistics Austria should be considered as a remarkable exception.

[2] In our case we have used regional data based on their availability. That means that we do not necessarily cover the key regional multipliers and linkages.

based on different regional IOT for the same country. As a result, nothing can be said about the behavior of different regional IOT in particular models. Also, we are not aware of studies that evaluate the accuracy of hybrid and ready-made regionalization approaches. This paper compares two different approaches to compile regional IOT – algorithm based approach that mechanically regionalizes national IOT using a predefined set of regional variables and hybrid approach that uses as much regional data as possible. We aim at verifying whether a use of a given approach has a significant impact on computable general equilibrium (CGE) simulation results. Hence, we compile regional IOT for Austria applying ready-made model (Horridge algorithm) and hybrid (partial survey) approach using the 2011 data. Then we calibrate The Enormous Regional Model (TERM)[3] using two different sets of regional IOT. Finally, we run a number of different simulations an analyze the differences in the results.

The remainder of the paper is organised as follows. Section 2 reviews prior literature on the regionalization of national input-output tables. Section 3 covers the research methodology, description of the data and an analysis of database differences. Section 4 presents the results of simulations based on two different methodological approaches to regionalization problem. Finally, section 5 offers some concluding remarks.

2. Literature review

Hybrid or partial-survey methods are nowadays considered as the most cost-effective in preparing regional input-output tables (e.g. Lahr, 1993). On one hand, they are much cheaper and less time consuming than survey based methods. On the other, they outperform simple non-survey based techniques (e.g. Bonfiglio and Chelli, 2008). Most of the hybrid methods differ only in terms of non-survey techniques used to adjust for regional trade patterns (e.g. Boomsma and Oosterhaven, 1992; Greenstreet, 1989; West, 1990). Hence, the choice of the most adequate non-survey approach is considered as a critical in terms of the quality of regionalization process. Four main strands can be distinguished in the literature concerning the non-survey[4], trade-adjust technology based IOT regionalization methods:

---

[3] For more details about TERM model see Horridge (2011).

[4] Actually, the term "non-survey" can be misleading since some of the above approaches actually require the use of some survey-based data (e.g. Lahr, 1993).

- Methods based on location quotient such as simple location quotient (SLQ), purchases-only location quotient (PLQ), cross-industry location quotient (CILQ), Flegg location quotient (FLQ) or industry-specific Flegg location quotient (SFLQ);
- Commodity balance approach and its extensions such as the cross-hauling adjusted regionalization method (CHARM) or modified CHARM (MCHARM);
- Regional purchase coefficient (RPC) approach;
- Bi-proportional methods such as RAS, entropy or mathematical programming models.

In recent years, there can be observed a growing body of methodological analyses devoted to the application of location quotient approach in the regionalization of national IOT. Here, the authors tend to compare the performance of existing methods or propose new approaches and adjustments. For instance, McCann and Dewhurst (1998) show that although there is no theoretical relationship between the size of the region and the values of the expenditure coefficients estimated using location quotients, there seems to exist a certain empirical relationship. They underline, however, the problem related to the inflation of the coefficients as the size of the region falls. Bonfiglio and Chelli (2008) apply Monte Carlo analysis and prove that FLQ[5] and its augmented version perform better than their predecessors such as SLQ. They also show that both methods tend to over/underestimate the impact with respect to the chosen values of parameter δ that influences the relative size of the regions within the formulae. The correct value of parameter δ has been a subject of other papers by Tohmo (2004), Bonfiglio (2009), Flegg and Tohmo (2013a, 2016), Kowalewski (2015), Zhao and Choi (2015) or Flegg et al. (2016). In general, most of the authors claim the location quotient approach to be one of the best non-survey alternatives to estimate regional IOT. Still, there is no agreement whether the FLQ or SFLQ is the best in terms of accuracy.

Commodity balance approach is an older methodology[6] that has been lately modified and extended within the CHARM framework. Following Kronenberg (2007), CHARM allows to improve the performance of location quotient approach in terms of estimating regional trade. By accounting for cross-hauling, it deals with the problem of underestimation of regional trade flows that leads in turn to overestimation of regional multipliers. Under the criticism by Flegg and Tohmo (2013b), initial CHARM approach was refined by Többen and Kronenberg (2015)

[5] See Flegg et al. (1995) for more details.
[6] Its roots can be found in the paper by Isard (1953).

in order to account for interregional trade[7]. Fujimoto (2018) claims that his version of the approach (called MCHARM) performs better in estimating regional trade flow in the case of Japan, as compared to LQ, FLQ and standard CHARM.

Regional purchase coefficient (RPC) approach was first proposed by Stevens et al. (1983). It estimates the proportion of regional demand fulfilled from regional production using regional economic and interregional transportation data. As shown by Brucker et al. (1990), RPC has been implemented in a number of ready-made models for the US such as IMPLAN. Here, the software offered as a part of bigger package is able to estimate regional I-O tables, multiplicators or linkages. Yet, several authors claim that RPC, is in fact, one of the weakest points of such a models (e.g. Lazarus et al., 2002).

The existing literature on RAS method shows that, in general, RAS produces satisfactory results (e.g. Miller and Blair, 2009; Morrison and Smith, 1974; Riddington et al., 2006; Harris and Liu, 1997; Oosterhaven and Escobedo, 2011; Wiebe and Lenzen, 2016). Actually, RAS can be applied alone, but most hybrid methods use RAS as an intermediate step in the procedure. For instance, Flegg and Tohmo (2013) suggest the combined use of FLQ and RAS to improve the accuracy of estimation. In the first step, national coefficients are modified by standard FLQ; then the table is constrained to be given a regional frame by RAS. The resulting regional tables are considered as a reasonably good representation of true regional I-O relations as compared to other non-survey methods (e.g. Szabó, 2015).

Other bi-proportional methods are less often discussed. Lahr and de Mesnard (2004) provide a review of different bi-proportional methods that in general share many RAS features. Canning and Wang (2005) implement a mathematical programming model to estimate interregional IO tables based on an interregional accounting framework and initial information of interregional shipments. Finally, there exist several papers applying entropy econometrics to the estimation of input-output tables (e.g. Fernandez Vazquez, 2015; Fernandez Vazquez et al., 2015). Still, they do not have a regional dimension.

There exists a large number of studies comparing the outcomes of different non-survey methods, beginning with the early works by Czamanski and Malizia (1969), Schaffer and Chu (1969), or Hewings (1969, 1971) and ending with the more recent papers by Oosterhaven et al.

[7] Kronenberg (2009) defined cross-hauling as a proportion to the sum of demand and output.

