EUROMOD TECHNICAL NOTE SERIES

EMTN 2.0

EUROMOD, CEQ, LIS and TAXSIM: a note on overlaps, differences and complementarities

Alari Paulus and Holly Sutherland

October 2016

EUROMOD Technical Note Series

This series includes documents and papers about EUROMOD or specific uses of it that are suitable for sharing with model users or others interested in technical background and for citing in papers using EUROMOD. It complements the EUROMOD Working Paper Series which mainly contains substantive papers making use of EUROMOD for research or policy analysis, prior to final publication elsewhere.

If you would like to submit a Technical Note to the series please send to euromod@essex.ac.uk.

The Technical Note series includes documents that are regularly updated and this is reflected in the numbering and, where relevant, revision information in the following table:

Title	**How revised?**	**EMTN No.**	**Date of revision**
EUROMOD, CEQ, LIS and TAXSIM: a note on overlaps, differences and complementarities	First version for Technical Note series	2.0	October 2016

Citation

Please cite this document as EUROMOD, 2016, EUROMOD, CEQ, LIS and TAXSIM: a note on overlaps, differences and complementarities, EUROMOD Technical Note EMTN 2.0, Colchester: Institute for Social and Economic Research, University of Essex.

EUROMOD, CEQ, LIS and TAXSIM: a note on overlaps, differences and complementarities

Alari Paulus and Holly Sutherland[1]

ISER, University of Essex

October 2016

Introduction

This note aims to clarify the differences and similarities between EUROMOD and two other major international projects (CEQ and LIS) that are highly relevant for comparative research on household income distribution. The comparison also includes the widely-used tax microsimulation model for the US (TAXSIM) to provide further illustrations.

1. **EUROMOD** – a multi-country tax-benefit microsimulation model, based at the University of Essex, www.euromod.ac.uk

2. **Commitment to Equity** (CEQ) – a methodological framework for analysing the impact of taxation and social spending on inequality and poverty (fiscal incidence) in individual countries, based at Tulane University, www.commitmentoequity.org

3. **Luxembourg Income Study Database** (LIS) – a database of harmonised income microdata collected from multiple countries over a period of decades, based at the LIS Cross-National Data Center, www.lisdatacenter.org/our-data/lis-database/

4. **TAXSIM** – a tax simulator for the US, based at the National Bureau of Economic Research (NBER), http://users.nber.org/~taxsim/

There are many ways of comparing the functionality and capabilities of research infrastructure projects. In our view a valid comparison needs to recognise differences in primary purpose and scope, and to avoid comparing "apples and oranges". Here we focus on the types of resource made available by the four projects to the wider research and user community. Recognising that other research infrastructure projects might choose to start from different points of comparison, ours are from the perspective of EUROMOD.[2] We are not attempting to provide a comprehensive or balanced overview.

[1] Thanks to Daria Popova, Olga Rastrigina and Iva Tasseva for their helpful suggestions. The usual disclaimers apply.

[2] The information on the other three projects makes use of information that is in the public domain at the time of writing. If there are errors or important omissions, please let us know by contacting hollys@essex.ac.uk.

The resources made available by the four projects are summarised in the following table:

Type of resources	EUROMOD	CEQ	LIS	TAXSIM
Methodological guidelines	√	√	√	√
Harmonised/enhanced micro-datasets (with documentation)	√		√	
Modelling software/platform	√			
Tax-benefit simulation models (with documentation and validation)	√			√
Analytical tools	√	√		
Teaching courses and materials	√		√	
Statistical summary indicators and reports	√	√	√	√
Research publications (working paper series)	√	√	√	

As is evident from this table, the "packages" made available to the respective "users" of each infrastructure vary considerably in scope. The core elements of EUROMOD are its software and simulation models (i.e. coded rules) for all 28 EU countries with prepared input data; TAXSIM provides only a tax calculator for a single country (albeit a major one with variation in tax policy across many states); CEQ focuses on common methodological guidelines and analytical tools; while LIS harmonises existing household micro-data for a large selection of countries world-wide and makes these available. We discuss each type of resources in turn with further details.

Methodological guidelines

EUROMOD input datasets and simulation models are constructed following EUROMOD MODELLING CONVENTIONS (EMC)[3]. This is a key document, which

- outlines general requirements for model input datasets and the main data adjustments[4]
- outlines the variable naming convention and explains how variables are structured
- provides guidelines for constructing uprating factors[5]
- defines the scope of simulations[6]

[3] EUROMOD, 2017, EUROMOD Modelling Conventions, EUROMOD Technical Note EMTN 1, Colchester: Institute for Social and Economic Research, University of Essex.

