

CELSI Discussion Paper No. 40

IMPROVING THE VALIDITY OF MICROSIMULATION RESULTS: LESSONS FROM SLOVAKIA

August 2016

ZUZANA ŠIEBERTOVÁ

NORBERT ŠVARDA

JANA VALACHYOVÁ



IMPROVING THE VALIDITY OF MICROSIMULATION RESULTS: LESSONS FROM SLOVAKIA

CELSI Discussion Paper No. 40

August 2016

Zuzana Siebertová

Council for Budget Responsibility

Norbert Švarda

Council for Budget Responsibility

Jana Valachyová

Council for Budget Responsibility

The Central European Labour Studies Institute (CELSI) takes no institutional policy positions. Any opinions or policy positions contained in this Discussion Paper are those of the author(s), and not those of the Institute.

The Central European Labour Studies Institute (CELSI) is a non-profit research institute based in Bratislava, Slovakia. It fosters multidisciplinary research about the functioning of labour markets and institutions, work and organizations, business and society, and ethnicity and migration in the economic, social, and political life of modern societies.

CELSI Discussion Paper series is a flagship of CELSI's academic endeavors. Its objective is the dissemination of fresh state-of-the-art knowledge, cross-fertilization of knowledge and ideas, and promotion of interdisciplinary dialogue about labour markets or broader labour issues in Central and Eastern Europe. Contributions from all social science disciplines, including but not limited to economics, sociology, political science, public policy, social anthropology, human geography, demography, law and social psychology, are welcome. The papers are downloadable from <http://www.celsi.sk>. The copyright stays with the authors.

Central European Labour Studies Institute (CELSI)

Zvolenská 29
821 09 Bratislava
Slovak Republic

Tel/Fax: +421-2-207 357 67
E-mail: info@celsi.sk
Web: www.celsi.sk

ABSTRACT

Improving the Validity of Microsimulation Results: Lessons from Slovakia

This paper summarizes the lessons learned in the process of building a microsimulation tool tailored to country-specific conditions and involving a maximum degree of user control. The objective to construct a model useful in the process of budgeting and fiscal forecasting has been achieved by paying attention to policy simulation details as well as to the representativeness of the underlying micro-dataset. The validity of simulated results improved significantly after the input database sample has been reweighted in such a way that the new weights replicate, among other factors, the earned income distribution and selected age cohorts directly. Innovative approaches in bringing the model closer to legislation as well as data highlight the benefits of having more user control compared with standardized microsimulation tools.

Keywords: microsimulation, calibration, EUROMOD, tax and transfer policy, Slovakia

JEL Classification: C81, C83, C88, I38, H24

Corresponding Author:

Zuzana Siebertová
Council for Budget Responsibility
I. Karvaša 1, 813 25 Bratislava, Slovakia
email: siebertova@rrz.sk

1 INTRODUCTION

Microsimulation modelling techniques are increasingly used to study the effects of reform policies contributing thus both to policy debates and academic literature (Figari et al., 2015). This is largely possible due to availability of highly useful and user-friendly standardized tools. EUROMOD – a fine example of such a tool - has become a benchmark for conducting microsimulation-based policy analyses for European countries.

If the appeal of the microsimulation models is to go beyond providing guidance on the design of policy reforms and be used in the budgeting process as a tool to assess an actual or proposed policy reform, policy makers must have confidence in the simulated results. The fundamental prerequisites of such a credible microsimulation model are high level of precision in the simulated policies together with high degree of representativeness of underlying micro-data. To achieve that, the interaction of the user with the model might have to go beyond the functionalities offered by standardized tools.

This paper summarizes the lessons learned in the process of building a microsimulation tool tailored to country-specific conditions and involving a maximum degree of user control. The tool is characterized by fine attention to detail and increased accuracy in important categories. The latter has been achieved by application of a recently developed approach to sample weights calibration on the underlying dataset. It is shown that the new approach applied in general improves the fit between simulated output, underlying data and official statistics. Improved fit has been documented convincingly for the simulations of payroll taxes and majority of family related benefits. The SIMTASK (**S**imulation **M**odel of **T**Axes and transfers in **S**lovakia) model itself is used to evaluate the impact of legislative changes in the areas of taxes and benefits in Slovakia but the exercise contains valuable lessons for users of standardized tools such as EUROMOD wishing to obtain an enhanced degree of user control.

Whilst the model alone is designed to assess static effects of policy changes, the benefits of maximum user control through modelling in Stata become more apparent when incorporating the model into other labour supply (Siebertova et al., 2015) and general equilibrium models (Horvath et al., 2015). In those cases, it can be readily used for the evaluation of the long-run consequences of tax and benefit reform strategies too.

In SIMTASK, the emphasis has been put on details, namely on precise adaptation of valid legislation in order to achieve the highest precision in policy simulations. The modelling of benefits whose amount and duration is conditional on unobserved factors - such as the material needs, unemployment and maternity leave benefit is a particular strength and contribution of the work summarized in this paper.

