EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

University of Tübingen
Working Papers in
Business and Economics

No. 132

RIOTs in Germany - Constructing an interregional input-output table for Germany

by

Oliver Krebs

Faculty of Economics and Social Sciences
www.wiwi.uni-tuebingen.de

https://publikationen.uni-tuebingen.de/xmlui/handle/10900/95156

RIOTs in Germany - Constructing an interregional input-output table for Germany*

Oliver Krebs†
University of Tübingen

March 11 , 2020

Abstract

This paper shows how to adapt recent methodological advances to derive a shipment based interregional input output table for 402 German counties and 26 foreign partners for 17 sectors that is, for national aggregates, cell-by-cell compatible with the WIOD tables. It far outperforms the standard approach of applying unit values to interregional shipments in replicating observed regional statistics and can be used for improved impact analysis and CGE model calibration. It thereby mitigates the surprising but problematic lack of regional German trade data in the analysis of both, regional effects of aggregate shocks such as trade agreements as well as network effects of regional policies. Moreover, the paper takes an in-depth look at the derived German production structure and trade network at the county level finding a surprisingly vast heterogeneity with respect to specialization, agglomeration and trade partners.

JEL-Classification: R15, R12, F17

Keywords: Germany, regional trade, input-output tables, unit values, proportionality

*I thank Michael Pflüger, Wilhelm Kohler, Wolfgang Britz, Thomas Heckelei, Erwin Winkler, Johannes Többen, Yaghoob Jafari, Marcel Henkel, Markus Wiest and Seminar and Conference participants in Tübingen, Bonn and Göttingen for helpful comments and suggestions.

†Oliver Krebs, Faculty of Economics, University of Tübingen, Nauklerstraße 50, 72074 Tübingen, Germany, Tel + 49(0)7071 - 29 76017, e-mail: oliver.krebs@uni-tuebingen.de.

1 Introduction

Regions matter! On the one hand macroeconomic shocks have vastly different effects across regions: Brexit, TTIP or US tariffs, robotization and artificial intelligence all will affect Berlin differently than Munich, depending not only on each city's local conditions but also on its linkages with other locations. On the other hand, shocks in individual regions, such as inventions, bankruptcies or the attraction of a major production plant can, through trade and input-output linkages, magnify to aggregate effects of macroeconomic relevance. Despite their importance, surprisingly little is known about the trade and production networks within Germany and their connection to the international markets.

Baden Wuerttemberg is the only state (of 16) in Germany that has consistently published a state level input-output table for several years but has stopped data collection in 1993 due to financial limitations (cf. Kowalewski (2015)). Only a few authors have constructed other regional input-output tables (RIOTs) usually relying on so-called "non-survey" methods that break down national input-output tables based on some locally available measure such as sectoral GDP or employment.[1] For example, Kronenberg (2009) derives such a table for the state of North Rhine–Westphalia, Koschel et al. (2006) for the state of Hessen and Schröder and Zimmermann (2014) for the German coastal region of the Baltic sea. In even fewer cases authors use a "survey" or "hybrid" approach relying on detailed regional data to construct a RIOT. Kronenberg (2010) who constructs such a table for the state of Mecklenburg West Pomerania is a case in point, as is, for example, Stäglin (2001) who derive a RIOT for the city of Hamburg. In all of these cases, however, the authors construct regional instead of interregional input-output tables (IRIOT). In the former "exports" are just a further category of final demand without specifying the target location and, similarly, "imports" are specified as a supply without a source location.

This paper, in contrast, analyses the trade linkages between German counties making use of a unique data set constructed by Schubert et al. (2014) as part of the official "Forecast of nationwide transport relations in Germany 2030" on behalf of the German ministry of transport and digital infrastructure ("Bundesministerium für Verkehr und digitale Infrastruktur"). The data provides total shipments in tons by water, train or truck for the year 2010 between 402 German counties and their trade partners, disaggregated along 25 product categories.

I use this data together with further information from the German regional statistical offices and the world input-output database (WIOD) to construct an *interregional* input-output table for 17 sectors across 402 German counties and 26 international trading partners.[2] To

[1] Section 2 describes different approaches to constructing regional input-output tables in more detail.
[2] See Timmer et al. (2015) for details on the world input-output database.

the best of my knowledge I am the first to construct such a data set at the deep sectoral and spatial disaggregation level of German counties.[3] The construction method and strength of the underlying data sets allow the IRIOT to remain strongly anchored in observable data. In particular, it replicates officially reported local revenues, value added and consequently intermediate demand levels of county sectors.[4] Further, interregional trade networks are based on the data on interregional shipments and mirror observed international trade flows. For service sectors, where such shipments are naturally unavailable, I show that, with observed revenues and derived local demand values, a gravity system can be inverted to solve for location-sector specific importer and exporter fixed effects. These in turn can be used to predict interregional service trade. Finally, the national aggregates of the IRIOT are, cell by cell, perfectly consistent with the international tables from the WIOD, allowing for an integrated analysis with a 'closed' world wide input-output table.

The remainder of this paper is structured as follows. Section 2 provides background on the construction of regional input-output tables. Section 3 gives details on the used data sets and initial data preparation and presents a descriptive analysis of the German production structure and trade linkages. Section 4 explains the construction of the IRIOT and discusses the resulting table. The final section sums up the results.

2 Background

Data sources and previous literature detailing trade flows within Germany are scarce and only a few regional input-output tables have been produced by select authors for individual states or cities. Using survey based methods to directly derive input-output tables from collected data is usually too costly and time intensive for individual researchers to accomplish, but some approaches combine non-survey methods with detailed regional data and are therefore considered "hybrid" approaches. For example, Kronenberg (2010), in deriving the regional input-output table of the state of Mecklenburg West Pomerania uses data from the German consumer expenditure survey ("Einkommens- und Verbrauchsstichprobe") to establish unique regional consumption levels across industries. The majority of RIOTs in Germany are, however, constructed using non-survey methods, which can be broadly classified into location quotients approaches and commodity balance approaches.

The simplest form of the location quotient approach going back to Schaffer and Chu (1969)

[3] In his PhD-thesis Többen (2017) derives a hybrid type interregional input-output table for the year 2016 at the level of the 16 German federal states relying on a CHARM approach and using a partial table of bilateral transport relations.

[4] As described in detail in section 3 I scale data from all sources such that national aggregates match the values reported by the WIOD, hence regional data from other sources is only matched up to scale.

relies on a measure such as the number of workers or GDP that is regionally available at the sector level to approximate the relative size of each sector in a region. If the relative size of a sector in the region is equal to or larger than the national relative size it is assumed that the sector can meet the local demand and regional input coefficients remain the same as in the national tables. If the relative size is smaller than in the aggregate data however, imports from other regions become necessary to satisfy the regional demand for the sector and domestic input coefficients in the regional input-output table are adjusted downwards from the national values for the respective sector. Several variants of location quotients have been developed in the literature to account for further aspects when determining the adjustment factors for input coefficients. The most prominent examples consider relative industry sizes within a region (cross-industry location quotient), the overall size of a region (Flegg et al. 1995; Flegg and Webber 1997) and regional specialization (Flegg et al. 2000). An overview of different location quotients and their construction can be found in Flegg and Tohmo (2013). These approaches, however, treat imports and exports as residual values and can thus not capture cross-hauling, that is the simultaneous import and export of goods from the same industry. Moreover, input-output tables constructed this way are only constructed for a single region and do not capture where imports come from nor where exports end up. This is of central importance if one wants to use input-output tables to calibrate general equilibrium models that capture regional trade networks and a 'closed system' or world economy.

The commodity balance or supply-demand-pooling approach also attributes a share of the national sectoral revenue to a region based on a regionally available measure such as employment levels. Subsequently intermediate demand is derived by applying the national input coefficients to the regional production and final demand by scaling national final demand, for example by the regions share in total population or GDP. Having determined both total domestic supply and total demand the difference between the two, that is the net imports or exports, is interpreted as a regions total imports or exports. The basic commodity balance approach therefore also does not allow for cross-hauling of products from the same industry, which is in strong contrast to international trade flows and also to the data used in this paper. While Kronenberg (2009) introduces a method that imposes a certain amount of cross-hauling based on measures of product heterogeneity within an industry the trade structure remains completely non-survey based. Moreover, it also applies that this method can not capture the source of imports and destination of exports that are important to understand linkages with other regions and countries.

The accuracy of such "mechanical" approaches to deriving regional from national input-output tables has been discussed intensively in the literature, often by comparing them to survey based results (recent examples include Flegg and Tohmo 2019; Kowalewski 2015;

Flegg and Tohmo 2013).[5] However, comparison with survey based methods might be misleading as their construction also involves a substantial amount of uncertainty and decision making, implying that they are not error free. Overall, the earlier conclusion by Hewings and Jensen (1986) in the Handbook of Regional and Urban Economics that survey methods "remain generally regarded as 'preferred' tables in terms of accuracy, more so by analysts inexperienced in their construction" can still be considered valid.

Independent of the chosen method there are several different types of regional input-output tables that one can construct. For European Union members, including Germany, national tables follow the recommendations of the European System of Accounts (ESA). Regional tables in Germany, being derived from the national tables, therefore usually also follow this structure. As shown in figure 1 the sum of each row of these tables give the total (regional) use and the sum of each column the total (regional) supply of goods from a specific industry.[6] This means that no difference is made between domestic and imported goods in the rows of the table, with cells showing the aggregate use of domestic and imported goods as intermediate or in final demand. Similarly, as columns explain the total supply of goods from a specific industry, they only include the contribution of aggregate imports of goods from each industry.[7] Importantly, total supply in these tables is the sum of domestically produced and imported goods and cells in the first two rows show where this aggregate supply is used. This means, that imports used as intermediates are counted twice. Once in the top left quadrant contributing to domestic production and once directly as "imported supply". To see this, consider an economy that, without further factors, uses 1 Dollar of intermediates to produce 1 Dollar of output that is then consumed. Total (domestic) output and total consumption are equal to 1 Dollar, but total supply and total use are equal to 2 Dollars: 1 Dollar domestic supply and 1 Dollar of imports, as well as 1 dollar of intermediate use and 1 Dollar of final use.

In contrast to this aggregate view an input-output table can also be constructed showing the use structure of domestic and imported goods separately. In this case, as depicted in figure 2, the aggregate of a row gives either the total use of domestic production or of imports from a specific sector whereas columns sum to the domestic production of each sector, to

[5] Bonfiglio and Chelli (2008) provide an alternative approach relying on Monte-Carlo simulations.

[6] Supply tables show which products are supplied by which industries. Use tables show how much of each product is consumed and how much ends up as intermediates in each industry. Constructing input-output tables from these two tables one has to decided between a product-by-product or an industry-by-industry table. To derive the former one has to assume that each product is always produced in the same way, irrespective of the industry where it is produced. For the latter one assumes that each product serves intermediate and final demand with fixed shares, irrespective of which industry produces it. There is no clear advantage between the two approaches. Here the focus is on industry-by-industry tables as this is also the type derived in this paper.

[7] Following the ESA national input-output tables should be accompanied by two separate tables, an input-output table of domestic production and an input-output table of imports. This additional information is, however, usually not produced by the previously mentioned works on regional tables in Germany.