(2003), Tohmo (2004), Riddington et al. (2006), Miller and Blair (2009), Jiang et al. (2012) or Fujimoto (2018). In most of the cases regional IOT based on different LQ approaches and RAS are being compared, applying different statistical criteria. As claimed by Miller and Blair (2009), the results of the above analyses vary not only due to the obvious differences in methodological approaches but also due to statistics used in assessing their accuracy. Also, in most of the cases there do not exist a survey-based benchmark table to be compared with estimated tables. Note, that regional IOT are not only a basis for regional IO and econometric IO models but also for regional CGE models. Still, even though regional CGE models are becoming more and more popular we are not aware of any study that compare the results of CGE simulations based on regional IOT assembled with different disaggregation approaches. This is very surprising given the fact, that the differences in regional IOT may potentially have a significant impact on simulation results and the conclusions drawn upon these simulations.

3. Methodology and data

In this paper we apply 2011 national supply and use tables for Austria to compare the results of simulations based on two different regionalization methods: hybrid approach and ready-made model. Apart from the differences in disaggregation approach there also exists a fundamental difference in terms of calibration. In both cases the IOT are calibrated to fit TERM model that in turn builds on ORANIG database structure. Hybrid approach database is calibrated after regional disaggregation process. However, ready-made model first fits the national IOT into ORANIG format and then perform the disaggregation procedure. As a result, there are possible differences not only in terms of regional matrices but also small differences in national totals. Below, we describe in detail two regionalization approaches followed by an analysis of resulting database differences.

3.1 Compiling a multiregional make-use system for Austria: a hybrid approach

Compilation of a multiregional input-output system aimed at full consistency with the national make-use system for 2011 (which mirrors the national accounts) on the one hand and full consistency with the official regional accounts on the other hand. Furthermore, input-output relevant information at the regional level from several sources were utilized; these include data from the Structural Business Statistics, the Household Budget Survey and from various sources providing information on government consumption by region. In addition, the results of a survey on interregional trade, carried out by the Austrian Institute Economic Research in the year 2000, were used. The resulting regional tables may thus be characterized as hybrid: table compilation relied on the use of extensive amounts of primary and secondary regional data for some sections of the tables, while for others, in particular those that depict service-related industries and commodities for which regional data was scarce or did not exist at all, the structure of the corresponding sections of the national tables had to be retained and regional information was limited to column sums of the tables (i.e. output levels).

Table compilation proceeded in the following steps:

Step 1 - *Estimation of missing information by industry* (output, value added, employment etc. at NACE 2-digit level) when data were published at a higher level of aggregation or not disclosed for confidentiality reasons. RAS was applied to ensure the above-mentioned consistency with national and regional accounts for each variable.

Step 2 - *Estimation of a regional make matrices*. Production by activities was transformed into production by commodities using the corresponding shares from the national make-matrix; contrary to a previously estimated table for 2001 information on commodity production by industry and region was not being made available to derive a genuinely regional make structure.

Step 3 - *Estimation of regional intermediate and final use matrices*, valued at purchasers' prices, independent of the origin of the commodities used.

During this step regional commodity input values by industries resulted from multiplying total intermediate use values by industries and regions with the respective commodity shares. Total intermediate use values were calculated by deducting value added from total output values both estimated in step (1). Generic regional information about commodity shares in total intermediate use was available only for mining and manufacturing, both with respect to industries and commodities. Since data on the use of services across all industries and on any input-commodity use by service industries are missing, national commodity input shares by industries from the national intermediate use table were applied to fill those gaps. Again, RAS was utilized to ensure consistency with the national intermediate use matrix with respect to total intermediate commodity use.

For private household consumption the information from Statistic Austria's Household Budget survey was used to identify differences in regional consumption patterns (e.g. with respect to the use of public transport services which is higher in urban than in rural areas of the country).

This data, however, only measures consumption at the place of residence. Total private consumption of the national input-output table, however, also includes consumption by foreign and domestic tourists at their respective destinations. While the value of total consumption by foreigners is published in the national make-use matrices, total consumption by domestic tourists had to be estimated. Information from the national and a few regional tourism satellite accounts for Austria on consumption by broad groups of goods and services was complemented with various other data to arrive at a complete commodity vector for foreign and domestic tourism.

National <u>public consumption expenditures</u> were regionalized with respect to each commodity either directly by using regional public consumption data provided by Statistics Austria or indirectly by applying different regional indicators based on a place of consumption concept; adhering to the latter is particularly important (but also tricky) since the production of government services at the federal level is highly concentrated in Vienna, the capital city of Austria, which implies that regional consumption of those services deviates significantly from regional production patterns. Specifically, shares of regional population in national population were used as indicators for commodities that could be classified as public goods; these included, for instance, national defence and part of national government services. Education services were regionalized by the number of students at different levels of education, counted at the location of the educational institution. Public expenditures on health services and pharmaceuticals were first allocated to different (partly regional) health insurance carriers based on the number of insured persons and then further regionalized if necessary. Since employees and their dependants are assigned to health insurance carriers based on the location of their employer and furthermore often stay in hospitals outside their home region adjustments for commuting (based on census data) and out-of-province hospitalization (based on data on regional hospital occupancy and the assumption of equal cost per occupied hospital bed across all regions) had to be made in order to comply with the place of consumption concept.

<u>Regional investment</u> was derived from the corresponding matrix of the Austrian input-output table, which shows investment values by industries and commodities, by applying national commodity shares to regional investment totals for each industry and a RAS procedure for balancing.

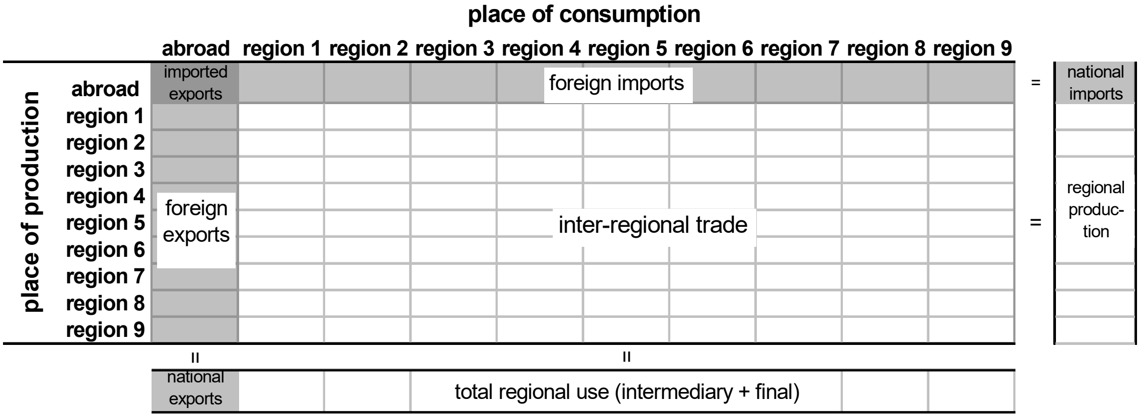
<u>Regional foreign exports</u> were computed by utilizing information on the location of the exporter contained in the national external trade statistics database. The main drawback of this data is its

unit of observation: it is the company level, whereas a meaningful regional input-output table compilation requires the establishment level since especially larger companies tend to have several establishments located in different regions with one single business unit at their headquarters which is responsible for managing exports and imports for the whole company. This results in a heavy bias towards the location of the headquarter (in many cases Vienna). Considerable effort and additional data (e.g. sales tax statistics) were used to correct for this problem. The regional foreign exports derived in this way serve only as first estimates and are revised when interregional trade is added to balance the multiregional input-output table system (see below).

