PBL Netherlands Environmental
Assessment Agency

Integrated Regional Europe:
European Regional Trade Flows in 2000

Corresponding author
mark.thissen@pbl.nl

Author(s)
Mark Thissen (PBL)
Dario Diodato (PBL)
Frank G. van Oort (Utrecht University)

PBL Netherlands Environmental Assessment Agency is the national institute for strategic policy analyses in the fields of the environment, nature and spatial planning. We contribute to improving the quality of political and administrative decision-making, by conducting outlook studies, analyses and evaluations in which an integrated approach is considered paramount. Policy relevance is the prime concern in all our studies. We conduct solicited and unsolicited research that is both independent and always scientifically sound.

Contents

Abstract

International regional competitiveness recently was developed into smart, sustainable and inclusive growth objectives of the Europe 2020 policy programme, as envisaged for the cohesion policy reform after 2013 (EC, 2004). Currently, place-based development policies are proposed for future cohesion policy (Barca, 2009) and have taken the shape of smart specialisation strategies. These strategies are based on a systems way of thinking about innovation and growth (McCann and Ortega-Argilés, 2011) and place a large emphasis on regional network data and analysis. However, crucial economic data on trade between regions are noticeably missing from European regional databases.

This paper, therefore, proposes a new methodology to determine interregional trade and present a unique data set on trade between 256 European NUTS2 regions, for the year 2000. The methodology stays close to a parameter-free approach as proposed by Simini et al. (2012), and deviates therefore from earlier methods based on the gravity model that suffer from analytical inconsistencies. Unlike a gravity model estimation, our methodology stays as close as possible to observed data without imposing any geographical trade patterns. The resulting data can therefore be used as such in other research.

According to our findings, most trade takes place within one and the same region, although there are large sectoral differences. The data showed that globalisation in trade in the year 2000 was still limited with only a very small percentage of goods and services being traded with countries outside Europe. The European regional economy, therefore, was found to be strongly interconnected, which is part of the explanation of the strong regional spillovers of the present economic crisis in Europe.

1. Introduction

During the last two decades, regional economic phenomena received increasing attention from researchers and policymakers. If, before, countries were the standard unit of analysis, it is now common for economic science to focus on a more detailed geographical level, namely that of cities, urban agglomerations or regions. This tendency is clearly identifiable in the flow of publications on the new economic geography (NEG) that started with the seminal paper by Krugman (1991). More traditional economic geography (a branch of human geography) with contributions such as by Porter (1990), Saxenian (1994) and Florida (2002), also emphasises a detailed geographical level with an increasing trend towards rigorous regional analysis.

International regional competitiveness has been propagated by European policymakers since the introduction of the Lisbon Agenda in 2000, and included into the current (smart, sustainable and inclusive) growth objectives of the Europe 2020 policy programme that are also central in the envisaged cohesion policy reform after 2013. The European Commission (2004, viii) envisages a common future for competitiveness and cohesion policies, stating that 'strengthening regional competitiveness throughout the Union and helping people fulfil their capabilities will boost the growth potential of the EU economy as a whole to the common benefit of all'. Currently, place-based development policies are proposed for future cohesion policy. The proposed place-based development strategy (Barca, 2009), recently, was expanded to include a smart specialisation concept based on a systems way of thinking about innovation and growth. It emphasises issues of economic potential, allowing for the complexity of regional systems (McCann and Ortega-Argilés, 2011). A 'smart' specialisation strategy is one that is based on the smartest data available for the specific regional context and is used in the smartest possible way of policy-making, given the challenges to be faced. Smart policy-making explicitly builds on these data and allows the most appropriate choices to be made given the particular regional challenges.

The result of these trends is a growing demand for data at a more detailed geographical level, and statistical offices have responded to this request. In Europe, Eurostat publishes key regional statistics for the European Union (EU) and other important non-EU countries. However, crucial economic data on trade between regions are notably missing from European regional databases. There is no data set that describes complete interregional trade flows, divided into product categories. Some regional trade flows, such as those for agriculture, may be available for a specific region but there is no comprehensive matrix of all trade between European regions.

The trade data presented here offers the possibility to develop a place-based smart specialisation strategy that is underpinned by a regional network of trade between European regions. The aim of the presented data and methodology, therefore, is to fill a vacuum and provide researchers in regional sciences and economics with an innovative data set with numerous potential applications. We inferred the most likely network of trade flows among regions in Europe, using all the available information. Our efforts have resulted in a trade matrix that includes interregional flows for 59 product categories including services (European Statistical Classification of Products by Activity (CPA), 2002), between 256 NUTS2 regions, belonging to 25 European countries, for the year 2000.

Although for many typical products global trade may be of primal importance, our data shows that the European regional economy in the year 2000 was strongly interrelated and had only limited links with the rest of the world. Most of what was produced was consumed in the same or a nearby region. The spreading of the present European economic crisis over Europe and its persistence, therefore, could partly be explained by the strong economic interdependence of European regions.

In the construction of the data set, no specific model was used to estimate trade patterns. This is different from earlier partial attempts that based their estimates on the gravity model, based on many regionally specific parameters and resulting in model outcomes and not in data. It is impossible to use model outcomes in unrelated empirical studies, as information is contained only in the parameters of the estimated gravity model itself and the related data set is the result of this information. For instance, research on the validity of the gravity model based on data generated by the gravity model will, by definition, confirm the validity of the model. The widespread use of the model is despite its well-known notable limitations and analytical inconsistencies (see Simini et al.

for an overview). Our methodology was developed to fit the information available, without pre-imposing any geographical structure on the data. With the exception of the cross-hauling estimation described in Section 2.2.3, the method is 'parameter-free' and therefore in line with more universal methodologies as proposed by Simini et al. (2012). Simini et al. argue that these parameter-free methodologies sometimes outperform sophisticated parameter estimations, stay closer to the actual data and are far less data demanding. More specifically, the unique new data set documented in this paper, was constructed using four main steps as described below.

First, we built a consistent international trade matrix of flows in goods and services between all the distinguished countries and, divided into several blocks, with the rest of the world. International trade in goods was based on the data collected by Feenstra et al. (2005). Data on services were based on Eurostat trade statistics taken from the balance of payments (Eurostat, 2009). These two sources were the best available for international trade. However, they were not always consistent with the national Supply and Use Tables, which provided information on total imports and total exports, per product. Since there are reasons to believe that national accounts are more accurate, trade flows for our matrix were constrained to these totals. The final estimated international trade flows between countries are therefore consistent with the national accounts and are as near as possible to the trade flows in Feenstra et al. (2005) and Eurostat (2009). In this first step, corrections were taken into account for inconsistencies with CIF (cost, insurance and freight) and FOB (free on board), direct purchases abroad and re-exports.

Second, for each of the 256 regions in our study, we assembled regional Supply and Use Tables by disaggregating national accounts. This operation, which in Input-Output literature is known as regionalisation of national tables, was carried out by combining non-survey techniques with Cambridge Econometrics (2008) data on regional production, investment and consumption. Particular attention was dedicated to solving the problem of cross-hauling – simultaneous export and import of the same type of goods. The outcome of the regionalisation consisted of regional figures on imports and exports, per product. These figures added up to the national accounts figures on Exports and Imports.

Third, freight transport data from the Dutch Ministry of Infrastructure and the Environment (2007) and business flight data from MIDT (2010) were used to estimate regional trade flows. The freight transport data was used to estimate the trade in goods, while the business flight data was used to estimate the trade in services. The presence of transport hubs within transport data was identified and accounted for. In the third step, we produced two distinct estimates for each of the 59 CPA goods and services. The first estimate was obtained by distributing regional export figures according to the outward transport pattern. The second was obtained using the inverse procedure: regional imports were distributed following the inward transport pattern.

In the fourth and final step, all the information available was combined to arrive at a final estimation of trade flows between 256 European NUTS2 regions, for the aggregated 59 products and services. The information gathered in the preceding steps is present in two different estimates of the interregional trade flows between European regions: 1) the regional total exports and total imports (according to CPA) consistent with the national accounts; 2) international trade flows consistent with the national accounts. The presented final trade matrix minimises the distance between these two estimates, given the international trade flows and constrained to regional total imports and exports.

This paper is divided into two main sections. The first section describes the data set and the second carefully explains each of the four steps of our methodology.

2. Methodology

The central principle in our methodology inferring European regional trade flows from different sources of information is increasing data reliability by imposing consistency with available statistics. Regional trade flows need to be consistent with statistics on production and consumption per region, which, in turn, must be in line with national data on production and consumption. These regional flows must also be consistent with international trade statistics, on a national level. Furthermore, international trade statistics also must be consistent with national data on production, consumption, imports and exports. Finally, trade statistics should be mutually consistent. That is, exports from a region or country A to a region or country B should equal the opposite flow of imports received by region or country B. All these consistency checks provide additional information and therefore add to the quality of the estimated trade flows.

Data were collected from various sources. International trade between countries was taken from Feenstra et al. (2005) for goods and from Eurostat (2009) for services. Information on national production, consumption, imports and exports was obtained from the Supply and Use Tables in the national accounts (Eurostat, 2009b). These Supply and Use Tables were not available for 2 (Latvia and Greece) of the 25 considered countries. The Supply and Use Tables from the year 1998, therefore, were updated using the commonly applied RAS method (or bi-proportional updating method). The necessary row and column sums for the Supply and Use Tables of Latvia and Greece were taken from Eurostat. With respect to regional data, statistics on regional production, investment and consumption were taken from Cambridge Econometrics (2008). There may have been large discrepancies between the different sources. For instance, the sum of regional statistics did not always match the national totals, and neither did the sums of exports per country of destination (registered as imports) match the countries' export totals.

The national accounts were central in our analysis because they were the most reliable statistics of all sources available to us. A large amount of information is used in the construction of national accounts, consisting of the combination of data from many different sources. We therefore controlled our estimated trade matrix to be consistent with the national accounts using constraint (non-)linear optimisation. First data on international trade was made consistent with figures on exports and imports from the national accounts (Section 2.1). Section 2.2 describes a regionalisation of Supply and Use Tables, using regional data from Cambridge Econometrics (2008). In this way, we made sure that we had reliable data on regional imports and exports that would be consistent with the national accounts. Section 2.3 presents the exports distribution per country of destination and imports per country of origin. We used data on freight between European regions from the Ministry of Infrastructure and the Environment (2007) as well as first-class and business-class flight data from MIDT (2010) to determine origin and destination of the different trade flows. In this step, we obtained two separate estimates of regional trade flows; one from the export point of view and one from the import point of view. These two estimates neither were consistent with each other, nor with the other accounts. In the last step, explained in Section 2.5, we produced the final estimate of regional trade flows. This estimate was as close as possible to the previous two estimates (in Section 2.4), but was controlled for consistency with regional accounts, national accounts and international trade flows.