		Use				
		Intermediate use		Final consumption	Exports	Total use
		Sector 1	Sector 2			
Supply	Sector 1	Input coefficients		Consumption of domestically produced and imported goods	Including reexports	Total demand of domestically produced and imported goods
	Sector 2					
	Imports	Imported supply				
	Value added					
	Total supply	Total supply				

Figure 1: Input-Output table ESA standard

aggregate final demand and to total exports.[8]

		Use				
		Intermediate use		Final consumption	Exports	Total demand
		Sector 1	Sector 2			
Domestic Supply	Sector 1	Domestic input coefficients		Consumption of domestically produced goods	Exports of domestically produced goods	Total demand of domestically produced goods
	Sector 2					
	Imports	Import coefficents		Final use of imports	Reexport	Total demand of Imports
	Value added					
	Total Output	Domestic output				

Figure 2: Input-Output table with import structure

While the second type of input-output table contains additional information on the underlying production structure both types are regional, that is imports and exports only appear aggregated across all sources or destinations, respectively. In contrast, an interregional input-output table captures the full interregional trade networks as demonstrated by the simplified two country table in figure 3.[9]

The great advantage of an IRIOT over a RIOT is that it distinguishes both imports and exports geographically. Since columns contain all possible trade partners including the country itself, the sum of each row equals the total use of goods produced in one sector in one location. This value must be equal to the respective column sum which includes all intermediates, domestic and imported, as well as value added in one sector in one region and

[8] It is also important to note, that the interpretation of input coefficients that can be derived in the upper-left quadrant of both types of input-output tables differ. In their seminal handbook article Hewings and Jensen (1986) refer to the former as "technical" and to the latter as "trade" coefficients but criticize that the literature on regional input-output tables does not use consistent terms to distinguish these coefficients and indeed often erroneously confounds the two when applying non-survey methods.

[9] Previous literature is not consistent in its use of the term "interregional input-output table" and some authors further differentiate between "interregional" and "multiregional" input-output tables (see, for example, Hewings and Jensen 1986). Figure 3 exemplifies the meaning of the term in this paper. The input-output table provided by the WIOD is a further example of this case, albeit being international and not interregional.

			Use				
			Intermediate use		Final consumption		
			Location 1	Location 2	Location 1	Location 2	Total use
			Sector 1 Sector 2	Sector 1 Sector 2			
Supply	Location 1	Sector 1 Sector 2	Domestic input coefficients	Foreign input coefficients	Consumption of location 1 goods in location 1	Consumption of location 1 goods in location 2	Total use of location 1 goods
	Location 2	Sector 1 Sector 2	Foreign input coefficients	Domestic input coefficients	Consumption of location 2 goods in location 1	Consumption of location 2 goods in location 2	Total use of location 2 goods
		Value added					
		Total Output	Output Location 1	Output Location 2			

Figure 3: interregional Input-Output table

hence represents the region's sectoral output. The IRIOT captures not only the sectoral but also the geographic component of a production network and can consequently also be used to study how economic shocks effect non-treated locations through spatial linkages. In contrast to all previous input-output tables for German regions this paper constructs an IRIOT which, being cell-by-cell consistent with the WIOD in terms of the national aggregate, even includes world-wide input-output data.

As interregional trade data is usually unavailable there is also few literature that discusses construction methods of RIOTs or even IRIOTs that rely on this type of data. An important exception is Wang and Canning (2005) who suggest a mathematical programming method that is similar to the multidimensional RAS method applied in this paper and that allows to derive IRIOTs based on initial estimates of trade flows and technical coefficients combined with further statistics at the region sector level, such as sectoral output and demand.[10] In contrast to their approach however, I do not observe the final demand structure, or any data about trade in service sectors and must approximate these values. As explained in the next two sections I instead rely on a two step process, treating sectors with known and those with unknown trade flows separately. Moreover, Wang and Canning (2005) apply their estimation method to an artificially created aggregate region consisting of several countries to test the validity of their approach, whereas this paper aims to calibrate an actual county level IRIOT for further applications.

In terms of the underlying data set Nitsch and Wolf (2013) rely on similar shipment data as I do - albeit at a much higher level of aggregation - to study the persistence of a border

[10] The RAS algorithm (Stone and Brown 1962; Bacharach 1965) is widely used in input-output analysis and, under different names such as proportionate fitting or matrix scaling, in a range of different fields. I discuss the algorithm in detail in section 2. Its name is not an abbreviation but originates in variable names (r, A, s) used by Stone and Brown (1962).

effect from German separation over time.[11] Lameli et al. (2015) use the same data as Nitsch and Wolf (2013) to derive the effect of dialects on intra-national trade. In both cases the authors rely on an empirical gravity approach, that is they estimate the effects of specific variables on aggregate trade flows. However, they use simple unit values from the national German export statistics to aggregate trade flows over all product categories in all regions and do not derive input-output linkages as in this paper.

In contrast the data treatment in this paper goes much further developing a full IRIOT strongly rooted in local and bilateral data and allowing to answer questions that require a deeper understanding of spatial and sectoral production networks. Krebs (2018), for example, relies on an earlier version of the data set produced in this paper, including full input-output linkages, to study regional and sectoral spill-over effects of productivity shocks within Germany, including unemployment effects. Becker and Henkel (2020) have recently also derived an input-output table based on the same underlying transport data but relying on simple unit values and aggregating the relevant data to the three general sectors agriculture, mining and manufacturing. They, similarly, apply it to studying German regional and sectoral productivity shocks but focus on identifying key regions for German aggregate output effects. Krebs and Pflüger (2019), as a further example, use an earlier version of the data set derived in this paper together with commuting data to analyze the linkages of local labor markets in Germany through trade and commuting.[12]

3 Data and descriptive analysis

This section discusses the different data sources used to inform the final interregional input-output table as well as initial data processing steps. Section 3.1 describes the source of international trade data and international input-output tables. Section 3.2 explains the derivation of regional and sectoral output values and section 3.3 shows how interregional trade flows are determined. Finally, section 3.4 provides a descriptive analysis of the obtained county level production structure and trade network.

3.1 WIOD

I use the World Input Output database (WIOD) as my main data source for the national production structure and international trade flows. This data set provides a time-series of

[11]They consider two data sets, one with 10 product categories and 101 regional entities ("Verkehrsbezirke") and another with 24 product categories but only 27 regions ("Verkehrsregionen").

[12]Seidel and Wickerath (2019) study rush hour effects in Germany in a similar model, relying on the same transport data but again applying simple unit values to arrive at interregional trade values.

world input-output tables compiled on the basis of officially published input-output tables in combination with national accounts and international trade statistics. The world input-output table for the year 2010 for which my subnational shipment data is available covers data from 56 industries in 44 countries, including one artificial “rest of the world” (ROW) country. To match this data to the sectors and countries for which shipment data is available I aggregate it to the 17 industries and 27 countries listed in tables 1 and 2.[13] In order to be able to use the final IRIOT produced in this paper to calibrate static economic trade models it is usually necessary to abstract from dynamic features such as inventories. Therefore, positive inventory changes in the WIOD will be included in final demand and negative inventory changes treated as if they had been produced in 2010 as well. The details of this process are laid out in Krebs and Pflüger (2018) and summarized in appendix B. Throughout the paper any values derived from the WIOD are taken from this inventory-adjusted WIOD.

Table 1: List of sectors

#	Description
1	Agriculture
2	Mininig
3	Food, Beverages, Tobacco
4	Textiles, Leather
5	Wood, Paper, Printing
6	Petroleum, Coke
7	Chemicals, Pharmaceuticals
8	Non-Metallic Minerals
9	Metal
10	Machinery, Electrical Equipment
11	Transport Equipment
12	Other Manufacturing
13	Utilities
14	Construction
15	Trade, Communication, IT
16	Financial, Insurance, Business
17	Government, Education, Health

Table 2: List of countries

ISO3	Name	ISO3	Name
AUT	Austria	NLD	Netherlands
BEL	Belgium	POL	Poland
BGR	Bulgaria	PRT	Portugal
CHE	Switzerland	ROU	Romania
CZE	Czech Republic	RUS	Russia
DEU	Germany	SVK	Slovakia
DNK	Denmark	SVN	Slovenia
ESP	Spain	SWE	Sweden
EST	Estonia	TUR	Turkey
FRA	France		
GBR	United Kingdom		
HRV	Croatia		
HUN	Hungary		
ITA	Italy		
LTU	Lithuania		
LUX	Luxembourg		
LVA	Latvia		

3.2 Production data

Unfortunately, revenue data for the 402 counties in Germany is not published at the level of sectoral disaggregation employed in this paper (see table 1) and therefore has to be derived from several sources. As different data is available for mining and manufacturing sectors compared to agricultural, construction and service sectors the process is reported separately for the two groups.

[13]The matching of the 56 sectors in the WIOD to these 17 industries is shown in table C.1 in the appendix.

Firstly, for the mining and manufacturing sectors (2-12) revenue and value added data in each county, $i \in \{1, ..., 402\}$, is only available as a sectoral aggregate ($R_{i,manufac}$) from the German regional statistical offices.[14] To derive specific county sector revenues I construct a matrix as depicted in figure 4 with one row for each of the 402 German counties, one column for each of the 11 sectors in question and with individual entries $\tilde{R}_{ij}$ denoting initial estimates of the revenue generated in a mining or manufacturing sector $j \in \{2, ...12\}$ in county $i \in \{1, ..., 402\}$. These estimates are obtained by distributing the German sectoral revenue taken from the WIOD across counties based on county employment shares in the particular industry, that is I set $\tilde{R}_{ij} = R_j^G \cdot \frac{L_{ij}}{\sum_i L_{ij}}$, where R_j^G is the national revenue in sector j and L_{ij} the number of workers employed in sector j in county i obtained from the German Federal Institute for Employment Research (IAB).[15]

	$j = 2$	$\cdots$	$j = 12$	$\sum_{j \in 2,...,12}$
$i = 1$				$\neq R_{1,manufac}$
$\vdots$		$\tilde{R}_{ij} = R_j^G \cdot \frac{L_{ij}}{\sum_i L_{ij}}$		$\vdots$
$i = 402$				$\neq R_{402,manufac}$
$\sum_{i \in 1,...,402}$	R_2^G	$\cdots$	R_{12}^G	

Figure 4: Matrix of county sector revenues

The column sums of these initial estimates equal, by construction, the national sectoral revenues obtained from the WIOD. However, the construction method counterfactually assumes that workers in each industry produce an equal amount of revenue across all counties. Consequently, county level aggregates across sectors, that is row sums, will not (necessarily) match the sectoral aggregates $R_{i,manufac}$ collected from the regional statistical offices. Instead, if a county produces a higher than average revenue per worker row sums will be too small and vice versa. To make use of the additional information contained in the observed county level sectoral aggregates I apply an RAS algorithm. This simple method iteratively scales rows and columns to match the given margin constraints. Specifically, the algorithm derives new estimates of the matrix entries by scaling each row i with a single factor ($R_{i,manufac} / \sum_{j \in 1,...,12} \tilde{R}_{ij}$), such that row sums match their target values $R_{i,manufac}$. Of course, having scaled each row by an individual value the column sums will no longer add up to the given margins (R_j^G). The algorithm then scales each column with a single value such that the column sums are again correct, but leaving the row constraints violated again. An iterative repetition of this process of row and column scaling approaches a set of new values R_{ij} that deliver the correct row and column sums. Interestingly, this simple method

[14] I scale county level revenue data for the aggregated mining and manufacturing sector such that the sum across all counties equals the national revenue level reported in the WIOD.