Step 4 - For the estimation of *foreign imports* by region national external trade statistics are less useful: additional to the company-establishment problem imports are often declared by the transporting company instead of the company the imported good is intended for. For that reason, foreign imports were regionalized using national import ratios by commodity and industry (final demand category) Again, these results only served as starting values for a balancing mechanism described below.

Step 5 - The estimation of *interregional trade flows* is based on a RAS procedure which is set up using an accounting identity; initial values are derived from a firm survey on interregional trade conducted in the year 2001. The accounting identity claims that for each region and each commodity, the value of total use of a commodity by firms and households within this region plus the value of regional and foreign exports must equal the total value of a commodity available in the region, i.e. the value of production by regional firms and the value of imports from other regions or from abroad. In other terms, whatever is consumed within the region or is exported must be produced somewhere, either in the region itself or in other regions or abroad. Equivalently, for each region and each commodity it must hold that the total value of production is equal to the total value of use of a regionally produced commodity within the region (by firms and households) plus the value of exports of regional production to other regions and abroad. In other terms, whatever is produced within the region must be consumed somewhere. Accounting for commodity supply and use in that manner provides values for the sums across rows and columns of the following matrix:

Figure 1. Balancing of interregional trade



Source: authors preparation

With row and column sums and initial values given, a RAS balancing procedure is carried out for each commodity resulting in final values for interregional trade and foreign exports and imports by region.

Applying this method allows for "cross-hauling": a commodity can at the same time be bought and sold by each region (instead of assuming that only "surplus production" is exported and only that part of demand is imported which cannot be satisfied out of regional production, respectively). One major drawback concerns the fact that "trans-shipping" is disregarded: this is the case when a commodity is imported into region 1 from region 2 and sold – unchanged – to region 3. A regional miss-allocation of trade (and transport) margins may result. Another and even more important drawback of the methodology applied lies in the implicit assumption of uniform import shares across all consuming industries and final consumption categories; the procedure estimates interregional and international trade patterns without discriminating between the use of the commodities.

Given interregional and international trade patterns from the RAS procedure described above, the final step of table compilation consisted of computing, for each region, matrices depicting intermediate and final use with respect to commodities produced within the region itself and with respect to commodities imported from abroad. Since there the sum of individual cells of foreign import matrices summed up over regions need not be consistent with the values in the corresponding cells of the national foreign import matrices, additional balancing is required.

3.2 Compiling a multiregional make-use system for Austria: a ready-made algorithm

As mentioned above, regionalization procedure based on the ready-made algorithm by Horridge is preceded by calibration of national IOT into the ORANIG format. At this stage additional regional data is provided to be used during the regionalization process. It includes regional industry shares, occupation shares by industry, investment shares or choice of commodities provided locally. In our case regional industry shares were calculated using sectoral employment data from both Structural Business Statistics and Employment by NUTS2 regions provided by EUROSTAT. The data used to calculate occupation shares also comes from EUROSTAT. The data on investment shares comes from Statistics Austria. All data covers the year 2011.

Similar as in the case of hybrid approach, regionalization procedure was divided into several steps:

Step 1 - *Formal verification that the national database* used in the regionalization process is compatible with the ORANIG format and that the database adds up. Here, it is also verified that the sum of each matrix is not negative and that there are no margins and taxes where flows equal 0. Finally, preliminary estimation of regional GDP is performed.

Step 2 - *Initial database is reformatted from ORANIG to TERM format*. This requires for instance to add (distribute) tariff revenue to the TAX matrix subtracting equivalent amounts from the BASIC matrix (ORANI-G values imports at tariff-paid prices, TERM values imports at CIF). Additionally, preliminary estimate of average distance travelled by goods based on provided distance matrix (travel time between capitals of difference regions) is performed.

Step 3 – *Application of regional shares to split user columns according to destination*. The user specifies what share of imports enters by a given region. Hence, in our case it was assumed that regions without external border would have a very tiny share. On other hand, in the EU this may not be relevant since there are no custom duties for intra-European trade. The user also determines a subset of local commodities – these are the ones the we believe to be supplied almost in 100% locally (e.g. primary education). He also defines the set of distance related margins and he specifies the exponential of the gravity formula. Finally, trial trade matrices are constructed to be used later in the RAS procedure.

Step 4 – *Interregional trade estimation*. In this step RAS procedure, is applied in order to scale TRADE matrix and margins to meet control totals. The procedure is repeated three times: first using conventional scaling approach, then using linear system approach and finally using again conventional scaling approach.

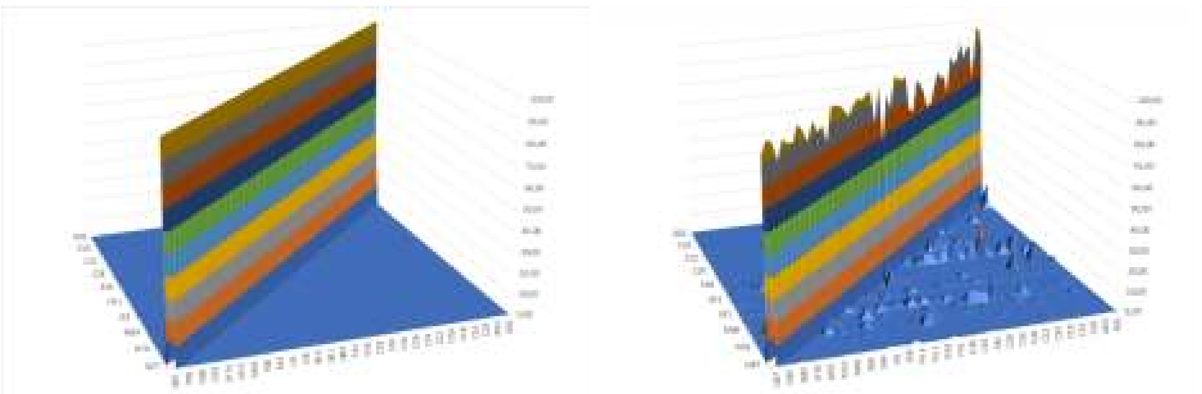
Step 5 - Final step consists in checking whether trade data after RAS sums to targets and combining regional IO tables and trade data into one file. Additionally the database can be aggregated for a given set of regions or industries before any simulation. Note, that regionalization algorithm assembles the multiproduction matrix (MAKE) as a diagonal one. Regionalization algorithm allows also to prepare additional matrices used in a recursive version of TERM model. However, in our case simulations were performed applying standard static approach.

