European Commission

JRC TECHNICAL REPORTS

Data Needs for Regional Modelling

A Description of the Data used in Support of RHOMOLO

Lesley Potters
Andrea Conte
d'Artis Kancs
Mark Thissen

2014

Report EUR 25955 EN

Joint Research Centre

European Commission
Joint Research Centre
Institute for Prospective Technological Studies

Contact information
Lesley Potters
Address: Joint Research Centre, calle Inca Garcilaso, 3 – 41092 – Seville, Spain
E-mail: lesley.potters@ec.europa.eu
Tel.: +34 954 480 582
Fax: +34 954 488 326

http://ipts.jrc.ec.europa.eu/
http://www.jrc.ec.europa.eu/

This publication is a Technical Report by the Joint Research Centre of the European Commission.

JRC80845

EUR 25955 EN

ISBN 978-92-79-29737-3 (PDF)

ISSN 1831-9424 (online)

doi:10.2791/10249

Luxembourg: Publications Office of the European Union, 2014

Printed in Spain

1 Introduction

RHOMOLO is a Computable General Equilibrium (CGE) model being developed under the Regional Modelling (REMO[1]) action of the JRC-IPTS on behalf of DG REGIO. The objective of the model is to provide scientific support to EU policymaking by evaluating the possible impacts of policy instruments available under the Cohesion Policy toolkit.

Cohesion Policy is one of the major investment tools at the Community level. Roughly a third of EU budget goes into this policy domain which aims at supporting job creation and enhancing competitiveness, economic growth, improved quality of life and sustainable development – in line with the delivery of the Europe 2020 strategy.

Current regional funding programmes will run until 2013. At the same time, there are on-going negotiations between the Commission, the European Council and the Parliament on the next seven-year Multi Annual Financial Framework (MFF) and, therefore, on the new shape of cohesion policy for the period 2014-2020.

The aim of this paper is to provide an overview of the data requirements to develop a regional model such as RHOMOLO. In doing so, the objective is to highlight important areas of analysis while stressing the methodological choices which are required in such a context.

The paper starts with a general overview of the model structure (Section 2) and then it describes Social Accounting Matrices (Section 3) – a crucial information framework for running a model such as RHOMOLO. Subsequently, data are organised by policy domains (Section 4) by looking at their availability and discussing the methodology to properly estimate regional and – where applicable – sector/regional information. Section 5 presents how the inter-regional trade flows have been estimated. Section 6 indicates the model parameters – mostly obtained from relevant literature and/or in-house econometric estimates /calibration. Finally, Section 7 will highlight the direction of research in terms of data needs and the currently established procedures for a timely update of the data used in the model.

2 Data sources and Model Structure

Eurostat (ESTAT) database represents the primary source of data for building an EU-wide system of national and regional statistical information. A hierarchical approach has been chosen in terms of

---

[1] See http://ipts.jrc.ec.europa.eu/activities/research-and-innovation/regional_economic_modelling.cfm.

data collection[2]. In particular, only when information is not available from ESTAT, other international data sources are consulted - such as OECD and/or IMF. Finally, data from National Statistical Offices and ad-hoc data from international consortia (such as those obtained from FP-related studies) and private companies are consulted when specific data needs arise.

However, data are not fully complete when exhaustive sectoral (NACE Rev. 1.1) and regional (NUTS2) information is needed. In turn, data constraints have led some variables to be *imputed* (such as in the case of missing values), *proxied* according to available information (i.e. sectoral R&D data at the regional data) or *fully constructed* (i.e. capital stocks and inter-regional trade data) according to some underlying hypotheses on data distribution and time variation. This paper tries to explain where the bottlenecks are in building such an exhaustive data infrastructure whereas it discusses underlying choices made in the construction of available indicators.

The model covers 267 EU NUTS2 regions[3] with 6 NACE Rev. 1.1 macro-sectors (see Table 1).

Table 1 RHOMOLO Macro Sectors

CODE	**SECTOR**
AB	Agriculture, hunting and forestry
CDE	Mining and quarrying + Manufacturing + Electricity and Gas
F	Construction
GHI	Wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods + Hotels and restaurants + Transport and Communications
JK	Financial intermediation, real estate and business services
LMNOP	Non-Market Services

The adopted base year is currently 2007 – although updated time series are built for relevant variables. Data storage is organised on a central location. Standard data formats are used (such as Excel and STATA) while GAMS is RHOMOLO language code.

---

[2] Indeed, the European System of National and Regional Accounts (ESA95) implies a data quality check in terms of consistency and comparability over time and across countries – an issue relatively more important when looking at sub-national data.

[3] See http://www.ec.europa.eu/eurostat/ramon. Four remote French regions (FR91, FR92, FR92 and FR94) have been removed due to a lack of data. The current version of the model does not include Croatia – which is set to become the 28th Member State of the EU on 2013, July 1st.

3 Social Accounting Matrices

The Social Accounting Matrix (SAM) represents the basic structure of the model. It is a matrix-based representation of the National Accounts where the cells are the flows of economic transactions within an economy: payment flows from column accounts to row accounts. A SAM represents a macro-economic equilibrium where aggregate demand equals aggregate supply.

For the construction of all EU27 countries' SAMs, the recently published data of the EU funded World Input Output Database (WIOD) project have been used[4]. This database provides World Input-Output tables, International and National Supply and Use tables, National Input-Output tables, and Socio-Economic and Environmental Accounts covering all EU27 countries and 13 other major countries in the world for the period from 1995 to 2009 (RHOMOLO takes 2007 as the base year). The availability of WIOD data allows verifying the consistency between different data sources and facilitated the construction of reliable and consistent SAMs. For instance, interregional trade data are now also based on WIOD data (in turn, consistent with ESTAT), just like the SAMs. The database ensures now a higher level of consistency compared to previous attempts where interregional trade data were estimated by using other external trade data. Moreover, an additional advantage of WIOD data is that re-exports are subtracted from normal exports to arrive at the final value of exports. This is necessary because re-exports do not undergo any value-added processes, so cannot be counted towards a nation's or region's exports. Data are available for 59 NACE Rev. 1.1 sectors, but are aggregated in an early stage into the six macro-sectors used in RHOMOLO.