[15] Throughout this paper variables pertaining to Germany as a whole are marked by a superscript "G" to differentiate them from variables pertaining to counties or foreign countries.

delivers the same results as an entropy maximizing approach (McDougall 1999). Specifically, its solution is equivalent to that of solving $\max\left\{-\sum_i \sum_j R_{ij} \log\left(R_{ij}/\tilde{R}_{ij}\right)\right\}$ subject to the constraints $\sum_{j\in 1,...,12} R_{ij} = R_{i,manufac}$ and $\sum_{i\in 1...,402} R_{ij} = R_j^G$. Intuitively, deviation from the initial matrix is penalized and hence the method preserves as much of the initial matrix structure as possible while ensuring that both the observed national sectoral revenues and county level aggregate revenues across sectors are replicated by the resulting values. Particularly appealing to my application is that in the process of iteratively scaling rows and columns the bilateral relative sizes of industries are kept constant, that is $\frac{\tilde{R}_{ij}/\tilde{R}_{ik}}{\tilde{R}_{nj}/\tilde{R}_{nk}} = \frac{R_{ij}/R_{ik}}{R_{nj}/R_{nk}}$ for non-zero revenues.

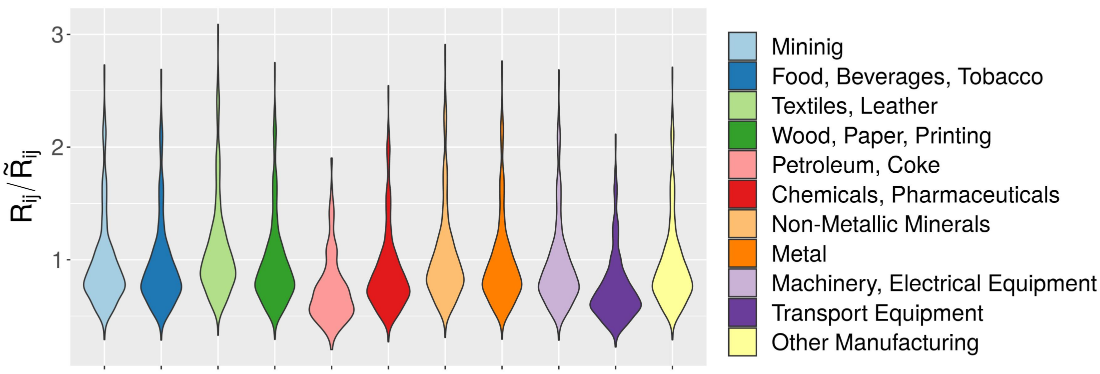


Figure 5: Effects of RAS algorithm on revenues

Figure 5 depicts the density distribution of relative county level revenues in sectors 2 through 12 before and after the application of the RAS algorithm. The matrix balancing approach distorts the initial revenue values to account for differences in revenue per worker across locations, while keeping the national aggregate of sector revenues constant. Only active location-sectors are included as those with 0 output remain unadjusted throughout the RAS procedure. Clearly the initial assumption of an equal revenue per worker in each sector can not be upheld. Instead revenue per worker has to be strongly adjusted upwards (values above 1) in a few counties and slightly lowered (values below 1) in a large number of counties.[16] The same process is applied to derive county sector level value added (V_{ij}) from the national sectoral value added given by the WIOD (V_j^G) and the county level aggregates across sectors from the regional statistical offices, defining the initial matrix entries as $\tilde{V}_{ij} = V_j^G \cdot \frac{R_{ij}}{\sum_i R_{ij}}$.

Secondly, for agriculture, utilities, construction and service sectors the regional statistical offices directly provide value added data at the county level.[17] In these cases I use sectoral value added shares to split national sectoral revenues across counties, that is for sectors

[16] Differences in revenue per worker do not necessarily imply a higher productivity. The highest average difference between initial and final revenues, for example, is observed in Hamburg. This result is partly due to Hamburg being a trading hub with a particular high share of intermediates in production and hence a higher revenue per worker.

[17] Again, I scale county level value added data for each sector such that the sum across all counties equals the national revenue level reported in the WIOD.

$j \in \{1, 13, ..., 17\}$ I set $R_{ij} = R_j^G \cdot \frac{V_{ij}}{\sum_i V_{ij}}$, where V_{ij} denotes value added in sector j in location i.[18]

Having calculated all county sector revenues R_{ij} and value added V_{ij}, aggregate intermediate demand M_{ij} in each sector and county can also easily be derived as the difference between the two, that is $M_{ij} = R_{ij} - V_{ij}$. Descriptive statistics for all results are provided in section 3.4.

3.3 Shipment data

My transport data stems from Schubert et al. (2014) as part of the official "Forecast of nationwide transport relations in Germany 2030" on behalf of the German ministry of transport and digital infrastructure ("Bundesministerium für Verkehr und digitale Infrastruktur"). The data set gives the total shipments in tons by water, train or truck for 2010 between German counties and their trade partners, disaggregated along 25 product categories.[19]

The trade partner can be either a further German county (including the county itself), one of 153 foreign regions aggregating into 41 third countries (of which 29 are also in the WIOD Database), or a major German or international port.[20] The latter two appear as origin or destination whenever the actual origin or final destination is unknown or not in the explicit country sample, for example, shipments to and from Japan. Moreover, the data thus differentiates between shipments to/from, e.g. Hamburg and Hamburg port. I assign all shipments to and from international ports as well as shipments to and from countries not in the WIOD to ROW.

The data on rail and river transport is based on data sets from the federal statistical office specially compiled to publicly unavailable levels of spatial and sectoral disaggregation. Data on truck shipments relies, firstly, on a similar special report at the county level prepared by the department of motor vehicles ("Kraftfahrtbundesamt") from a one week 0.5‰ mandatory sample of German registered trucks with a gross vehicle weight rating above 3.5 tons and, secondly, on complementary NUTS-3 level shipment data for foreign owned trucks from Eurostat.

[18]For the "utilities" sector (13) county level value added data is only available combined with the mining sector. I still opt to use this aggregate to split sectoral revenues across counties, since the possible alternative, that is spreading the national sectoral revenue across counties by county sector employment shares, produces several counties with revenue values smaller than the reported value added.

[19]Air transport is not included in the data set. However, air transport only makes up about 0.1 percent of total transported weight in Germany (4.2 mio tons compared to 3.7 billion tons, cf. Schubert et al. 2014) and only about 1 percent of the value of total foreign trade (212 billion Euros compared to 2050 billion Euros in 2014, Source: "Bundesverband der deutschen Luftverkehrswirtschaft").

[20]The data set includes 43 third countries, but Iceland and Cyprus have no recorded shipments to Germany, that is shipments from these countries are recorded with a German international port as origin.

Shipments of goods from their source to their destination often occur via several "sub-shipments" with potential changes in the mode of transport, for example, a supplier delivering goods by truck to a container terminal where they are loaded onto a boat together with other goods, transported to another terminal and then sent to their final destination via truck. In these cases the product category in the data set of the first and/or last part of the route will be a specific category, while the middle part might be of type "unknown" or "mixed".[21]

Similarly, if complete trucks are transported via train across the Alps, as is common in German-Italian shipments, the weight of the truck will be added to the transported weight for the middle part of the shipment and the weight in the first and/or last part gives the true weight of the transported commodity.

Of the 25 product categories 18 can be directly matched to my agriculture, mining and manufacturing sectors 1 to 12 as shown in table C.1 in the appendix.[22] In two cases, "mining" and "petroleum, coke" several product categories are matched with the respective industries. In these cases I weight transported tons with unit values from the German trade statistics before aggregating them. Three categories have no match in my data ("mail", "moving items, not-for-market items", "Equipment and material for transportation, packaging") and are dropped. The remaining three categories that can occur in the data are "mixed", "unknown" and "other" goods. These are used to scale trade in all other sectors for the respective pair of trade partners.[23] Finally, while the category "Secondary raw materials; municipal wastes and other wastes" would match to the sector "utilities" of this paper, it only makes up for a small share of that sector. The much larger share, that is electricity, gas and steam supply, as well as, water treatment, collection and supply, is (usually) transported by means not captured in the shipment data. Consequently, I do not use the category to proxy for trade in sector "utilities". Instead I drop the category from the shipment data and treat the "utilities" sector as the service sectors below.

Overall I obtain trade flows in terms of weight between the 402 German counties and 26 foreign partners, including ROW, in 17 sectors.[24] These flows include own trade, that is

[21] In the case of such "intermodal" shipments Schubert et al. (2014) use data from container terminals to match shipments to the source container terminal with shipments from the destination container terminal. In some instances, however, a clear match is impossible, for example, if a truck delivers a specific product category to a boat but only "mixed" product trucks leave the ship's destination terminal. In these cases matches might ultimately be assigned randomly and the product category of the first and last part of a shipment can diverge with one being unspecific. In these cases I assume that the specific product category holds for the complete shipment as matched in the data set.

[22] Shipments are given in terms of product categories whereas employment, revenue and value added data are for industries and I therefore have to assume that each industry produces only goods from the matched product category.

[23] Some select reporter-partner pairs only have shipments in the category "unknown". In these cases I assume that these shipments consist of the exporter's average export mix.

[24] Ireland, Greece, Finland and Norway show a large number of zero trade flows compared to other coun-

goods that are produced and used in the same location. Thus, total weight flows with the same origin county must, in each sector, add up to the weight of the total production in the sector. To calculate the value of trade flows $X_{n,ij}$ from sector j goods in a German county $i \in \{1, ..., 402\}$ to location n I therefore multiply the weight share with the county sector revenue. Specifically, I set $X_{n,ij} = R_{ij} \cdot \frac{W_{n,ij}}{\sum_n W_{n,ij}}$, where $W_{n,ij}$ denotes the weight of flows from sector j in location i to location n. In the case of foreign countries exporting to Germany, I split the national level trade flows from the WIOD across counties according to weight shares, that is for $i \in \{403, ..., 428\}$ the trade value is $X_{n,ij} = X_{ij}^G \cdot \frac{W_{n,ij}}{\sum_{n=1,...,402} W_{n,ij}}$, where X_{ij}^G are German imports from sector j in location i. In a final step I rescale all counties intra-national flows and exports to foreign locations such that the aggregate German exports to foreign locations match the values given in the WIOD.[25] Compared to the alternative approach of using national unit values to translate weight flows into value flows my method accounts for the fact that goods in the same sector but from different counties can have very different values per ton. I turn to a descriptive analysis of the final trade network in the next subsection.[26]

3.4 Descriptive analysis

Production. Table 3 provides an overview of the derived production structure in Germany. As shown in the last column, almost all counties are active in the production of almost all sectors with strong exceptions in the "mining" and "petroleum, coke" industries. The three service sectors are by far the largest sectors in the German economy. Adding up the respective values in columns 5 and 6 of table 3 their combined share in total revenue is 0.57 and their share in value added is 0.68. The largest manufacturing sectors in terms of revenue are "machinery, electrical equipment", "petroleum, coke" and "chemicals, pharmaceuticals", the smallest ones are "mining", "textiles, leather" and "non-metallic minerals". The unweighted mean of value added in output across counties is constant in "agriculture", "construction" and service sectors by assumption but varies profoundly in the remaining sectors with a range from 0.05 to 0.97, albeit the mean being relatively similar around 35% to 45% in most sectors.

tries, likely due to the fact that a large share of trade with these countries occurs via international ports. For this reason I chose to aggregate these countries with ROW.

