PortableDyme – A Simplified Software Package for Model Building

Anett Großmann[1,*], Frank Hohmann[1,**]

[1]Gesellschaft für wirtschaftliche Strukturforschung mbH, Osnabrück, Germany

[*]Corresponding author. Tel.: +49 541 40933 180, Fax +49 541 40933 110,

[**]Corresponding author. Tel.: +49 541 40933 130, Fax +49 541 40933 110,

E-Mail: grossmann@gws-os.com, hohmann@gws-os.com

Abstract

Macroeconomic multisectoral models have proven to be powerful tools for forecasting, simulation and policy analysis. Such models are based on huge datasets, contain hundreds to thousands of (interdependent) equations and therefore tend to be inherently complex. Building such models is by no means a simple task. Model builders not only have to have a profound knowledge of economic theory but also have to get acquainted with software packages and programming languages.

PortableDyme (Großmann, Hohmann, Wiebe 2012) has been designed to make the whole model building process easier especially for beginners. From the technical point of view, PortableDyme is a complete software package for model building. It is based on the powerful INFORUM set of programs and has been pre-configured to run on any Microsoft Windows computer without installation.

To assist beginners in model building, PortableDyme also contains a basic but extensible macroeconomic multisectoral model. The structure of the model reflects the different steps of model building – starting from collecting historical data up to performing scenario analysis. Most of the data is taken from the WIOD (World Input Output Database) which provides a harmonized dataset for 40 countries. For these countries, adapting the model can be accomplished within a few hours.

The aim of this paper is to introduce the main buildings blocks of the PortableDyme programming environment and to show how this software can be used to build large-scale econometric input-output models.

The PortableDyme model will be presented as part of the software session conducted by Douglas Meade (INFORUM, USA).

CONTENT

1   Introduction

Macro-econometric models have proven to be invaluable tools for forecasting and scenario analysis. Sophisticated economic modelling is achieved by rich datasets, combining econometric-statistical analysis with input-output analysis and a high degree of endogenization.

Beginners in model building not only have to have profound knowledge of economics and statistics but also have to cope with issues related to information technology such as data file formats and programming languages. Furthermore, they have to develop a conceptual model design which allows for evaluation, maintenance, extensions and updates of the model.

PortableDyme[1] (Großmann, Hohmann, Wiebe 2012) addresses these issues. It provides a complete model building framework containing everything that is needed to build comprehensive macro-econometric input-output (IO) models.

The software is based on the INFORUM (see http://www.inforum.umd.edu) set of programs in combination with open-source and other freely available software and has been pre-configured to run on Microsoft Windows computers without installation.

The framework not only contains the necessary software but also comes with a step-by-step approach for the model building process. Users may either choose to adopt the included small macro-econometric IO model or – if they are more experienced – may start from scratch by selecting the vanilla version which does not contain any data, regressions and programming statements.

The small macro-econometric IO model included in PortableDyme is based on the WIOD (World Input-Output Database) dataset. Due to the harmonized WIOD data structure, the model may be adopted to other countries within a few hours.

Section 2 describes PortableDyme's core software components and the step-by-step modelling approach which is reflected by its directory structure.

In Section 3, the building blocks of the macro-econometric IO model template are discussed in detail. The section contains an example for applying the model for scenario analysis by performing and comparing variations of population forecasts for Germany.

Section 4 draws some conclusions and gives hints for future developments.

---

[1] The PortableDyme model building framework is available free of charge for INFORUM members.

2 PortableDyme Model Building Framework

2.1 *Framework Components*

From a technical point of view, the code for a macro-econometric IO model could be developed in almost any computer language as long as this language provides loop and branch statements, file access and sub-routines. For input-output modelling and regression analysis, the language should support time-series data and vector/matrix algebra.