2.1 Consistent national trade

This section describes how the exports and imports from the national accounts statistics were divided over country of destination and country of origin using international trade statistics on goods (Feenstra et al., 2005) and services (Eurostat, 2009). However, before comparing imports and exports, valuation differences have to be addressed and direct purchases abroad have to be taken into account. The valuation differences occurred because Supply and Use Tables report on exports valued 'free on board' (FOB, i.e. the value before transportation), while imports are valued including 'cost, insurance and freight' (CIF – inclusive of trade costs). We chose to value both exports and imports FOB in order to enable price comparison. Statistics on international trade in goods per country also included the direct purchases by these countries' citizens abroad and the direct purchases of foreign residents within the countries concerned. Therefore, these direct

purchases were included in the exports and imports of goods and services. Here, again, corrections had to be made to enable comparison between the different data sources. Section 2.1.2 provides a detailed description of these corrections.

These trade data also were made internally consistent. Thus, the statistics on export trade flows equal the same flow in the opposite direction registered as imports. With these corrections implemented, trade statistics from Feenstra et al. (2005) and Eurostat (2009) could be made consistent with national Supply and Use Tables. First, the different product classifications needed to be matched. The maximum level of detail was the Nace 2-digit product classification, because we used the national Supply and Use Tables as the starting point for our analyses. Trade in goods (Feenstra et al., 2005) was based on the 4-digit SITC and, therefore, required conversion and aggregation. Trade in services, in contrast, was divided in only four macro-categories and required disaggregation. Subsequently, we were left with two prior estimates on international trade for each of the 59 product categories. We used linear programming to achieve a final estimate of international trade. We wanted this final estimate to be as near as possible to the two prior estimates, but also to be consistent with the (corrected) national accounts. The whole procedure is illustrated in Section 2.1.2.

Section 2.1 ends with a further correction. Neither total trade from national accounts nor the trade patterns took re-exports into account. For goods exported from country A to country B via country C, in many cases, the flow would be recorded twice (first A to C, then C to B). This generates two inconvenient results: one, it inflates the value of exported goods, and, two, it misreports the true countries of origin and destination of the traded goods. Therefore, we used an iterative procure to try and solve this problem (see Section 2.1.3).

2.1.1 Correction for CIF/FOB valuations and direct purchases abroad

This section describes two corrections made to the trade figures in the national account Supply and Use Tables, in order to enable comparison between export and import figures. The first important correction concerned the valuation of exports with respect to trade costs, while the second dealt with direct purchases abroad.

The problem of the valuation of trade costs arose from the fact that they are generally recorded according to two methods. In the system of national accounts, the export of goods usually is recorded, by the customs office, when goods leave the producing country. Imports, however, are recorded when they enter the country of destination (by its customs office). The former way of registering exports is called 'free on board' (FOB), and the latter is called 'cost, insurance and freight' (CIF).

When two trading countries have a different method of recording (FOB or CIF), obviously, the value of the same trade flow would also differ per country. The international guidelines for the construction of Supply and Use Tables (the UN System of National Accounts SNA1993 and the European System of Accounts ESA95) recommend FOB reporting of the value of trade, for both imports and exports. However, for imports it would be much easier to use a CIF valuation, since imports are observed by national customs. For this reason, the ESA95 system, the methodology used in Europe for national accounts (Eurostat, 2008), allows some flexibility in this respect, and prescribes FOB valuation only for total imports. Therefore, the Supply Tables in many countries in Europe report on imports in CIF per product, with an additional row presenting CIF-to-FOB corrections such that total imports are valued in FOB. To express imports in FOB at product level, the same correction factor is then applied for total imports to the various product categories.

The most accurate information available on how to apply the correction factor to the products is in the column of trade and transport margins in the Supply Tables of the individual country. This column gives an indication of the trade and transport costs according to the various product categories. These data would refer to domestic transportation, but could be used as a proxy for imported goods. We also found that the cost component for transportation from country of origin (the exporter) to country of destination (the importer) incurred by foreign companies was not included and, hence, not accounted for in the correction factor. Transport statistics on the Netherlands indicate that 35% of domestic transportation is executed by foreign companies. We

8

used this percentage for all countries and increased transport margins by an additional 35/65 = 0.53%.

The following equations describe the adjustment. $TR_{c,g,d}$ is the total transport costs for country c, in product g. Supply and Use Tables of many European countries, for the year 2000, distinguished between products traded with EU15 partners and those traded with the rest of the world (ROW). We included the index d, area of destination, to retain this information. $\tau_{c,g,d}$ is defined as the share of transport costs according to product

$$\tau_{c,gnr,d} = \frac{TR_{c,gnr,d}}{\sum_{gnr',d'} TR_{c,gnr',d'}}$$

(1)

where gnr stands for all products except trade and transport services gr. CF_c is the correction factor for the CIF/FOB adjustment, κ is the ratio between foreign and domestic transportation ($\approx 53\%$) and $I_{c,g,d}$ refers to imports. The correction on imports of products other than transportation services is

$$I_{c,gnr,d} = I_{c,gnr,d} - \tau_{c,gnr,d}(1+\kappa)CF_c$$

(2)

We derived imports valued at FOB using (0) to apply the correction factors to the various products. However the accounts still needed to be corrected for transport services. The share of transport service, taken from the different services categories gr (e.g. air transport, road transport) equals

$$\tau_{c,gr,d} = \frac{TR_{c,gr,d}}{\sum_{gr',d'} TR_{c,gr',d'}}$$

(3)

Imported transport services were increased by the expected contribution by foreign companies to the transportation of imported goods, with respect to the correction factor. In fact, some of the value of CF, the correction factor that was subtracted from imports, actually represents imported transport services. The adjustment was carried out as follows:

$$I_{c,gr,d} = I_{c,gr,d} + \tau_{c,gr,d}(\kappa)CF_c$$

(4)

The remaining component of CF subsequently was evaluated as exports of transport services supplied by domestic companies on national territory. For this reason, we applied the last correction and removed this value of transport services from exports X.

$$X_{c,gr,d} = X_{c,gr,d} - \tau_{c,gr,d}CF_c$$

(5)

The second issue addressed in this section is the one of direct purchases abroad and direct domestic purchases by non-residents, which for some countries was a substantial part of their exports and imports. Similar to the CIF/FOB valuation, Supply and Use Tables also included correction rows for such purchases. These correction rows were adjusted for total exports and total imports, but did not provide information on the types of products or services purchased. Therefore, the values in these rows had to be distributed over the various product categories.

Most of the direct purchases abroad could likely be attributed to tourism. Thus, we applied the correction factor to those services that were most likely to have been consumed by tourism, namely hotels and restaurants, recreational, cultural and sporting services. For most countries, the adjustment rows were distributed using the share of final demand in those service categories. Nevertheless, in Hungary and Luxemburg, purchases by non-residents were also distributed over

their shares in food and real estate service. From the Supply Tables of these countries for 2000, it is clear that these two product categories are the only ones with a large enough production to cover the total amount of the direct purchase correction.

Dp_c represents the direct purchases abroad by residents of country c and Pd_c the domestic purchases in country c by non-residents. As previously in the text, X and I represent exports and imports and g different products. Target products, those services that the direct purchases will be distributed to, are indicated by tg. Lastly, γ_{tg}, is the share of these services with respect to household consumption $HC_{c,tg}$, and $\eta_{c,tg,d}$ represents the share of imports.

$$\gamma_{c,tg} = \frac{HC_{c,tg}}{\sum_{tg'} HC_{c,tg'}} \quad and \quad \eta_{c,tg,d} = \frac{I_{c,tg,d}}{\sum_{tg'} I_{c,tg',d}}$$

(6)

Once again, imports were divided according to their country of origin, either from the EU15 or the rest of the world (ROW). We applied the following adjustments

$$X_{c,tg,d} = X_{c,tg,d} + \gamma_{c,g} \frac{X_{c,tg,d}}{X_{c,tg,eu15} + X_{c,tg,row}} Pd_c$$

(7)

and

$$I_{c,tg,d} = I_{c,tg,d} + \eta_{c,tg,d} Dp_{c,d}$$

(8)

For countries whose tables did not distinguish between EU15 and ROW trade destinations, we simply used the total per country of destination to distribute the direct purchases abroad.

2.1.2 Estimation of consistent international trade

The adjustments described in the previous section resulted in data on exports and imports in 59 Nace 1.1 categories that were valued using the same prices. We determined the origins and destinations of trade flows and made them internally consistent using these comparable figures for exports and imports. We started with creating two estimates of trade flows: the 'priors'. One prior was taken from the export point of view, the other from the import point of view. Subsequently, we searched for a final estimate that 1) would stay nearest to the two priors; 2) met the requirement of the export of a product from country A to country B matching the imports to country B from country A; and 3) was consistent with total exports and imports as reported in the (corrected) national accounts.

The priors for goods and services were obtained from different sources. The priors of the trade in goods were obtained using Feenstra et al. (2005) data on the year 2000. First, this data set was converted to the product classification used in this research, which implies an aggregation from 4-digit SITC classification to a 2-digit Nace1.1. The concordance was achieved by following the tables of the Eurostat RAMON website[1]. Then, Feenstra et al. (2005) data were used to create shares of exports per country of destination and shares of imports per country of origin. By multiplying the former shares by total exports and the latter by total imports, we obtained two priors for the trade in goods.

The origin and destination shares on a detailed product level were more difficult to obtain with respect to the trade in services. For the year 2000, only four broad categories of services are available: transportation, travel, other business services, and other services (Eurostat, 2009). Moreover, the data on the year 2000 has many missing values. Therefore, we pooled data from 2000 to 2004 to obtain a full matrix of the trade in services. This matrix was subsequently used to

---

[1] **h**ttp://ec.europa.eu/eurostat/ramon/

calculate the share of exports and imports per country of destination and origin, respectively. To account for the difference in the classification detail of services, we distributed imports and exports of 2-digit Nace1.1 services, by using the share of the Eurostat' macro-sector to which the 2-digit services belong.

This distribution of the export and import of goods and services over destinations and origins resulted in priors for both trade patterns. The export priors $X_{i,g,j}^{prior}$ from country i, to country j, in good or service g and the import priors $I_{i,g,j}^{prior}$ to country i, from country j. These two priors were the starting point for the final estimate of the trade matrix $T_{i,g,j}$. The values of T were found by use of constrained minimisation. We minimised the absolute value of the relative distance between T and the two priors. The minimisation was constrained to be consistent with total national export and import values, and totals for the EU15 were taken from the national accounts. We awarded more weight to the error on imports, since – following the literature on constructing consistent international trade statistics (Feenstra, 2005; Oosterhaven et al., 2008; Bouwmeester and Oosterhaven, 2009) – import statistics are more reliable because they are used for tariff and registration purposes. In mathematics, the optimisation is written as follows:

$$Min\ Z = \left| \frac{X_{i,g,j}^{prior} - T_{i,g,j}}{\frac{1}{4}X_{i,g,j}^{prior} + \frac{3}{4}I_{j,g,i}^{prior}} \right| + 3 \left| \frac{I_{i,g,j}^{prior} - T_{j,g,i}}{\frac{1}{4}X_{j,g,i}^{prior} + \frac{3}{4}I_{i,g,j}^{prior}} \right|$$

$$+ \left| \frac{X_{i,g,eu15}^{prior} - T_{i,g,eu15}}{X_{i,g,eu15}^{prior}} \right| + 3 \left| \frac{I_{i,g,eu15}^{prior} - T_{eu15,g,i}}{I_{i,g,eu15}^{prior}} \right|$$

$$s.t.$$

$$1)\quad T_{i,g,eu15} = \sum_{j=eu15} T_{i,g,j}$$

$$2)\quad T_{eu15,g,i} = \sum_{j=eu15} T_{j,g,i}$$

$$3)\quad X_{i,g,total} = \sum_{j=} T_{i,g,j}$$

$$4)\quad I_{j,g,total} = \sum_{i=} T_{i,g,j} \tag{9}$$

2.1.3 Correction for re-exports

Although the matrix of international trade was consistent after optimisation, we further improved the quality of our data set by removing re-exports. The existence and size of the re-export problem is illustrated by the phenomenon of the value of export numbers for certain products being larger than the value of their total production levels. In other words, according to official statistics, countries appear to have exported more than they produced. According to the definition in the main international guidelines on the construction of national accounts (SNA, 2008): *Re-exports are foreign goods (goods produced in other economies and previously imported with a change of economic ownership) that are exported with no substantial transformation from the state in which they were previously imported*. If a good goes from country A to B, making an intermediate transit in country C, the international guidelines recommend to register a double trade flow (import from A to C, and export from C to B) if a resident of country C acquires ownership of this good. Under the current system it is possible that countries are registered to export a certain good without actually producing it. This problem is in fact acknowledged in SNA (2008): *because re-exported goods are not produced in the economy concerned, they have less connection to that economy than other exports.* In our view, re-exports are problematic in at least two ways: first, because national statistics over-report the total amount of trade, and, second, because they misreport the origin–destination pattern of products.