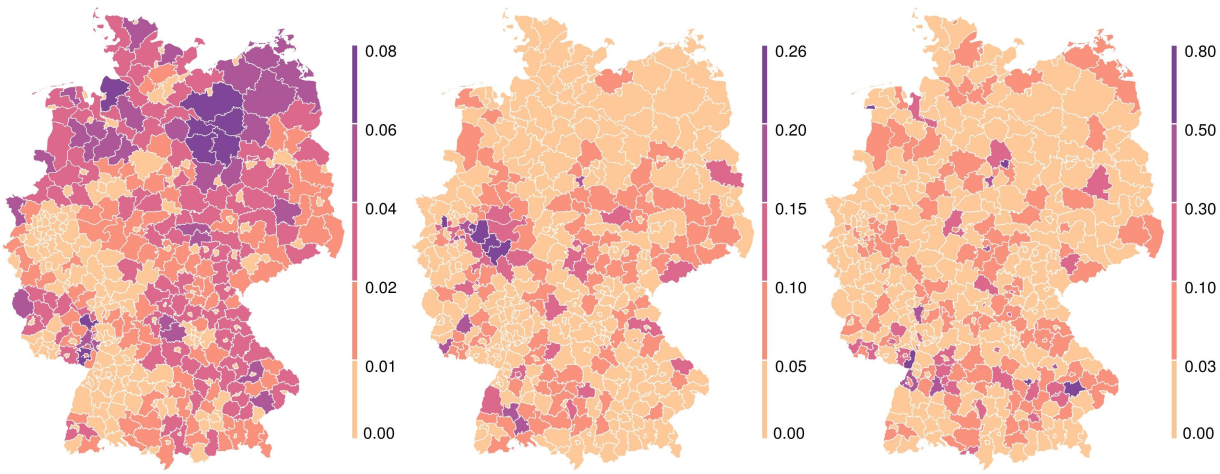
[25]In the "petroleum, coke" sector there are two countries, Latvia and Portugal, to which no German county reports exports, despite the WIOD reporting a country to country flow. Again, this is likely due to this sector relying heavily on pipeline transport. In these case I therefore assume that all producers of "petroleum, coke" export equal shares of their output to these countries.

[26]It should be noted that in contrast to the full IRIOT calculated in section 4, the interregional trade flows derived in this section contain no information about their use category, that is, whether they serve as intermediates in a specific sector or as final consumption at their destination.

Table 3: Sectoral production structure across 402 Germancounties.

	R_{ij}						V_{ij}/R_{ij}								
Sector	min	median	mean	max	$\sum R_{ij}$	$\sum V_{ij}$	min	median	mean	max	HHI	KS	Gini	MI	$R_{ij} > 0$
Agriculture	1	120	153	1026	61611	24900	0.40	0.40	0.40	0.40	0.005	1.02	0.52	0.42	402
Mininig	0	8	42	904	16761	7789	0.00	0.62	0.53	0.66	0.021	1.26	0.79	0.12	357
Food, Beverages, Tobacco	16	350	526	7491	211327	54305	0.07	0.28	0.29	0.58	0.006	0.53	0.36	0.25	402
Textiles, Leather	0	30	80	1021	32194	11117	0.00	0.35	0.35	0.71	0.009	0.96	0.64	0.27	400
Wood, Paper, Printing	6	164	276	3176	110812	34954	0.08	0.34	0.34	0.69	0.007	0.54	0.39	0.09	402
Petroleum, Coke	0	0	207	19559	83259	13817	0.00	0.00	0.07	0.50	0.092	1.53	0.94	0.05	120
Chemicals, Pharmaceuticals	2	349	804	24201	323250	129320	0.11	0.44	0.44	0.89	0.014	0.65	0.47	0.04	402
Non-Metallic Minerals	1	80	130	958	52261	19920	0.10	0.41	0.41	0.83	0.006	0.81	0.55	0.22	402
Metal	6	357	671	8594	269571	88873	0.08	0.35	0.35	0.72	0.008	0.72	0.43	0.35	402
Machinery, Electrical Equipment	18	613	1207	18825	485167	205554	0.11	0.45	0.45	0.91	0.009	0.51	0.38	0.25	402
Transport Equipment	0	213	1123	35187	451391	141551	0.00	0.36	0.36	0.73	0.029	0.98	0.68	0.07	396
Other Manufacturing	16	160	272	5882	109458	47557	0.11	0.47	0.48	0.97	0.009	0.52	0.38	0.14	402
Utilities	15	281	637	12429	256199	115320	0.05	0.46	0.44	0.48	0.011	0.56	0.40	0.09	402
Construction	84	588	766	9579	307845	136702	0.44	0.44	0.44	0.44	0.005	0.38	0.23	0.22	402
Trade, Communication, IT	280	1729	3211	73891	1291013	648894	0.50	0.50	0.50	0.50	0.011	0.25	0.17	0.20	402
Financial, Insurance, Business	370	1849	3418	66457	1373878	847507	0.62	0.62	0.62	0.62	0.012	0.25	0.15	0.20	402
Government, Education, Health	393	1597	2436	53477	979179	712082	0.73	0.73	0.73	0.73	0.008	0.30	0.20	0.18	402

The Herfindahl-Hirschman-Index (HHI) provides a measure of concentration of production.[27] It is strongest in "petroleum, coke", "mining" and "transportation equipment". In the first two cases this is driven by the availability of necessary resources, in the latter case it mirrors the strong concentration of the industry among a few large German car producers. Concentration is lowest in the "agriculture", "construction" and "food, beverages, tobacco" sectors. As an absolute measure of concentration, however, the HHI is influenced by the large size differences of counties in Germany, that is the large size of Berlin, Hamburg and Munich in most sectors increases the HHI and their low significance for agriculture greatly reduces it in this sector. In contrast the Krugman (1991) specialization index (KS) and the spatial Gini coefficient provide measures of relative specialization, comparing the relative county level specialization to the national relative specialization.[28] For both measures "petroleum, coke" and "mining" continue to exhibit the highest level of concentration, followed by "agriculture" and "transport equipment".



(a) Agriculture (b) Metal (c) Transport Equipment

Figure 6: Shares of different industries in regional total production

These simple measures of concentration still hide important aspects of production patterns.

[27]Here the HHI is measured as $HHI_j = \sum_i \left(\frac{R_{ij}}{\sum_i R_{ij}} \right)^2$.

[28]These indices are calculated as:

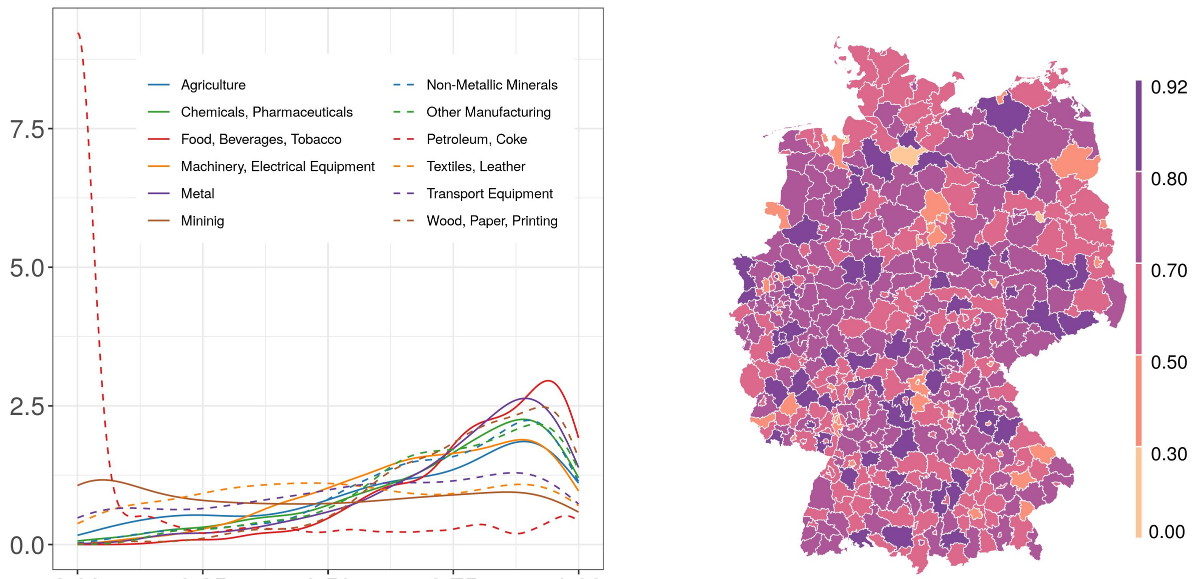
$$KS = \sum_i | \frac{R_{ij}}{\sum_i R_{ij}} - \frac{\sum_j R_{ij}}{\sum_i \sum_j R_{ij}} |$$

$$Gini = \frac{2}{402^2 \bar{LQ}_j} \sum_i r_{ij} \left(LQ_{ij} - \bar{LQ}_j \right)$$

$$LQ_{ij} = \frac{\frac{R_{ij}}{\sum_j R_{ij}}}{\frac{\sum_i R_{ij}}{\sum_i \sum_j R_{ij}}}$$

where LQ_{ij} is a location quotient for region i in sector j, $\bar{LQ}_j$ the mean location quotient in sector j and r_{ij} the rank of county i with respect to location quotients in sector j.

Figure 6 exemplifies this by showing the relative share of "agriculture", "metal" and "transport equipment" in each county's total output. All three show some agglomeration in the KS and Gini coefficient. However, since these indices do not account for distances between counties they fail to capture agglomerations that do not conform to administrative borders. Clearly, the "metal industry" industry shows a strong agglomeration in the Ruhr-area of Germany, albeit spread over several counties. Similarly, agriculture is strongly agglomerated in the north and north-east of Germany, whereas single counties highly specialized in "transport equipment" can be found spread out across the map. Moran's I ($-1 < MI < 1$, see Gibbons et al. 2015) tries to capture this by measuring the strength of spatial correlation in industry location, that is whether counties specialized in a sector are more closely located to similarly specialized counties (positive values) or further away (negative values).[29] With this measure "agriculture" and "metal" are reported as the most strongly agglomerated industries whereas "transport equipment" with its randomly spread production centers drops to the third last position.[30] Further important aspects, such as the clear and intuitive difference between cities and rural counties in the production of agricultural goods, can only be captured through individual observation or by using additional data.



(a) Density of sectoral GL-Indices

(b) Aggregate GL-Index

Figure 7: Intra-industry trade

[29] Moran's I is calculated as:

$$MI = \frac{402}{\sum_i \sum_l w_{il}} \frac{\sum_i \sum_l w_{il} (s_{ij} - \bar{s}_j)(s_{lj} - \bar{s}_j)}{\sum_i (s_{ij} - \bar{s}_j)^2}$$

where $s_{ij} = R_{ij} / \sum_j R_{ij}$ is sector j's share in the total output of location i and w_{il} are elements of a 402 by 402 matrix that take the value 1 if counties i and l have a common border and 0 otherwise (or if $i = l$).

[30] Due to the lack of firm level data, more evolved distance based agglomeration measures such as the Duranton-Overman (Duranton and Overman 2005) index can not be derived here.

Trade. To gain an overview over the derived interregional trade matrix I begin by looking at the strength of intra-industry trade, as measured by the Grubel-Lloyd (GL) index in figure 7.[31] The left hand panel shows the density distribution of Grubel-Lloyd indices across all 402 German counties for the 12 manufacturing sectors for which trade data was derived.[32] Clearly, one way trade is exceptionally prominent in the “petroleum, coke” sector, which is inline with the previous result of a limited number of counties active in this industry. “mining”, “textiles, leather” and “transport equipment” include both counties with strong inter-industry and intra-industry trade whereas the remaining sectors have GL indices above 0.5 in most counties.

The GL index for aggregate trade in each location is depicted in the right hand panel of figure 7. It is above 0.5 for most counties indicating a strong prevalence of intra-industry trade for German counties. Some exceptions exists in and around the cities of Munich, Frankfurt and the largest VW producer Wolfsburg, as well as a handful of further counties.