3.3 Differences in estimated interregional matrices

Preliminary analysis of differences in the main interregional matrices shows that these differentials are relatively tiny at the aggregated level. However, they become much more significant once we disaggregate the matrices into particular regions or particular industries.

Figure 2 shows the differences between the multiproduction matrices (MAKE). The ready-made model matrix is simply diagonal, while hybrid-approach based matrix is almost diagonal. In most of the cases the values of cells outside diagonal are very small. Most important exceptions include industries such as Food, beverage and tobacco, Manufacture of non-metallic mineral products, Manufacture of basic metals, Manufacture of machinery and equipment n.e.c. or Trade.

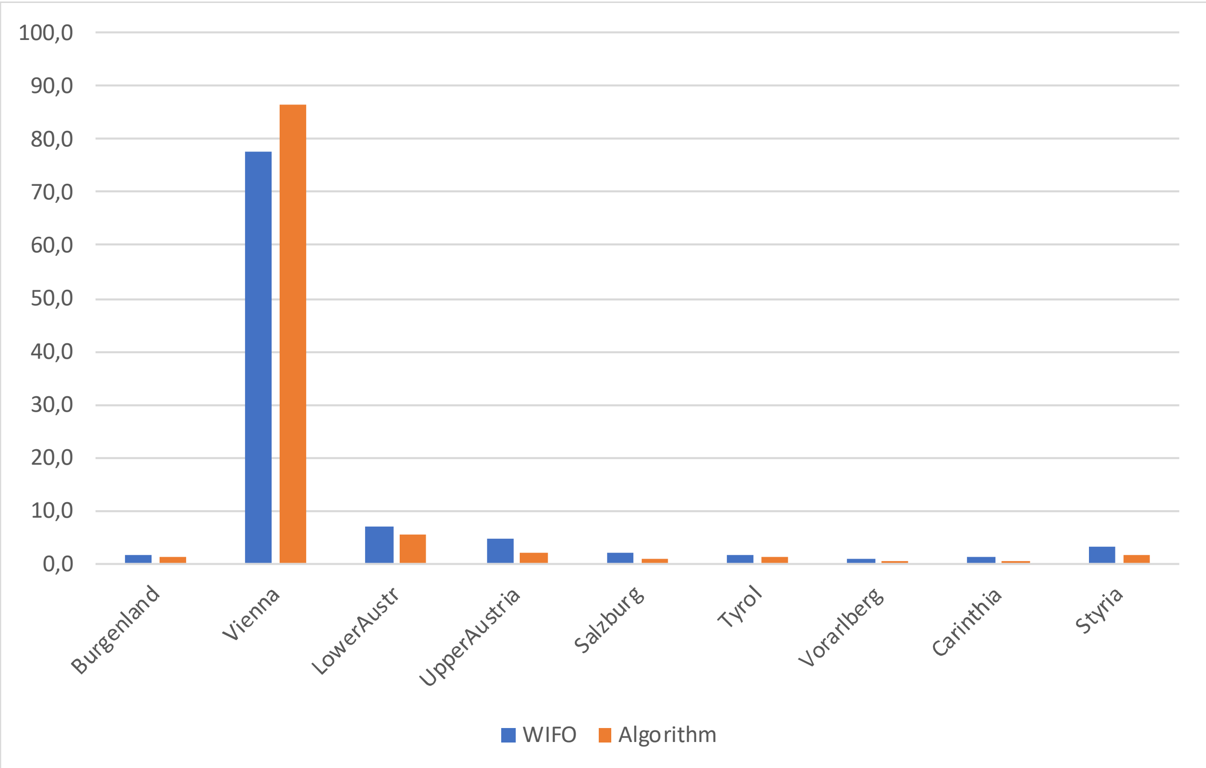
Figure 2. Differences in multiproduction matrices – left panel ready-made model, right panel hybrid approach



Source: authors preparation

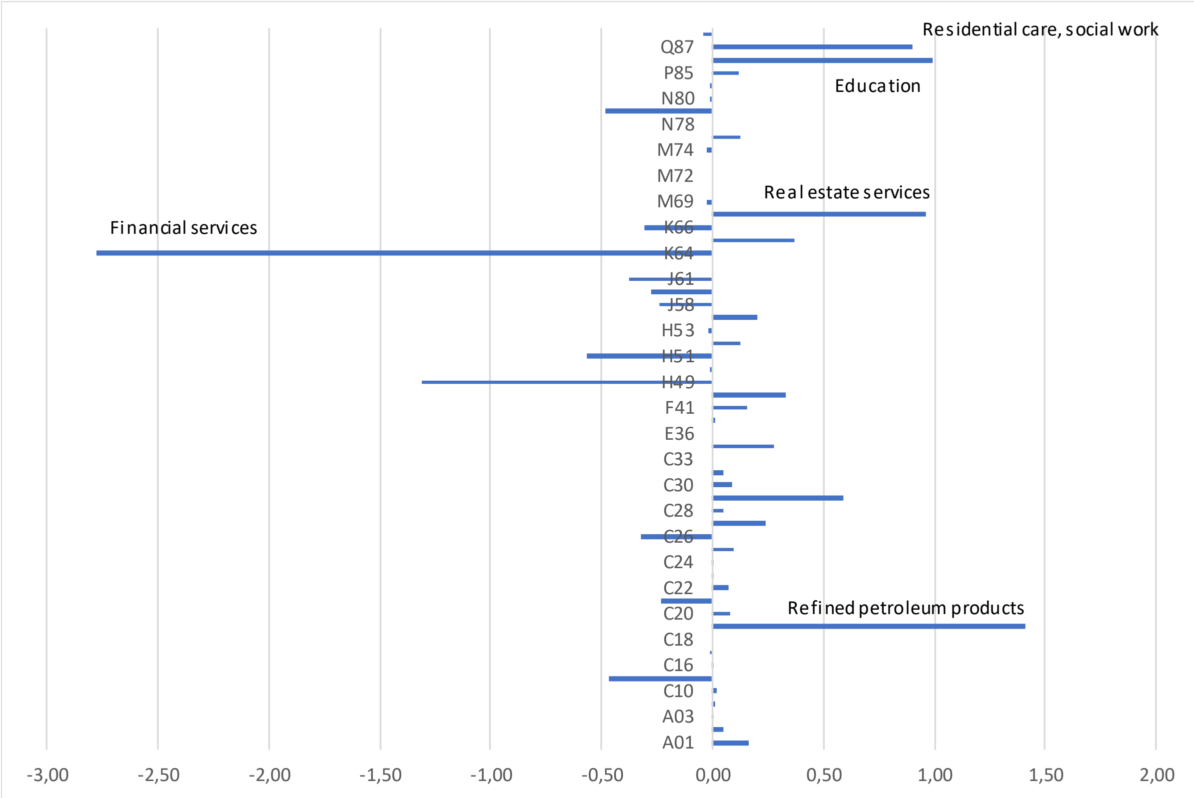
Difference between databases are rather small also once we analyse regional sourcing of domestic commodities or household consumption in Vienna that is the most important Austrian region in terms of regional economy. It can be observed on Figure 3 that the algorithm-based matrix overestimate local sourcing of domestic commodities in Vienna (as compared to the hybrid-based matrix). On the other hand the differences in shares of household consumption by commodities in Vienna lays in most of cases within +/- 0.5% range (see Figure 4). The biggest difference is found in Financial services (almost 3%), followed by Coke and refined petroleum products (almost 1.5%) and Land transport services (almost 1.4%).