[4] The list of minimum compulsory variables and variable definitions are given in a template DATA REQUIREMENT DOCUMENT (DRD) and template Stata do-files for deriving datasets from the EU-SILC. Every country's input data typically includes additional country-specific (income) variables needed for detailed simulations of taxes and benefits.

[5] Uprating factors are used to adjust non-simulated monetary variables (i.e. income components taken directly from data – for example, employment income and public pensions) over time.

[6] The EUROMOD standard approach covers the components of household disposable income (direct taxes, employee/self-employed social contributions and cash benefits) plus employer SIC. For selected countries (and waves), non-cash benefits (public housing, education, health) have also been imputed in the past but these are not

- explains how policy systems and databases are configured in the model
- describes how simulated policies should be organised in the model
- sets out the key assumptions for simulating contributory (unemployment) benefits
- defines default standard income concepts and the common content of micro-output files
- explains what micro- and macro-validation need to cover.

Further general information on EUROMOD is provided in Sutherland and Figari (2013).[7]

CEQ methodological guidelines are set out in Lustig (2016). This provides a step-by-step guide to applying the incidence analysis, from generating data to analysing outputs:

- definitions of income concepts[8]
- methodological assumptions relevant for constructing income components not available directly from household surveys
- methodological assumptions relevant for adjusting income components with external information
- suggestions for allocating (imputed) income components at the household level
- definitions for the central scenario and variant sensitivity scenarios
- description of indicators for assessing the distributive impact, progressivity and effectiveness of taxes and transfers
- links to a Stata ado file which calculates the outcome indicators.

LIS provides the following guidelines:

- LIS DATA TEMPLATE, explaining their general approach to standardising definitions and variable names, how variables are structured, how documentation is organised etc.
- HARMONISATION GUIDELINES, describing in detail the rules, practices and definitions applied when the data are prepared into a common template.
- LIS VARIABLE LIST, a quick reference guide to person and household-level variables.
- VARIABLE DEFINITION LIST, which contains detailed definitions of the variables including content, warnings and recommendations.

TAXSIM provides only online instructions for preparing data for the model. General information on TAXSIM is provided in Feenberg and Coutts (1993).

---

made generally available, currently. In addition, there is work in progress to extend the standard model with indirect tax simulations at the household level for EU countries.

[7] See also Sutherland (2014) on multi-country microsimulation and Figari et al. (2015) on microsimulation modelling in the context of income distribution analysis.

[8] The CEQ approach seeks to construct a measure of final income, defined as disposable income plus indirect subsidies and in-kind transfers minus indirect taxes, co-payments and user fees.

Harmonised/enhanced micro-datasets (with documentation)

EUROMOD users can get access to micro input datasets for 28 EU countries, which have been derived from the EU-SILC cross-sectional household data and are ready for use in combination with its tax-benefit microsimulation model. Each input dataset is documented separately: general information about original data sources and derivations and imputations are included in EUROMOD COUNTRY REPORTS and further details on each variable in DATA REQUIREMENT DOCUMENTS. The input datasets available from December 2016 are derived from the following waves of EU-SILC (or its national versions): 2006-2008, 2010, 2012 and 2014. Anyone with access to the EU-SILC data approved by Eurostat can obtain EUROMOD input datasets.[9] Without (real) micro data, the EUROMOD model can still be applied to analysing hypothetical households (similar to the OECD tax-benefit model, for example).[10]

LIS provides users with harmonised household income micro-data, i.e. a common set of harmonised variables. These have been derived from various household income and budget surveys (including national versions of EU-SILC) for a wide range of countries (47) and time periods (within the window 1967 to 2013). Micro-data are available to registered users through a remote access system. Each dataset has own documentation, which contains general information on original data sources, a codebook with variable value labels and a description of relevant taxes and transfers. There is also a table summarising which variables are available across datasets.

Activities of TAXSIM and CEQ do not involve centrally organised production or enhancement of micro datasets for users. Household income data derived by researchers following their guidelines are generally not made available to external users. In the case of TAXSIM, however, there are ready-built codes available (http://users.nber.org/~taxsim/to-taxsim/) for creating TAXSIM datasets from commonly used survey datasets, e.g. CES, CPS, PSID.