The data sources we used are based on the national accounts and, therefore, include re-exports. In addition, the trade patterns mostly also include re-exports. Feenstra et al. (2005) dedicated a large

amount of work to dealing with the very high re-export figures of Hong Kong, but they did not correct this phenomenon for other countries. It therefore was still necessary to correct national trade data for re-exports within Europe. Fortunately, information on re-exports was available in the Import tables belonging to the national accounts. These tables were obtained from the statistical offices of most of the 25 countries studied.

Below, an outline is presented of the method we used for removing re-exports from trade matrix T. This technique can be applied independently to different product categories. For this reason and to simplify the text, we left out all references to goods and services categories from the respective indices in the equations. First, we defined the export destination shares e_{ij}^c from country i to country j,

$$e_{ij}^c = \frac{T_{ij}}{\sum_j T_{ij}}, j \neq c$$

(10)

and the imports shares m_{ij} that country i received from country j, so that

$$m_{ij} = \frac{T_{ij}}{\sum_i T_{ij}}$$

(11)

The total re-exports (per product) RE_i, for country i, were taken from the Input–Output (import) tables. We used the minimum non-zero product re-export share observed among other countries for those countries where no information on re-exports was available. We estimated the pattern of re-exports R_{ij}^c, from i, to country j, via country c as

$$R_{ij}^c = RE_c m_{ci} e_{cj}^i$$

(12)

When re-exports were identified in the intermediate country, they needed to be redistributed to a different destination. The country of origin was excluded in equation (0) because it does not make sense to redistribute the trade flow back to the country of origin. As mentioned before, re-exported goods generally do not receive substantial treatment in the intermediate country (they are repackaged, at the most) and therefore could be redistributed in this way.

Once the 'true' pattern was identified, we needed to adjust the trade matrix accordingly. The trade flow between country of origin i and the intermediate country c needed to be removed. The same is true for a flow of the same size from the intermediate country c to the destination country j. This trade flow (which was removed twice, from i to c and from c to j) then was added in the form of exports from country of origin i to country of destination j. In mathematics:

$$T_{ij} = T_{ij} + \sum_c R_{ij}^c - \sum_c R_{ic}^j - \sum_c R_{cj}^i$$

(13)

The methodology had to take the following three issues into account to be successful:

1) The method produces results that are independent from the order of countries to which it is applied. This is certainly an advantage, but it comes with the disadvantage of adjusted export flows incidentally becoming smaller than zero. Although an alternative methodology is available (where the adjustments are done for all country sequentially), this would present the opposite problem: it would not create negative flows, but be dependent on the order of the countries to which the methodology would be applied. For our study, we preferred the first methodology because the outcomes were easily reproducible and not affected by the random choice of the order of countries to which the method was applied. The (small) negative export flows were corrected by changing them into positive imports.

2) After the procedure, some countries still had export flows larger than production levels. We corrected this by redefining the excess of exports over production as re-exports (*RE*) and then reapplying the procedure.
3) Issues 1 and 2 may interact with each other. The correction in 1) may cause re-exports to become larger than production and the correction in 2) may cause some export flows to become negative. In such cases, the procedure can be reapplied as many times as needed until both issues are solved.

2.2 Regionalisation of Supply and Use Tables

In the next step, described in this section, the obtained national Supply and Use Tables were regionalised. The main aim of this second step was to obtain, for every region, total exports and total imports, per product. This would be achieved by constructing consistent regional Supply and Use Tables. These tables were organised according to the examples in Figures 1 and 2.

Figure 1. Use Table

	input industry 1	input industry 2	...	input industry s	household demand	non-profit demand	government demand	fixed capital formation	changes in inventories	changes in valuables	exports	total
product 1												
product 2												
..												
product p												
total												

Figure 2. Supply Table

	input industry 1	input industry 2	...	input industry s	trade and transport margins	+ taxes –subsidies on production	imports	total
product 1								
product 2								
..								
product p								
total								

These regionalised Supply and Use Tables are conform national Supply and Use Tables. Thus, total use in the region would match total supply, which implies that the row of totals of the Use table equals the row of totals of the Supply table. This equality is the regional version of the more familiar macroeconomic condition that production is equal to consumption plus exports and minus imports. The totals columns of the regional Supply and Use Tables are also equal because total output of every (regional) industry equals this industry's total input and value added. The regional Supply and Use Tables give boundaries to the total regional exports and imports and are therefore crucial to infer regional export and import patterns.

This section presents the available techniques that were used to build the regional Supply and Use Tables that we needed to infer regional imports and exports, according to product (CPA). More precisely, we employed the approach known as Commodity Balance (CB) method, first suggested by Isard (1953). National Supply and Use Tables were crossed with regional data from Cambridge Econometrics (2008) on total consumer demand, sectoral added value and investment. These data provided relevant information on regional totals, without distinguishing between the various products. All this information was used to obtain reliable column totals of the regional tables. To disaggregate these totals into different products (the rows), the national Supply and Use Tables were used. The structure of the national Supply and Use Tables was assumed to give a good approximation for the regional tables. More formally, consumers were assumed to have homogenous preferences throughout the country concerned, homogenous government spending was assumed over the regions, and industries were assumed to use the same technology irrespective of their location within the country concerned. Under these assumptions, regional household demand (*HHD*) per product (CPA, marked with index *g*) was obtained according to the following equation:

$$HHD_g^r = \frac{HHD^r}{\sum_{k=1}^{R} HHD^k} HHD_g^N$$

(14)

Where, *N* stands for the country and *R* for total number of regions within that country. In a comparable way regional production was determined.

Since no information was available on regional total exports and imports, we used one more assumption to obtain these key variables. We assumed exports to have originated from the producing regions and imports to have been transported to using regions. Exports and imports, therefore, were divided according to production and consumption shares. To exemplify, exports would be:

$$X_g^r = \frac{Y_g^r}{\sum_{k=1}^{R} Y_g^k} X_g^N$$

(15)

where *X* is exports and *Y* is production per CPA region. This would be a first estimate of the international trade of regions, based on fixed shares. This first estimate subsequently was improved on (see Section 2.4). The Commodity Balance (CB) approach has the advantage that it automatically guarantees national consistency for every item in the Supply and Use Tables; thus, the summed government demand for product *g* per region results in the national government demand for that product.

Following this operation, the regional macroeconomic condition of production equalling consumption plus exports and minus imports would no longer hold. This is not surprising, because an important part of the puzzle is missing, namely the domestic (intranational) trade.[2] The data obtained on regional exports and imports refer to international trade, in which domestic interregional trade was not included. Information on this type of regional trade was needed to complete the regional macroeconomic condition.

Data on intranational trade was needed to construct fully consistent and reliable regional Supply and Use Tables. We needed information on cross-hauling in order to determine intranational trade. Cross-hauling means the simultaneous trade in products of the same product category (CPA) among two regions of the same country. Only recently the existence and importance of cross-hauling was recognised with respect to the regionalisation of Supply and Use Tables (see Kronenberg, 2009). Unfortunately, there were no procedures readily available to determine the cross-hauling for a consistent set of regions within one country. Therefore, we used a new methodology, which is explained further in this section. Section 2.3.1 explains the organisation of the Supply and Use Tables with respect to production and final demand. In Section 2.3.2 explains the procedure used to construct regional exports and imports per product. Finally, Section 2.3.3 presents the methodology we developed to solve the issue of cross-hauling.

2.2.1 Production and consumption

In order to regionalise the national Supply and Use Tables, we first divided production and final demand over the regions. The regional demand in the Use Tables was divided into intermediate demand (input by industry), household demand, government demand, demand from non-profit organisations, gross fixed capital formation, changes in inventories, and changes in valuables. With

---

[2] In this report we have chosen to use intranational instead of domestic trade. The reason is that domestic trade is often confused in the literature with domestic sales. Accordingly, regional domestic trade is often confused with all the trade within the own region and not between the regions. We hope that our terminology that emphasises the national borders will help our readers in understanding the methodology.

respect to the Supply Tables, regional output had to be determined according to industry, trade and transport margins and net taxes.

First, the data on production was determined per industry as well as the intermediate demand. At the European NUTS2 (regional) level, although there was no information available on output, we did obtain data on value added (*VA*) for 15 economic sectors, made available by Cambridge Econometrics (2008). Table 1 presents the classification for these 15 economic sectors.

Table 1. Industry classification in 15 sectors

S1	Agriculture
S2	Mining, quarrying and energy supply
S3	Food, beverages and tobacco
S4	Textiles and leather
S5	Coke, refined petroleum, nuclear fuel and chemicals etc.
S6	Electrical and optical equipment
S7	Transport equipment
S8	Other manufacturing
S9	Construction
S10	Distribution
S11	Hotels and restaurants
S12	Transport, storage and communications
S13	Financial intermediation
S14	Real estate, renting and business activities
S15	Non-market services

Maintaining the index *g* for products and introducing the index *s* for sectors, regional output and input were constructing using information on value added in the following way:

$$input_{gs}^{r} = \frac{VA_s^r}{\sum\limits_{k=1}^{R} VA_s^k} input_{gs}^{N}$$

(16)

$$output_{gs}^{r} = \frac{VA_s^r}{\sum\limits_{k=1}^{R} VA_s^k} output_{gs}^{N}$$

(17)

This approach is based on two main assumptions. First, the technology through which a sector transforms inputs into outputs is assumed to be homogeneous throughout the country considered. The value added represents a good proxy for the magnitude of the industries inputs and outputs. It must be noted that this method maintains consistency per product, with the sum of output (or

input) per regional product equalling its national output, and per industry, with the regional shares of *VA* fixed at the original proportions.