SAMs are first constructed at the national level – based on the SUTs and expanded with ESTAT data – and afterwards regionalized in the model code. The structure of a Social Accounting Matrix used in RHOMOLO is shown in Table 2.

The task of balancing SAMs - where columns equal rows - is one of the most challenging ones. WIOD facilitates the balancing process due to the fact that both Supply and Use Tables (SUTs) and Input-Output Tables (I/O) are consistent and based on ESTAT National Accounts. In turn, this makes both the international comparison and the completion of the SAMs with other ESTAT data more reliable than other CGE modelling databases.

The **use row** of the Commodities corresponds to the WIOD Use Table and contains the activities' Intermediate Use of each commodity. The **Commodities column** is the inverted WIOD Supply table and represents Domestic Supply, Taxes less Subsidies on Products, total Imports, and the total Trade and Transport margins (of the tradable commodities and services paid to Transport). The

[4] See Timmer et al. (2012, ed.) and http://www.wiod.org. Please notice that Croatia is currently not included in WIOD.

Domestic supply matrix in RHOMOLO has been diagonalized[5], which means that activities only produce one commodity in the same category – so Manufacturing activities only produces manufacturing commodities[6]. For **each activity**, the intermediate use comes, as mentioned above, from the WIOD Use table. Value added consists of employees' compensation (wages + employees social security contribution) and the net operating surplus (a residual variable for the difference between value added and labour compensation). The assumption that net operating surplus is paid from activities to the production factors means that firms are considered as a production factor. Capital consumption was not available in the WIOD SUT framework and was therefore obtained from ESTAT (consumption of fixed capital). For the **factor** column, the wages and salaries as paid to Households, Employer's Social Contributions (SC) and Taxes paid to the Government, the Operating Surplus flow to the Households are simply the sum of the corresponding row in the Activities column. Employees' SC, Personal Income Tax and Corporate Income Tax are obtained from ESTAT.

In the **Household** column, Household Savings is represented by a residual cell that balances the column and row and can be calculated after income and consumption (and tax payments) are known Expenditures Abroad are represented by Resident's purchases abroad from the WIOD Input-Output tables.

[5] This has been done by taking the sectoral gross outputs as given and adjusting accordingly the diagonal entries in the supply matrix.

[6] This is due to the standard assumption of mono-product firms in the model. In practical terms, the Domestic supply matrix is the gross output of each of the sectors diagonalized into a matrix.

Table 2 Social Accounting Matrix structure

	Commodities	**Activities**	**Factors**	**Households**	**Government**	**Savings and investment**	**RoW**	**Total**
Commodities		Intermediate demand		Consumption spending	Government spending	Investment demand	Intra- and extra-EU exports	Total use in purchasers' prices
Activities	Domestic supply							Activity income
Factors		Wages Employees' SSC Net Operating Surplus (-corp tax)						Total Factor income
Households			Wages Employees' SSC Net Operating Surplus (-corp tax)		Social transfers		Expenditure of tourists	Total Household income
Government	Taxes less subsidies on products	Net producer taxes Corp tax	Wages Employers' SSC Corporation tax Net producer taxes Taxes less subsidies on products	Employees' SCC Personal income tax			Net lending and borrowing	Government income
Savings and investment		Capital consumption		Household savings	Fiscal surplus	Net change in inventories	Capital transfers	Total savings
RoW	Intra- and extra-EU imports			Household expenditure abroad	Government transfers to abroad			Foreign exchange outflow
Trade and Transport Margins	Trade and Transport Margins							
Total	Total supply at purchasers' prices	Gross output	Factor expenditures	Total household spending	Government expenditures	Total investment spending	Foreign exchange inflow	

The expenditures of the **government** consist of consumption of goods and services, social transfers to households (as a residual cell) and a fiscal surplus investment. Government income consists of Social Contributions and Taxes.
The Savings and Investment column represents the investments in commodities and changes in inventories.

A country's Current Account balance shows up in the flow of money from Intra/Extra EU exports to the Savings Account. It is a residual of the total imports minus the total exports. Intra EU and Extra EU exports were calculated with the International Supply and Use tables. The international Supply table provided the detailed imports per country, separated by country of origin. Combining this information with the total exports per country from the International Use tables, the Intra EU vs. Extra EU imports and exports could easily be calculated.
Since RHOMOLO is a regional model a separate SAM for each region is required. National SAMs have been consistently regionalized by taking into account regional data on value added, employment and other regional and sectoral information in order to make regional SAMs more heterogeneous.

Future improvements of the SAMs include an increased sectoral breakdown (59 NACE Rev. 1.1 sectors) and a more sophisticated method of regionalization.

4 Data Domains for RHOMOLO Model

4.1 Regional GDP

Although regional GDP data are readily available from ESTAT Regional Economic Accounts, additional checks were executed in the model. By following ESA95, the different methods for calculating GDP were compared to assure consistency in the dataset, as well as a comparison with the WIOD data was performed.
Standard macro-economic GDP data – measured as the sum of final consumption, gross capital formation and exports, minus imports of goods and services - are available at national and regional data. GDP data were double-checked with available sector gross value added (GVA) data – namely the sum of gross value added and taxes, less subsidies on products – in order to build consistent and complete sector/regional figures.

Finally, additional checks on GDP and national account figures were made by looking at both DG ECFIN AMECO[7] and EU KLEMS[8] databases.

4.2 Migration and Employment Data

The current setting of RHOMOLO model assumes that net aggregate migration between any two regions is a function of expected income and distance – which, in turn, are considered proxies of the tangible and intangible costs of migration. This modelling approach to aggregate flows is rooted in the Discrete Choice theory[9]. ESTAT Regional Migration Statistics dataset provides data on *within country* migration. Key Statistics on migration in OECD countries provides data on international migration. Parameter estimates for RHOMOLO model will be based on estimating a micro-model of migration using a constructed dataset combining these two data sources.