The computer language is only one component of a complete model building environment though: Source data is usually spread over dozens or even hundreds of input files. Combining them into a unified database is essential not only for performance reasons but also for consistency checks and evaluation purposes. A regression program is required to integrate behavioural equations. An advanced editor program with project management features is an important prerequisite to keep track of source files, variables and their usage. For evaluation and policy analysis, the model building environment has to provide tools to generate tables and graphs, ideally by connecting to spreadsheet programs like Microsoft Excel.

PortableDyme is built on top of the INFORUM package of programs, which have been in use for more than two decades to develop sophisticated macroeconomic multisectoral models. The core of the INFORUM software is written in the programming language C++ which ensures outstanding performance and scalability. The Interdyme programming library defines the building blocks of a model – time series, vectors, matrices, equations – which are almost as easy to use as built-in types and is capable of dealing with millions of time series efficiently (INFORUM 2008). The G7 regression software is seamlessly integrated into the Interdyme programming library and offers read/write access to Microsoft Excel workbook files (INFORUM 2011).

In addition to the INFORUM software, PortableDyme comes with a free C++ compiler, an advanced editor with project management features, a lightweight Microsoft Excel-based database manager and some other productivity tools. To help beginners to keep things organized, the editor comes with a predefined project template which reflects the different steps of the model building process. The PortableDyme model building framework does not require any installation and is pre-configured to run on any Microsoft Windows computer (Windows XP or later): "Portable" means that sharing a model between different persons/computers is accomplished by copying the contents of the main PortableDyme directory.

2.2 Step-by-step modelling approach

PortableDyme models come with a pre-configured model structure which is reflected both on the storage media and within the editor software (see Figure 1):
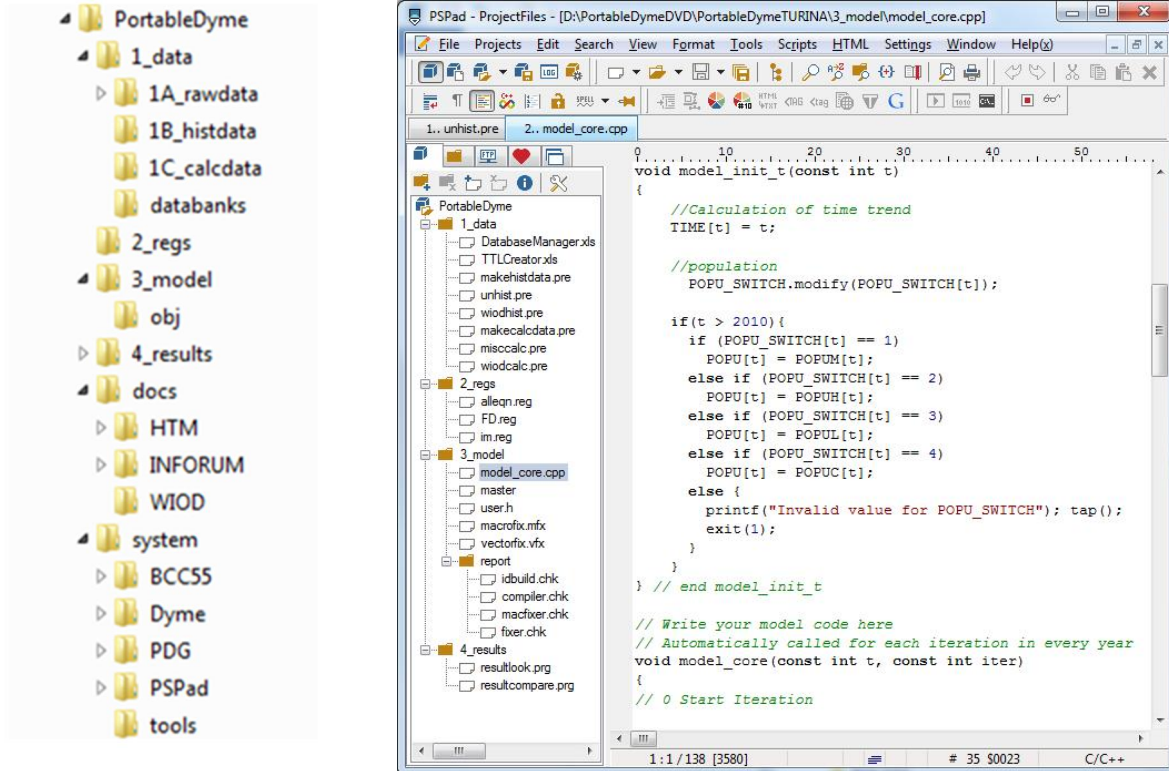