Second, the preferences of final consumers in each region are assumed to reflect preferences nation-wide. The equation used for this part of the regionalisation is that presented in the introduction of this section and is as follows.

$$HHD_g^r = \frac{HHD^r}{\sum\limits_{k=1}^{R} HHD^k} HHD_g^N$$

$(14)^3$

Regional household demand was aggregated with the demand from non-profit organisations, because there was no regional information available on non-profit organisations. As their demand, on average, would be less than 2.5% of household demand, the latter makes up the main component. Therefore, this aggregated category subsequently is referred to as household demand (*HHD*).

In order to regionalise the total government demand per region, we used data on value added non-market services (sector 15). This sector includes typical government activities, such as public administration, defence, education and health. It is reasonable to assume that this sector reflects how national government budgets are allocated. The demand pattern per product is assumed not to differ between regions, and is thus:

$$GOVD_g^r = \frac{VA_{S15}^r}{\sum\limits_{k=1}^{R} VA_{S15}^k} GOVD_g^N$$

(18)

Gross capital formation is divided into three items: gross fixed capital formation, changes in inventories and changes in valuables. With respect to the first category, we proceeded as for government demand. Investments were considered per region, therefore:

$$INV_g^r = \frac{INV^r}{\sum\limits_{k=1}^{R} INV^k} INV_g^N$$

(19)

Changes in inventories and valuables required more effort. In order to maintain full consistency between the regional and national accounts, the data in these two fluctuating and unpredictable columns had to be regionalised. The following observations were important in the regionalisation of these two accounts: 1) Changes in inventories are a more common phenomenon than changes in valuables. For valuables the changes are often small or even not included in the accounts of some countries; and 2) The changes in inventories and valuables were found to follow a similar pattern for the various goods and are combined in a subtotal under the name 'changes in inventories and valuables' in the ESA95 format. Given these observations, we decided to merge the two categories in a new aggregated column under the heading of inventories. Positive and negative changes in inventories would reflect supply excess or shortages, within the time period of one particular year. We assumed the demand for a certain product to have the same degree of fluctuation on a national level; therefore, producers per region were assumed to face excess or shortages of supply proportional to their production levels. Hence, changes in inventories and valuables (*CIV*) were defined as:

$$CIV_g^r = \frac{VA^r}{\sum\limits_{k=1}^{R} VA^k} CIV_g^N$$

(20)

---

[3] This equation is a copy of that on page 14.

In a similar fashion, this last consideration could also be applied to the two remaining columns in the Supply table: trade and transport margins (*TTM*) and taxes and subsidies (*TAX*). Since their regional variation is also assumed to be proportional to production, we defined them as:

$$TTM_g^r = \frac{VA^r}{\sum_{k=1}^{R} VA^k} TTM_g^N \quad \text{and} \quad TAX_g^r = \frac{VA^r}{\sum_{k=1}^{R} VA^k} TAX_g^N$$

(21)

2.2.2 Exports and Imports

In Section 2.2.1, all variables of the Supply and Use Tables are defined, with the exception of exports and imports. As mentioned before, we did not have any information available on regional exports and imports. However, we did have accurate information on regional production and consumption, per type of product, following the regionalisation. We used this information for a proportional allocation of exports to regional production and imports to regional demand. A simple example may illustrate the rationale behind this operation. Imagine that in a country *N,* industry *s* is agglomerated in one region (region *r*). From the national supply table it is known that this industry is providing a certain mix of products *g* as output. If country *N* is an exporter of these products, then exports must be brought in from the region where they are produced – in this case,region *r*. Naturally, full agglomeration of one industry almost never happens, but this reasoning, nonetheless, can be extended to less extreme cases. It therefore may safely be assumed that the largest share of exports of a given product is produced in the region where most of these types of products are produced. A similar reasoning results in the following comparable conclusion with respect to imports. The largest share of imports of a given product is used in the region that uses the most of this product in intermediate or final use.

Hence, for every product *g*, we determined the regional production and consumption. These two figures, subsequently, were used to allocate national exports and imports to the different regions. Defining regional production and consumption is not as straightforward as it may seem. The general rule is that consumption (*D*) equals use minus exports (*X*), and production (*Y*) is supply minus imports (*I*). Hence:

$$D_g^r = Use_g^r - X_g^r = \sum_s Input_{gs}^r + HHD_g^r + GOVD_g^r + INV_g^r + CIV_g^r$$

and

(22)

$$Y_g^r = Supply_g^r - I_g^r = \sum_s Output_{gs}^r + TTM_g^r + TAX_g^r$$

However, there are some small, but important, corrections that must be taken into account. A change in inventories should be added to the production (in this period) although the demand does not take place until in the next period. Hence, negative entries under 'changes in inventories and valuables' are part of the supply and not a negative correction of the demand. In the same way, negative entries in trade and transport margins, as well as negative tax values (i.e. subsidies), are a positive demand instead of a negative supply. Once these corrections were applied, regional exports and imports were defined as follows:

$$X_g^r = \frac{Y_g^r}{\sum_{k=1}^{R} Y_g^k} X_g^N \quad \text{and} \quad I_g^r = \frac{D_g^r}{\sum_{k=1}^{R} D_g^k} I_g^N$$

(23)

These regional exports and imports are a first estimate. Section 2.3 presents an adjustment using information on interregional and international transport flows. It must be emphasised that the exports and imports in Equation (0) refer to international trade. Products that are sold outside the producing region but within the same country are also exports from the region's perspective. We refer to this type of exports as intranational exports (*IX)* to avoid possible confusion. Similarly, we

use intranational imports (*II*) and, more generally, intranational trade. Section 2.2.3 is dedicated to intranational exports and intranational imports.

2.2.3 Cross-hauling and intranational trade

Information about the national share of products consumed and produced domestically was available, as data on re-exports were cleaned from the Supply and Use Tables (see Section 2.1.3). If the national production that is supplied to the national market is called *own production* (*OY*) and the national consumption of these products is *own consumption* (*OD*), we can write:

$$OY_g^N = Y_g^N - X_g^N = D_g^N - I_g^N = OD_g^N$$

(24)

Own production and *own consumption* can be derived in two ways: either from production minus exports or from consumption minus imports. The two values, of course, are in fact the same.[4]

At the regional level, exports do not include intranational exports; therefore, the following slightly different relationship exists:

$$IY_g^r = OY_g^r + IX_g^r = Y_g^r - X_g^r \quad \text{and} \quad ID_g^r = OD_g^r + II_g^r = D_g^r - I_g^r$$

(25)

with *IY* being production for domestic use and *ID* being demand for domestic products. Thus, *IY* is the production within a region, the resulting products of which either are sold within the same region (*OY*) or in other regions of the same country (intranational exports, *IX*). *ID* is the demand within a region, which is either being met by the production within the same region (*OD*) or by the production in other regions of the same country (intranational imports, *II*). Equation (0) shows that *IY* can be derived by subtracting exports (*X*) from production (*Y*), and *ID* by substracting imports (*I*) from consumption (*D*). Although production for domestic use and demand for domestic products are known, we did not know how many of the produced goods would remain in the producing region. Hence, we were unable to separate *IY* into its components *own production* (*OY*) and *intranational exports* (*IX*). Similarly, *ID* could also not be separated into *own consumption* (*OD*) and *intranational imports* (*II*).

Although, by definition, *own production* equals *own consumption* (*OY=OD*), it is likely that intranational exports, in actual practice, deviate from intranational imports. We therefore defined the domestic trade balance (*IZ*) as:

$$IZ_g^r = ID_g^r - IY_g^r$$

(26)

which is also equal to the difference between intranational imports (*II*) and intranational exports (*IX*). A full regionalisation of Supply and Use Tables may be obtained assuming that there is no cross-hauling or trade between regions; this is described for the internal trade balance (*IZ*). However, cross-hauling is an important and sizeable empirical phenomenon that invalidates this type of regionalisation (see Kronenberg, 2009). Therefore, here, a theoretical model is presented that allows for cross-hauling. We subsequently used constraint nonlinear optimisation techniques to estimate the amount of cross-hauling in the 256 distinguished NUTS2 regions.

The Krugman (1991) model is the only theoretical international macromodel derived from microeconomic behaviour that allows for cross-hauling. This approach is to be preferred[5] because,

---

[4] This happens automatically due to the way Use and Supply tables are constructed. In fact, as can be seen from Equations (0), Y = Supply - I and D = Use - X. In combination with Equation (0), then OY = Supply - I - X and OD = Use - X - I. Since Supply = Use, OY = OD. The corrections we applied in the previous section (on negative changes in inventories, negative trade and transport margins and subsidies) do not interfere with this line of reasoning.

[5] See Kronenberg (2009) for a discussion of different approaches.

if it is applied to all regions within a country, it (1) guarantees consistency with the national accounts and (2) is rigorously derived from microeconomic theory. For a more extensive explanation of the model, see Diodato and Thissen (WP, 2012). A concise description of the approach is given below.

The core of the model was built on the assumption that consumers love variety, and the demand for a variety of goods can be described by a Dixit-Stiglitz-Krugman demand function. This demand function is a CES function (constant elasticity of substitution) including (iceberg) transport costs to deal with the demand at different locations. This function describes that consumers are aware of small differences between products within the same product category, which they perceive as imperfect substitutes. Even though producers in every region use the same technology, the model differs from that of perfect competition, because consumers identify every variety as a unique product and, consequently, every producer has a certain degree of monopolistic power. The basic Dixit-Stiglitz monopolistic competition (Dixit and Stiglitz, 1977) has been extended by Krugman (1991) with iceberg-type transport costs (Samuelson, 1954) to build a spatial two-region model. We used this basic theoretical model to derive cross-hauling.

It is important to emphasise that we used a two-region model because we only wanted to determine the amount of cross-hauling per region. We did *not* want to use the model to determine trade patterns. As explained before, using the model to estimate trade patters would make the data set useless for any further data analysis. The use of the Krugman model to determine only the amount of cross-hauling, therefore, was crucial in the presented approach.

In the two-region model, there was always one region under investigation (the focus region *r*) and the rest of the country *h* (the second region). This second region, therefore, differs for every focus region. Producers of product *g* were distributed over both regions. Consumers not only pay interregional transport costs if products are imported from elsewhere in the country, but transport costs are also payable for products produced locally.[6] Given the theoretical model we therefore have intranational exports from region *r* to the rest of the country *h* equal to

$$IX^r = \frac{(P^r T_{rh})^{1-\sigma}}{n^h (P^h T_{hh})^{1-\sigma} + n^r (P^r T_{rh})^{1-\sigma}} ID^h n^r$$

(27)

And intranational imports into region *r* from the rest of the country *h* as:

$$II^r = \frac{(P^h T_{hr})^{1-\sigma}}{n^r (P^r T_{rr})^{1-\sigma} + n^h (P^h T_{hr})^{1-\sigma}} ID^r n^h$$

(28)

Where *P* represents price, *n* the number of firms, *σ* the elasticity of substitution and *T* the iceberg transport costs. The index *g* (product category) was omitted to simplify the equations.

With homogeneous technology this is:

$$n^r = \frac{IY^r}{\alpha P^r}$$

(29)

With α equalling optimal output per variety, which in the model is the same for every producer, irrespective of location. Substitution of the number of varieties (0) in both Equations (0) and (0) led to a simplification of the intranational exports and imports. Combining Equations (0) and (0) resulted in two expressions that, after rearranging, resulted in expressions for intranational exports and intranational imports expressed as a function of value and not of price or quantity.