Foreign trade plays a relatively large role for all counties in Germany. The top row of figure 8 depicts the share of foreign trade in each counties exports and imports respectively. Overall these values are very high, with maximum foreign shares of 0.91 for exports, 0.98 for imports and respective (unweighted) means of 0.5 and 0.44. Surprisingly, counties with higher foreign trade shares are not necessarily located closer to the border. One explanation for this is that a lot of trade occurs via international ports and water ways as witnessed by the high values in the north of Germany. To support this claim the bottom row depicts the share of the nine neighboring countries of Germany in each county’s total exports and imports respectively.[33] As these countries are a subset of all foreign countries the trade shares are obviously reduced. However, it is now clearly visible that being close to the border has a much larger influence on the trade share compared to trade with all foreign countries. Importantly, trade shares of northern counties that either host large international ports or are connected to them via waterways are strongly reduced, as this mode plays a reduced role in trade with immediate neighbors. The differences in these patters are a stark reminder for the necessity to better understand regional trade networks. The IRIOT developed in this

[31] The Grubel-Llyod index for an individual sector j in location i is calculated as

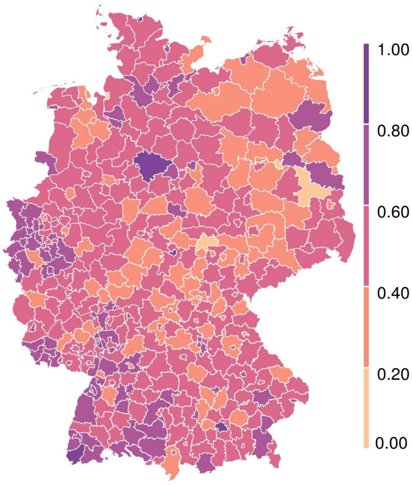
$$GL_{ij} = 1 - \frac{|Exports_{ij} - Imports_{ij}|}{Exports_{ij} + Imports_{ij}}$$

and for the aggregate economy of location i as

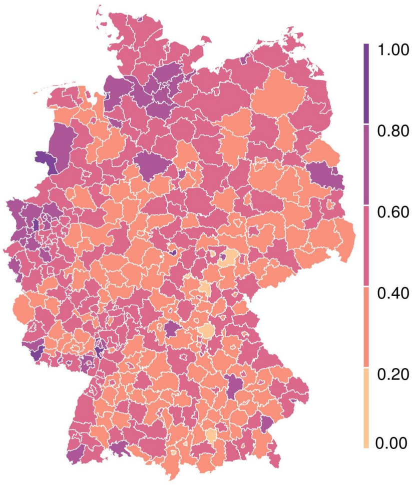
$$GL_i = 1 - \frac{\sum_j |Exports_{ij} - Imports_{ij}|}{\sum_j (Exports_{ij} + Imports_{ij})}$$

[32] As noted, this figure, but also the remainder of this section refers only to trade in the agriculture, mining and manufacturing sectors 1-12 for which shipment data is available.

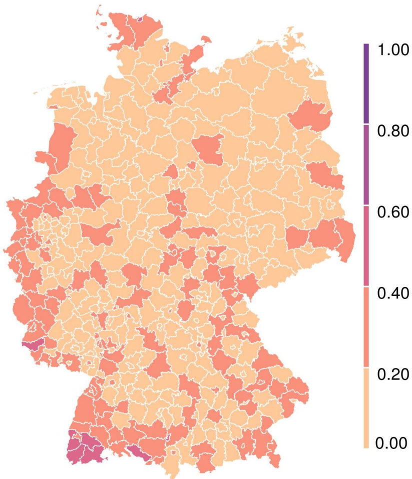
[33] These countries are Denmark, Belgium, the Netherlands, Luxembourg, France, Switzerland, Austria, the Czech Republic, and Poland.



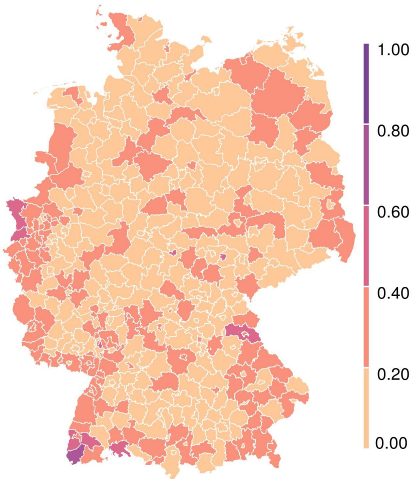
(a) Foreign share in total exports



(b) Foreign share in total imports



(c) Foreign neigbors' share in total exports



(d) Foreign neighbors' share in total imports

Figure 8: Foreign trade shares

paper captures such features and can hence help to understand for example the heterogeneity in effects of international trade agreements signed with different partners.

Lastly, I estimate the effect of internal distance on sectoral trade flows through a gravity estimation. Specifically, trade is approximated by exporter and importer fixed effects as well as by a measure of physical distance between counties.[34] Hence, for imports of county n

[34] Distance is often measured as the distance between the centroids of two counties. However, a particularity of German counties is that often an independent "city county" is surrounded by a roughly ring shaped county, implying centroids that fall very close together. To circumvent this problem I measure distance by drawing 100 random points in each county and calculating the mean of the resulting 10,000 pairwise distances between two counties. A further benefit of this procedure is that it can also be used to derive an internal distance for each county, which allows to include own trade flows in the gravity estimation.

from sector j in location i I estimate

$$X_{n,ij} = \frac{Im_{nj} \cdot Ex_{ij}}{dist_{ni}^{\theta_j}} \ , \tag{1}$$

where Im_{nj} and Ex_{ij} denote importer-sector and exporter-sector specific effects, $dist_{ni}$ is the physical distance between two locations and θ_j the elasticity of trade flows with respect to distance.

Table 4: Sectoral gravity estimates

	OLS			PPML		
Sector	estimate	se	R^2	estimate	se	R^2
Agriculture	-1.90	0.01	0.53	-2.00	0.01	0.74
Mininig	-2.78	0.01	0.69	-2.87	0.02	0.91
Food, Beverages, Tobacco	-1.81	0.01	0.43	-1.62	0.01	0.66
Textiles, Leather	-1.24	0.01	0.54	-1.47	0.01	0.66
Wood, Paper, Printing	-1.49	0.01	0.49	-1.43	0.00	0.71
Petroleum, Coke	-2.46	0.04	0.73	-1.75	0.01	0.85
Chemicals, Pharmaceuticals	-1.64	0.01	0.48	-1.70	0.01	0.65
Non-Metallic Minerals	-2.15	0.01	0.58	-2.25	0.01	0.83
Metal	-1.43	0.01	0.44	-1.55	0.01	0.68
Machinery, Electrical Equipment	-1.43	0.01	0.48	-1.49	0.01	0.74
Transport Equipment	-1.29	0.01	0.65	-1.35	0.00	0.85
Other Manufacturing	-1.07	0.01	0.44	-1.24	0.01	0.68

The first three rows of table 4 report the results of a log-linear estimation of equation (1) including only intra-national trade. All results are highly significant and within in the range usually found in the literature for other countries. Interestingly, having accounted for exporter and importer fixed effects distance suffices to explain a large share of the observed variance in trade flows as witnessed by the relatively high R^2. The inclusion of the fixed effect implies that residuals must be driven by sector specific bilateral factors such as, for example, plants in two counties belonging to the same company. Turning to the actual estimates distance effects are strongest in "mining", "petroleum, coke" and "non-metallic minerals" and weakest in "other manufacturing", "textiles, leather" and "transport equipment".

As the OLS estimator is potentially biased under heteroscedasticity and due to the large number of zero-trade flows that have to be excluded when log-linearizing the gravity equation I re-estimate the model in multiplicative form using PPML (Santos Silva and Tenreyro 2006). With the exception of the "petroleum, coke" sector with its particularly high number of zero trade flows this has only a limited effect on the estimated coefficients but the explanatory power of distance (and fixed effects) increases even further. Overall, this is strong support for my use of gravity estimation for trade flows in the remaining sectors.

4 The IRIOT

The approach to constructing the interregional input-output table in this paper is unique as it relies on intra-national trade flows that are usually unavailable in the construction of such tables. Using these trade flows I proceed in two steps.

First, for agriculture, mining and manufacturing sectors for which interregional trade flows were derived in the previous section I balance an IRIOT based on national input-output coefficients from the WIOD adapting them to the given trade flows. This process also leads to approximations of each county's demand, independent of origin, for "utilities", "construction" and service sectors.

Second, as no shipment data is available in the "utilities", "construction" and service sectors, I rely on a gravity model to estimate interregional trade flows based on county level demand and production, as well as physical distance between locations. Final input-output coefficients for each exporter importer pair in these sectors are derived employing a multi-dimensional extension of the RAS matrix balancing approach and constraining the result to all previously derived trade flows, demand and production levels.[35] Throughout I define $X_{nk,ij}$ as the flow from industry j in location i to use category k in location n, where the use category can be one of the 17 industries (where the flow is used as an intermediate input) or final demand.[36] Moreover, for ease of notation denote as $\Omega_g \equiv \{1, ..., 402\}$ the set of all German counties in the $N = 428$ total locations.

Flows $X_{nk,ij}$ between foreign countries, that is for $i \notin \Omega_g$ and $n \notin \Omega_g$, in any sector and for any use are taken from the WIOD and remain unchanged.

4.1 Agriculture, Mining and Manufacturing

To match the previously derived total flows from sectors 1 through 12 between each German county and each foreign partner to the specific use category I apply the proportionality assumption that the use shares of these flows are constant for all exporting counties. They can then be immediately recovered from the WIOD as the use shares of German exports in each foreign country. Hence, for $j \in \{1, ..., 12\}$, $i \in \Omega_g$ and $n \notin \Omega_g$ flows are derived as $X_{nk,ij} = X_{n,ij} \cdot U^G_{nk,j}$, where $U^G_{nk,j}$ is the share of total German exports of sector j to foreign

[35] Holý and Safr (2017) introduce the multidimensional RAS method in an input-output analysis context and prove the convergence of the iterative procedure. They apply it for the Czech-Republic albeit to a different specific problem with viewer dimensions and constraints. Section A in the appendix explains the multidimensional extension of the RAS approach.

[36] This notation allows for an easy interpretation of all other variables which are simply the sum of these flows over the missing indices. For example, revenue R_{ij} is the sum of flows $X_{nk,ij}$ over n and k; trade flows $X_{n,ij}$ are the sum of flows $X_{nk,ij}$ over all use categories k and so forth.

importer n used in category k.