For a microsimulation model to provide trustworthy assessment of budgetary and distributional effects of tax and transfer system reforms, it is crucial that the underlying data are representative with respect to the income distribution. As survey data rarely comply with this requirement, some input data corrections are shown to be beneficial. In addition, for the precise simulation of family related benefits, the correct representation of children in the corresponding age cohorts is essential. Therefore, we choose the approach of recalibrating the sample weights.

The importance of retaining information from the original weights is often stressed in the literature, there are, however, situations where re-weighting is required because the original weights supplied with the data do not adequately represent key analytical groups required for the analysis (O'Donoghue and Loughrey, 2014). This, we believe, is very relevant in our case. Moreover, there is no guarantee that sample weights calibrated to match demographic population totals produce appropriate revenue, expenditure and income distribution results (Creedy, 2004).

In our re-calibration, compared to original sample weights, the new weights allow for more detailed control of small children age categories and the earned income distribution is taken into account as a calibration factor directly. We show that this approach improves the match between simulated output and official statistics.

The paper is structured as follows. Section 2 describes the micro-level data used and the re-weighting method applied to sample weights of the underlying input dataset. Section 3 summarizes the tax and benefit system in Slovakia. Section 4 gives an overview of SIMTASK's development process and describes major differences compared to the existing Slovak EUROMOD modules. Section 5 presents validation and provides a discussion of the simulation results. We also report the implications of these methodological improvements for income distribution and inequality indicators. Section 6 concludes.

2 DATA

A necessary precondition for the development of a microsimulation model is the existence of a suitable micro-dataset containing information preferably both on individuals and households. Usually, household survey data are used for these types of analyses; use of the administrative (or census) data is rather scarce.

The national version of the EU-SILC survey, abbreviated as SK-SILC, was selected as a base dataset for the tax-benefit microsimulations. Currently, it does best at meeting the data requirements for a microsimulation model of tax and transfer system when compared to other datasets that are available. In contrast to the EU-SILC, the SK-SILC dataset includes more variables that are country specific. The EU-SILC is an annual survey that has been conducted in Slovakia since 2004, it is collected by the Statistical Office of the Slovak Republic on behalf of EUROSTAT. Survey questions are focused on the income and living conditions of different types of households, as well as on the individual demographic characteristics, education, health status, employment, housing conditions and deprivation measures.

2.1 Re-weighting of the input SK-SILC dataset

Microsimulation tax-benefit system models are frequently used for the assessment of the effects of actual system reform policies as well as ex-ante simulations of reforms. Therefore, the input dataset should reflect the actual economic and social conditions as closely as possible. As survey data are available with a time lag (usually 2-3 years), the reference period of the input dataset and the baseline tax-benefit system might not refer to the same period. The approach frequently applied in the literature is to uprate the original market income by appropriate growth factors and to re-weight the sample to account for selected demographic and labour market changes. For a survey on current methodologies, see O'Donoghue and Loughrey (2014) and Figari et al. (2015).

The SK-SILC dataset is calibrated (such that sample weights are adjusted to match the known population totals in selected categories) and integrated weights (such that cross-sectional household weights and personal weights equal) are provided by the Statistical Office of the Slovak Republic. The calibration is an optimization procedure undertaken at two levels (household and individual) by using CALMAR2, a SAS macro developed by INSEE (LeGuennec and Sautory, 2002). By employing this macro, calibration uses 21 different categories in one strata and is performed on a number of household members (5 categories), 6 age categories divided by gender (together 12

categories) and 4 variables describing labour market status of a person (employees, unemployed, self-employed and pensioners). Stratification is based on NUTS3 level (8 regions). Condition of having integrated weights implies that simultaneous calibration should be applied, but there is no guarantee that the calibration process converges and the result is a kind of an approximate solution (as argued by Glaser-Opitzova et al., 2014 an exact simultaneous solution has never been found).

As it is frequent in most survey data, SK-SILC dataset does not correctly represent the income distribution (both labour and non-labour) when compared to the official statistics. The most important component of individual labour income are gross earnings from employment (it constitutes more than 85% of aggregate income share in SK-SILC dataset and around 90% in the official data). When the distribution of gross earnings from employment of individuals implied by SK-SILC (using original weights) is compared to the official statistics that can be retrieved from the official administrative data (in Slovakia it is a Social Security Agency (SSA) database), low-income groups and high-income groups are under-represented (the latter in fact missing) and incomes around the average wage substantially over-represented. Gross income from employment is the most important element that enters to the computation of the income tax base and its correct representation is highly important for the validity of any tax and transfer microsimulation model. Given this argument, it is beneficial to use the gross earned income as an additional calibration factor. Income as a calibration factor has been used in Slovenian EU-SILC, where calibration included also employee cash or near cash income (Inglic et al., 2013). A similar idea has been described also by Creedy and Tuckwell (2003) who calibrated the weights of the New Zealand Household Economic Survey to take into account the number of recipients of several social transfers directly.