Figure 1: **Structure of PortableDyme models**

Source: own illustration

For each step, the model building framework contains essential source code files which may be tailored by the user. A detailed explanation of each step and its contents is beyond the scope of this paper and can be found in the PortableDyme documentation (Großmann, Hohmann, Wiebe 2013). Basically, the model building process can be divided in four steps:

1.   Setting up and completing the historical database (1_data, see Figure 1)

2.   Performing regressions based on the historical database (2_regs)

3.   Building and running the model by combining behavioral equations and definitions (3_model).

4.   Performing scenario analysis and analyzing the results (4_results)

The folder *docs* includes documentation about the dataset, Interdyme software and PortableDyme.

The folder *system* includes the software framework which is necessary to build and run the model, e. g. the Dyme software package, C++ compiler etc.

3 PortableDyme model template

3.1 Overview

The PortableDyme template for a basic, *dynamic macro-econometric input-output forecasting and simulation model* serves as a blueprint for all 40 countries that are part of the WIOD (see Figure 2).

Code	Coutry	Code	Coutry	Code	Coutry	Code	Coutry
AT	Austria	EE	Estonia	JP	Japan	RO	Romania
AU	Australia	FI	Finland	KR	Korea	RU	Russia
BE	Belgium	FR	France	LV	Latvia	SK	Slovak Republic
BR	Brazil	DE	Germany	LT	Lithuania	SI	Slovenia
BG	Bulgaria	GR	Greece	LU	Luxembourg	ES	Spain
CA	Canada	HU	Hungary	MT	Malta	SE	Sweden
CN	China	IN	India	MX	Mexico	TW	Taiwan
CY	Cyprus	ID	Indonesia	NL	Netherlands	TR	Turkey
CZ	Czech Republic	IE	Ireland	PL	Poland	GB	United Kingdom
DK	Denmark	IT	Italy	PT	Portugal	US	United States of America

Figure 2: WIOD country list

Source: WIOD

A macroeconomic model should consider the economic circuit from production to consumption and the monetary flows between different sectors (e. g. private households, government, corporations). This information is part of the national accounts at an aggregated level (*top-down* approach). Detailed information at industry level can be derived from the IO tables (*bottom-up* approach). So, the model template combines *bottom-up* with *top-down* approaches (Almon 1991).

Furthermore, PortableDyme models combine *IO analysis* with *econometric-statistical analysis*. They are based on the IO framework. Production, intermediate demand and value added are derived from demand side by using the Leontief inverse. This approach allows for considering direct and indirect effects of demand side driven impulses.

National account variables can be derived from aggregated IO variables such as production, private consumption, etc. because the System of National Accounts and Balancing Items (SNAB) is fully consistent with the ESA (European System of National and Regional Accounts), which encompasses the IO framework, too. Other variables e. g. all balancing items like disposable income or net lending/net borrowing are deduced from interrelations within the national accounts. The double entry accounting also requires consistency for a transaction. For example, households and firms pay taxes which the government receives.

Other variables which cannot be forecasted by definitions have to be estimated. In PortableDyme models, most of the regression equations are based on causal regression analysis. Here, observed relationship between the dependent variable and independent variable(s) are validated with statistical methods. It is assumed that reactions which were observed for the past are also effective in the future. In contrast to time series analysis, regression analysis is not used to predict changes between dependent and independent variables over time. It hast to be pointed out that time series analysis can be applied to capture autonomous trends as well.