Figure 3. Regional sourcing of all domestic commodities used in Vienna.



Source: authors preparation

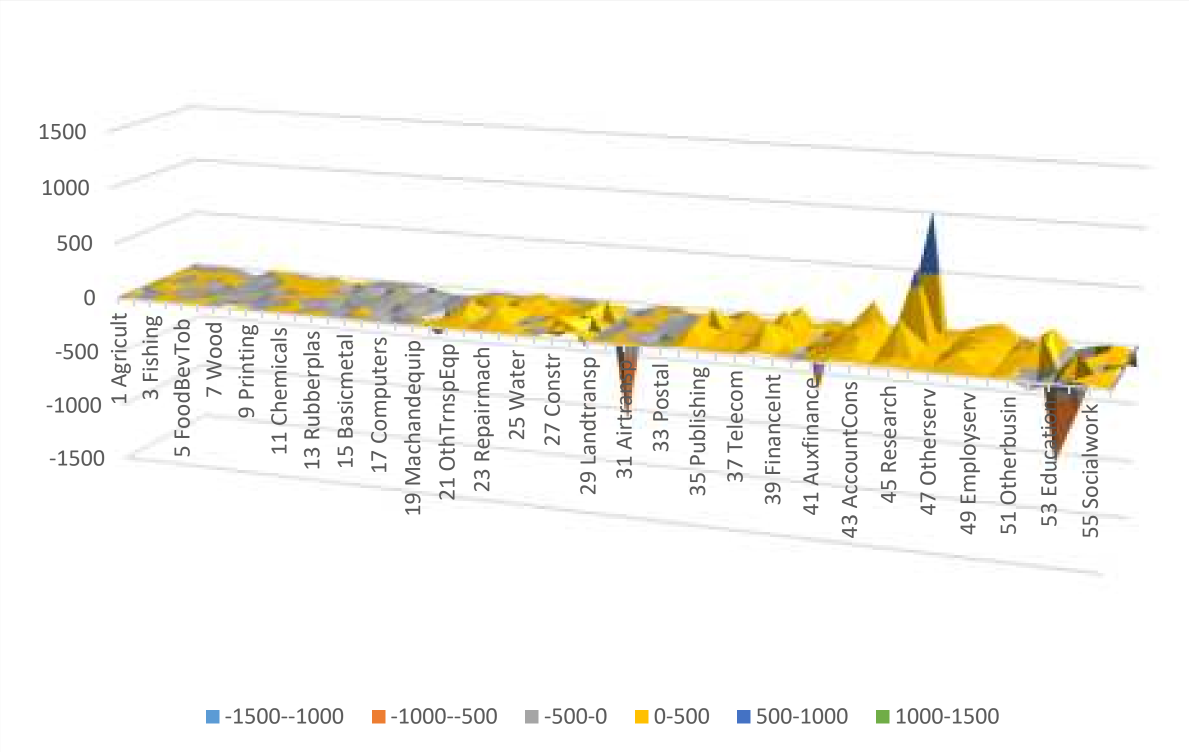
Figure 4. Household consumption in Vienna by commodities (difference in shares)



Source: authors preparation

Significant differences can be found in the case of wage matrix. Certain industries, in particular in the service sector, show dissimilarities valued in hundreds of euro (see Figure 5). The disparities top over 1 billion euro in the case of Trade, Scientific research and development and Public administration – all in Vienna.

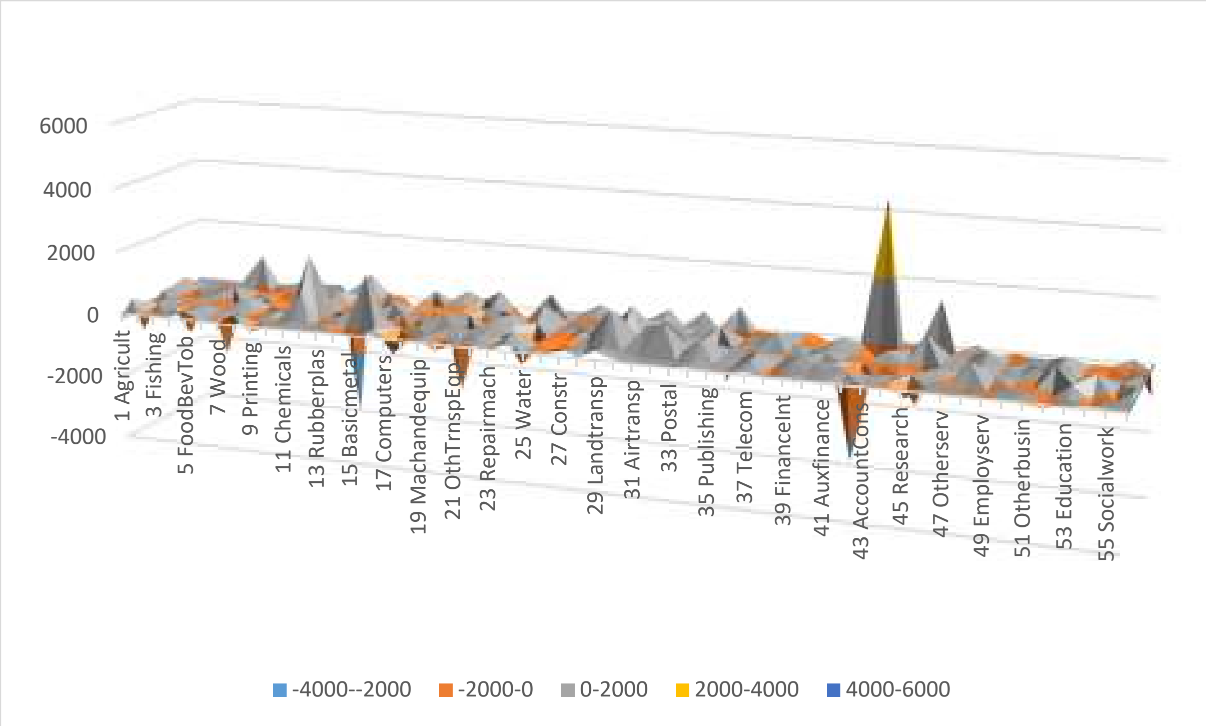
Figure 5. Wage matrices differences by industry and region in EUR million