Modelling software/platform

EUROMOD provides stand-alone and custom built software for modelling tax-benefit systems. It features a specific modelling "language" (i.e. generic routines for performing various calculations to determine tax liabilities and benefit entitlements) and a user-friendly interface with extensive integrated help. EUROMOD software can be used in combination with other software (by supporting command-line execution), which enables the integration of EUROMOD calculations into other statistical packages or automated systems as well as linking EUROMOD tax-benefit models with other type of models (e.g. labour supply, general equilibrium and macro models). The software is free for registered users.[11]

TAXSIM, CEQ and LIS do not provide (executable) software programmes. Nonetheless, TAXSIM includes two interfaces (a web-based and a Stata interface) and, similar to EUROMOD, supports a low level access to the model (a raw TCP/IP interface in this case) so that it can be used with other software.

[9] http://ec.europa.eu/eurostat/documents/203647/771732/How_to_apply_for_microdata_access.pdf

[10] http://www.oecd.org/social/benefits-and-wages.htm

[11] https://www.euromod.ac.uk/using-euromod/access/step-by-step

Tax-benefit simulation models (with documentation and validation)

EUROMOD provides fully developed and validated microsimulation models for 28 EU countries.[12] Tax-benefit rules are coded using own developed modelling "language" with users having full access to and control over the coded rules – meaning they can change the tax-benefit calculations in a flexible way and apply alternative rules or assumptions as they wish. The latest publicly available model covers policy years 2005-2016. The model (in combination with input micro-data) produces output micro-data containing simulated income components (individual policy instruments or sub-components) and standard income definitions (e.g. total means-tested benefits, employee social security contributions, and disposable income).

Using its "what if" functionality, EUROMOD can calculate indicators of work incentives at the individual level; a tool is provided for the estimation of Marginal Effective Tax Rates.

The tax-benefit simulations are documented in EUROMOD COUNTRY REPORTS, containing (i) the description of the tax-benefit system; (ii) detailed rules for simulated instruments; (iii) macro-validation results (e.g. simulation results compared with external statistics on total expenditure/revenue and number of beneficiaries/contributors by instrument; inequality and poverty indicators). There is also a general log file documenting changes in country models in every model release.

TAXSIM is a FORTRAN program, which calculates (federal and state) income tax liabilities from individual data. It can be used to calculate average and marginal income tax rates and income tax elasticities. The program is hosted in a server and the code itself is not released or visible to the user. There are no validation results or documentation for coded rules provided.

CEQ and LIS do not provide users with ready-made statistical, econometric or simulation models.

Analytical tools

Output micro-data produced by the EUROMOD model can be analysed with any statistical software package (e.g. Stata or R). In addition, EUROMOD software includes tools for calculating various summary statistics (e.g. inequality and poverty measures, average household income by income components and decile groups) as well as the Policy Effects Tool (PET) for estimating and analysing the (first-order) effects of policy changes.

CEQ provides Stata routines for analysing household income data in the context of fiscal incidence, using specifically the CEQ income concepts.

The LIS remote-execution system (LISSY) allows researchers to submit their own programs written in Stata, R, SAS or SPSS, which are run on micro-data in LIS secure servers and can report back summarised or aggregate results. There is also LIS WEB TABULATOR (see below) for generating descriptive tables based on the information in the LIS datasets.

[12] There are also "spin-off" models for non-EU countries (e.g. Serbia, Macedonia, Russia, South Africa, Namibia, and Australia), which have been built by other researchers using the EUROMOD software platform and methodology, separate from the main model and not necessarily made available to other users.

Teaching courses and materials

EUROMOD team provides regular 2.5-day training courses on how to use the model (https://www.euromod.ac.uk/events) with hands-on sessions. Courses are free. Course exercises are also available online (https://www.euromod.ac.uk/using-euromod/user-resources/training). There is a contact email for general enquiries.

LIS runs a one-week intensive summer course (http://www.lisdatacenter.org/news-and-events/workshop/), which includes lectures on comparative research and hands-on sessions with LIS data. LIS has also online self-teaching materials (for the four statistical packages supported in LISSY) together with sample files to learn how to work with the harmonised micro-data (http://www.lisdatacenter.org/resources/self-teaching/). Additional online learning materials include best practices and FAQs.

TAXSIM and CEQ do not provide open courses for external researchers.