Moreover, ESTAT information on Household Income and Active Population will be included in the analysis. Together with data on regional/sector unemployment and wages – micro-estimated via the ESTAT labour force survey – this information will provide the necessary input to the analysis of labour market and migration features in RHOMOLO model.

4.3 Innovation expenditures

R&D intensities (measured over GDP and/or Gross Value Added) are available at the national and regional level from ESTAT database (Science and Technology Indicators). National NACE sector data are available while sector within-region information is missing. In order to fill this gap, it is necessary to introduce some distributional assumptions on the R&D intensities across NACE sectors within NUTS2 regions for use in the RHOMOLO model.

First, we start from a theoretical baseline condition where the same regional R&D intensity is given to each sector within that specific region. In this way, the only source of variability is given by the regional dimension whereas the within-region sector dispersion is equal to zero by construction.

Second, we introduce distributional assumptions across NACE sectors based on the data currently available. In particular, we start from the broader taxonomy available from the ESTAT database in order to distinguish broader sectors of performance – namely governments, higher education institutions, business sector and private non-profit organizations. Given the sector taxonomy adopted in RHOMOLO (see Table 1), most of these sectors fall under the broader business sector. Then, we correct sector distribution within the broad business sector category by looking at the intersection between two further dimensions – NACE sector R&D data available at the *national* level

---

[7] See http://ec.europa.eu/economy_finance/db_indicators/ameco/index_en.html

[8] See http://www.euklems.net

[9] For more background information see Persyn, Brandsma and Kancs (2012).

corrected for *regional* gross fixed capital formation by NACE sector. In this way, we aim at identifying the best available proxy for NACE sector R&D spending at the regional level – given the current data availability[10].

Planned extension of the innovation module in RHOMOLO with additional features beyond R&D includes three elements. First, JRC-based regional patent statistics and citations offer valuable information on technological proximity across regions in Europe. Second, the inclusion of the micro-estimated data from the Community Innovation Survey may help to identify a broader set of regional innovation features – closely related to the policy domains identified in the current taxonomy of cohesion policy investments.

4.4 Human Capital Accumulation

Human capital accumulation will be introduced in the model and contribute to the analysis of both regional innovation and labour market performance. Human capital will be proxied by 3 different levels of education:

- isced0_2: Pre-primary, primary and lower secondary education - levels 0-2 (ISCED 1997)
- isced3_4: Upper secondary and post-secondary non-tertiary education - levels 3-4 (ISCED 1997)
- isced5_6: Tertiary education - levels 5-6 (ISCED 1997)

Wages will be differentiated on basis of the education levels in order to internalize the investment decision of the households on the amount of patents to be bought in each period. Data for this come from the Labour Force Survey (LFS) and the EUKLEMS database. The wage will enter the equations as income deciles.

4.5 Capital Stock

Comparable data amongst the EU27 countries on capital stock data at the regional/sectoral level of disaggregation are not available from any statistical data source.

Data are therefore built in-house by following the Perpetual Inventory Method proposed in Derbyshire *et al.* (2010). This approach begins by looking at the initial stock by country and industry (regionalized by GVA share) in the year 1995 and calculates the final capital stock by region and by

[10] Gross fixed capital formation is found to be highly correlated with R&D spending at the sectoral level. A methodological approach linking these two types of investments is described by Conte and Vivarelli (2011).

industry in 2007 by adding the yearly capital investments and depreciating the existing stock every year according to the following formula:

$$KS_{t+1} = (1 - \delta)\, KS_t + GFCF_t$$

For this approach we need the following data:

- Gross fixed capital formation (GFCF) on NUTS 2 level and by industry (A+B, C-E, F, G-I, J+K, L-P) in current prices for the years 1995-2007
- Price deflators for converting the GFCF into constant prices
- Initial stocks (KS_0) for calculating the net capital stocks for each year applying the Perpetual Inventory Method (PIM) from the EU KLEMS database. These data are available on national level, but regionalized by the GVA share.
- Depreciation rates: for applying the PIM, a depreciation rate of the fixed capital in stock is needed. This rate has been calculated by weighing the average service life of each of the six types of assets (EAS95 classification) per country.
- Gross Value Added: although not directly needed for constructing the Capital Stocks, it is necessary for filling in missing data by applying the investment-output ratio. For the years with observations, this ratio (GFCF/GVA) was calculated and growth ratio was calculated and extrapolated to be able to fill in years with missing GFCF values.

For constructing the capital stocks, the yearly investments and the initial stocks are needed. As expected, ESTAT represents the main source for investments (on NUTS 2, industry and type of asset level) data. However, the data are not complete and thus additional data sources have been consulted or data have been filled in using the investment-output ratio (on industry level). Additional databases that have been consulted in the construction process of the Capital Stocks are:

- The EU KLEMS database for the initial capital stocks, price deflators and initial capital stocks in some countries;
- AMECO database for calculating the growth rate of asset investment over total investment;
- OECD STAN Database for price deflators and initial capital stocks for some countries (Bulgaria and Luxembourg);
- National Statistical Offices: for some countries, there were no data available on especially the price deflators. In these cases, the NSOs have been consulted and price deflators have been calculated with data on volumes.

In theory, it is possible to arrive to negative capital stocks by applying PIM, if depreciation of stock is greater than the investment. Therefore, we apply an additional check in the model to indicate this.

4.6 Transport costs

Cohesion Fund investments in infrastructure (such as TEN-T) can be translated into bilateral transport costs per sector and are therefore an important part of RHOMOLO. This information is provided externally from the TRANS-TOOLS model[11], an EC-based model covering freight and passenger movements around Europe. The transport costs are measured in Euros and have been divided by the value of bilateral regional trade to which they relate to calculate the transport margins.