An improvement along these lines is provided by the calibration software “Calif” that has been recently developed by the Slovak Statistical Office. This tool allows to consider more categories in the calibration procedure due to the fact that it is able to find an approximate solution of the optimization problem, it works with several optimization methods, takes into account stratification and computes integrated weights. For the detailed description, mode of use and the documentation on Calif, see Vlacuha and Frankovic (2015). As an input to calibration procedure when using Calif, the number of individuals in the defined categories in the underlying SILC dataset and corresponding population totals obtained from administrative sources are needed. External administrative statistics on demographic categories, labour market status and household

composition comes from the Statistical Office of the Slovak Republic, while income distribution is matched to the individual dataset of the SSA. Details are documented in Table 1.

Table 1 Sources of external controls for calibration

Calibration category	Source of the official statistics
Age cohorts (0, 1-3, 4-16, 17-25, 26-45, 46-retirement age, over retirement age). Categories over 16 years considered separately for gender (male, female)	Statistical Office of the Slovak Republic (Demographic Balances of Population)
Labour market status (employees, unemployed, self-employed)	Statistical Office of the Slovak Republic (Labour Force Survey)
Household size (members 1, 2, 3, 4, 5+)	Statistical Office of the Slovak Republic (Population and Housing Census 2011)
Gross income from employment distribution (deciles 1-2, 3-5, 6-8 and 9-10)	Database of individual records from Social Security Agency (SSA)

Majority of calibration categories that are used when computing original sample weights were also considered when re-calibration using Calif is performed. The difference between the two is in the definition of age cohorts and that the earned income has been directly taken into account as an additional calibration factor in the latter approach. When controlling for the labour market status, we use 3 categories (employees, unemployed and self-employed). We left out the category of pensioners (used in the original weighting scheme), since this is highly correlated with the corresponding old-age cohort, see also Kump and Navicke (2014).

In the calibration procedure, 7 age categories (0, 1-3, 4-16, 17-25, 26-45, 46-retirement age, over retirement age) were considered. Age cohorts over 16 years are matched to population totals such that also gender has been considered, i.e. separately for males and females. The idea of using extended age categories was to correctly represent newborn (age 0), small children (1-3) and youth (up to 25) population sub-groups that are essential controls for the simulation of family related benefits (child's birth grant, maternity benefit, child benefit and parental allowance).

Based on the earned income distribution of individuals identified in SILC dataset and income distribution of individuals that can be retrieved from administrative source SSA, extra categories for calibration were formed. To make the corresponding income categories comparable, in both underlying datasets the gross income from employment has been considered. For the calibration procedure two pieces of information were needed. First, decile points of income distribution given by data in SSA were computed. These decile points were used as threshold values also in SK-SILC dataset. In the second step, in every income decile (defined by decile points computed in SSA

dataset) number of individuals were counted both in SSA and SK-SILC datasets and 4 additional calibration categories (in 8 strata) were constructed by grouping several deciles together (deciles 1-2, 3-5, 6-8 and 9-10).

The weights were re-calibrated separately for every strata, i.e. independently for 8 regions by using a linear bounded optimization method. When using the linear bounded method, the upper and lower bounds for the exit rates should be set. We decided to start initially with wider bounds (0.1 for the lower bound and 10 for the upper) and to gradually reduce them (similarly like Creedy, 2004 or Kump and Navicke, 2014). At the same time, we checked whether calibrated weights match the population totals and control for the standard deviation of weights (that should be low). Based on these three criteria, we discussed the estimated results and chose the set of calibrated weights. Consequently, newly estimated calibration weights correct the earned income distribution in a way that it sufficiently matches the official statistics (see Figure A1 in the Appendix).

The SK-SILC dataset corresponding to income reference period 2011 reports 15,440 individuals living in 5,291 households and SK-SILC referring to 2012 contains 15,426 individuals in 5,402 households. Table 2 presents descriptive statistics of the grossing-up weights and population estimates of the samples weighted by original weights and using weights computed with a new calibration tool. In addition, in Table A1 in the Appendix we present the descriptive statistics of main demographic and income related variables.