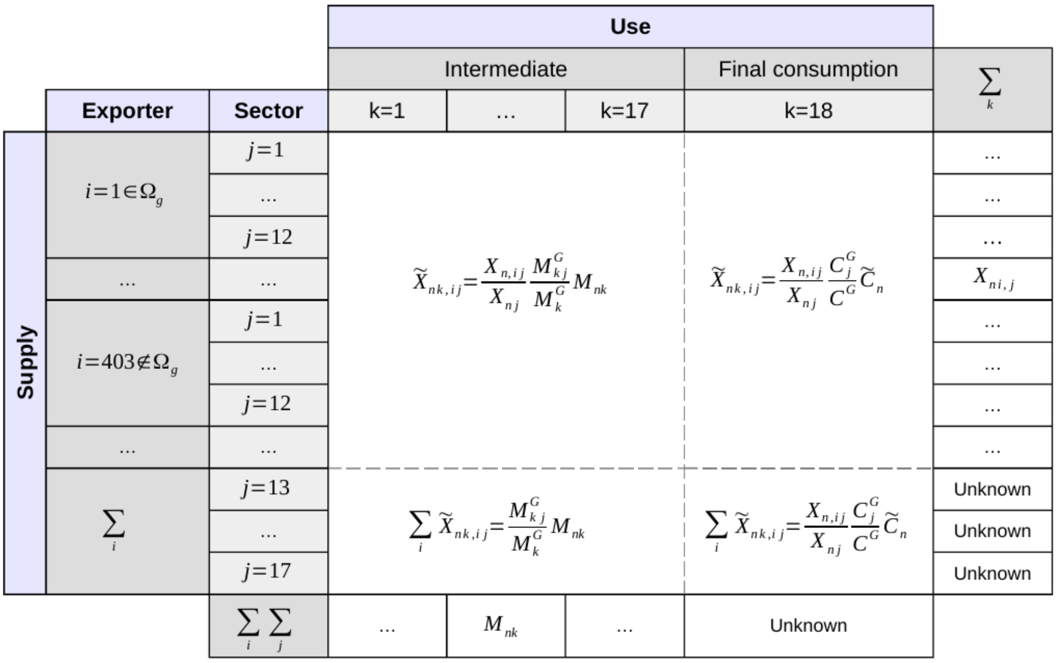


Figure 9: Matrix slice of initial RAS array for any importer $n \in \Omega_g$

For any flow in agriculture, mining or manufacturing where a German county is the importer I derive flows to different use categories through a multidimensional extension of the RAS method.[37] Specifically, for each importer $n \in \Omega_g$ I consider a matrix as shown in figure 9. For exporting sectors $j \in \{1, ..., 12\}$ the matrix depicts detailed flows from exporter to importer by use category. For the remaining 5 sectors for which no trade data exists the aggregate imports (from all potential exporters) in county n and use category k are displayed in the last 5 rows. All initial flows $\tilde{X}_{nk,ij}$ are constructed as follows:

Firstly, I use residential income data from the regional statistical offices to derive an initial estimate of aggregate consumption demand $\tilde{C}_n$ in each county $n \in \Omega_g$ by distributing the WIOD German national demand across counties based on their share in national income.

Secondly, I split these consumption demands and the previously derived total intermediate demand M_{nk} of each sector in each county across different industries based on the proportionality assumption that national shares, that is the size of industry j in the total intermediate demand of sector k $\left(M^G_{jk}/M^G_k\right)$ or in consumption $\left(C^G_j/C^G\right)$, hold at the county level. In sectors for which trade data exists the demand for sector j intermediate or consumption goods in each use category k in county n is then split across potential exporters i based on their share in the total imports of sector j goods by county n (that is $X_{ni,j}/X_{nj}$). This last step shows the great advantage of having shipment data available and is what allows to construct a data based interregional instead of just regional input output database.

[37] For the multidimensional extension of the RAS method see section A in the appendix.

Having constructed initial flows the final step applies a multidimensional RAS balancing algorithm to the array keeping its structure as close to the initial values as possible while satisfying all previously derived margins. In particular these constraints can be expressed as the following conditions:

1. For sectors $j \in \{1, ..., 12\}$ summing flows across all use categories k for a specific importer n and exporter i must equal total trade flows from i to n in sector j as derived from the shipment data, that is $\sum_k X_{nk,ij} = X_{n,ij}$ for $j \in \{1, ..., 12\}$.

2. For the remaining sectors $j \in \{13, ..., 17\}$ further summing the already aggregated flows in the matrix across all use categories and importing counties must result in the total national demand for goods from sector j as reported by the WIOD.[38]

3. Summing flows across all exporters i and sectors j for a fixed importer and use category $k \in \{1, ..., 17\}$ must result in the previously calculated aggregate intermediate demand levels, that is $\sum_i \sum_j X_{nk,ij} = M_{nk}$.

4. For a given foreign exporter $i \notin \Omega_g$, sector j, and use category k summing flows across all importers $n \in \Omega_g$ must be equal to the international flows from sector j in country i to use category k in Germany as reported by the WIOD.

5. Finally, summing flows across all importing counties and all exporters for a specific sector j and use category k gives the use of intermediate j in the national German use category k as reported by the WIOD.

In the “petroleum, coke” sector and a few further instances almost exclusively in the sector “mining” matching the derived trade flows (as imposed by the first constraint) and observed use structure (as imposed by the last three constraints) simultaneously is not possible. Specifically, from the WIOD more than 50% of the intermediate inputs of the national “petroleum, coke” sector come from the “mining” sector which includes crude oil. However, there are only a few counties that are important producers of refined petroleum and coke and even if all flows of mining goods in the shipment data to these locations were attributed as inputs to the petroleum industry the national share in inputs of over 50% could not be achieved. Similarly, the national usage of “petroleum, coke” sector goods as inputs in the same industry can also not be matched given the shipment data. The likely reason for this is that the shipment data unfortunately do not contain pipeline transports which are the major mode of transport for both crude oil and petroleum. To solve this problem the first condition is not enforced in sector 6 and in some international flows, mainly in sector 2,

[38] For sectors $j \in \{1, ..., 12\}$ the first condition already ensures that summing across all use categories k and importers n equals the total national demand for sector j goods from i.

representing for example German imports of Russian crude oil and gas. In these cases trade flows $X_{n,ij}$ are allowed to adapt during matrix balancing as long as the observed aggregate flows to Germany $\sum_{n\in\Omega_g} X_{n,ij}$ remain constant (cf. footnote 38).

It is also important to note, that the third condition is only binding for use categories $k \in \{1, ..., 17\}$. Hence, aggregate consumption in each county is allowed to change from the initial value estimated through income shares. The reason for this is twofold. Firstly, as there is substantial mobility of consumers and commuting at the county level (see Krebs and Pflüger 2018) residential household income can only be an imperfect estimate of the actual final demand of goods in each county. Secondly, the use of shipment data comes with the great benefit that not only total intermediate demand in each county is directly derived from the data but, importantly, also the supply of goods from sectors 1 through 12 to each county. Therefore, if, for example total intermediate demand is small but the value of shipments to the specific location is relatively large the county must either use and consume less than the national average from sectors 13 through 17 or aggregate consumption must be higher. Making use of this the balancing procedure is allowed to adapt values along both margins.

4.2 Utilities, Construction and Services

Gravity. In the remaining five sectors only country level trade data is available from the WIOD but no county level trade data.[39] For these sectors I rely on a gravity approach to establish county level trade flows. Specifically, for $n \in \Omega_g$ and/or $i \in \Omega_g$ trade flows from location i to location n in sector j are expressed as

$$X_{n,ij} = \frac{Im_{nj} \cdot Ex_{ij}}{d_{ni,j}} \;, \tag{2}$$

where Im_{nj} and Ex_{ij} denote importer sector and exporter sector specific effects and d_{nij} is a sector specific trade barrier for flows from location i to location n. As I only need to derive trade flows where either the importer or the exporter is a German county, the exporter fixed effects and importer fixed effects of foreign countries comprise any international border effects, such as having a common currency or language. As a consequence and as in the previous section I apply the common assumption that trade barriers are a log linear function of the distance between locations, that is $d_{ni,j} = dist_{ni}^{\theta_j}$. The parameter values θ_j are taken from Anderson et al. (2016) who are among the view that derive the effects of interregional distance on service trade.[40] Their values, which are mostly in the range of 0.91 to 1.38, are

[39] As explained above transport data on "secondary raw materials; municipal wastes and other wastes" is available but only makes up for a small share in the sector "utilities" and is therefore not used here.

[40] I use the aggregate service sector coefficient for sectors 13 and 14, the unweighted average of their transport, wholesale, accommodation and communication sectors for sector 15, of finance and business for sector 16 and of education and health for sector 17.

however similar to distance coefficients derived for international service trade and aggregate trade flows in the literature.

A novelty and particular strength of my two step approach is that to derive the levels of exporter and importer fixed effects I can rely on the previously calculated county level sectoral revenue and demand data. Denoting location n's total demand, that is the sum of intermediate and final demand, for sector j goods as D_{nj} and its total demand for sector j goods produced in any country in Germany as D^G_{nj} it must hold that $\sum_i X_{n,ij} = D_{nj}$ for all $n \in \Omega_g$ and $\sum_{i\in\Omega_g} X_{n,ij} = D^G_{nj}$ for all $n \notin \Omega_g$. Plugging equation (2) into these constraints allows to solve for importer fixed effects as

$$Im_{nj} = \frac{D_{nj}}{\sum_i Ex_{ij} dist_{ni}^{-\theta_j}} \quad \forall n \in \Omega_g \tag{3}$$

$$Im_{nj} = \frac{D^G_{nj}}{\sum_{i\in\Omega_g} Ex_{ij} dist_{ni}^{-\theta_j}} \quad \forall n \notin \Omega_g \tag{4}$$

Similarly summing a specific exporter's sectoral trade flows across all importers (including the exporter itself) yields the exporter's sectoral revenue, that is $\sum_n X_{n,ij} = R_{ij}$, and summing across all German importer's gives the location's total exports to Germany, that is $\sum_{n\in\Omega_g} X_{n,ij} = X^G_{ij}$. Plugging the gravity equation 2 and the derived importer fixed effects into these constraints allows to derive exporter fixed effects as

$$Ex_{ij} = \frac{R_{ij}}{\sum_n Im_{nj} dist_{ni}^{-\theta_j}} = \frac{R_{ij}}{\sum_{n\in\Omega_g} \frac{D_{nj} dist_{ni}^{-\theta_j}}{\sum_l Ex_{lj} dist_{nl}^{-\theta_j}} + \sum_{n\notin\Omega_g} \frac{D^G_{nj} dist_{ni}^{-\theta_j}}{\sum_{l\in\Omega_g} Ex_{lj} dist_{nl}^{-\theta_j}}} \quad \forall i \in \Omega_g \tag{5}$$

$$Ex_{ij} = \frac{X^G_{ij}}{\sum_n Im_{nj} dist_{ni}^{-\theta_j}} = \frac{X^G_{ij}}{\sum_{n\in\Omega_g} \frac{D_{nj} dist_{ni}^{-\theta_j}}{\sum_l Ex_{lj} dist_{nl}^{-\theta_j}}} \quad \forall i \notin \Omega_g \tag{6}$$

Normalizing one location's exporter fixed effect to 1 in each sector allows to numerically solve this system for all remaining exporter and subsequently importer fixed effects.

Finally, plugging fixed effects and my parameterization of trade costs into the gravity equation 2 allows to calculate all bilateral trade flows $X_{n,ij}$ in sectors $j \in \{13, ..., 17\}$.

Input-Output structure. Having estimated all bilateral trade flows I can assign them to use categories using a similar approach as for manufacturing sectors above. Specifically, as before, trade flows from German counties to foreign countries are distributed across use categories in the foreign country using the proportionality assumption together with use shares from the WIOD.

For all flows in sectors 13 through 17 where a German county is the importer I construct a set of initial flows $\tilde{X}_{nk,ij}$ by also imposing the proportionality assumption that counties have equal use shares for imports independent of the source. Of course, this implies that aggregate sectoral flows from foreign countries to all German counties ($X^G_{k,ij}$) will not (necessarily) match the values given in the WIOD. To ensure that these flows are replicated while keeping constant the total demand for each service sector in each county I rely on a final RAS balancing step imposing $\sum_{n\in\Omega_g} X_{nk,ij} = X^G_{k,ij}$ for all $i \notin \Omega_g$ and $\sum_i X_{nk,ij} = M_{nk,j}$ for all i.