4.7 Cohesion Policy Investments by EU regions

In line with the on-going negotiations on the new MFF, we are currently building theoretical budgetary allocation to each region (or group of regions) for the period 2014-2020 under each broad policy domain (very similar to what is indicated, as an example, in Table 3). Together with quantitative policy targets – such as those identified in the Europe 2020 strategy – this information will help to prepare the ground for the simulation exercise to be conducted over the following months.

[11] See http://energy.jrc.ec.europa.eu/TRANS-TOOLS/TT_model.html.

Table 3 Allocation of cohesion budget by categories & territories 2007-2013 (% shares)

	Lagging Regions and countries	Non-lagging Regions	Territorial cooperation	Total
Research and technological development (R&TD) and innovation(1)	13.0	21.9	16.3	14.5
Support to firms' investments	4.0	4.5	0.4	4.0
Information Society	4.4	4.2	7.1	4.4
Transport	25.7	4.5	13.2	22.0
Energy	3.1	3.3	4.3	3.1
Environmental protection and risk prevention	16.0	6.3	16.1	14.6
Tourism	1.8	1.5	7.4	1.8
Culture	1.7	1.5	6.0	1.7
Urban and rural regeneration	2.9	3.5	2.2	3.0
Adaptability of workers and firms, enterprises and entrepreneurs	3.0	10.4	1.8	4.2
Access to employment and active and preventive labour market measures	5.0	15.3	2.5	6.6
Social inclusion of less-favoured persons(2)	1.8	9.1	0.8	2.9
Human capital (education, life-long training, high-level studies in R&TD)	7.6	8.0	3.9	7.6
Social infrastructure(3)	5.5	1.5	5.7	4.9
Partnership and networking	0.2	0.7	2.8	0.4
Institutional capacity at national, regional and local level	1.1	0.2	3.5	1.0
Reduction of additional costs of outermost Regions	0.1	0.5	0.1	0.2
Technical assistance	3.1	3.1	5.9	3.1
Total	100.0	100.0	100.0	100.0

Source: Barca Report (2009)

5 Estimation of inter-regional trade flows

Regional general equilibrium models, such as Rhomolo, require bilateral inter-regional trade flows data for the base year. Such data are not (yet) available from statistical sources. Therefore, bilateral inter-regional trade flows need to be estimated. Two types of data are of particular interest for estimating inter-regional trade flows: prior information on inter-regional trade and macro-constraints on production, consumption and trade.

For most of the EU countries (all EU Member States except Bulgaria and Romania) we base the trade flows estimation on prior information derived from the PBL (2012) dataset. The prior information is combined with the available macro-constraints, such as the distribution of production and consumption over the EU regions, given national SAMs constructed from the WIOD tables in combination with additional national accounts and international trade information.

Where no prior information on inter-regional trade is available (Bulgaria and Romania), first, the priors of inter-regional trade have to be estimated using a gravity model of trade and the available macro-constraints. The estimation of inter-regional trade priors is described in section 5.1. The inter-regional trade flows, which are consistent with other parts of Rhomolo's data base, are estimated by solving an optimisation problem minimising the error of estimated and actual trade given the available macro-constraints (Section 5.2).

5.1 *Estimation of inter-regional trade priors*

5.1.1 Available data from statistics

In order to illustrate the estimation of inter-regional trade priors, we use a simple 2x3 example (for a formal discussion see Ivanova, Kancs and Stelder 2010). The simple example consists of two countries, c1 and c2, and the rest of the word (RoW). Each country has three regions reg1, reg2 and reg3 (c1), reg4, reg5 and reg6 (c2). There is only one sector, therefore no sectoral notation is used.

SAMs (country-level data)

- The value of goods produced and consumed in each country. In our example c1 produces and consumes domestically 35, c2 – 25, and RoW – 94 units.
- The SAMs contain information also about transport and trade margins, trade and transport margin. In our example 6 for c1, 7 for c2, and 0 for RoW.

International trade data (country-level data)

- The bilateral trade flows between countries. In our example C1 exports 4 to c2, and 11 to RoW; C2 exports 10 to c1, and 5 to RoW; RoW exports 5 to c1, and 11 to c2.
- The value of goods imported into each country. In our example 56 in c1, 47 in c2, and 110 in RoW.

Regional production (regional data)

- The regional production (value added). In our example 12 in reg1, 15 in reg2, 23 in reg3, 8 in reg4, 11 in reg5, 21 in reg6.

Table 4 Available macro-constraints

	reg1	reg2	reg3	**c1**	reg4	reg5	reg6	**c2**	**RoW**	**Tot FOB**
reg1										12.0
reg2										15.0
reg3										23.0
c1				**35.0**				**4.0**	**11.0**	**50.0**
reg4										8.0
reg5										11.0
reg6										21.0
c2				**10.0**				**25.0**	**5.0**	**40.0**
RoW				**5.0**				**11.0**	**94.0**	**110.0**
Tot FOB				**50.0**				**40.0**	**110.0**	**200.0**
TT marg				**6.0**				**7.0**	**0.0**	**13.0**
Tot CIF				**56.0**				**47.0**	**110.0**	**213.0**

Note: Regional data Country data

5.1.2 Estimation of domestic sales/exports (from region to country)

Assume that the share of regional production which is sold in the country of production and which is exported to each trading partner is the same as at the national level. This implies that reg1 is exporting the same share of its output to c2 as regions reg2 and reg3. Analogously, the shares reg1, reg2 and reg3 sell in c1 are uniform, as are the shares exported to Row.

Multiplying the above shares with domestic production sold domestically and international trade flows yields domestic sales and exports from each region to each country. In our example reg1 sells 8.4 in c1, exports 1.0 to c2, and 2.6 to RoW.

Note that imposing this assumption is not a conceptual requirement. However, it puts more structure to the data. This is important, when little data is available and/or quality of the data is questionable, as in our case.