Table 2 Descriptive statistics of grossing-up weights in SK-SILC samples

Policy year	2011	2011	2012	2012	2013	2014
Underlying SILC dataset	2011	2011	2012	2012	2012	2012
Grossing-up sample weight	original	calibrated	original	calibrated	calibrated	calibrated
Mean	349.5	350.0	350.4	350.0	351.2	351.8
Std. Dev.	126.0	468.7	132.0	482.5	521.3	581.1
Minimum	108.7	10.9	119.9	12.0	12.0	12.0
Maximum	1,226.1	5,505.1	1,083.9	4,481.0	5,974.1	8,961.1
Dataset characteristics						
Individuals	15,440	15,440	15,426	15,426	15,426	15,426
Households	5,291	5,291	5,402	5,402	5,402	5,402
Projected population	5,395,519	5,396,355	5,404,664	5,398,917	5,417,340	5,427,220
Projected households	1,911,664	1,909,248	1,852,059	1,852,027	1,852,059	1,853,409

Source: Authors' calculations using SK-SILC.

In order to test the predictive accuracy of the SIMTASK when the income reference period of the underlying input dataset and the simulated tax and transfers system refer to different time periods, we applied a two-step nowcasting method. As an underlying dataset we used the latest SILC survey available to us at the time of writing with the income reference period 2012. In the first step, we updated income variables (including all labour and non-labour income variables listed in Table A1 in the Appendix) in the dataset by the corresponding growth factors. In the next step, we applied a variant of static ageing technique and re-weighted the input dataset to account for changed population structure (both demographic and labour-market status). Calibration of weights has been performed by comparing the data from the updated dataset to the external statistics from the target policy year. In the last two columns of Table 2 we present the new calibration weights that were estimated and that are later used in simulating taxes and transfers to test for the accuracy of SIMTASK in policy years 2013 and 2014.

2.2 SK-SILC versus official statistics

The SK-SILC dataset is largely representative of the country population. However, as it is frequent in survey data, SK-SILC might also over-represent or under-represent certain population groups. Particular limitations are inspected in details below, in such a way that SK-SILC data are compared to the appropriate official statistics using both the original and calibrated weights. Tables displayed below suggest that in most aspects the newly calibrated weights helped to improve the fit. These comparisons are also highly instructive in later assessment of simulations.

Table 3 presents the ratios of the number of individuals in the selected age cohort in the input SK-SILC database to external benchmark. While 2011 and 2012 SK-SILC datasets weighted with original weights underestimate the number of new-born (age 0) and small children (under 3 years), using the calibrated weights where we directly control for the number of children in certain age groups leads to almost perfect fit. For the prime age and retirement age cohorts, datasets using calibrated weights match demographic statistics closely both in 2011 and 2012.

Table 3 Validation of weighting schemes: representation of gender and age cohorts

	original weights		calibrated weights	
	2011	2012	2011	2012
Female	1.00	1.00	1.00	1.00
Age cohort				
0	0.55	0.78	1.00	1.04
0 - 3	0.63	0.84	1.00	1.01
0 - 16	0.82	0.96	1.00	1.00
0 - 26	0.98	1.04	1.00	1.00
Prime age	1.03	1.01	1.00	1.00
Retirement age	1.08	1.04	1.02	1.01

Source: Authors' calculations using SK-SILC and Slovak Statistical Office

Note: Ratios display number of individuals in SK-SILC dataset (weighted) to population by gender and in the respective age cohort.

Prime age: 15-64 years. Retirement age: males 62+, females 58+ in 2011 and 59+ in 2012.

Data on representation of the economic activity of Slovak population is shown in Table 4. The reported ratios document that based on these criteria, SK-SILC dataset reflects the official statistics very well, the only exception being the group of employees. Comparing the two weighting schemes, the number of employees is originally significantly oversampled, but when calibrated weights are used the number of employed gets well closer to the official statistics (in both years).

Table 4 Validation of weighting schemes: representation of economic activity

	original weights		calibrated weights	
	2011	2012	2011	2012
Employed	1.01	0.97	0.96	0.96
Employee	1.20	1.15	1.08	1.08
Self-employed	1.01	0.96	1.02	1.01
Unemployed	0.97	0.94	0.99	0.99
Economic active pop.	1.01	0.97	0.97	0.97
Economic inactive pop.	0.99	1.03	1.04	1.04
Population total	1.00	1.00	1.00	1.00

Source: Authors' calculations using SK-SILC and LFS.

Note: Ratios display number of individuals in SK-SILC dataset (weighted) to LFS in the respective category.

In Table 5, the different sources of income reported in SK-SILC are related to the official statistics given by SSA. A comparison is provided with respect to the ratios of reported aggregate amounts of income as well as in terms of the ratios of the number of individuals receiving certain type of income.

Table 5 Validation of weighting schemes: representation of income

Weights	Amounts (ratios)				Individuals (ratios)			
	original		calibrated		original		calibrated	
Income from	2011	2012	2011	2012	2011	2012	2011	2012
Employment	1.16	1.02	0.97	0.95	1.08	1.05	0.93	0.95
Agreements	0.43	0.34	0.43	0.34	0.36	0.38	0.33	0.36
Self-employment*	2.84	2.73	2.98	2.78	1.34	1.25	1.36	1.31
Employment and agreements	1.13	0.99	0.94	0.91	0.97	0.94	0.97	0.94

Source: Authors' calculations using SK-SILC and SSA.