In contrast to some CGE models or optimizing models, agent's behaviour are tested by using econometric methods instead of using elasticities from literature and calibrating the model to a base year dataset (Capros et al. 2013). In PortableDyme models, coefficients are empirically validated. A prerequisite is a data set with many observations for variables that are included in regressions. Other econometric models such as E3ME from Cambridge Econometrics and the Austrian model e3.at also apply empirical methods to derive country-specific elasticities (Cambridge Econometrics 2014, Stocker et al. 2014).

PortableDyme models are driven by exogenous impulses (e. g. population dynamics, world trade imports from global forecasting models such as the IMF forecast) and therefore could be classified as *dynamic* models*:* Such an impulse activates responses in the interdependent and non-linear equation system by using the Gauss-Seidel technique which solves the left hand side of any equation using current values for the right hand side variables. The equation system is solved if a certain convergence criterion is met. This criterion has to be an endogenously calculated variable. PortableDyme uses the endogenously calculated production variable as the criterion: The model has converged if the percentage deviation is smaller than 0.01 %. This solution algorithm is applied year by year for a given time span.

The outcome of the model solution process is the so-called 'business-as-usual' (BAU) scenario. The forecasted values are based on empirically validated agents behavior and definitions for variables represented in the model as well as exogenous assumptions.

Scenario analysis is carried out to examine the impact of changes to variable values (e. g. population forecasts, world import demand) on the economy. All other variables are affected endogenously according to the modelling context. The results of a *simulation* can be compared to the BAU scenario and differences between both scenarios can be interpreted as responses of the exogenous impulse(s) ('what if' - analysis). PortableDyme models are equipped with some helpful evaluation tools to analyze data for one model run or to compare results from different model runs.

3.2 Dataset

The PortableDyme model template is based on data from WIOD (http://www.wiod.org) and the United Nations (UN, http://data.un.org):

- National input-output tables in nominal values and previous year prices (WIOD),
- Output prices for the respective country and USA to calculate real exchange rates (WIOD),
- Population data (UN),
- SNAB (UN).

The aforementioned data have to become part of the unified historical model database (see section 2.2, step 1_data). Each variable has to be extracted from the downloaded Excel-files, stored in the model database and be given a name. Although variables can be named in any manner, it is preferable to use some sort of naming convention which helps with understanding what a certain variable name actually means. For SNAB data the official SNA item code (see Statistical Commission 2008[2]) should be used. An additional character identifies cash flow or balancing item (u - used, r - received, b - balancing item). Another character describes the sector code: t - total, h - households, g- government, etc.

For data without any official code, the variable name can be build by combining the first letter of every noun plus a character for nominal, real values or prices. A concept like "household consumption expenditures" becomes "hce".

3.3 Model template

The core of PortableDyme models are IO tables. Final demand by components is part of the system as well as intermediate demand, value added and output by industries. The SNAB is part of the model context to encompass the whole economy and to calculate e. g. disposable income and net lending/net borrowing.

IO tables and the SNAB are, for example, connected through consumption and investments. Industry data from IO tables are added up and are equal to the respective value at the macroeconomic level as stated in the SNAB. Within the SNAB, all balanced items like disposable income are calculated by definition. The disposable income is used for estimating real household consumption expenditures.