Generally, it is possible to skip step 2, and proceed directly with the estimation of gravity model of trade (step 3). In such an unconstrained gravity model, all inter-regional trade flows would solely depend on regional characteristics. Empirically, however, such an unconstrained gravity model will yield hardly explainable results, because many unobservable regional characteristics, such as differences in consumer preferences, language, culture, geo-political history, etc., are omitted from the simple gravity model (1). For example, according to the unconstrained gravity model, in relative terms a region in South UK would trade with Germany many times more than a region in North UK, because the South UK region has considerably lower trade costs to all German regions than North UK. Given that such results are not supported by the data, in the case of Rhomolo it is highly recommended to impose a structure on the shares of regional production which are sold in the country of production and which are exported to each trading partner.

Table 5 Domestic sales/exports from region to country

	reg1	reg2	reg3	**c1**	reg4	reg5	reg6	**c2**	**RoW**	**Tot FOB**
reg1				8.4				1.0	2.6	12.0
reg2				10.5				1.2	3.3	15.0
reg3				16.1				1.8	5.1	23.0
c1				**35.0**				**4.0**	**11.0**	**50.0**
reg4				2.0				5.0	1.0	8.0
reg5				2.8				6.9	1.4	11.0
reg6				5.3				13.1	2.6	21.0
c2				**10.0**				**25.0**	**5.0**	**40.0**
RoW				**5.0**				**11.0**	**94.0**	**110.0**
Tot FOB				**50.0**				**40.0**	**110.0**	**200.0**
TT marg				**6.0**				**7.0**	**0.0**	**13.0**

	reg1	reg2	reg3	c1	reg4	reg5	reg6	c2	RoW	Tot
Tot CIF				**56.0**				**47.0**	**110.0**	**213.0**

5.1.3 Estimating a gravity model of inter-regional trade

The third step consists of estimating a gravity model of inter-regional trade. According to Anderson and van Wincoop (2003), if consumers have CES preferences with common elasticity of substitution σ among all goods, the gravity equation can be expressed as:

(1)
$$P_{ij} = X_i C_j / Y_w \, (t_{ij} / \prod_i \prod_j)^{1-\sigma}$$
$$\Pi_j^{1-\sigma} = \Sigma_i \, \Pi_i^{\sigma-1} \, y_i \, t_{ij}^{1-\sigma}$$
$$\Pi_i^{1-\sigma} = \Sigma_j \, \Pi_j^{\sigma-1} \, y_j \, t_{ji}^{1-\sigma}$$

where P_{ij} are the inter-regional trade priors we are interested in, X_i and C_j are production and consumption in origin region *i* and destination region *j*, respectively, y is the corresponding share, Y_w is total production of all trading partners, t_{ij} are trade costs between *i* and *j*, and $\prod_i$ and $\prod_j$ are price indices, which often are referred to as multilateral resistance.

Data on regional production and consumption, X_i and C_j, are available from the regional SAMs, inter-regional trade costs, and t_{ij}, are readily available from the TRANS-TOOLS data base. Values for the elasticity of substitution, σ, are taken from the literature (Anderson and van Wincoop 2003; and Feenstra 2002).

Table 6 The estimates of inter-regional trade priors

	reg1	reg2	reg3	**c1**	reg4	reg5	reg6	**c2**	**RoW**	**Tot FOB**
reg1	2.8	2.9	2.7	8.4	0.3	0.3	0.3	1.0	2.6	12.0
reg2	2.8	3.6	4.2	10.5	0.4	0.4	0.4	1.2	3.3	15.0
reg3	5.9	6.3	3.9	16.1	0.7	0.6	0.5	1.8	5.1	23.0
c1	**11.5**	**12.8**	**10.7**	**35.0**	**1.5**	**1.3**	**1.2**	**4.0**	**11.0**	**50.0**
reg4	0.7	0.7	0.5	2.0	2.2	1.5	1.3	5.0	1.0	8.0
reg5	0.8	1.1	0.8	2.8	2.1	1.9	2.9	6.9	1.4	11.0
reg6	1.8	1.8	1.7	5.3	3.5	5.8	3.9	13.1	2.6	21.0
c2	**3.3**	**3.6**	**3.0**	**10.0**	**7.8**	**9.2**	**8.1**	**25.0**	**5.0**	**40.0**
RoW				**5.0**				**11.0**	**94.0**	**110.0**
Tot FOB				**50.0**				**40.0**	**110.0**	**200.0**
TT marg				**6.0**				**7.0**	**0.0**	**13.0**
Tot CIF				**56.0**				**47.0**	**110.0**	**213.0**

According to the gravity model estimates, in our example reg1 sells 2.8 locally, exports 2.9 to reg2, 2.7 to reg3, 0.3 to reg4, 0.3 to reg5, and 0.3 to reg6 (Table 3). Note that in our one sector economy expenditure is equal to output.

Note that the above gravity model is only one among many potential gravity models that can be used to estimate inter-regional trade flows. This particular gravity model, however, is well founded theoretically and is widely used in the literature, which explains our choice.

5.1.4 Regionalising the trade margins and imports from RoW

The next step involves regionalisation of imports from RoW, and splitting the national trade and transport margins (from SAMs) by importing region.

The most natural way how to disaggregate the imports from RoW is to assume the same regional import shares as of regional imports from other countries. According to Table 4, in our example reg1 imports 1.6, reg2 imports 1.8, and reg3 imports 1.5 from RoW.

Analogously, the most natural way how to disaggregate the national trade and transport margins (from SAMs) by importing region is to proportionate these to the total import shares. According to Table 4, in our example reg1 pays 2.0, reg2 pays 2.2, and reg3 pays 1.8 for transportation services of importing goods.