* Validation of income for self-employed is only indicative. SK-SILC reports for self-employed the value of profit/loss in the current year, while the SSA database reports the assessment base which is based on the value of return in the year t-2 (inconsistency both in variable and time).

Note: Amounts (ratios) display aggregate amount of income of individuals in SK-SILC dataset (weighted) to aggregate income computed by using records from SSA. Individuals (ratios) display aggregate number of individuals in SK-SILC dataset (weighted) to aggregate computed by using records from SSA.

The overall picture does not differ in 2011 and 2012; the number of people reporting an income from employment is only slightly undersampled and matches relatively well with the administrative data from SSA. Those declaring an income from agreements (temporary employment contracts) are significantly under-represented and this applies to both original and calibrated weights. On the other hand, the number of self-employed individuals compared to SSA statistics is substantially oversampled. It should be noted that comparing the number of self-employed to the statistics of SSA is not completely correct. SSA database is primarily a dataset of paid social insurance contributions providing information on gross income. In the case of self-employed persons, SSA dataset captures only those individuals who pay SIC (social insurance contributions) which is a subset of the total number of registered self-employees. However, it is instructive to show these ratios, since our simulations of taxes and social security contributions are validated against the statistics provided by SSA.

Aggregate income from employment approximately matches the aggregate amount documented by SSA, while the income from agreements is substantially underreported. Since the volume of agreements makes approximately only 5% when compared to the income from employment, the total effect of employment and agreements matches SSA dataset well.

Note that aggregate income from self-employment should be validated with caution and the results proposing substantial over-reporting in the input data are only indicative. The reason is that SK-SILC reports for the self-employed the value of profit/loss in the income reference period, while the SSA database reports the legislatively correct assessment base which is based on the value of

declared return in the year t-2 (i.e. there is an inconsistency both in variables that are equated and time aspect). However, relative weight of self-employed in the labour market is rather low, as they constitute only 7% of the total population.

The main non-simulated benefits and pensions, which serve as an input to later simulations, are inspected in Table 6. Maternity benefit recipients are substantially undersampled when the original weighting scheme has been applied. Using the calibrated weights makes the number of recipients to match well in 2011, but overestimate in 2012. Since the eligibility for the maternity benefit is up to approximately 7 months after the child's birth, the reported ratios match with the good fit of the youngest age cohort of new-born children in SK-SILC in 2011 and small oversampling in 2012 as it is documented in Table 3.

Table 6 Non-simulated benefits and pensions

Weights	Amounts (ratios)				Recipients (ratios)			
	original		calibrated		original		calibrated	
Benefits	2011	2012	2011	2012	2011	2012	2011	2012
Maternity	0.63	0.60	1.22	1.27	0.57	0.61	1.01	1.28
Sickness	0.31	0.34	0.50	0.59	n.a.	n.a.	n.a.	n.a.
Pensions								
Old age	1.16	1.09	1.09	1.08	1.10	1.04	1.04	1.03
Disability	0.87	0.88	1.03	1.10	0.84	0.86	1.03	1.10
Widow/er	0.98	1.05	0.97	1.06	1.06	0.99	1.01	0.97
Orphans	0.83	0.82	0.93	0.57	0.69	0.68	0.79	0.52

Source: Authors' calculations using SK-SILC and SSA.

On the other hand, the demographic group of elderly is represented well in both input samples. This subsequently mirrors in the share of old-age pension beneficiaries close to one. Orphans are undersampled in the input data when both weighting schemes are used, while disability pensioners are slightly underestimated/overestimated when original/calibrated weights are applied. Widows and widowers well approximate the figure addressed by SSA.

Table 6 summarizes also the information on the aggregate amounts of paid benefits and pensions: data in input datasets are compared to the external statistics recorded by SSA. Not surprisingly, old-age pension payments are slightly overestimated, but match relatively well. Other non-simulated benefit and pension payments are in general underestimated when original weighting has been used. The gap between official records and input data is extreme in the case of sickness benefits, where aggregate payments reported in SK-SILC reached around 30% of the official

statistics using the original weighting scheme. The gap has been slightly reduced with the calibrated weights to around 50%. Maternity benefit payments represent around 60% of the official SSA records with the original weights, while using the calibrated weights leads to overestimation. Both these ratios are in line with the number of recipients reported above.

3 THE TAX AND BENEFIT SYSTEM IN SLOVAKIA

3.1 Taxes and social insurance contributions

The Slovak tax system is largely unified; all important components are set at the state level. Taxation of income is conducted at an individual level and it is levied on gross income including wages, income from business activities, fringe benefits, capital incomes (dividends excluded), interest and rental income. Joint taxation of married couples is not possible. Social and health insurance contributions and social benefits are exempt from the tax base, i.e. the tax base is given as gross earnings net of employee social and health insurance contributions.