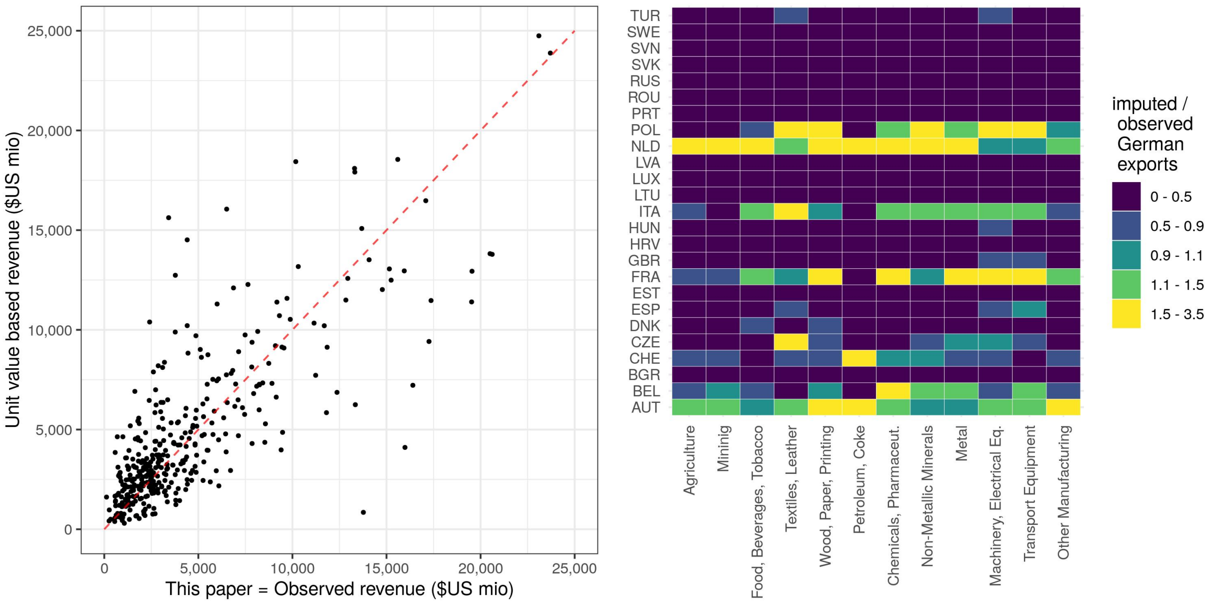
4.3 The final IRIOT

Combining the derived data for all sectors results in the final input-output table for 402 German counties and 26 foreign countries, including ROW, across 17 sectors and 18 use categories. Summing values across all German counties leads to a matrix with 27 countries that exactly replicates all flows as given in the inventory-adjusted WIOD. In all German counties sectoral revenues, value added and, consequently intermediate demand equal the values reported by the regional statistical offices.[41] Interregional trade flows, and international trade flows with German counties in agriculture, mining and manufacturing sectors match the export shares implied by shipment data in weights, with some exceptions in the petroleum and mining sector that are necessary to replicated the national input-output structure from the WIOD. Regional consumption spending is based on residential household income as reported by the regional statistical offices and adjusted using a compromise between keeping sectoral intermediate good spending shares on services constant across locations and accounting for the observed trade imbalances in agriculture, mining and manufacturing sectors. Trade in the remaining sectors is represented by flows derived from a gravity equation with slight adjustments necessary to match the national input-output structure given by the WIOD.

The strong replication of observed data points is the key benefit of the IRIOT developed in this paper compared to regionalizations that are purely based on porportionality assumptions and/or unit value approaches. To exemplify this point I construct a set of trade flows, including within county trade, based on the regional shipment data in weights presented above and relying on national unit values. Specifically, I scale transported weights in each sector with unit values such that the sum of the resulting values across all regions equals the respective sectoral German production value in the WIOD. I then aggregate the implied sectoral revenue of each region for mining and manufacturing and compare it to the observed county level revenue in mining and manufacturing (see section 3.2).

Figure 10(a) depicts the resulting imputed and the observed revenues. While the observed aggregate mining and manufacturing revenue of each region is exactly replicated by the

[41] Up to a scaling factor that matches national aggregates to the values given in the WIOD.



(a) revenues (b) exports

Figure 10: Comparison of unit value based approach to this paper

IRIOT constructed in this paper the unit value approach clearly implies a substantial level of error. About a third of all imputed revenues lie more than 50 percent below or above their observed value and only about 15 percent of revenues deviate by less than 10 percent from their observed value. Overall the R-squared between imputed and observed regional revenues reaches only 0.52 and the standard deviation of the *ratio* of calculated to observed regional revenues is as high as 1.06.

To further show the shortcomings of a national unit value based IRIOT I scale the unit value based regional foreign exports such that their aggregate matches the observed German sectoral exports from the WIOD. Figure 10(b) compares the resulting unit-value based, imputed German exports by sector and partner to the observed data by showing the respective ratios. While again the IRIOT developed in this paper replicates these flows exactly by construction, the unit value approach, in most sectors, vastly overpredicts trade with Germany's neighbors, especially Austria, the Netherlands, Poland, France and Belgium, as well as with Italy and underestimates trade with other countries. This bias is particularly strong in the "Petroleum, Coke", "Agriculture", "Mining" and "Other Manufacturing" sectors. Overall, only 16 of 300 imputed German sectoral exports to specific countries fall within a 10 percent range of the observed data.

The strong connection to the observed data in combination with the high level of sectoral and regional disaggregation that the IRIOT developed in this paper exhibits, is a large step forward in economists' ability to quantify regional economic effects in Germany. In particular the table can be used both for improved on impact analysis as well as to better calibrate regional CGE models capturing the vast heterogeneity across regions within Germany. For this purpose all code necessary to construct the final IRIOT is available from the author upon request.[42]

5 Conclusion

This paper analyzed the trade and production structure across German counties based on a unique data set of county level shipments in Germany and between German counties and foreign partners, as well as on a number of further regional data sets. The heterogeneity across locations is vast along a wide variety of agglomeration measures, as well as specialization and trade indices. For this reason it is important to account for regions and their role in the German production network when analyzing the effects of international shocks on Germany. Similarly, the effects of regional shocks both on other locations and within the treated counties can only be understood in the context of this network. To provide the necessary data to perform these types of studies I adapt several recent methods in the construction of input-output tables to the specific raw data availability in Germany and show how to use shipment data to construct a county level IRIOT for Germany. Keeping the national aggregates of this table cell-by-cell compatible with the WIOD allows to embed foreign trade into the IRIOT and to provide a fully specified input-output table for German counties in the world economy that can serve as the basis for CGE modelling and on impact analysis. Finally, I show that the obtained table is far superior to simple, national unit value or proportionaly based approaches, thus allowing for better future quantifications and simulations of German regional effects of international and regional shocks.

References

Anderson, J., M. Vesselovsky, and Y. Yotov (May 2016). “Gravity with scale effects”. **Journal of International Economics** 100, 174–193.

Bacharach, M. (1965). “Estimating Nonnegative Matrices from Marginal Data”. **International Economic Review** 6 (3), 294–310.

Becker, R. and M. Henkel (2020). “The role of key regions in spatial development”. *DICE Discussion Papers* 331. Düsseldorf Institute for Competition Economics (DICE).

Bonfiglio, A. and F. Chelli (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.

[42]The underlying shipment data can be obtained from http://daten.clearingstelle-verkehr.de.

Duranton, G. and H. Overman (2005). “Testing for Localization Using Micro-Geographic Data”. **The Review of Economic Studies** 72 (4), 1077–1106.

Flegg, A. and T. Tohmo (2013). “Regional input-output tables and the FLQ formula: A case study of Finland”. **Regional Studies** 47 (5), 703–721.

Flegg, A. and T. Tohmo (2019). “The regionalization of national input-output tables: A study of South Korean regions”. **Papers in Regional Science** 98 (2), 601–620.

Flegg, A. and C. Webber (1997). “On the appropriate use of location quotients in generating regional input output tables: reply”. **Regional Studies** 31, 795–805.

Flegg, A., C. Webber, and M. Elliott (1995). “On the appropriate use of location quotients in generating regional input output tables”. **Regional Studies** 29, 547–561.

Flegg, A., C. Webber, and M. Elliott (2000). “Regional size, regional specialization and the FLQ formula”. **Regional Studies** 34, 563–569.

Gibbons, S., H. Overman, and E. Patacchini (2015). “Spatial Methods”. In: **Handbook of Regional and Urban Economics**. Ed. by G. Duranton, V. Henderson, and W. Strange. Vol. 5. Elsevier. Chap. 3, 115–168.

Hewings, G. and R. Jensen (1986). “Regional, Interregional and Multiregional Input-Output Analysis”. In: **Handbook of Regional and Urban Economics**. Ed. by P. Nijkamp. Vol. 1. Elsevier. Chap. 8, 295–355.

Holý, V. and K. Safr (2017). “The Use of Multidimensional RAS Method in Input-Output Matrix Estimation”. mimeo.

Koschel, H., U. Moslener, B. Sturm, U. Fahl, B. Rühle, and H. Wolf (2006). “Integriertes Klimaschutzprogramm Hessen - InKlim 2012”. *Endbericht*. ZEW, IER, ISI, ifeu, ZIV, IWU, ISET, HLUG.

Kowalewski, J. (2015). “Regionalization of National Input–Output Tables: Empirical Evidence on the Use of the FLQ Formula”. **Regional Studies** 49 (2).

Krebs, O. (2018). “Shocking Germany — A spatial analysis of German regional labor markets”. *BGPE Discussion Paper* 183.

Krebs, O. and M. Pflüger (2018). “How Deep Is Your Love? A Quantitative Spatial Analysis of the Transatlantic Trade Partnership”. **Review of International Economics** 26 (1), 171–222.

Krebs, O. and M. Pflüger (2019). “On the Road (Again): Commuting and Local Employment Elasticities in Germany”. *Revised version of IZA Discussion Paper 12257*.

Kronenberg, T. (2009). “Construction of Regional Input-Output Tables Using Nonsurvey Methods”. **International Regional Science Review** 32 (1), 40–64.

Kronenberg, T. (2010). “Erstellung einer Input-Output-Tabelle für Mecklenburg-Vorpommern”. **AStA Wirtschafts- und Sozialstatistisches Archiv** 4 (3), 223–248.

Krugman, P. (1991). “Increasing Returns and Economic Geograpy”. **The Journal of Political Economy** 99 (3), 483–499.

Lameli, A., V. Nitsch, J. Südekum, and N. Wolf (Aug. 2015). "Same Same But Different: Dialects and Trade". **German Economic Review** 16 (3), 290–306.

McDougall, R. (1999). "Entropy Theory and RAS are Friends". *GTAP Working Papers* 300. Center for Global Trade Analysis.

Nitsch, V. and N. Wolf (2013). "Tear down this wall: on the persistence of borders in trade". **Canadian Journal of Economics** 46 (1), 154–179.

Santos Silva, J. M. C. and S. Tenreyro (2006). "The Log of Gravity". **The Review of Economics and Statistics** 88 (4), 641–658.

Schaffer, W. and K. Chu (1969). "Nonsurvey Techniques for Constructing Regional Interindustry Models". **Papers in Regional Science** 23 (1), 83–104.

Schröder, A. and K. Zimmermann (2014). "Erstellung regionaler Input-Output-Tabellen. Ein Vergleich existierender Ansätze und ihre Anwendung für die Deutsche Ostseeküstenregion". *RADOST-Berichtsreihe* 33. Regionale Anpassungsstrategien für die deutsche Ostseeküste.