---

[2] http://unstats.un.org/unsd/statcom/doc08/SNA-Annex1.pdf

An overview of the model template with all relations between the components is shown in Figure 3. All variables except the exogenous variables such as population, exports and consumer price indices are calculated endogenously, that means they are dependent on other model variables.
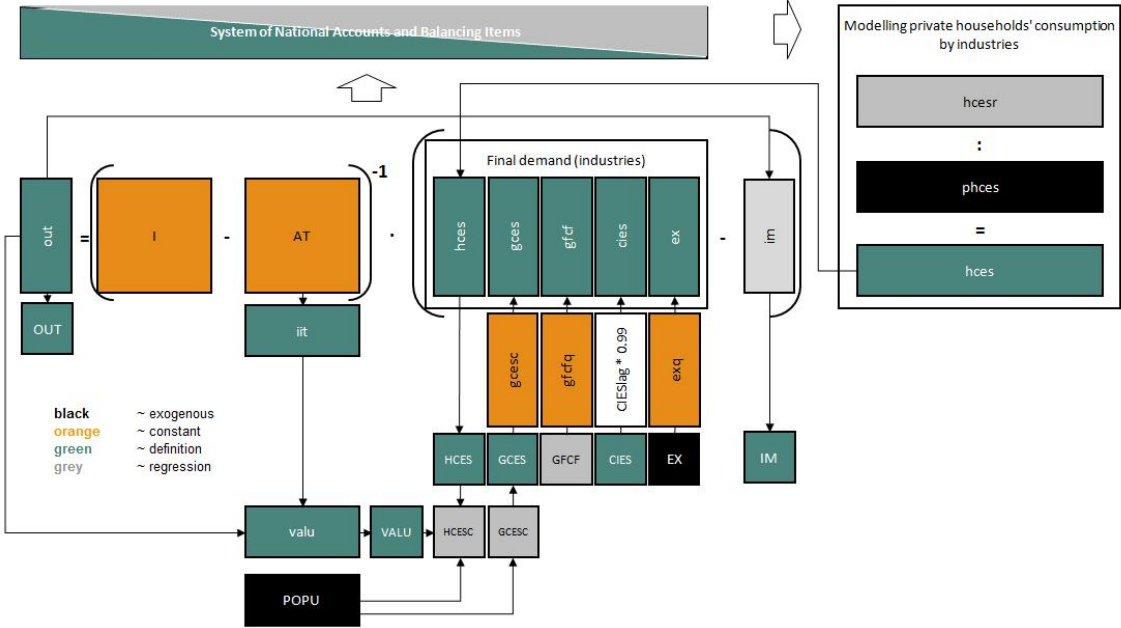

Figure 3:   PortableDyme model template

Source: own illustration

So far, the basic model is based on nominal values with one exception: private consumption expenditures are expressed in nominal values, year over year values and price indices. The integration of prices and real values is a first step to enhance the model.

3.3.1 Input-Output Calculations

The PortableDyme template is a demand-side driven model. Final demand components are estimated first (see section 2.2, step 2_regs). Exports *EX*, gross fixed capital formation *GFCF* and government consumption expenditures *GCES* are determined at an aggregated level and then broken down to industries (top-down).

Government consumption expenditures per capita are estimated with a time trend as explaining variable and then multiplied by population to calculate total government consumption expenditures. Gross fix capital formation is dependent from output. Exports are set exogenously.

For household consumption expenditures, changes in inventories and imports a bottom-up approach is applied. Household consumption *hces* and imports *im* by industries are estimated and added up to overall household consumption respectively imports. For changes in inventories *cies,* it is assumed that they will tend to zero, because storage is not for free and firms trying to minimize their costs.

Imports by industries are estimated with output by industries. Explanatory variables for private consumption are real disposable income and relative prices.

Increasing real disposable income will affect consumption positively. The relation between relative prices and real consumption expenditures is negative. Relative prices show the relation between an average consumer price index on industry level and the overall consumer price index.

Consumer price indices *phces* need to be set exogenously. In this simplified model template, a regression approach is not possible due to missing unit costs which are needed to explain price dynamics. Future price development by industries is calculated by trend extrapolation. In other words: A more or less constant cost structure and a similar change in cost components (e. g. capital, labor, energy) is expected to be effective in the future. For sectors that are strongly influenced by world market prices (e. g. coke and refined petroleum as well as agriculture and food sector) forecasts of the International Energy Agency (IEA) regarding energy prices and food prices from Food and Agriculture Organization (FAO) are used.