All relevant parameters needed to compute personal income tax (PIT) are available in the SK-SILC data - both those which are related to individual and household level. During the years 2009 to 2012 PIT amounts to a 19% flat tax rate with a non-taxable allowance. From 2013, two tax brackets were introduced and incomes exceeding the threshold are taxed by 25% rate.

Tax expenditures that are deducted from the tax liability in the PIT and that are incorporated in SIMTASK include:

- (a) Basic tax allowance: tax allowance each individual can apply, the amount of the allowance is based on the legally defined minimum subsistence level. A progressive reduction in basic tax allowance is applied when annual gross earnings exceed about 18,000 euros (approximately twice the Slovak average yearly gross wage) and it influences around top 10% of tax payers.
- (b) Spouse tax allowance: an individual may be entitled to a spouse tax allowance if the income of spouse satisfies certain conditions (earnings under a certain level).
- (c) Employee tax credit (ETC): the amount depends on employee's income and on the period he has been working (at least 6 months). It is targeted at low-income groups who have to pay health and social insurance contributions.

- (d) Child tax credit: one spouse may claim an allowance for each child in the household if the child satisfies certain conditions (e.g., aged under 18 or aged under 26 and in full time education or aged under 26 when physically or mentally disabled and not receiving disability pension). This tax credit can be received, if the parent annually earns at least 6 times the minimum wage. If the credit exceeds the tax liability, the excess is paid to the taxpayer.

The Slovak social insurance system is made up of two components; namely social insurance contributions and health insurance contributions. The assessment base for contributions is narrower compared to the PIT base since capital income is not considered. Up to 2012 maximum assessment base differed based on the type of insurance and employment contract. Effective from 2013, assessment bases for social and health insurance contributions of employees were unified. For the self-employed, the computation of the assessment base was redefined.

- (a) Social insurance contributions (SIC)

Both employers and employees pay an unemployment, sickness, disability, and an old age insurance, but different percentages from the social insurance assessment base. In addition, employers also pay contributions to a reserve solidarity fund, accident insurance and guarantee insurance. The self-employed are treated differently; they pay sickness, disability and old age insurance and contributions to the reserve solidarity fund.

- (b) Health insurance contributions (HIC)

These contributions are paid by employers, employees and self-employed. The percentage to be paid is different for the three categories of payers.

3.2 The social system

The Slovak benefit system consists of three components, termed as contributory, social assistance and poverty, and state social support.

- (a) Contributory benefits include old-age pension, early old-age pension, disability pension, widow's and widower's pension, orphan's pension, sickness cash benefit, benefit for nursing a sick relative, equalization allowance, maternity benefit, and unemployment insurance benefit.

- (b) Social assistance program covers material need benefit.

(c) State social support includes several programs, namely child birth grant, additional birth grant, multiple birth benefit, child benefit, additional child benefit, parental allowance, funeral benefit, scholarships for pupils in elementary school, scholarships for students in secondary school, and social scholarships for university students.

4 TAX AND BENEFIT SIMULATIONS

4.1 Existing models

EUROMOD has been the only model available for the Slovak tax-benefit system microsimulations, which could be used equally by government agencies and the academic community. It is an EU-wide tax-benefit microsimulation model that can simulate individual and household tax liabilities and benefit entitlements according to policy rules valid in the respective EU states. EUROMOD is developed and maintained by the Institute for Social and Economic Research at the University of Essex, in collaboration with national teams. For its current state and details of the project, see Sutherland and Figari (2013). EUROMOD for Slovakia is well documented in the EUROMOD Country Report, for a detailed overview of application rules and payable eligibility, see Porubsky et al. (2013) or Strizencova and Hagara (2014). In this analysis, EUROMOD version G2.0+ is used.

The Slovak EUROMOD runs on SK-SILC data and the simulated policies currently include personal income tax, all health and social insurance contributions paid by employers, employees and self-employed. Benefits that are fully simulated include family related programs, namely child birth grant, child benefit including additional child benefit and parental allowance. Means-tested material needs benefit and contributory unemployment insurance benefit are simulated partially under simplifying assumptions. Simulations of other benefits, which may impact both individual and household incomes, are not included due to the lack of information on previous employment and contribution history. In particular, these include sickness benefits and disability pensions. Old-age pensions are not simulated since there is no information on contribution period.

4.2 SIMTASK

A new microsimulation model SIMTASK has been developed such that the setup of the EUROMOD model has been taken as a template and an independent program that runs in software Stata has been created. It is important to stress that a primary intention has not been to replace the existing Slovak EUROMOD, which is a simple and transparent static tax-benefit calculator with the advantage of cross country comparability and that can be also linked to other

models. Rather, the objective has been to expand its use and to tailor it directly to the principal demand of having a simulation tool that can be easily incorporated into other models to provide an accurate-enough evaluation of measures for the process of budgeting and fiscal forecasting. Besides the considerations about the type of microsimulation model that was needed in terms of its capability to include behavioural responses, the mode of operation, i.e. how easy it is to incorporate and handle with it in such a model setup, where the convergence could be achieved only after numerous iterations, has been an issue too.