Statistical summary indicators and reports

EUROMOD provides for 28 EU countries, with updates each year:

- EUROMOD COUNTRY REPORTS, which include aggregate simulated values by instrument (total expenditure/revenue and number of beneficiaries/contributors) for a given policy year and estimates of the latest year's tax-benefit policy changes on the income distribution. The latter are also consolidated in a separate report "Effects of tax-benefit policy changes across the income distributions" (https://www.euromod.ac.uk/research/reports) with additional comparative figures.
- (Online) STATISTICS ON THE DISTRIBUTION AND DECOMPOSITION OF DISPOSABLE INCOME (www.euromod.ac.uk/using-euromod/statistics), which contain the latest estimates of income components (taxes and benefits) by household income decile group as well as inequality and poverty indicators before and after taxes and benefits, and marginal effective tax rates. The statistics are further summarised and discussed in EUROMOD BASELINE RESULTS reports (see Leventi and Vujackov, 2016, for the most recent).

CEQ provides CEQ STANDARD INDICATORS (www.commitmentoequity.org/indicators.php), which include estimates based on CEQ approach for 6 North and South American countries – summarised in terms of the distribution of income components by income decile groups, inequality and poverty indicators before and after taxes and benefits, and progressivity measures of taxes and benefits.

LIS offers

- LIS KEY FIGURES (http://www.lisdatacenter.org/data-access/key-figures/): inequality and poverty indicators for disposable income and socio-economic characteristics (population composition, employment rates and work hours) from all LIS datasets.
- LIS WEB TABULATOR (http://www.lisdatacenter.org/data-access/web-tabulator/), which allows tabulating standardised components of household disposable income against household characteristics. It is freely available to anyone.

The TAXSIM website makes available key summary outputs and estimates for the US: income shares, marginal and average tax rates, elasticity of income tax.

Research publications (working paper series)

EUROMOD, CEQ and LIS maintain a working paper series to collect and disseminate work done by internal and external researchers using their data, model, software or methodology:

- https://www.euromod.ac.uk/publications/
- http://www.commitmentoequity.org/publications/index.php
- http://www.lisdatacenter.org/working-papers/

There is no separate working paper series for TAXSIM based work, though it is likely that bulk of it has featured in NBER Working Papers (http://users.nber.org/papers.html).

Conclusions

Focussing on EUROMOD, CEQ and LIS which all aim to facilitate the analysis of incidence of public policies across the income distribution and comparisons of such analysis across countries, there are naturally some overlaps. There are also complementarities and potentially useful combinations. For example, EUROMOD could be used to generate some data for CEQ-style indicators for EU or other countries; with some technical IT barriers to overcome, LIS data could be used as input data for EUROMOD.

It is important, however, to recognise differences in primary purpose and the scope of the resources provided. One of the main distinctions to be made is between (i) establishing a methodology for the provision of harmonised indicators, and (ii) providing the basis for user-defined analysis using harmonised inputs. CEQ does the former, with its main output being standardised definitions and methodological assumptions. EUROMOD and LIS cover both grounds by additionally providing users with micro-level information with which to do their own analysis. This has clear implications for the amount of own effort needed by users.

In understanding the different capabilities of the three infrastructures it is important to bear these distinctions in basic philosophy in mind. CEQ aims to make its detailed methodology transparent and replicable for new countries, with the main consequence being the public availability of standardised indicators for an increasing number of countries. The "public good" that LIS and EUROMOD provide is a harmonised and micro-level starting point for analysis of many kinds. In the case of EUROMOD, the infrastructure further provides a structured platform for making *reforms* to public policies and analysing the effects, as well as a methodology (analogous to CEQ) and a software tool for building the infrastructure for further countries.

Readers wanting more detailed information about any of the infrastructures are encouraged to visit their websites and read the cited background material.

References

Feenberg, D. and Coutts, E. (1993), "An Introduction to the TAXSIM Model", *Journal of Policy Analysis and Management*, 12(1).

Figari, F., Paulus A. and Sutherland H. (2015), "Microsimulation and Policy Analysis", in *Handbook of Income Distribution*, Vol. 2, A. B. Atkinson and F. Bourguignon (eds.), Elsevier.

Leventi, C. and Vujackov. S. (2016), "Baseline results from the EU28 EUROMOD (2011-2015)", EUROMOD Working Paper EM3/16.

Lustig, N. (ed.) (2016), *CEQ Handbook: A Guide to Estimating the Impact of Fiscal Policy on Inequality and Poverty*. Brookings Institution Press and CEQ Institute.

Sutherland, H. (2014), "Multi-Country Microsimulation", in *Handbook of Microsimulation Modelling*, Contributions to Economic Analysis, Vol. 293, C. O'Donoghue (ed.), Emerald Group Publishing Ltd, pp. 77-106.

Sutherland, H. and Figari F. (2013), "EUROMOD: the European Union tax-benefit microsimulation model", *International Journal of Microsimulation*, 6(1), 4-26.