Table 4. Regionalising trade margins and imports from RoW

	reg1	reg2	reg3	**c1**	reg4	reg5	reg6	**c2**	**RoW**	**Tot FOB**
reg1	2.8	2.9	2.7	8.4	0.3	0.3	0.3	1.0	2.6	12.0
reg2	2.8	3.6	4.2	10.5	0.4	0.4	0.4	1.2	3.3	15.0
reg3	5.9	6.3	3.9	16.1	0.7	0.6	0.5	1.8	5.1	23.0
c1	**11.5**	**12.8**	**10.7**	**35.0**	**1.5**	**1.3**	**1.2**	**4.0**	**11.0**	**50.0**
reg4	0.7	0.7	0.5	2.0	2.2	1.5	1.3	5.0	1.0	8.0
reg5	0.8	1.1	0.8	2.8	2.1	1.9	2.9	6.9	1.4	11.0
reg6	1.8	1.8	1.7	5.3	3.5	5.8	3.9	13.1	2.6	21.0
c2	**3.3**	**3.6**	**3.0**	**10.0**	**7.8**	**9.2**	**8.1**	**25.0**	**5.0**	**40.0**
RoW	1.6	1.8	1.5	**5.0**	3.5	4.0	3.5	**11.0**	**94.0**	**110.0**
Tot FOB	**16.5**	**18.3**	**15.3**	**50.0**	**12.8**	**14.4**	**12.8**	**40.0**	**110.0**	**200.0**
TT marg	2.0	2.2	1.8	**6.0**	2.2	2.5	2.2	**7.0**	**0.0**	**13.0**
Tot CIF	**18.4**	**20.5**	**17.1**	**56.0**	**15.0**	**17.0**	**15.1**	**47.0**	**110.0**	**213.0**

Regional trade margins and regional imports from RoW were the last missing entries for completing the full inter-regional trade matrix. Having this information, one can estimate the total consumption and imports for each region in CIF prices. According to Table 4, in our example the total consumption of reg1 is 18.4, 20.5 of reg2, and 17.1 of reg3.

5.2 The estimation of inter-regional trade flows

5.2.1 The estimation problem

The prior information on inter-regional trade, P_{ij}, which is equal to the trade from region i to region j, is obtained from the PBL data base for 25 EU Member States. For Bulgaria and Romania we estimate the inter-regional trade priors according to methodology described in section 5.1.

In order to estimate the trade flows between the different regions given the prior information, first, we introduce the following two new priors that give the *relative* trade information such that the procedure takes possible differences in the overall totals of regional production and consumption into account. These relative priors Px_{ij} and Pc_{ij} are taken relative to the production and demand, respectively:

$$Px_{ij} = \frac{P_{ij}}{\sum_j P_{ij}} \qquad Pc_{ij} = \frac{P_{ij}}{\sum_i P_{ij}} \tag{1}$$

The objective function used to estimate the trade flows T_{ij} is given by the following equation:

$$Min\,Z = \sum_{ij}\left[\left(Px_{ij} - \frac{T_{ij}}{X_i}\right)^2 + \left(Pc_{ij} - \frac{T_{ij}}{C_j}\right)^2\right] + \frac{1}{\left(\frac{\sum_i X_i}{\#i}\right)^2}\sum_{ij}\left(X_i Px_{ij} - T_{ij}\right)^2 + \frac{1}{\left(\frac{\sum_j C_j}{\#j}\right)^2}\sum_{ij}\left(C_j Pc_{ij} - T_{ij}\right)^2 \tag{2}$$

, where $\#i$ is the number of regions. This objective function consists of three parts. The first part describes the quadratic relative error of the final trade matrix in relation to the prior information. The second and third parts describe the absolute errors which are rescaled such that they have the same weight in the objective function as the relative errors. Note that these trade data include the diagonal and therefore take the cross-hauling of trade into account as well.

We have chosen a quadratic objective function, because it is convenient in solving very large optimisation problems. The quadratic function can be solved as either a conic or a quadratic program which are much faster to solve than the nonlinear logarithmic objective function. An entropy based optimisation problem would involve changing the objective function into a logarithmic function. This is easy to change from a programming perspective, but cumbersome from a computational perspective.

In order to obtain the trade matrix consistent with the regional consumption and production figures, we add the following two constraints to the minimisation problem:

(3) $$X_i = \sum_j T_{ij}$$

(4) $$C_j = \sum_i T_{ij}$$

Note that one of these constraints can be skipped for one region because the system would otherwise be over-determined.

5.2.2 Additional constraints

Given the national SAM, next we add additional constraints where needed. For instance, we add the country constraints such that the regional trade adds up to the country trade T_{od}^c between origin country o and destination country d:

(5) $$T_{od}^c = \sum_{i \in o} \sum_{j \in d} T_{ij}$$

Note, however, that this also adds information on the national flows, which if not properly taken into account would result in a bias in the estimate. The estimate should also be on the national level if the international trade is taken as given. The complete optimisation problem together with constraint (5) changes to:

(6) $$Min\, Z = \sum_{ij} \left[\left(PxN_{ijn} - \frac{T_{ij,j \in n}}{X_{i,n}} \right)^2 + \left(PcN_{ijn} - \frac{T_{ij,i \in n}}{C_{n,j}} \right)^2 \right] + \frac{1}{\left(\frac{\sum_i X_{i,n}}{\# i} \right)^2} \sum_{ij} \left(X_{i,n} PxN_{ijn} - T_{ij} \right)^2 + \frac{1}{\left(\frac{\sum_j C_{j,n}}{\# j} \right)^2} \sum_{ij} \left(C_{j,n} PcN_{ijn} - T_{ij} \right)^2$$

, where subscript n refers to the set of nations. Together with (5) we have the following set of additional constraints:

(7) $$X_i = \sum_j T_{ij}$$

(8) $$C_j = \sum_i T_{ij}$$

(9) $$X_{i,n} = \sum_{j \in n} T_{ij}$$

(10) $$C_{n,j} = \sum_{i \in n} T_{ij}$$

The relative probability priors Px_{ij} and Pc_{ij} are:

$$PxN_{ijn} = \frac{P_{ij}}{\sum_{j \in n} P_{ij}}$$

$$PcN_{ijn} = \frac{P_{ij}}{\sum_{i \in n} P_{ij}}$$

As a result, a fully completed, consistent and balanced inter-regional trade matrix is obtained, which can be readily used as data input in Rhomolo. Note that this estimation procedure needs to be performed only once (not for each model run).