All tax and benefit instruments in the SIMTASK model are simulated in the same order as in Slovak component of EUROMOD (further referred as a baseline model). In addition, SIMTASK includes the simulation of the length of the eligibility period to a maternity benefit and a substantial extension of simulation of material needs benefit.

In the baseline model setup all benefit instruments are simulated on a yearly basis. Based on predefined eligibility requirements, it is tested if an individual is entitled to receive a certain benefit. An assignment is provided if the predefined conditions are met and subsequently the corresponding amount is simulated. For example, conditional eligibility to an unemployment benefit is checked (among other conditions, an individual should not receive parental allowance) and parental allowance is simulated prior to unemployment benefit. In other words, subsequent entitlement to certain transfers is ruled by the order of simulation policies. However, this procedure does not take into account possible variability that can occur during the whole period of one year – such that an individual might be eligible for several transfers that are available to him/her subsequently, if these transfers are paid for shorter period than one year.

In order to allow an individual to receive different benefits during the annual period in SIMTASK, eligibility to selected transfers is simulated on a monthly basis depending on the predefined requirements. This approach could be applied thanks to the fact that information on month of birth of an individual is recorded in SK-SILC dataset. The simulations of benefits for shorter periods than one year are already available for some countries within EUROMOD for example in Estonia (Vork and Paulus, 2014) and it would be possible to implement this approach also in Slovak EUROMOD. Consequently, knowing the month of the year when a child was born, it is possible to accurately allocate family related benefits. This applies particularly to family related and unemployment benefits, which are simulated in the following order:

- maternity benefit: the length of the eligibility period is simulated, which is 8 months (or 10 months in case of multiple births, or 9 months for lonely parent). The amount of benefit is presently not simulated because of lack of information on contribution history to health insurance.
- parental allowance: the length of the eligibility period is simulated, entitlement ends when the child reaches 3 years of age. Entitlement is possible up to 6 years in case of child's unfavourable health condition, but there is a lack of information to simulate this. The amount needs not to be simulated - it is a fix payment.
- unemployment benefit: the length of the eligibility period is simulated, maximum is 6 months.

Minor modifications of tax-benefit system simulations used in SIMTASK (as compared to Slovak EUROMOD) are detailed in Siebertova et al. (2014). Two major modifications were implemented and these apply to the simulation of material needs benefit and unemployment benefit.

The material needs benefit (MNB) is a means tested transfer that is intended for families with income below the minimum subsistence level. The actual benefit amount is calculated as a difference between the eligible maximum of MNB - composed of social benefit, health care allowance, housing allowance, activation and protection allowance - and the income of individuals living in a household. In our simulation, we include a more precise specification of the assessed income computation (compared to baseline). Social benefit and health care allowance are set as fixed amounts, these are not simulated. Furthermore, we include a different computation of the protection allowance: in our implementation, it is based on the set of predefined eligibility conditions. The essential is the change in the definition of an individual allocation to the activation allowance. In the baseline model, activation allowance is assigned to all those, who are not eligible to receive protection allowance. However, this approach is not based on a valid legislation and as a result, it largely overestimates the assignment of the activation allowance. On the contrary, in our approach we define a set of eligibility conditions that an individual needs to fulfil in order to be entitled to draw this allowance. This gives us a set of people who potentially might take part in activation works. In the next step, we randomly draw from this predefined group a subset of individuals (who will be finally assigned to activation works participation), such that the ratio of those who participate in activation works to the total number of those who receive MNB equals,

when compared to the official statistics. In 2014 this “random draw” procedure is applied also to the basic allowance due to legislation changes.

The unemployment insurance benefit is a contributory transfer aimed to compensate temporarily for the income loss due to unemployment. In our detailed adaptation (as compared to the current version of Slovak EUROMOD) we provide a more precise simulation of eligibility period on a monthly basis, this is possible also thanks to the more precise simulation of the length of the maternity benefit.

5 MODEL AND VALIDATION OF SIMULATION RESULTS

Validation of model outputs, i.e. comparison of computed results with reality, is a useful approach to test the overall relevance and weak points of the microsimulation model. There are several possible approaches how to validate results produced by a microsimulation model. We adopt an approach frequently used in the academic literature, where baseline systems are validated and tested at aggregate macro level such that simulated outputs are compared to the external official statistics. In this section we also show the validation of the predictive performance of SIMTASK. Finally, we provide an overview how the whole income distribution, inequality and poverty indices are affected by the new weights and refined simulation.