---

[6] This is known in the literature as the assumption of mill pricing.

$$IX^r = f(T, \sigma, ID^r, ID^h, IZ^r) \equiv II^h$$

and (30)

$$II^r = f(T, \sigma, ID^r, ID^h, IZ^h) \equiv IX^h$$

Intranational imports (*II*) and exports (*IX*) are now functions of only two unknown quantities: transport costs and the elasticity of substitution[7]. Intranational trade can only be calculated if these two parameters are known. Since they were not, the problem had to be approached in a different way. The elasticity of substitution σ was taken from the literature, where it is commonly assumed to equal 1.5 (McKitrick, 1998). Transport costs were estimated (together with intranational trade) using non-linear programming.

Given the theoretical model described by Equation (0), we determined the optimal value for transportation costs t_{rh} between regions *r* and *h* in the non-linear optimisation, such that the national transport costs per product would be as close as possible to the national accounts' data on trade and transport margins (TTM). In the optimisation we assumed a common transport cost function that is declining in transported distance according to a logarithmic relationship. The cross-hauling, or the total intranational trade per region, was endogenously determined in the procedure.

Unfortunately, the data on trade and transport per product were limited to the national level, and the data were not always of a very high quality. We therefore expanded the methodology by including a second objective in the non-linear optimisation procedure based on cross-hauling, derived from freight data from the Dutch Ministry of Infrastructure and the Environment (2007). The data on cross-hauling derived from freight data is described by the share *ChS* of the total of the goods that are traded within regional borders divided by the goods sold to other regions. The data on services needed correction, since cross-hauling for this category is expected to be less important. The share for services is calculated as follows:

$$ChS_r^{services} = ChS_r^{goods} \frac{\displaystyle\sum_{g \in services} Supply_g^N \Big/ \sum_{g \in services} \left(X_g^N + Supply_g^N \right)}{\displaystyle\sum_{g \in goods} Supply_g^N \Big/ \sum_{g \in goods} \left(X_g^N + Supply_g^N \right)}$$

(31)

The correction was based on a division that describes the relative propensity of exporting services in comparison to goods.

All described elements taken together lead to the following, non-linear minimisation problem that had to be solved to determine the amount of cross-hauling or, in other words, the share of products that were produced and used within the same region.

---

[7] The explicit functional forms is:

$$IX_r = \frac{-(IZ_r t_{rr} t_{hh} + ID_h t_{rh}^2 - IZ_r t_{rh}^2 + ID_r t_{rh}^2) + \sqrt{\begin{array}{c}[IZ_r t_{rr} t_{hh} + ID_h t_{rh}^2 - IZ_r t_{rh}^2 + ID_r t_{rh}^2]^2 + \\ -4[t_{rr} t_{hh} - t_{rh}^2][ID_h IZ_r t_{rh}^2 - ID_h ID_r t_{rh}^2]\end{array}}}{2(1 - t_{rh}^2)} \equiv I_h$$

where $t_{ij} = T_{ij}^{1-\sigma}$. See Diodato and Thissen (2011) for more details.

$$Min \quad Objective = Z_1 + Z_2$$

$$s.t.$$

$$Z_1 = \sum_c \left[TTM_c - \sum_{r \in c} (T_{rh} - 1) IX^r - \sum_{r \in c} (T_{rr} - 1)(IY^r - IX^r) \right]^2$$

$$Z_2 = \sum_r \left[ChS_r IY_r - (IY^r - IX^r) \right]^2$$

$$t_{rh} = \left(1 + \ln\left[\beta + \gamma dist_{rh} \right] \right)^{1-\sigma}$$

$$(t_{rr}t_{hh} - t_{rh}^2)IX_r^2 + (IZ_r t_{rr}t_{hh} + ID_h t_{rh}^2 - IZ_r t_{rh}^2 + ID_r t_{rh}^2)IX_r + (ID_h IZ_r t_{rh}^2 - ID_h ID_r t_{rh}^2) = 0 \qquad (32)$$

In Equation (0) there are seven free variables, which needed to be determined: the objective variables (Z_1 and Z_2), the transportation variables (β, γ, t and T) and intranational exports (IX). All other elements are fixed parameters, whose values were taken from the available data sources. We used the conopt3 solver of the gams software to solve the problem described by Equation (0).

2.3 The origins and destinations of regional trade flows

This section describes how the origins and destinations of regional trade were determined. Data on goods and services were mainly derived from freight transport data and airline ticket information on first class and business class travel. The trade flows were determined by distributing the trade over the regions, given the amount produced and consumed in every region.

The determination of the amounts of goods and services produced and consumed per region is presented in the previous section. This provides the diagonal of the trade matrix. The number of products and services leaving and entering a region are also recorded in the regionalised Supply and Use Tables. These regional 'exports' were divided into international exports and intranational exports (destined for different regions within the same country).

In this section, the complete trade network between all distinguished NUTS2 regions is determined, given the intranational and international exports of the different regions. To determine export destinations, we used a simplified transport model, based on the probabilities of trade flows between different regions. These probabilities were derived from data on airline business trips (compare Derudder and Witlox, 2005), while for goods transport destinations the probabilities were based on freight transport data.

2.3.1 Transport hubs and the estimation of trade from transport data

The existence of transport hubs makes the derivation of trade data from transport data a complex procedure, as goods that are transported may be going to a transport hub instead of to their final destination. Therefore, there is a large difference between transport and trade data. We found that only 40% of *all* goods traded is being transported directly from origin to final destination. For the remaining 60%, at the least one transport hub is used before the final destination is reached. Especially in the case of international trade, it is likely that more than one hub is used: one in the country of origin and another in the country of destination.

The methodology used is based on a combination of two estimates on international trade between NUTS2 regions. The first estimate concerns the export of goods (the destination) and the second relates to their importation (the origin). We awarded both estimates a weight of 0.5 and minimised the quadratic difference between the final trade matrix and the two estimates. Below, the methodology is first described for the destination (exports) of goods and services, followed by a description of the methodology to determine the trade based on the origin of imported goods. Both methods are similar, but for the latter destination probabilities are replaced by origin probabilities

and consumption is distributed over regions of origin instead of production being distributed over regions of destination.

The methodology to determine the two estimates consisted of three steps. In the first step, the direct flows were determined. These are the traded goods and services that are directly transported from the region of origin to the region of destination without the use of transport hubs. In the second step, indirect trade flows were determined. These indirect trade flows are goods and services for which at least one transport hub was used between producer and consumer. The third and final step determined the higher order indirect flows. Higher order trade flows consist of traded goods or services that are transported via more than one transport hubs before they reach their final destination. These three steps are elaborated below, followed by a description of the determination of the final trade matrix.

2.3.2 Estimation of interregional trade flows: The export perspective

The first estimation was used to distribute the production in every region over the regions of destination. In this trade-flow estimation, the distribution of intranational and international export trade flows were not predetermined with respect to their destination regions. Only the overall international trade flows were predetermined.

Step 1 – Direct trade flows: export perspective

The direct trade flows concern goods and services that are directly transported from the production location (origin) to the consumption location (destination) without being reloaded at a transport hub. In order to determine these direct trade flows, we needed the direct probability $P_{i,k}^{0}$ of goods being exported from a region i to a region k. This probability can be described as:

$$P_{i,k}^{0} = \frac{T_{i,k}}{\sum_{k} T_{i,k}}$$

(33)

where $P_{i,k}^{0}$ is equal to the probability of a good being transported from *i* to *k* without using a hub and $T_{i,k}$ is the data on the quantity of goods transported from origin i to destination k.[8] Given the exported goods (intranational and international) X_{j} that are exported from region *j* we can describe the direct flows of exports X_{ik}^{0} from origin i to destination k in the following way:

$$X_{ik}^{0} = \lambda P_{i,k}^{0} X_{i}$$

(34)

where λ is the fraction of the goods that, on average, are transported directly to their final destination (around 40% of all goods). However, in some regions and especially at transport hubs, according to the regionalisation in Section 2.2, fewer goods were imported (I_{k}) than were exported to these regions, according to this methodology. Therefore, the following parameter Z_{k}^{0} was introduced.

$$Z_{k}^{0} = \max\left(-I_{k} + \lambda \sum_{i} P_{i,k}^{0} X_{i}, 0\right)$$

(35)

---

[8] The quantity of goods is approximated by the number of trips, due to data limitations.

The exports to an area k are larger than the imports in this area if Z_k is positive. In those cases, Equation (0) must be adjusted to determine the trade flows as follows:

$$X_{ik}^0 = \lambda P_{i,k}^0 X_i - \frac{Z_k^0}{\lambda \sum_i P_{i,k}^0 X_i} \lambda P_{i,k}^0 X_i$$

(36)

Equations (1) and (0) are sufficient to determine the direct trade flows between all regions.

Step 2 – Indirect trade flows using a transport hub: export perspective

The following step was used to determine the indirect trade flows of goods that are reloaded at a transport hub between production and consumption location. The probability $P_{i,k}^1$ that a good is transported from i to k using such a hub, the amount of exported goods X_j^1 that yet have to be distributed over the various regions of destination, and the amount of imported goods I_k^1 that yet have to be imported into the region were calculated as follows:

$$P_{i,k}^1 = \sum_j P_{i,j}^0 P_{j,k|k \neq i}^0$$

(37)

$$X_i^1 = X_i - \sum_j X_{ij}^0$$

(38)

$$I_k^1 = I_k - \sum_i X_{ik}^0$$

(39)

where a condition was imposed in such a way that goods could not be transported back to their production region, nor could they be reloaded. Analogue to the direct trade flows, it was determined whether the estimation of trade in this second step generated more exports to a specific region than there were imports into that region, according to the regionalisation in Section 2.2.

$$Z_k^1 = \max\left(-I_k^1 + \sum_i P_{i,k}^1 X_i^1, 0\right)$$

(40)

The exports to an area k are larger than the imports into this area if Z_k is positive. In those cases, the exports were adjusted in such a way that they were always smaller or equal to the imports into that region. Indirect trade flows were therefore determined by the following equation:

$$X_{ik}^1 = P_{i,k}^1 X_i^1 - \frac{Z_k^1}{\sum_i P_{i,k}^1 X_i^1} P_{i,k}^1 X_i^1$$

(41)

Step 3 – Higher order indirect trade flows using multiple hubs: export perspective

Higher order indirect trade flows use multiple transport hubs in the transportation of goods from production to consumption locations. Here, the same methodology can be used as described for indirect flows that use a single transport hub. Therefore, the higher order probabilities and trade flows can be described by the following set of equations:

23

$$P_{i,k}^2 = \sum\nolimits_{j,l|l\neq i} P_{i,j}^0 P_{j,l}^0 P_{l,k|k\neq i,k\neq j}^0 \tag{42}$$

$$X_i^2 = X_i - \sum\nolimits_j X_{ij}^1 - \sum\nolimits_j X_{ij}^0 \tag{43}$$

$$I_k^2 = I_k - \sum\nolimits_i X_{ik}^1 - \sum\nolimits_i X_{ik}^0 \tag{44}$$

$$Z_k^2 = \max\left(-I_k^2 + \sum\nolimits_i P_{i,k}^2 X_i^2, 0\right) \tag{45}$$

$$X_{ik}^2 = P_{i,k}^2 X_i^2 - \frac{Z_k^2}{\sum\nolimits_i P_{i,k}^2 X_i^2} P_{i,k}^2 X_i^2 \tag{46}$$

Any subsequent transport hub along the transportation route was treated in the same way, up to a maximum of 5 hubs. The remaining X_k^n was bi-proportionally distributed such that total export would match total import. Total interregional trade was then the aggregate over all subsets of trade using n different numbers of hubs. Thus, the following equation describes the total exports between regions:

$$X_{ik}^{prior} = \sum\nolimits_n X_{i,k}^n \tag{47}$$

2.3.3 Estimation of interregional trade flows: The import perspective

A procedure similar to the one used for estimating the export trade flows can be applied to import trade flows. Thus, the origin of products consumed in a region can be determined in the same way as the destination of regionally produced products. The procedure for imports is mostly the same as for exports and, therefore, only a concise description is provided below.