6 Parameters

In a CGE model setting, there are generally more parameters to calibrate than observations. Hence some of the parameters need to be fixed exogenously. In RHOMOLO many of the model elasticity parameters are estimated econometrically using either time-series or panel data techniques, typically based on European data from 1995 and onwards. This includes the parameters of the sectoral production function as well as equations related to modelling of inter-regional trade, migration and capital flows/investments. Other parameters are taken from literature, such as the elasticity parameter for the CES utility function. The level parameters have been set exogenously to replicate the base year.

7 Future data requirements

The RHOMOLO will remain updated with new data regularly. Some future update for the longer term consist of a new base year (most likely 2009 - when sufficient reliable data are available) and the update from NACE Rev. 1.1 to Rev. 2. An example of a mid-term adjustment is the introduction of three different types of labour skills.

The future membership to the European Union of Croatia is also a point of interest. Unfortunately, no SAM is available for this country, which complicates the introduction. In addition, the coverage of other variables used in the model is very scarce for Croatia and there is decided not to include it yet. However, the final objective is to include Croatia to RHOMOLO in due time.

References

Anderson, James E. and Eric van Wincoop (2003), “Gravity with Gravitas: A Solution to the Border Puzzle,” American Economic Review, 93, 170-192.

Anderson, James E. and Eric van Wincoop (2004), “Trade Costs,” Journal of Economic Literature, 42(3), 691-751.

Brandsma, A., Kancs, D., Monfort, P. and Rillaers, A. (2013). Rhomolo: A Regional-based Spatial General Equilibrium Model for Assessing the Impact of Cohesion Policy. JRC-IPTS Working Paper Series JRC2013, European Commission, DG Joint Research Centre.

Conte, A. & Vivarelli, M. (2011), Imported Skill-Biased Technological Change in Developing Countries, *Developing Economies,* **2011***, Vol. 49*, pp. 36-65

Derbyshire, J. and Gardiner, B. and Waights, Sevrin (2013) Estimating the capital stock for the NUTS2 regions of the EU27. Applied economics, 45 (9). pp. 1133-1149.

EU KLEMS (2008). EU KLEMS Growth and Productivity Accounts: March 2011 Update. Productivity in the European Union: A Comparative Industry Approach (EU KLEMS2003) funded by the European Commission, Research Directorate General http://www.euklems.net/index.html (Accessed December 2011).

European Commission. (2008). Ameco database. DG Ecfin, Brussels. (Accessed December 2011).

European Commission. (1996). European System of Accounts ESA 1995. Eurostat, Brussels-Luxembourg.

European Commission. (2009). Barca Report. An Agenda for a Reformed Cohesion Policy, Brussels.

European Commission (2010). TRANS-TOOLS. Funded by the JRC (IPTS) and DG TREN. (Accessed August 2012).

Eurostat (2008) Eurostat Manual of Supply, Use and Input-Output Tables (Luxembourg, Eurostat).

Feenstra, Robert C. (2002), “The Gravity Equation in International Economics: Theory and Evidence,” Scottish Journal of Political Economy, 49(5), 491-506.

Ivanova, O., D. Kancs, D. Stelder (2010), Modelling Inter-Regional Trade Flows: Data and Methodological Issues in Rhomolo, ERSA Conference Papers 10/500, European Regional Science Association.
www-sre.wu.ac.at/ersa/ersaconfs/ersa10/ERSA2010finalpaper500.pdf

JRC (2012), RHOMOLO Model Manual – Technical User Version 2. Institute for Prospective Technological Studies, Seville, Spain.

Kapur, J.N., H.K. Kesavan (1992), “Entropy Optimization Principles with Applications.” Academic Press, New York, 1992.

OECD (1999). The OECD STAN database for industrial analysis, OECD, Paris. (Accessed January 2012).

OECD (2011). International Migration Outlook: Sopemi. OECD.

Persyn, D., Kancs, d., and Brandsma, A. (2012), "Modelling Migration and Regional Labour Markets: an application to the regional CGE model RHOMOLO", mimeo.

Planbureau voor de Leefomgeving (PBL) 2012. “Inter-regional trade matrices for Europe,” The Hague, Netherlands.

Thissen, M.J.P.M., F.G. van Oort, D. Diodato and A. Ruijs (2013 forthcoming), Regional Competitiveness and Smart specialization in Europe: Place based Development in International Economic networks, Edward Elgar, Cheltenham: UK.

Timmer, M. (2012). The World Input-Output Database (WIOD): Contents, Sources and Methods. WIOD Background document available at www.wiod.org.

WIOD Project (2012). World Input-Output Database. A project funded by the European Commission, Research Directorate General as part of the 7th Framework Programme, Theme 8: Socio-Economic Sciences and Humanities. (Accessed in May 2012).

Annex I: Model Variables[12]

Variable Name	Description
$AUXV_{i,r}$	Auxiliary variable
$C_{th,l,r}$	Demand for consumer goods and leisure
$CBUD_{th,r}$	Consumer expenditure commodities
$CG_{i,r}$	Intermediate public demand for goods of national government
$CGR_{i,r}$	Intermediate public demand for goods of regional government
$CV_{th,r}$	Welfare change (compensating variation) as a percentage of households income
$CV\_ENV_{th,r}$	Monetary value of improvements in ecosystem
$EEU27_{i,r}$	Exports to EU27
$ENER_{i,r}$	Energy input
ER_c	Euro exchange rate
$EROW_{i,r}$	Exports to RoW
ET_c	Total exports
GDP_C_c	Gross domestic product (real)
$GDPC\_C_c$	Gross domestic product (nominal)
$GDPDEF\_C_c$	Country level GDP deflator
$GDPR_r$	Gross regional product (real)
$GDPRC_r$	Gross regional product (nominal)
$I_{i,r}$	Demand for investment goods private
$INDEX_r$	Consumer price index
$INDEXE_c$	Price index for exports
$INDEXE_m$	Price index for imports
$INV_{i,r}$	Sectoral investment
IT_c	Total private investment
$K_{i,r}$	capital input
$KL_{i,r}$	Capital-labour bundle
$KLE_{i,r}$	Energy-capital-labour bundle
KS_r	Capital endowment (exogenous)
$L_{i,ed,r}$	Labour input
$LMIG_{th,r}$	Net labour migration from or to region
$LROW_{ed,r}$	Labour supplied to RoW (exogenous)
$LS_{ed,r}$	Labour supply

[12] For a more detailed account of the variables in the model, see RHOMOLO Model Manual (JRC, 2012).