Source: authors preparation

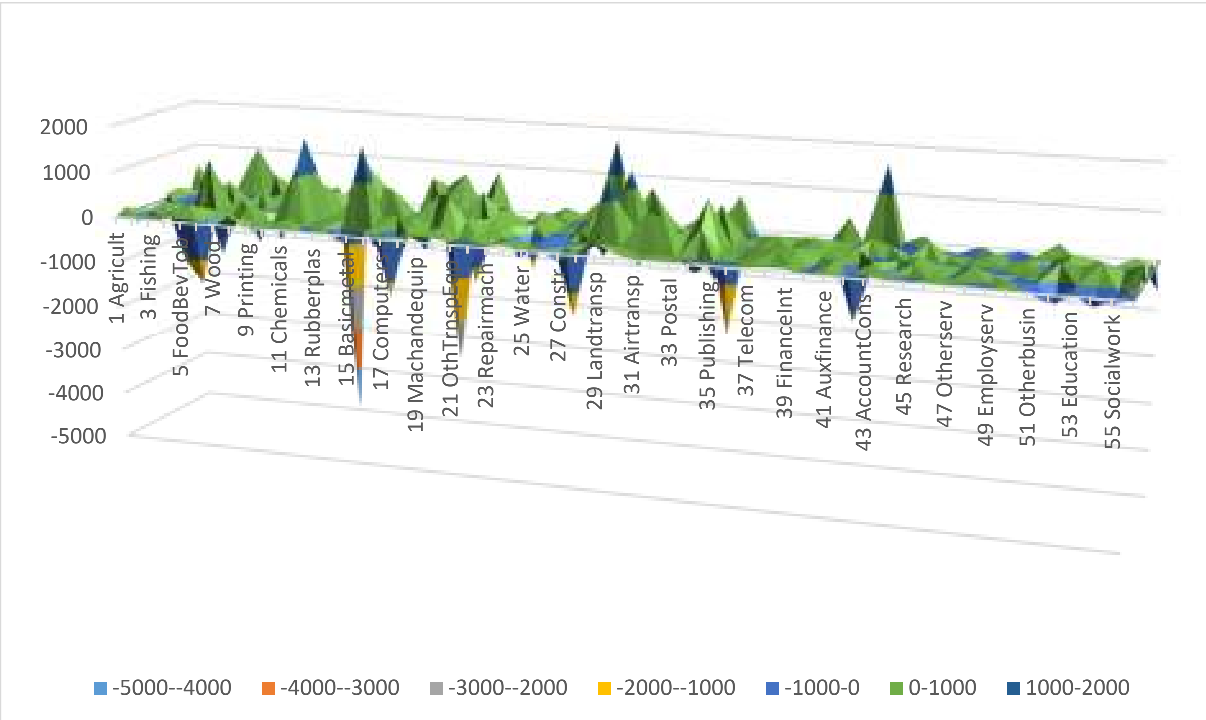
Trade matrices also display significant dissimilarities with Vienna as a region with the highest existing discrepancies. In the case of origins these are Real estate activities, Chemical products and Telecommunications are among the sectors with the highest absolute differences (see Figure 6). In the case of destinations these are Chemical products, Motor vehicles and Real estate activities (see Figure 7).

Figure 6. Trade matrices differences by industry and region in EUR million - origins



Source: authors preparation

Figure 7. Trade matrices differences by industry and region in EUR million - destinations



Source: authors preparation

4. Simulation results

The databases constructed in accordance to procedures described in detail in the previous section were used to perform both short-run and long-run simulations with TERM model. In all of the cases we have applied exactly the same structural parameters (e.g. elasticity of substitution between the regions of margin production), structural framework and estimation methods (Euler). We have conducted many different experiments to verify the extent that the use of a given database significantly influences the results of simulations. However, we present the results of selection of simulations that cover different aspects of macroeconomic policies.. Note, that the we have specified large values of shocks in order to maximize possible differences in simulation results. This particularly affects the simulations concerning change in tax policy and productivity improvements in manufacturing sector. Below we discuss the results of following simulation experiments:

a. Construction of a new hospital in Vienna (144.7% increase in investment in healthcare sector)
b. Increase in labor supply in Vienna (11% positive shock)
c. Drought in Lower Austria (10% negative productivity shock in agriculture)
d. Reduction of production taxes for hotel- and restaurant-services across all regions (10% reduction in tax rate)
e. 10% manufacturing productivity improvements in all regions

In the case of the first and second experiment we focus on Vienna since it is the biggest Austrian region from the economic point of view. First experiment consists in assessing the economic impact related to the construction of new hospital. This involves an increase in investment in the healthcare sector in Vienna by over 140%. Figure 8 shows the differences in simulation results between the ready-made model database and hybrid-approach database. On the left panel we can see that the differences in aggregate real GDP growth do not exceed 0.04 percentage points. They are only slightly larger in the case of aggregate employment growth where they exceed 0.06 p.p. in the long run in the case of Upper Austria and Styria. On average, however, those differences are almost negligible.

Figure 8. Hospital simulation – aggregate results in percentage points

Source: authors preparation

More detailed results for the healthcare sector only are shown on Figure 9. In this case the differences are even smaller – they hardly exceed 0.01 p.p. in the case of GDP growth and do not reach 0.02 p.p. in the case of employment growth. This is not really surprising though, given the fact that the shock affects the non-tradable service sector. Furthermore, the direct effect is limited to a particular region only.

Figure 9. Hospital simulation – sectoral results in percentage points

Source: authors preparation

Second experiment assumes an increase in labor supply in Vienna by 11% due to immigration. The differences in aggregate results are presented on Figure 10. They are similar in magnitude to the hospital simulation. So, they do not exceed 0.015 p.p. in the case of GDP growth and do not reach 0.03 p.p. in the case of employment.

Figure 10. Labor supply simulation – aggregate results in percentage points

Source: authors preparation

Figure 11 shows differences in sectoral results of labor supply simulation. The sector chosen for detailed analysis is trade since it is an industry with the highest number of employed in Vienna. It can be observed that the differences in results at sectoral level are around ten times higher than for aggregate ones. However, they are still rather small. The differences in GDP growth slightly exceed 0.2 p.p. in Vienna and are much lower in the case of remaining regions. The differences in employment growth almost reach 0.4 p.p. in Vienna and do not achieve 0.2 p.p. in any other region.

Figure 11. Labor supply simulation – sectoral results in percentage points

Source: authors preparation

Our third experiment is a simulation of drought in Lower Austria that leads to a 10% productivity decrease in agriculture sector. The differences in aggregate results are again very small (see Figure 12). The highest difference in GDP growth hardly exceeds 0.2 p.p. in Lower Austria while all other regions show differences below 0.05 p.p.. In the case of employment growth the disparities in results are even lower – they exceed 0.06 p.p. only in the case of Vorarlberg in the log-run simulation.