Schubert, M., T. Kluth, G. Nebauer, R. Ratzenberger, S. Kotzagiorgis, B. Butz, W. Schneider, and M. Leible (2014). "Verkehrsverflechtungsprognose 2030". *Schlussbericht.* Bundesministerium für Verkehr und digitale Infrastruktur.

Seidel, T. and J. Wickerath (2019). "Rush Hours and Urbanization". *CESifo Working Paper* 7587.

Stäglin, R. (2001). "A step by step procedure to regionalized input-output analysis". In: **Regional Input-Output-Analysis - Conceptual Issues, Airport Case Studies and Extensions**. Ed. by W. Pfähler. Vol. 66. Nomos.

Stone, R. and A. Brown (1962). **Programme for Growth: A Computable Model of Economic Growth Part 1**. London: Chapman and Hall.

Timmer, M., E. Dietzenbacher, B. Los, R. Stehrer, and G. de Vries (2015). "An Illustrated User Guide to the World Input-Output Database: the Case of Global Automotive Production". **Review of International Economics** 23, 575–605.

Többen, J. (2017). "Effects of energy and climate policy in Germany: A multiregional analysis". *PhD-Thesis.* University of Groningen.

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

Appendix: RIOTs in Germany - Constructing an interregional input-output table for Germany

A Multidimensional RAS

As explained for the revenue matrix in the main text, the simple RAS approach takes an $I \times J$ matrix $\tilde{A}$ with elements $\tilde{A}_{ij}$ and transforms it into a matrix A with elements A_{ij} that satisfies given margin constraints, that is, constraints for all row and column sums. The process consists of a simple iterative scaling of rows and columns. Specifically, given target row sums T_i^I for each row i the first step of the algorithm scales each row by a factor $T_i^I / \left(\sum_j \tilde{A}_{ij}\right)$ to obtain new values $\tilde{A}_{ij}^{t=1}$. In the second step, given the target column sums T_j^J each column is scaled by a factor $T_j^J / \left(\sum_i \tilde{A}_{ij}^{t=1}\right)$ resulting in new estimates $\tilde{A}_{ij}^{t=2}$. These two steps are repeated until the actual margin sums match the targets up to a given precision.

Importantly, one can also apply the RAS approach partially, that is, without a full set of row sum and column sum constraints. In this case the unconstrained rows and columns are simply left unscaled in the appropriate steps.[43]

The multidimensional procedure applies the same process of iterative scaling of margins to meet given constraints but for a multidimensional array. As an example, consider a three dimensional extension of the matrix $\tilde{A}$ to an $I \times J \times K$ array with elements $\tilde{A}_{ijk}$. Given a target sum T_i^I for margin I, that is, imposing $\sum_j \sum_k A_{ijk} = T_i^I$, the algorithm scales each matrix slice i by $T_i^I / \left(\sum_j \sum_k A_{ijk}\right)$ and equivalently for margins J and K. Repeating these scaling steps iteratively again leads to an array A that matches the target sums up to a given precision.

In contrast to the simple RAS approach, constraints in the multidimensional case can also be applied to combination of margins. To see this, consider again the three dimensional $I \times J \times K$ array $\tilde{A}$. This time, however, we are interested in a target array A that satisfies the two constraints $\sum_i A_{ijk} = T_{jk}^{JK}$ and $\sum_j A_{ijk} = T_{ik}^{IK}$ where T^{JK} and T^{IK} are matrices of size $I \times K$ and $J \times K$ respectively. Similar to the above we begin by scaling all elements of the array $\tilde{A}$ for a fixed j and k with a factor $T_{jk}^{JK} / (\sum_i A_{ijk})$. Cycling through all combinations of j and k we gain new array elements $\tilde{A}_{ijk}^{t=1}$. In the second step all elements of the new array $\tilde{A}^{t=1}$ for a fixed i and k are scaled by the factor $T_{ik}^{IK} / \left(\sum_j A_{ijk}\right)$. Again cycling through all combinations of i and k we obtain the new array elements $\tilde{A}_{ijk}^{t=2}$. As before, these two steps are repeated until the target margins are met up to a given precision.

Importantly, any combination of such margin constraints can be imposed upon the initial estimate $\tilde{A}$ and, again, any number of elements of each margin can be left unconstrained by simply skipping the appropriate scaling step.

[43]Leaving, for example, row i unconstrained does not mean that values in this row remain unchanged, as elements A_{ij} are still affected from column scaling. However, the size of the sum of all elements in row i is allowed to change freely.

Figure 9 in the main text presents one matrix slice of an initial array with the appropriate margin constraints. The full set of constraints in this case is given in the list in section 4.1.

B Initial WIOD preparation

As explained in the main text, I use the world input output database (WIOD) as my main data source for the national production structure and international trade flows aggregating its content to the 17 industries and 27 countries listed in tables 1 and 2.[44] The WIOD includes inventory changes as a final use category and these can sometimes be of a substantial magnitude and also, of course, negative. If I were to calculate final demand by simply summing over consumption, investment, government spending and inventory changes given in the WIOD I would end up with a negative final demand in some cases. Therefore, I directly include positive inventory changes in final demand but treat negative inventory changes as if they had been produced (and consumed) in the current period. To correctly capture all the intermediate products that would have been necessary to produce this additional output I construct a Leontief-inverse from the WIOD's input-output table which captures the intermediate requirements of final goods production. I then recalculate total production with final demand increased by the negative inventory changes and use the resulting input-output table to calibrate my model. The details of this process are laid out in Krebs and Pflüger (2018).

[44]The matching of the 56 sectors in the WIOD to these 17 industries is shown in table C.1 in the appendix.

C Industry Correspondence

County level data (WZ08)		Shipment data (NST07)		WIOD		This paper	
01	Crop and animal production, hunting and related service activities	10	Products of agriculture, hunting, and forestry; fish and other fishing products	1	Crop and animal production, hunting and related service activities	1	Agriculture
02	Forestry and logging			2	Forestry and logging		
03	Fishing and aquaculture			3	Fishing and aquaculture		
05	Mining of coal and lignite	21	Coal	4	Mining and quarrying	2	Mining
		22	Lignite				
06	Extraction of crude petroleum and natural gas	23	Crude petroleum and natural gas				
07	Mining of metal ores	31	Metal ores				
08	Other mining and quarrying	32	Fertilizers				
		33	Stones and earth, other mining products				
09	Mining support service activities						
10	Manufacture of food products	40	Food products, beverages and tobacco	5	Manufacture of food products, beverages and tobacco products	3	Food, Beverages, Tobacco
11	Manufacture of beverages						
12	Manufacture of tobacco products						
13	Manufacture of textiles	50	Textiles and textile products; leather and leather products	6	Manufacture of textiles, wearing apparel and leather products	4	Textiles, Leather
14	Manufacture of wearing apparel						
15	Manufacture of leather and related products						

	County level data (WZ08)		Shipment data (NST07)		WIOD		This paper
16	Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials	60	Wood and products of wood and cork (except furniture); articles of straw and plaiting materials; pulp, paper and paper products; printed matter and recorded media	7	Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials	5	Wood, Paper, Printing
17	Manufacture of paper and paper products			8	Manufacture of paper and paper products		
18	Printing and reproduction of recorded media			9	Printing and reproduction of recorded media		
19	Manufacture of coke and refined petroleum products	71	Coke	10	Manufacture of coke and refined petroleum products	6	Petroleum, Coke
		72	Refined Petroleum				
20	Manufacture of chemicals and chemical products	80	Chemicals, chemical products, and man-made fibers; rubber and plastic products ; nuclear fuel	11	Manufacture of chemicals and chemical products	7	Chemicals, Pharmaceuticals
21	Manufacture of basic pharmaceutical products and pharmaceutical preparations			12	Manufacture of basic pharmaceutical products and pharmaceutical preparations		
22	Manufacture of rubber and plastic products			13	Manufacture of rubber and plastic products		
23	Manufacture of other non-metallic mineral products	90	Other non metallic mineral products	14	Manufacture of other non-metallic mineral products	8	Non-Metallic Minerals
24	Manufacture of basic metals	100	Basic metals; fabricated metal products, except machinery and equipment	15	Manufacture of basic metals	9	Metal

County level data (WZ08)		Shipment data (NST07)		WIOD		This paper	
25	Manufacture of fabricated metal products, except machinery and equipment			16	Manufacture of fabricated metal products, except machinery and equipment		
26	Manufacture of computer, electronic and optical products	110	Machinery and equipment n.e.c.; office machinery and computers; electrical machinery and apparatus n.e.c.; radio, television and communication equipment and apparatus; medical, precision and optical instruments; watches and clocks	17	Manufacture of computer, electronic and optical products	10	Machinery, Electrical Equipment
27	Manufacture of electrical equipment			18	Manufacture of electrical equipment		
28	Manufacture of machinery and equipment n.e.c.			19	Manufacture of machinery and equipment n.e.c.		
29	Manufacture of motor vehicles, trailers and semi-trailers	120	Transport Equipment	20	Manufacture of motor vehicles, trailers and semi-trailers	11	Transport Equipment
30	Manufacture of other transport equipment			21	Manufacture of other transport equipment		
31	Manufacture of furniture	130	Furniture; other manufactured goods n.e.c.	22	Manufacture of furniture; other manufacturing	12	Other Manufacturing
32	Other manufacturing						
33	Repair and installation of machinery and equipment			23	Repair and installation of machinery and equipment		
35	Electricity, gas, steam and air conditioning supply	140	Secondary raw materials; municipal wastes and other wastes	24	Electricity, gas, steam and air conditioning supply	13	Utilities
36	Water collection, treatment and supply			25	Water collection, treatment and supply		

	County level data (WZ08)		Shipment data (NST07)		WIOD		This paper
37	Sewerage			26	Sewerage; waste collection, treatment and disposal activities; materials recovery; remediation activities and other waste management services		
38	Waste collection, treatment and disposal activities; materials recovery						
39	Remediation activities and other waste management services						
F	Construction			27	Construction	14	Construction
G-J	Trade, transportation, storage, accommodation and food service activities, information and communication.			28	Wholesale and retail trade and repair of motor vehicles and motorcycles	15	Trade, Communication, IT
				29	Wholesale trade, except of motor vehicles and motorcycles		
				30	Retail trade, except of motor vehicles and motorcycles		
				31	Land transport and transport via pipelines		
				32	Water transport		
				33	Air transport		
				34	Warehousing and support activities for transportation		
				35	Postal and courier activities		
				36	Accommodation and food service activities		
				37	Publishing activities		

County level data (WZ08)		Shipment data (NST07)		WIOD		This paper
			38	Motion picture, video and television programme production, sound recording and music publishing activities; programming and broadcasting activities		
			39	Telecommunications		
			40	Computer programming, consultancy and related activities; information service activities		
K-N	Financial, insurance, business, real estate activities		41	Financial service activities, except insurance and pension funding	16	Financial, Insurance, Business
			42	Insurance, reinsurance and pension funding, except compulsory social security		
			43	Activities auxiliary to financial services and insurance activities		
			44	Real estate activities		
			45	Legal and accounting activities; activities of head offices; management consultancy activities		
			46	Architectural and engineering activities; technical testing and analysis		
			47	Scientific research and development		
			48	Advertising and market research		
			49	Other professional, scientific and technical activities; veterinary activities		
			50	Administrative and support service activities		

County level data (WZ08)		Shipment data (NST07)		WIOD		This paper
O-T	Public and other services, education, health, private households		51	Public administration and defense; compulsory social security	17	Government, Education, Health
			52	Education		
			53	Human health and social work activities		
			54	Other service activities		
			55	Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use		
			56	Activities of extraterritorial organizations and bodies		

Table C.1: Industry correspondence