Variable Name	Description
$LSH_{ed,r}$	Human capital stock
$LT_{i,r}$	Composite labour input
$MEU27_{i,r}$	Imports from EU27
$MEU27\_share_{i,r}$	Share of imports from rest of EU
$MROW_{i,r}$	Imports from RoW
$MROW\_share_{i,r}$	Share of imports from rest of the world
MT_c	Total imports
$NF_{i,r}$	Equilibrium number of monopolistic firms in an industry
$P_{i,r}$	Domestic sales prices of commodities and price of leisure
$PCV_{th,r}$	Compensating variation price index
$PD_{i,r}$	Domestic producer prices of commodities
$PDC_{i,r}$	Monopolistic competition prices of commodities
$PDDT_{i,r}$	Composite price of domestic goods
PI_r	Price of private investments
$PK_{i,r}$	Price of capital
$PKL_{i,r}$	Price of composite labour-capital bundle
$PKLE_{i,r}$	Price of composite energy-labour-capital bundle
$PL_{i,ed,r}$	Domestic price of labour (labour compensation per employee)
PLROW	Price of labour supplied to RoW (exogenous)
$PLT_{i,r}$	Price of composite labour bundle
$POLL\_CO2_{i,r}$	Energy-related CO2 pollution in mil tons
$POLL\_CO2\_PROD_{i,r}$	Output-related CO2 pollution in mil tons
$POLL\_GHG_{emis,i,r}$	Energy-related GHG (non-CO2) pollution in mil tons
$POLL\_GHG\_PROD_{emis,i,r}$	Output-related GHG (non-CO2) pollution in mil tons
$POLL\_NGHG_{emis,i,r}$	Energy-related non GHG pollution in mil tons
$POLL\_NGHG\_PROD_{emis,i,r}$	Output-related non GHG pollution in mil tons
$PROFITS_{i,r}$	Profits of the sectors
PTM_c	Composite price of trade and transport margin
$RD_{r???}$	R&D expenditure???
RGD_c	Nominal interest rate
$RK_{i,r}$	Return to capital
$RORLAG_{i,r}$	Lagged rate of return
S_c	National savings
$SEU27_c$	Savings of or from EU27 (exogenous)
SG_c	Government savings on national level (exogenous)
SGR	Government savings on regional level (exogenous)
$SH_{th,r}$	Household savings

Variable Name	Description
$SHWTREAT_{r,wtreat,tw}$	Share of waste by type of treatment
$SII_{th,r}$	Compensating variation budget
$SROW_{c}$	Savings of or from RoW (exogenous)
$SUBS_{c}$	Total subsidies of national government
$SUBSR_{r}$	Total subsidies of regional governments
$SV_{i,r}$	Changes in stocks
$SWF\_country_{c}$	Country level welfare function
$SWF\_ENV\_country_{c}$	Country level welfare function related to changes in ecosystem
$TAXR_{c}$	Tax revenues of national government
$TAXRR_{r}$	Tax revenues of regional governments
$TMX_{i,r}$	Commodity consumed for prod of transport and trade margins
$REU27_{c}$	Total transfers to government from EU27 (exogenous)
$TREU27R_{r}$	Total transfers to regional governments from EU27 (exogenous)
$TRF_{th,r}$	Total transfers from national government to households (exogenous)
TRFNGv	Transfers from national government to regional ones
$TRFR_{th,r}$	Total transfers from regional governments to households (exogenous)
$TRFRG_{r}$	Transfers from regional governments to national one
$trmV_{rr,r,i}$	Rreight transport costs
$U_{th,r}$	Regional utility
$UNEMP_{ed,r}$	Regional unemployment level
$UNRTE_{ed,r}$	Regional unemployment rate
$Walras_{c}$	Walras law
$X_{i,r}$	Domestic sales (domestic and foreign origin)
$XD_{i,r}$	Gross domestic output
$XDDE_{i,r,rr}$	Inter-regional trade flows
$XDDE_{i,r,rr}$	Inter-regional trade shares
$Y_{th,r}$	Household income

Source: Model Manual

European Commission
EUR 25955 EN – Joint Research Centre – Institute for Prospective Technological Studies

Title: **Data Needs for Regional Modelling**

Author(s): Lesley Potters, Andrea Conte, d'Artis Kancs, Mark Thissen

Luxembourg: Publications Office of the European Union

2014 – 23 pp. – 21.0 x 29.7 cm

EUR – Scientific and Technical Research series –ISSN 1831-9424 (online)

ISBN 978-92-79-29737-3 (PDF)

doi:10.2791/10249

Abstract

RHOMOLO is a Computable General Equilibrium (CGE) model aimed at investigating the impact of policy instruments developed according to the objectives of Cohesion Policy. In doing so, this model attempts to cover a broader set of economic variables mostly related to an economic geography setting while aiming at a complete regional and sectoral coverage for the EU 27 – by means of a dataset covering 267 EU NUTS2 regions with 6 NACE Rev. 1.1 sectors. This working paper aims at describing the data used in the EU27 RHOMOLO Model.

LF-NA-25955-EN-N

JRC Mission

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new methods, tools and standards, and sharing its know-how with the Member States, the scientific community and international partners.

Serving society
Stimulating innovation
Supporting legislation

Publications Office

doi:10.2791/10249
ISBN 978-92-79-29737-3