Figure 12. Drought simulation – aggregate results in percentage points

Source: authors preparation

Once we focus on sectoral results for agriculture we find that the differences in simulation results are far more significant. As shown on Figure 13 differences in GDP growth are below 0.2 p.p. in most of the regions. However, in the case of Vienna they exceeds 1 p.p. In both cases simulations indicate an increase in agricultural production in Vienna. Still, expected increase is lower once ready-made model based database is applied (0.17% growth versus 1.25% growth). The differences are even more visible once employment growth is analyzed. They reach over 2 p.p. in Lower Austria and over 1 p.p. in Salzburg, Tyrol, Vorarlberg and Vienna. In general, algorithm-based simulations show lower impact of productivity shock on employment. Both in the case of employment fall (Lower Austria) and employment increase (the rest of regions).

Figure 13. Drought simulation – sectoral results in percentage points

Source: authors preparation

Our fourth simulation concerns 10% reduction in tax rate for tourism (hotels and restaurants) in all regions. Differences in aggregate simulation results are shown on Figure 14. These differences are very small in the case of GDP growth – in most of the cases they do not reach 0.1 p.p. Bigger, although still small differences are observed in the case of employment growth.

In Tyrol they exceed 0.3 p.p. with algorithm-based database simulations showing smaller impact of tax reduction on employment than hybrid-based database simulations.

Figure 14. Tax reduction in tourism sector simulation – aggregate results in percentage points

Source: authors preparation

Still, far more significant differences can be found in the case of sectoral results that focus on hotels and restaurants (see Figure 15). Here, differences in results related to GDP growth are particularly noticeable in Lower Austria (almost 1.5 p.p.) and Vienna (over 2 p.p.). In the rest of the regions they do not exceed 1 p.p. Similar situation is observed in the case of employment growth with Lower Austria and Vienna being the regions with the highest discrepancies in results. Once more, simulations applying algorithm-based database show smaller impact of tax cut on both GDP and employment growth.

Figure 15. Tax reduction in tourism sector simulation – sectoral results in percentage points

Source: authors preparation

Our final experiment consists in improving productivity in the entire manufacturing sector by 10%. Here, for the first time aggregate results show significant discrepancies (see Figure 16).

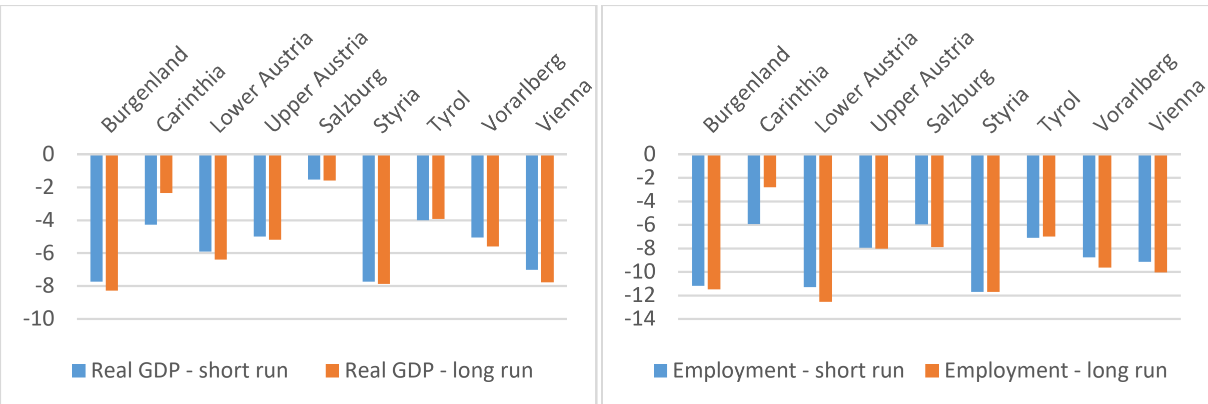
For instance, the differences in GDP growth exceed 3 p.p. in Vorarlberg and 2 p.p. in Carinthia, Upper Austria and Styria. In the case of employment growth they almost achieve 2 p.p. in Vorarlberg (long-run simulations) and exceeds 1 p.p. in Burgenland, Carinthia, Upper Austria and Styria. For almost all of the regions simulations using algorithm-based database seem to underestimate the effects of productivity shock. The exceptions are Lower Austria (GDP growth) and Vienna (employment growth).

Figure 16. Manufacturing productivity simulation – aggregate results in percentage points

Source: authors preparation

Figure 17 shows the differences in sectoral results where the analysed industry is the Food, beverage and tobacco manufacturing. In the case of GDP growth the differences reach almost 8 p.p. in Burgenland (or even exceed 8 p.p. in long-run simulations), Styria and Vienna. They are even greater in the case of employment growth where they exceed 10 p.p. in Burgenland, Lower Austria and Styria. In all of the regions simulations using algorithm-based database strongly underestimate the impact of productivity improvement. We have also analysed the results for other manufacturing industries and the differences are also more than significant.

Figure 17. Manufacturing productivity simulation – sectoral results in percentage points



Source: authors preparation

The experiments discussed above show that, on average, the differences in results of simulations using ready-made model database and hybrid-model database are negligible at the aggregate level. The only exception is the experiment concerning productivity increase in manufacturing. Differences in sectoral results seem to be more significant, although usually it refers to particular regions (with the exception of productivity increase experiment). We have tried to verify which coefficients (or the difference in coefficients between databases) are to the greatest extent responsible for observed differentials in simulation results. What we find is that the main reason behind those differentials is not necessarily the disaggregation approach itself but rather the fact that the ready-made algorithm prepares a database with diagonal multiproduction matrix. This is particularly important for manufacturing where inter-industry linkages are very strong. In the case of non-diagonal MAKE any shock that influences a given manufacturing industry will also have an impact on other industries through production linkages. This relationship does not exist once diagonal MAKE is used. It may not significantly influence simulation results if the linkages are weak (e.g. most service industries). However, if the production linkages are strong the results can differ substantially not only at the sectoral level but also at the aggregate one.

In order to formally verify our assumptions we prepared hybrid-based database with diagonal MAKE matrix and repeated simulations focusing on productivity shock in manufacturing sector. We find that in the case of sectoral results for Food, beverage and tobacco manufacturing the differences has decreased both in the case of GDP growth and employment growth (see Figure 18). In the case of the former they are halved in Burgenland or Styria and almost disappear in Upper Austria. In the case of the latter they are below 5 p.p. with the exception of Lower Austria and Vienna. In all regions the differences in simulation results are much lower than once non-diagonal MAKE is applied although still remain significant. We are unable to find any other particular coefficient that would explain them.

Figure 18. Manufacturing productivity simulation with diagonal MAKE – sectoral results in percentage points

Source: authors preparation

5. Conclusion

References:

Bonfiglio, A. (2009). On the Parameterization of Techniques for Representing Regional Economic Structures. *Economic Systems Research* 21, 115–127.

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(3), 243–258.

Boomsma, P., Oosterhaven, J. (1992). A Double Entry Method for the Construction of Biregional Input-Output Tables. *Journal of Regional Science* 32, 269-284.