Disposable income is calculated within the SNAB as a residual variable (see section 3.3.2). It is deflated with the overall consumer price index and real disposable income can be derived.

After calculating all final demand components, sectoral output values *out*, sectoral intermediate input *iit* and also sectoral value added *valu* can be computed using the total coefficient matrix *AT*, and, hence, also total gross value added *VALU*, total output *OUT* and total intermediate consumption *IIT*.

3.3.2 System of National Accounts and Balancing Items

The SNAB is a system of consistent accounts and balance sheets that systematically describes the complete economy at the different stages of the economic process and records the various activities (transactions) between the institutional sectors such as households, government, corporations and rest of the world (see Figure 4). In the current version of the PortableDyme model template, total economy values of the SNAB are used.

	ESA code	Total Economy (S1)	
		Uses	Resources
Production account			
Gross Output	P1		
Taxes less subsidies on products	D.21-D.31		
Intermediate consumption	P2		
Gross domestic product/ gross value added	B.1g		
Consumption of fixed capital	K1		
Net domestic product/ net value added	B1n		
Generation of income account			
Net domestic product/ net value added	B1n		
Compensation of employees	D1		
Taxes on production and imports minus subsidies	D23		
Gross operating surplus and mixed income	B2/3g		
Primary distribution income account			
Gross operating surplus and mixed income	B2/3g		
Compensation of employees	D1		
Taxes on production and imports minus subsidies	D23		
Property income	D4		
Gross primary income	B5g		
Secondary distribution income account			
Gross primary income	B5g		
Current taxes on income, wealth	D5		
Social contributions	D61		
Benefits other than social transfers in kind	D62		
Other current transfers	D7		
Gross national disposable income	B6g		
Use of disposable income account			
Gross national disposable income	B6g		
Adjust. for the change in net equity of households in pension funds	D8		
Final consumption expenditure	P3		
Gross saving	B8g		
Capital account			
Gross saving	B8g		
Capital transfer	D9		
Consumption of fixed capital	K1		
Cross capital formation	P5		
Acquisitions less disposals of non-financial non-produced assets	K2		
Net lending/net borrowing	B9		

Figure 4:   SNAB - simplified overview

source: http://circa.europa.eu/irc/dsis/nfaccount/info/data/ESA95/en/een00555.htm

Projecting the transactions within the SNAB system comprises different approaches. Basically, there are two options for consistently linking the SNAB to the rest of the model:

(1) By definition

(2) By estimation

The **first approach** is used for *all balancing items* to assure consistency *within the SNAB*. For example, the disposable income is calculated by adding up primary income, received current taxes, social contributions and other transfers and subtracting paid taxes, social contribution and current transfers. Additionally, variables such as production and consumption can be calculated by definition with the corresponding variables from the IO system.

The **second approach** is used for all variables that cannot be forecasted by using definitions.

It is very important that the order of regressions within the SNAB is built on the sequence of the SNAB (see Figure 4) E. g. gross domestic product *B1GT[t]* is a result of the production account and used in calculating operating surplus and gross mixed income *B23GT[t]*. It is obvious that *B1GT[t]* has to be calculated before *B23GT[t]*.

All transactions within the SNAB are forecasted either by approach (1) or (2). As a result from the SNAB, disposable income is calculated and used in the consumption regression (section 3.3.1).

The aforementioned tasks are carried out in step 3_model (see section 2.2). Programming the model is an iterative process and is finished if results are plausible and consistent. During the evaluation process adjustments in regressions and model code could be required.