Step 1– Direct trade flows: import perspective
The main difference between the import and export estimations is that, here, to allocate the imports to their respective regions of origin, we used the probability $Q_{i,k}^0$ that a good is imported in region k directly from region i without the use of a transport hub. The following equation was used for determining the probability $Q_{i,k}^0$:

$$Q_{i,k}^0 = \frac{T_{i,k}}{\sum\nolimits_i T_{i,k}} \tag{48}$$

Thus, the flow of goods between region of destination j and region of origin i can be determined with the following set of equations:

$$I_{k,i}^0 = \lambda Q_{k,i}^0 I_i - \frac{ZZ_k^0}{\lambda \sum_i Q_{k,i}^0 I_i} \lambda Q_{k,i}^0 I_i$$

(49)

$$ZZ_k^0 = \max\left(-X_k + \lambda \sum_i Q_{k,i}^0 I_i, 0\right)$$

(50)

where the adjustment parameter ZZ_k is positive if the imports coming from region k are larger than the exports from this region.

Steps 2 and 3 – Indirect trade flows using a transport hub: import perspective

The following steps describe how trade flows that use one or more transport hubs were determined. Analogous to the export trade flows, those that use only one hub can be determined according to the following set of equations:

$$Q_{i,k}^1 = \sum_j Q_{i,j}^0 Q_{j,k}^0$$

(51)

$$I_i^1 = I_i - \sum_j I_{ji}^0$$

(52)

$$X_k^1 = X_k - \sum_i I_{k,i}^0$$

(53)

$$I_{k,i}^1 = \lambda Q_{k,i}^1 I_i^1 - \frac{ZZ_k^1}{\lambda \sum_i Q_{k,i}^1 I_i^1} \lambda Q_{k,i}^1 I_i^1$$

(54)

$$ZZ_k^1 = \max\left(-X_k^1 + \lambda \sum_i Q_{k,i}^1 I_i^1, 0\right)$$

(55)

where $Q_{i,k}^1$ is the probability that a good is transported from i to k using one hub, with the respective amounts of goods I_j^1 and X_j^1 still having to be imported into and exported from other regions.

Step 3 is completely analogous to the methodology used for estimating the export flows and is therefore not presented here. The total import flow into region k from region i is simply the following sum of all direct and indirect flows:

$$I_{i,k}^{prior} = \sum_n I_{i,k}^n$$

(56)

2.4 Final estimation

The previous section describes the distribution of exports and imports over regions of destination and origin by use of a logistic trade model. This method resulted in two estimates of interregional trade. The first estimate, $X_{i,j,g}^{prior}$, is the trade matrix from region i to region j of product g, from the export point of view. The second estimate, $I_{i,j,g}^{prior}$, is the trade matrix from region i to region j of

product *g,* from the import point of view. Following this estimation methodology, both international and intranational trade were distributed simultaneously. This implies that both 'compete' for a destination in Equation (0). International and intranational quantities, therefore, were not predetermined. They were endogenously determined in the process based on the probabilities of trade flows between regions. The resulting priors from the export point of view are therefore consistent with international exports from the regionalised Supply and Use Tables, but inconsistent with international imports. Vice versa, import priors are consistent with international imports from the regionalised Supply and Use Tables, but inconsistent with international exports. In addition, export and import priors are not entirely consistent with data on international trade at the country level.[9] The total international and intranational exports and imports are consistent with the regionalised Supply and Use Tables in both estimates.

The purpose of this last step was to achieve consistency between the two estimates of trade. The methodology applied was that of constrained optimisation, the same as described in Section 2.1 for the estimation of consistent trade flows at the country level. The final estimate of the regional trade matrix $T_{i,j,g}$ was determined as the trade matrix $T_{i,j,g}$, which is as near as possible to both priors ($X^{prior}_{i,j,g}$ and $I^{prior}_{i,j,g}$), while meeting the conditions for total exports, total imports and international trade.

The minimised objective function *Z* is the sum of the errors (absolute value of relative error) with respect to the two priors $X^{prior}_{i,j,g}$ and $I^{prior}_{i,j,g}$. This can be described in mathematics as follows:

$$\min Z = \sum_{i,j,g} \left(\frac{|X^{prior}_{i,j,g} - T_{i,j,g}| + |I^{prior}_{j,i,g} - T_{i,j,g}|}{X^{prior}_{i,j,g} + I^{prior}_{j,i,g}} \right)$$

s.t.

$$1)\quad X^{total}_{i,g} = \sum_{j} T_{i,j,g}$$

$$2)\quad I^{total}_{j,g} = \sum_{i} T_{i,j,g}$$

$$3)\quad T^{country}_{o,d,g} = \sum_{i \in o} \sum_{j \in d} T_{i,j,g}$$

(57)

Since all product categories are independent of each other, the same results could have been achieved by solving the minimisation in Equation (1) for each product, separately.

---

[9] In fact, it must have the following equivalence: the sum of exports from all regions *i* belonging to country *o* to all regions *j* belonging to country *d* needs to equal the export between country *o* and country *d*.

3. An illustration of the regional trade data

This section provides a concise illustration of several key elements of the data set, but does not provide a data analysis (e.g. on differences between national and regional trade or specific empirical regularities in regional trade), as this is outside the scope of this paper . The main aim of the presented research was to build a data set on regional trade for Europe that can be used in (smart specialisation) regional development analysis. As mentioned in the introduction, there is a growing interest among both academics and policymakers to study and analyse economic phenomena on a regional scale. The available data, however, are not always available or sufficiently detailed to permit a thorough analysis. For the data set, however, thorough analysis has been done in other studies, see Thissen et al. (in prep.), Thissen et al. (2011), Raspe et al. (2012a), Raspe et al. (2012b) and Diodato and Weterings (2012), and more analyses will be possible once the data become publicly available.

The novelty in the presented research is the proposed non-survey approach to generate regional trade data. Key elements of our approach are the following trade patterns: within regions, between regions in Europe and between European regions and the rest of the world. Trade by firms with customers in the local region was determined using the cross-hauling approach presented in Section 2.2.3. The trade with other regions in the same country and with those in Europe was mainly determined by the parameter-free methodology to determine trade flows, presented in Section 2.3. The international trade flows within Europe and the trade flows with the rest of the world were also largely determined as described in Sections 2.1 and 2.2, while all trade flows also were readjusted to make them consistent in the final estimation described in Section 2.4.

Table 2, therefore, presents the export destinations for three macro-sectors that account for most of the economic activity in Europe[10]. These export destinations were divided into: the same region, regions within the same country, the rest of Europe, and the rest of the world (excuding Europe). Total exports add up to 100% (over the rows of Table 2). The division showed the above described key elements of the approach used.

Table 2. Average Export destinations for all NUTS2 regions, in percentages

	Region	Country	Europe	World
Agriculture	66	23	9	2
Manufacturing	47	23	20	10
Services	75	18	4	3

From Table 2 it can be observed that the degree to which firms sell their products to customers within their own region, in relation to those in the rest of the country, directly relates to the cross-hauling problem discussed before. With the exception of our data set, there is little information on domestic trade. There are some figures available for a number of regions, but no comprehensive and full trade matrix. The persistent fallacy of missing data that could be used to assess the component pertaining to domestic trade is problematic for many regional economic issues, such as the debate on how 'local' or 'global' a regional economy actually is. Our data showed there to be a substantial sectoral difference in the amounts of produce that stayed in the producing region. Typical non-tradables, such as services, were predominantly produced and sold within the same region. This also applied to most agricultural products, which were either being sold to consumers directly, or to the food processing industry. And, although manufacturing, generally, is

---

[10] From this table, the resource based sector, construction and utilities are excluded.

characterised by a far higher trade level, almost half of its trade nevertheless took place within the producing region.

Table 2 shows that services consist not only of non-tradables, but in fact also have a small, but sizeable, international trade component (to Europe and the rest of the world), as well as a significant amount of domestic trade. The international trade component can perhaps be explained by the trade in services that can be provided electronically (e.g. internet services). National service trade is probably still driven by the constraint that production and consumption should be at the same location. This therefore requires a certain amount of travelling either by the customer or the supplier (Samson and Snape, 1985). This would also explain the small difference between regions' national trade levels and that with other European regions and with the rest of the world.

A comparison between agricultural and manufacturing sectors showed that the former appears to be subject to a much steeper distance decay in its sales; for the year 2000, with up to 66% of output sold within the producing region. This number is relatively impressive given that is not too far away from the 75% that we saw in services. This may be related to the actual physical decay that some agricultural products are subjects to, which may drive up transportation costs and push for spreading and self-sufficiency. If we compare services with the manufacturing sector, the distance decay in trade outside Europe was observed to be flatter for services than for manufacturing trade. This would support the earlier suggestion of the importance of the low transport costs involved in the trade in internet services, where production location hardly matters, compared to manufacturing locations.

A general conclusion from Table 2 may be that, although globalisation may be an important phenomenon for certain product groups, regional economies stay predominantly local. Most of what people produce stays in the same region or country. For Europe as a whole, this is even more so the case. European regions mainly trade with other European regions and the European regional economy is therefore mainly dependent on other European regions with strong interregional links.

4. Discussion

The presented trade data in this paper represents the most likely trade between European NUTS2 regions, on the basis of currently available information. This information was derived not only from data sources, but also from simple economic consistency and bookkeeping rules applied in the methodology. The most simple economic consistency rule was that if a product is exported to a region it should also be used in that region, while bookkeeping rules ensure that export to a certain country equal the imports into that country. Consistency and bookkeeping improved the quality of the data, but one has to keep in mind that the presented trade data were inferred from other data sources. The data was not measured as a flow from one region to another. Small errors in trade flows between regions, therefore, remained at the CPA 2-digit product level. The data, therefore, should preferable be used as network data without too much focus one each individual trade flow.

Several studies have already been completed using the presented trade data (Thissen et al., 2011; Raspe et al., 2012a; Raspe et al., 2012b; Diodato and Weterings, 2012). These studies are examples of how the data may be used to build up a smart specialisation strategy for different regions. In addition to the presented data, trade data were updated for the 2000–2010 period (*European regional trade flows: An update for 2000–2010*, Thissen et al. (in prep)). Finally, another study focuses on the competitiveness of regions and smart specialisation strategies, providing examples of how the trade data set may be used for regional economic development strategies (*European regional competitiveness and Smart Specialization* ,Thissen et al.(in prep.).