Brucker, S., Hastings, S., Latham III, W. (1990) The variation of estimated impacts from five regional input-output models. *International Regional Science Review* 13, 119-139.

Canning, P., Wang, Z. (2005). A Flexible Mathematical Programming Model to Estimate Interregional Input-Output Accounts. *Journal of Regional Science* 45(3), 539–563.

Czamanski, S., Malizia, E. (1969). Applicability and Limitations in the Use of National Input-Output Tables for Regional Studies. *Papers, Regional Science Association* 23, 65–77.

Fernandez Vazquez, E. (2015). Empirical Estimation of Non-Linear Input–Output Models: an Entropy Econometrics Approach. *Economic Systems Research* 27(4), 508–524.

Fernandez Vazquez, E., Hewings, G., Ramos Carvajal, C. (2015). Adjustment of Input–Output Tables from Two Initial Matrices. *Economic Systems Research*, 27(3), 345–361.

Flegg, A., Mastronardi, L., Romero, C. (2016). Evaluating the FLQ and AFLQ Formulae for Estimating Regional Input Coefficients: Empirical Evidence for the Province of Córdoba, Argentina. *Economic Systems Research* 28(1), 21-37.

Flegg, A., Tohmo, T. (2013a). Regional Input–Output Tables and the FLQ Formula: A Case Study of Finland. *Regional Studies* 47(5), 703–721.

Flegg, A., Tohmo, T. (2013b). A comment on Tobias Kronenberg's 'Construction of regional input–output tables using nonsurvey methods: The role of cross-hauling'. *International Regional Science Review* 36, 235–257.

Flegg, A., Tohmo, T. (2016). Estimating Regional Input Coefficients and Multipliers: The Use of FLQ is Not a Gamble. *Regional Studies* 50(2), 310–325.

Flegg, A., Webber, C. (2000). Regional Size, Regional Specialization and the FLQ Formula. *Regional Studies* 34(6), 563–569.

Flegg, A., Webber, C., Elliott, M. (1995). On the Appropriate Use of Location Quotients in Generating Regional Input–Output Tables. *Regional Studies* 29(6), 547–561.

Fujimoto, T. (2018). Appropriate Assumption on Crosshauling National Input–Output Table Regionalization. *Spatial Economic Analysis*, DOI:10.1080/17421772.2018.1506151.

Greenstreet, D. (1989). A Conceptual Framework for Construction of Hybrid Regional Input-Output Models. *Socio-Economic Planning Sciences* 23(5). 283–289.

Harris, R., Liu, A. (1997). Input-Output Modelling of the Urban and Regional Economy the Importance of External Trade. *Regional Studies* 32(9), 851–862.

Hewings, G. (1969). Regional Input-Output Models Using National Data: the Structure of the West Midlands Economy. *Annals of Regional Science* 3, 179–190.

Hewings, G. (1971). Regional Input-Output Models in the UK: Some Problems and Prospects for the Use of Nonsurvey Techniques. *Regional Studies* 5, 11–22.

Horridge, M. (2011). The TERM model and its database. Centre of Policy Studies, General Paper No. G-219.

Isard, W. (1953). Regional Commodity Balances and Interregional Commodity Flows. *American Economic Review* 43, 167-80.

Jiang, X., Dietzenbacher, E., Los, B. (2012). Improved Estimation of Regional Input–Output Tables Using Crossregional Methods. *Regional Studies* 46(5), 621–637.

Kowalewski, J. (2015). Regionalization of National Input-Output Tables: Empirical Evidence on the Use of the FLQ Formula. *Regional Studies* 49, 240-250.

Kronenberg, T. (2009). Construction of Regional Input–Output Tables Using Nonsurvey Methods: The Role of Cross-Hauling. *International Regional Science Review* 32, 40–64.

Lahr, M. (1993). A Review of the Literature Supporting the Hybrid Approach to Constructing Regional Input-Output Models. *Economic Systems Research* 5(3), 277–293.

Lahr, M., de Mesnard, L. (2004). Biproportional Techniques in Input–Output Analysis: Table Updating and Structural Analysis. *Economic System Research* 16(2), 115–134.

Lazarus, W., Platas, D., Morse, G. (2002). IMPLAN's Weakest Link: Production Functions or Regional Purchase Coefficients?. *Journal of Regional Analysis and Policy* 32(1), 1-17.

McCann, P., Dewhurst, J. (1998). Regional Size, Industrial Location and Input-Output Expenditure Coefficients. *Regional Studies* 32(5), 435–444.

Miller, R., Blair, P. (2009). *Input-Output Analysis: Foundations and Extensions.* Cambridge University Press, New York.

Morrison, W., Smith, P. (1974). Nonsurvey Input-Output Techniques at the Small Area Level: An Evaluation. *Journal of Regional Science* 14(1), 1–14.

Oosterhaven, J., Escobedo, F. (2011). A New Method to Estimate Input-Output Tables by Means of Structural Lags, Tested on Spanish Regions. *Papers in Regional Science* 90, 829–844.

Oosterhaven, J., van der Knijff, E., Eding, G. (2003). Estimating Interregional Economic Impacts: An Evaluation of Nonsurvey, Semisurvey, and Full-Survey Methods. *Environment and Planning A* 35(1), 5–18.

Riddington, G., Gibson, H., Anderson, J. (2006). Comparison of Gravity Model, Survey and Location Quotient based Local Area Tables and Multipliers *Regional Studies* 40(9), 1069–1081.

Stevens, B, Treyz, G., Ehrlich, D., Bower, J. (1983). A New Technique for the Construction of Non-Survey Regional Input-Output Models. *International Regional Science Review* 8(3), 271–286.

Szabó, N. (2015). Methods for Regionalizing Input-Output Tables. *Regional Statistics* 5(1), 44–65.

Schaffer, W., Chu, K. (1969). Nonsurvey Techniques for Constructing Regional Interindustry

Models. *Papers, Regional Science Association* 23, 83–101.

Többen, J., Kronenberg, T. (2015). Construction of Multi-Regional Input–Output Tables Using the Charm Method. *Economic Systems Research* 27, 487–507.

Tohmo, T. (2004). New Developments in the Use of Location Quotients to Estimate Regional Input-Output Coefficients and Multipliers. *Regional Studies* 38(1), 43–54.

West, G. (1990). Regional Trade Estimation: A Hybrid Approach. *International Regional Science Review* 13, 103-118.

West, G. (1995). Comparison of Input-Output, Input-Output + Econometric and Computable General Equilibrium Impact Models at the Regional Level. *Economic Systems Research* 7(2), 209-227.

Wiebe, K., Lenzen, M. (2016). To RAS or Not to RAS? What Is the Difference in Outcomes in Multi-Regional Input–Output Models? *Economic Systems Research* 28(3), 383–402.

Zhao, X., Choi, S-G. (2015). On the Regionalization of Input-Output Tables with an Industry-Specific Location Quotient. *Annals of Regional Science* 54, 901-926.