3.4 Scenario analysis

3.4.1 General Overview

Scenario analysis is a common technique to evaluate possible future events in a so-called 'what if' manner (e. g. Stocker et al. 2011, 2014, Ökoinstitut, Fraunhofer ISI 2014, Cambridge Econometrics 2013). A scenario describes possible future developments based on plausible and consistent assumptions. The BAU scenario is based on the assumption that past behavior is also effective in the future. Other scenarios do not necessarily reflect historical trends but can include current trends, changed economic conditions or policy frameworks.

For evaluating the effects two scenarios need to be compared. They are different with respect to policy variables (e. g. tax rates) or other exogenous variables such as population. These variables are adjusted. The Interdyme software provides the concept of so-called 'fixes' which allows for modifying the values of certain variables without having to alter the source code (see INFORUM 2009). The results for each scenario are stored in a separate folder (see section 2.2, step 4_results).

Differences in endogenous model variables between the BAU scenario and an alternative scenario are responses to exogenous impulse(s). A scenario can either be evaluated stand alone or can be compared to another scenario for a particular year (see Figure 5).
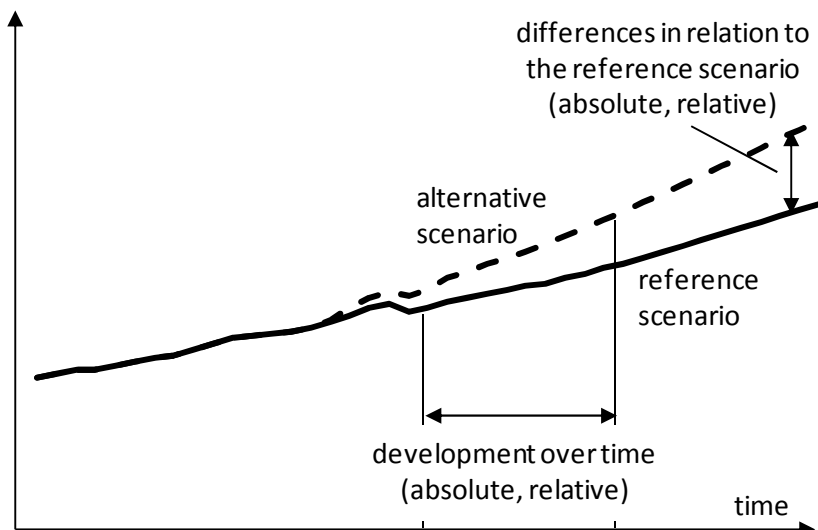
Figure 5:   Comparing scenarios

Source: own illustration

The G7 software which is part of the PortableDyme software may be used to export data to Microsoft© Excel sheets for further processing and evaluation.

3.4.2 Example: Variation of the Population Forecast for Germany

In the German PortableDyme model, the BAU-Scenario is driven by the medium-fertility variant population forecast from the UN. In this alternative scenario, it is assumed that population dynamics follow the high-fertility variant population forecast. In 2030, population is about 3.7 million higher than in the BAU scenario.

The results show that higher population has a positive impact on economic growth (see Table 1). Private consumption is still increasing but on a higher level than in the BAU-scenario. The increased demand affect production and value added. Due to an import share bigger than zero imports increase and foreign countries can improve their trade balance. Furthermore, government can profit from higher tax revenues.