The data showed that the European regional economy in the year 2000 was strongly interrelated, with only limited links with the rest of the world. Although for many typical products international global trade may have been of primal importance, we found that most of what was produced was consumed in nearby or even the same region. This strong interdependence of European regions is part of the reasons for the strong regional spillovers within the European Union and the spreading of the present economic crisis over Europe.

References

Barca F. (2009). An agenda for a reformed cohesion policy: a place-based approach to meeting European Union challenges and expectations. Report for the European Commission, Brussels.

Barca F, McCann P and Rodriguez-Pose A. (2012). The case for regional development intervention: place-based versus place-neutral approaches. *Journal of Regional Science* 52: 134–152.

Bouwmeester MC and Oosterhaven J. (2009). Methodology for the Construction of an International Supply-Use Table. Working Paper, University of Groningen.

Cambridge Econometrics (2008). Regional production, investment and consumption in Europe for year 2000. Data acquired in 2008.

Derudder B and Witlox F. (2005). An Appraisal of the Use of Airline Data in Assessing the World City Network: a Research Note on Data. Urban Studies, 42(13), pp. 2371–2388.

Diodato D and Thissen M. (2011). Towards a New Economic Geography based Estimate of Trade Elasticity and Transport Costs. Working Paper.

Diodato D and Weterings A. (2012). The Resilience of Dutch Regions to Economic shocks. Measuring the relevance of interactions among firms and workers. Utrecht University, Paper in Evolutionary Economic Geography.

Dixit AK and Stiglitz JE. (1977). Monopolistic competition and optimum product diversity, American Economic Review 67: 297–308.

Eurostat (2008). Eurostat Manual of Supply, Use and Input-Output Tables. Eurostat Methodologies and Working Papers.

Eurostat (2009). International trade in services. Data for years from 2000 to 2004, downloaded in 2009.

Eurostat (2009b). National use and supply tables. Data from year 2000 for Austria, Belgium, Czech Republic, Germany, Denmark, Estonia, Spain, Finland, France, Hungary, Ireland, Italy, Lithuania, Luxembourg, Malta, the Netherlands, Norway, Poland, Portugal, Sweden, Slovenia, Slovakia, United Kingdom. Data from year 1998 for Greece and Latvia. Downloaded in 2009.

Feenstra RC, Lipsey RE, Deng H, Ma AC and Mo H. (2005). World Trade Flows: 1962–2000. National Bureau of Economic Research.

Florida R. (2002). The Rise of the Creative Class. And How It's Transforming Work, Leisure and Everyday Life. Basic Books

Isard W. (1953). Regional commodity balances and interregional commodity flows. American Economic Review 43: 167–80.

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

Krugman PR. (1991). Geography and trade. Cambridge, Mass.: The MIT Press.

McCann P and Ortega-Argilés R. (2011). Smart Specialisation, Regional Growth and Applications to EU Cohesion Policy, DGRegio Website:

http://ec.europa.eu/regional_policy/cooperate/regions_for_economic_change/index_en.cfm#4

McKitrick R. (1998). The econometric critique of computable general equilibrium modeling: The role of functional forms, Economic Modelling 15: 543–574.

MIDT (2010). Regional flights, business and first class. Data for November 2000.

Ministry of Infrastructure and the Environment (2007). Interregional freight. Data for years from 2000–2004.

Oosterhaven J, Stelder D and Inomata S. (2008). Estimating international interindustry linkages: Non-survey simulations of the Asian-Pacific Economy. Economic Systems Research, 20, 395–414.

Porter M. (1990). The Competitive Advantage of Nations. New York, The Free Press.

Raspe O, Weterings A and Thissen M. (2012a; in Dutch). De internationale concurrentiepositie van de topsectoren [The international competitive position of Dutch top sectors]. The Hague, PBL Netherlands Environmental Assessment Agency.

Raspe O, Weterings A, Thissen M and Diodato D. (2012b; in Dutch). De concurrentiepositie van de topsectoren in Noord-Brabant [The competitive position of top sectors in the Dutch province of North Brabant]. The Hague, PBL Netherlands Environmental Assessment Agency.

Sampson G and Snape R. (1985). Identifying the issues in trade in services. The World Economy, 8, 171–81.

Samuelson PA. (1954). The Transfer Problem and Transport Costs, II: Analysis of Effects of Trade Impediments. Economic Journal 64: 264–289.

Saxenian A. (1994). Regional Advantage: Culture and Competition in Silicon Valley and Route 128. Cambridge, Massachusetts: Harvard University.

System of National Accounts, SNA1993 and SNA2008. United Nation website: http://unstats.un.org/unsd/nationalaccount/sna.asp

Simini F, González MC, Maritan A and Barabási A. (2012). A universal model for mobility and migration patterns, Nature, 484, pp. 96–100, doi:10.1038/nature10856.

Thissen M, Ruijs A, Van Oort FG and Diodato D. (2011; in Dutch). De concurrentiepositie van Nederlandse regio's. Regionaal-economische samenhang in Europa [The competitive position of Dutch regions. Regional-economic cohesion within Europe]. The Hague, PBL Netherlands Environmental Assessment Agency.

Thissen M, Van Oort F, Diodato D and Ruijs A (in prep.). European regional competitiveness and Smart Specialization; Regional place-based development perspectives in international economic networks, Cheltenham: Edward Elgar.

Thissen M, Diodato D. and Van Oort F. (in prep). European regional trade flows: An update for 2000–2010. The Hague, PBL Netherlands Environmental Assessment Agency.

Appendix A: The data set on interregional trade

The data set documented in this paper describes interregional trade flows between 256 European regions, for the time period of 2000. Export and import flows were measured in millions of euros, and divided into 59 product and service categories. These 256 regions are part of the EU25, with the exception of Cyprus and including Norway. This choice of regions was determined by data availability. The regional classification follows the second level of Eurostat's *Nomenclature of Statistical Territorial Units* (NUTS2), which in many cases in Europe is equivalent to a pre-existing countries' administrative division. Section 1.2 details the regional units in which data were divided. We used the *Classification of Products by Activity* (Nace1.1-CPA 2002), the division that is also used by Eurostat for the national accounts' Supply and Use Tables. Consistent with Eurostat's publications, we used the second level of this classification (2 digits), which distinguishes between 59 goods and services. This disaggregation of products is reported in Section 1.3. It must be noted that the data set not only provides information on international trade between regions, but also reports the trade between regions of the same country. Moreover, since for the whole research the emphasis was on maintaining consistency over all accounts, the data set also includes information on regional production for local use (the diagonal of the trade matrix).

Region and product classification

In the construction of the data set, we included 256 NUTS2 regions, which are part of 25 selected European countries: 24 EU countries and Norway, listed in Table 3.

Table 3. The European countries in the data set

L1	Austria	L11	Hungary	L21	Portugal
L2	Belgium	L12	Ireland	L22	Sweden
L3	Czech Republic	L13	Italy	L23	Slovenia
L4	Germany	L14	Lithuania	L24	Slovakia
L5	Denmark	L15	Luxembourg	L25	United Kingdom
L6	Estonia	L16	Latvia		
L7	Spain	L17	Malta		
L8	Finland	L18	Netherlands		
L9	France	L19	Norway		
L10	Greece	L20	Poland		

The data set also covers the trade between the distinguished European regions and the rest of the world. This rest of the world has been divided into main economic countries and groups of countries of less economic importance. These trading partners outside Europe are listed in Table 4.

Table 4. European trading partners outside Europe

L26	Rest of Europe	L35	Cyprus
L27	Africa	L36	Canada
L28	Asia	L37	China
L29	Japan	L38	Hong Kong
L30	Middle and South America	L39	Korea
L31	Australia and Oceania	L40	Singapore
L32	Northern America	L41	Switzerland
L33	Russia	L42	Turkey
L34	Rest of the World	L43	United States

Table 5 presents a list of all the NUTS2 regions in the data set. The first column refers to the code we used in our research (from R1 to R256), the second gives the corresponding NUTS2 code, and the third contains the name of the region.

Table 5. NUTS2 regions in the data set

R1	AT11	Burgenland	**R129**	GR30	Attiki
R2	AT12	Niederösterreich	**R130**	GR41	Voreio Aigaio
R3	AT13	Wien	**R131**	GR42	Notio Aigaio
R4	AT21	Kärnten	**R132**	GR43	Kriti
R5	AT22	Steiermark	**R133**	HU10	Közép-Magyarország
R6	AT31	Oberösterreich	**R134**	HU21	Közép-Dunántúl
R7	AT32	Salzburg	**R135**	HU22	Nyugat-Dunántúl
R8	AT33	Tirol	**R136**	HU23	Dél-Dunántúl
R9	AT34	Vorarlberg	**R137**	HU31	Észak-Magyarország
R10	BE10	Région de Bruxelles	**R138**	HU32	Észak-Alföld
R11	BE21	Prov. Antwerpen	**R139**	HU33	Dél-Alföld
R12	BE22	Prov. Limburg (B)	**R140**	IE01	Border, Midlands and Western
R13	BE23	Prov. Oost-Vlaanderen	**R141**	IE02	Southern and Eastern
R14	BE24	Prov. Vlaams Brabant	**R142**	ITC1	Piemonte
R15	BE25	Prov. West-Vlaanderen	**R143**	ITC2	Valle d'Aosta
R16	BE31	Prov. Brabant Wallon	**R144**	ITC3	Liguria
R17	BE32	Prov. Hainaut	**R145**	ITC4	Lombardia
R18	BE33	Prov. Liège	**R146**	ITD1	Provincia Bolzano
R19	BE34	Prov. Luxembourg (B)	**R147**	ITD2	Provincia Trento
R20	BE35	Prov. Namur	**R148**	ITD3	Veneto
R21	CZ01	Praha	**R149**	ITD4	Friuli-Venezia Giulia
R22	CZ02	Strední Cechy	**R150**	ITD5	Emilia-Romagna
R23	CZ03	Jihozápad	**R151**	ITE1	Toscana
R24	CZ04	Severozápad	**R152**	ITE2	Umbria
R25	CZ05	Severovýchod	**R153**	ITE3	Marche
R26	CZ06	Jihovýchod	**R154**	ITE4	Lazio
R27	CZ07	Strední Morava	**R155**	ITF1	Abruzzo
R28	CZ08	Moravskoslezko	**R156**	ITF2	Molise
R29	DE11	Stuttgart	**R157**	ITF3	Campania
R30	DE12	Karlsruhe	**R158**	ITF4	Puglia
R31	DE13	Freiburg	**R159**	ITF5	Basilicata
R32	DE14	Tübingen	**R160**	ITF6	Calabria
R33	DE21	Oberbayern	**R161**	ITG1	Sicilia