Table 1:   Scenario results

		Absolute differences compared to BAU (in respective values)					Relative differences compared to BAU (in percent)				
		2010	2015	2020	2025	2030	2010	2015	2020	2025	2030
POPU	Mln. Persons	0,0	0,6	1,5	2,7	3,7	0,00	0,74	1,91	3,32	4,69
GDP and components											
GDP	Bn. EUR	0,0	6,0	16,3	29,7	43,5	0,00	0,23	0,54	0,87	1,12
HCES	Bn. EUR	0,0	2,5	6,9	12,5	18,2	0,00	0,18	0,44	0,71	0,93
GCES	Bn. EUR	0,0	3,7	10,1	18,5	27,3	0,00	0,74	1,91	3,32	4,69
GFCF	Bn. EUR	0,0	0,7	1,9	3,4	5,0	0,00	0,15	0,38	0,62	0,81
IM	Bn. EUR	0,0	0,9	2,6	4,7	6,9	0,00	0,11	0,26	0,41	0,52
EX	Bn. EUR	0,0	0,0	0,0	0,0	0,0	0,00	0,00	0,00	0,00	0,00
Other economic variables											
OUT	Bn. EUR	0,0	9,9	26,7	48,6	71,3	0,00	0,20	0,47	0,74	0,95
VALU	Bn. EUR	0,0	6,0	16,3	29,7	43,5	0,00	0,23	0,54	0,87	1,12
D2131RT	Bn. EUR	0,0	0,4	1,0	1,8	2,6	0,00	0,13	0,33	0,55	0,74
D5UT	Bn. EUR	0,0	0,6	1,7	3,0	4,4	0,00	0,20	0,49	0,80	1,04

Source: own calculations

4 Conclusion and outlook

The PortableDyme model building framework aims at both beginners and intermediate users who want to become acquainted with econometric IO model building. With the self-contained software package, clear model structure and extensible model blueprint users can focus on the model building process itself. Due to the given unified dataset and naming convention, results may be easily compared across different country models.

The framework is still under active development (Großmann, Hohmann 2013). The feedback that has been gathered from workshops and model builders helps to optimize the framework. From the software side, especially model debugging and automatic integration of behavioural equations should be improved. The model blueprint has room for improvement as well. For example, the bottom-up approach for final demand components could be applied to investments, exports and government consumption. Additionally, the labour market could be explicitly modelled to capture e. g. effects from a continuously aged society.

References

Almon, C. (1991): The INFORUM Approach to Interindustry Modeling. Economic System Research 3/1, Maryland.

Capros, P., Regemorter, D. Van, Paroussos, L. & Karkatsoulis, P. (2013): GEM-E3 Model Documentation. JRC Technical Reports.

Cambridge Econometrics (2014): E3ME. Technical Manual, Version 6.0, Cambridge.

Cambridge Econometrics (2013): Modelling Milestones for Achieving Resource Efficiency: Economic Analysis of Waste Taxes. Draft Report for the European Commission (DG Environment), Cambridge.

Großmann, A., Hohmann, F., Wiebe, K.S. (2012): PortableDyme Documentation, September 2012, Osnabrueck.

Großmann, A., Hohmann, F., Wiebe, K.S. (2013): PortableDyme - A simplified software package for econometric model building. In: Bardazzi, Gezzi (Ed.) Macroeconomic modelling for policy analysis. Firenze University Press.

Großmann, A., Hohmann, F. (2013): PortableDyme Documentation Version 1.2. Documentation from the 2nd Model Building Workshop, June 17-21, 2013 in Florence, Italy.

INFORUM (2008): INTERDYME – A package of programs for building interindustry dynamic macroeconomic models. March 2000. Department of Economics, University of Maryland.

INFORUM (2009): Fixer – General Vector Fixing Program for Interdyme Models, September 2009, University of Maryland.

INFORUM (2011): The G7 User Guide and Reference for G7.3, August 2011. Department of Economics, University of Maryland.

Ökoinstitut, Fraunhofer ISI (2014): Klimaschutzszenario 2050. Studie im Auftrag des Bundesministeriums für Umwelt, Naturschutz, Bau und Reaktorsicherheit, Berlin, Karlsruhe.

Stocker, A., Großmann, A., Hinterberger, F., Wolter, Mi. I. (2014): A low growth path in Austria: potential causes, consequences and policy options. Empirica. DOI 10.1007/s10663-014-9267-x

Stocker, A., Großmann, A., Madlener, R. & Wolter, M. I. (2011): Sustainable energy development in Austria until 2020: Insights from applying the integrated model e3.at. Energy Policy 39 (2011), p. 6082-6099.