R34	DE22	Niederbayern		**R162**	ITG2	Sardegna
R35	DE23	Oberpfalz		**R163**	LT00	Lietuva
R36	DE24	Oberfranken		**R164**	LU00	Luxembourg
R37	DE25	Mittelfranken		**R165**	LV00	Latvija
R38	DE26	Unterfranken		**R166**	MT00	Malta
R39	DE27	Schwaben		**R167**	NL11	Groningen
R40	DE30	Berlin		**R168**	NL12	Friesland
R41	DE41	Brandenburg - NO		**R169**	NL13	Drenthe
R42	DE42	Brandenburg - SW		**R170**	NL21	Overijssel
R43	DE50	Bremen		**R171**	NL22	Gelderland
R44	DE60	Hamburg		**R172**	NL23	Flevoland
R45	DE71	Darmstadt		**R173**	NL31	Utrecht
R46	DE72	Gießen		**R174**	NL32	Noord-Holland
R47	DE73	Kassel		**R175**	NL33	Zuid-Holland
R48	DE80	Mecklen.-Vorpom.		**R176**	NL34	Zeeland
R49	DE91	Braunschweig		**R177**	NL41	Noord-Brabant
R50	DE92	Hannover		**R178**	NL42	Limburg (NL)
R51	DE93	Lüneburg		**R179**	NO01	Oslo og Akershus
R52	DE94	Weser-Ems		**R180**	NO02	Hedmark og Oppland
R53	DEA1	Düsseldorf		**R181**	NO03	Sor-Ostlandet
R54	DEA2	Köln		**R182**	NO04	Agder og Rogaland
R55	DEA3	Münster		**R183**	NO05	Vestlandet
R56	DEA4	Detmold		**R184**	NO06	Trondelag
R57	DEA5	Arnsberg		**R185**	NO07	Nord-Norge
R58	DEB1	Koblenz		**R186**	PL11	Lódzkie
R59	DEB2	Trier		**R187**	PL12	Mazowieckie
R60	DEB3	Rheinhessen-Pfalz		**R188**	PL21	Malopolskie
R61	DEC0	Saarland		**R189**	PL22	Slaskie
R62	DED1	Chemnitz		**R190**	PL31	Lubelskie
R63	DED2	Dresden		**R191**	PL32	Podkarpackie
R64	DED3	Leipzig		**R192**	PL33	Swietokrzyskie
R65	DEE1	Dessau		**R193**	PL34	Podlaskie
R66	DEE2	Halle		**R194**	PL41	Wielkopolskie
R67	DEE3	Magdeburg		**R195**	PL42	Zachodniopomorskie

R68	DEF0	Schleswig-Holstein		**R196**	PL43	Lubuskie
R69	DEG0	Thüringen		**R197**	PL51	Dolnoslaskie
R70	DK01	Hovedstadsreg		**R198**	PL52	Opolskie
R71	DK02	Øst for Storebælt		**R199**	PL61	Kujawsko-Pomorskie
R72	DK03	Vest for Storebælt		**R200**	PL62	Warminsko-Mazurskie
R73	EE00	Eesti		**R201**	PL63	Pomorskie
R74	ES11	Galicia		**R202**	PT11	Norte
R75	ES12	Principado de Asturias		**R203**	PT15	Algarve
R76	ES13	Cantabria		**R204**	PT16	Centro (PT)
R77	ES21	Pais Vasco		**R205**	PT17	Lisboa
R78	ES22	Com. Foral de Navarra		**R206**	PT18	Alentejo
R79	ES23	La Rioja		**R207**	SE01	Stockholm
R80	ES24	Aragón		**R208**	SE02	Östra Mellansverige
R81	ES30	Comunidad de Madrid		**R209**	SE04	Sydsverige
R82	ES41	Castilla y León		**R210**	SE06	Norra Mellansverige
R83	ES42	Castilla-la Mancha		**R211**	SE07	Mellersta Norrland
R84	ES43	Extremadura		**R212**	SE08	Övre Norrland
R85	ES51	Cataluña		**R213**	SE09	Småland med öarna
R86	ES52	Comunidad Valenciana		**R214**	SE0A	Västsverige
R87	ES53	Illes Balears		**R215**	SI00	Slovenija
R88	ES61	Andalucia		**R216**	SK01	Bratislavský kraj
R89	ES62	Región de Murcia		**R217**	SK02	Západné Slovensko
R90	ES63	Ciudad Autónoma de Ceuta		**R218**	SK03	Stredné Slovensko
R91	ES64	Ciudad Autónoma de Melilla		**R219**	SK04	Východné Slovensko
R92	ES70	Canarias		**R220**	UKC1	Tees Valley and Durham
R93	FI13	Itä-Suomi		**R221**	UKC2	Northumberland, Tyne and Wear
R94	FI18	Etelä-Suomi		**R222**	UKD1	Cumbria
R95	FI19	Länsi-Suomi		**R223**	UKD2	Cheshire
R96	FI1A	Pohjois-Suomi		**R224**	UKD3	Greater Manchester
R97	FI20	Åland		**R225**	UKD4	Lancashire
R98	FR10	Île de France		**R226**	UKD5	Merseyside
R99	FR21	Champagne-Ardenne		**R227**	UKE1	East Riding and North Lincoln
R100	FR22	Picardie		**R228**	UKE2	North Yorkshire
R101	FR23	Haute-Normandie		**R229**	UKE3	South Yorkshire

R102	FR24	Centre	**R230**	UKE4	West Yorkshire
R103	FR25	Basse-Normandie	**R231**	UKF1	Derby and Nottingham
R104	FR26	Bourgogne	**R232**	UKF2	Leicester, Rutland and Northants
R105	FR30	Nord - Pas-de-Calais	**R233**	UKF3	Lincolnshire
R106	FR41	Lorraine	**R234**	UKG1	Hereford, Worcester and Warks
R107	FR42	Alsace	**R235**	UKG2	Shrop and Stafford
R108	FR43	Franche-Comté	**R236**	UKG3	West Midlands
R109	FR51	Pays de la Loire	**R237**	UKH1	East Anglia
R110	FR52	Bretagne	**R238**	UKH2	Bedford, Hertford
R111	FR53	Poitou-Charentes	**R239**	UKH3	Essex
R112	FR61	Aquitaine	**R240**	UKI1	Inner London
R113	FR62	Midi-Pyrénées	**R241**	UKI2	Outer London
R114	FR63	Limousin	**R242**	UKJ1	Berks, Bucks and Oxford
R115	FR71	Rhône-Alpes	**R243**	UKJ2	Surrey, East and West Sussex
R116	FR72	Auvergne	**R244**	UKJ3	Hampshire and Isle of Wight
R117	FR81	Languedoc-Roussillon	**R245**	UKJ4	Kent
R118	FR82	Provence-Alpes-Côte d'Azur	**R246**	UKK1	Gloucester, Wilt and North Somerset
R119	FR83	Corse	**R247**	UKK2	Dorset and Somerset
R120	GR11	Anatoliki Makedonia, Thraki	**R248**	UKK3	Cornwall and Isles of Scilly
R121	GR12	Kentriki Makedonia	**R249**	UKK4	Devon
R122	GR13	Dytiki Makedonia	**R250**	UKL1	West Wales and The Valleys
R123	GR14	Thessalia	**R251**	UKL2	East Wales
R124	GR21	Ipeiros	**R252**	UKM1	North Eastern Scotland
R125	GR22	Ionia Nisia	**R253**	UKM2	Eastern Scotland
R126	GR23	Dytiki Ellada	**R254**	UKM3	South Western Scotland
R127	GR24	Sterea Ellada	**R255**	UKM4	Highlands and Islands
R128	GR25	Peloponnisos	**R256**	UKN0	Northern Ireland

Product categories

In this study, trade between European regions is detailed on a product level. Export and import flows are divided according to the 2-digit *Classification of Products by Activity* (CPA, 1996) and presented in Table 6. The original 1996 classification was revised a number of times; the last revision was on the version of 2002 (CPA, 2008). Nevertheless, to date, Eurostat has published national accounts which are in line with the classification of 1996. There is a total of 62 goods and services in the CPA 2002, but goods and services produced by households for own use and services provided by extraterritorial organisations (product numbers 96, 97 and 99) are not included in the Supply and use system of accounts, reducing the total number of products analysed in this study to 59.

Table 6. 2-digit classification of Products by Activity (CPA, 1996)

P1	AA01	Products of agriculture, hunting and related services
P2	AA02	Products of forestry, logging and related services
P3	BA05	Fish and other fishing products; services incidental of fishing
P4	CA10	Coal and lignite; peat
P5	CA11	Crude petroleum and natural gas; services incidental to oil and gas extraction excluding surveying
P6	CA12	Uranium and thorium ores
P7	CB13	Metal ores
P8	CB14	Other mining and quarrying products
P9	DA15	Food products and beverages
P10	DA16	Tobacco products
P11	DB17	Textiles
P12	DB18	Wearing apparel; furs
P13	DC19	Leather and leather products
P14	DD20	Wood and products of wood and cork (except furniture); articles of straw and plaiting materials
P15	DE21	Pulp, paper and paper products
P16	DE22	Printed matter and recorded media
P17	DF23	Coke, refined petroleum products and nuclear fuels
P18	DG24	Chemicals, chemical products and man-made fibres
P19	DH25	Rubber and plastic products
P20	DI26	Other non-metallic mineral products
P21	DJ27	Basic metals
P22	DJ28	Fabricated metal products, except machinery and equipment
P23	DK29	Machinery and equipment n.e.c.
P24	DK30	Office machinery and computers
P25	DL31	Electrical machinery and apparatus n.e.c.
P26	DL32	Radio, television and communication equipment and apparatus

P27	DL33	Medical, precision and optical instruments, watches and clocks
P28	DM34	Motor vehicles, trailers and semi-trailers
P29	DM35	Other transport equipment
P30	DN36	Furniture; other manufactured goods n.e.c.
P31	DN37	Secondary raw materials
P32	EA40	Electrical energy, gas, steam and hot water
P33	EA41	Collected and purified water, distribution services of water
P34	FA45	Construction work
P35	FA50	Trade, maintenance and repair services of motor vehicles and motorcycles; retail sale of automotive fuel
P36	GA51	Wholesale trade and commission trade services, except of motor vehicles and motorcycles
P37	GA52	Retail trade services, except of motor vehicles and motorcycles; repair services of personal and household goods
P38	HA55	Hotel and restaurant services
P39	IA60	Land transport; transport via pipeline services
P40	IA61	Water transport services
P41	IA62	Air transport services
P42	IA63	Supporting and auxiliary transport services; travel agency services
P43	IA64	Post and telecommunication services
P44	JA65	Financial intermediation services, except insurance and pension funding services
P45	JA66	Insurance and pension funding services, except compulsory social security services
P46	JA67	Services auxiliary to financial intermediation
P47	KA70	Real estate services
P48	KA71	Renting services of machinery and equipment without operator and of personal and household goods
P49	KA72	Computer and related services
P50	KA73	Research and development services
P51	KA74	Other business services
P52	LA75	Public administration and defence services; compulsory social security services
P53	MA80	Education services
P54	NA85	Health and social work services
P55	OA90	Sewage and refuse disposal services, sanitation and similar services
P56	OA91	Membership organisation services n.e.c.
P57	OA92	Recreational, cultural and sporting services
P58	OA93	Other services
P59	PA95	Private households with employed persons