5.1 Aggregate validation

In this section we demonstrate that refinement of simulations as well as re-weighting of the input dataset leads to improved validity of aggregate results. We show that different approach to the simulation of the material needs benefit (compared to EUROMOD) significantly improves accuracy of the results. Increased precision of the simulation of parental allowance, which is the most important transfer paid to families (in the sense of total volume of payments) is gained by using calibrated weighting scheme and SIMTASK that takes into account duration of benefit take-up less than one year. The same argument applies to the simulation of personal income and payroll taxes where the improved accuracy has been achieved mainly due to application of the calibrated weighting scheme.

Total expenditures and the number of beneficiaries of those transfers that are not simulated, but act as inputs to SIMTASK model, are compared to the official statistics in section 2.2 above. In the next step we look in detail at transfers that are simulated both by EUROMOD and SIMTASK and compare the simulation results to the official statistics in 2011 and 2012. To make simulation

results comparable, we use the same underlying SK-SILC datasets when running EUROMOD and SIMTASK.

When validating results with respect to the total number of people, a concept of “unique occurrence” has been used. This applies to the aggregate number of benefit recipients, tax payers, unemployed, employed, self-employed or persons with agreement contracts. By construction, the SK-SILC dataset should include every person receiving a given benefit, paying taxes or having an employment contract during the income reference period. Therefore, the statistics on “unique occurrence” should better correspond to the reality that is reflected in SK-SILC than the average monthly number, which is the statistics usually reported by administrative sources.

The choice of an appropriate external statistics has been considered also regarding the aggregate validation of estimates of tax and different contributions revenues. The official statistics on PIT, SIC and HIC revenues published by the Ministry of Finance mirrors the payments received during the income reference period, which might be distorted by the sum of unpaid contributions. Therefore, PIT, SIC and HIC revenues are calculated directly using the administrative Social Security Agency database that contains individual records of payments on monthly basis. Note that this corresponds better to simulated aggregates by SIMTASK that represent liabilities that should be paid, rather than actually received payments.

Finally, we provide a simulation exercise where the predictive ability of SIMTASK is tested. Based on 2012 input data we simulate tax and transfer systems valid in 2013 and 2014 and verify simulation results against the official statistics.

A summary on the aggregate validation of the main simulated benefits from EUROMOD and SIMTASK with original and calibrated weighting schemes against the external official statistics is depicted in Table 7. Comparing the results in columns “original” and “calibrated” shows the disparities that arise due to different weighting schemes used. On the other hand, comparing “EUROMOD” and “SIMTASK” (with the same weighting) document distinctions that appear due to refinements in simulations.

Table 7 Simulated benefits: Ratios of aggregate amounts

Model	EUROMOD		SIMTASK	
	original	calibrated	original	calibrated
Weights				
	2011			
Unemp. benefit	0.54	0.46	0.55	0.47
Parental allowance	0.78	1.18	0.64	0.99
Child benefit	0.87	1.00	1.08	1.16
Child birth grant	0.61	1.10	0.61	1.10
Material needs benefit	1.24	1.82	0.83	1.19
	2012			
Unemp. benefit	0.38	0.37	0.40	0.38
Parental allowance	1.09	1.07	0.87	0.96
Child benefit	0.97	0.98	1.17	1.14
Child birth grant	0.80	1.08	0.81	1.08
Material needs benefit	1.46	1.92	1.00	1.27

Source: Authors' calculations using SIMTASK and EUROMOD, official statistics SSA (unemployment benefit), other benefits COLSAF (Central Office of Labour, Social Affairs and Family).

Note: Numbers display ratios of aggregate number of payments to individuals computed by EUROMOD and SIMTASK to aggregate amounts referred by official statistics.

Table 7 shows that simulation results are substantially improved by using the calibrated weights. Aggregate validation of total amounts of family related benefits, namely parental allowance and child birth grant, shows that simulations using original weighting schemes are underestimated in EUROMOD and SIMTASK model when compared with the official statistics in 2011. Using the calibrated weights made the corresponding ratios get closer to one. The reported underestimation of these transfers directly mirrors undersampling of new-born and small children in SK-SILC with original weights. Since calibrated weights directly control for the correct number of children, this has led to the improved validation results.

The aggregate amount of child benefit payments is overestimated in SIMTASK. Using original weights payments are overestimated by 8% in 2011 and 17% in 2012, when using calibrated scheme overestimation is 16% in 2011 and 14% in 2012. This imprecision arises due to the broader definition of eligibility condition that is applied. According to valid legislation, also parents of university students (up to 26 years of age) studying in an internal form are eligible to receive child benefit. It is not possible to distinguish between internal and external form of university study in the input dataset. When we adjusted the simulated output and took into account that internal form of study applies to around 70% of university students, the resulting numbers approached the official statistics closely. In the EUROMOD simulation of child benefit, all university students irrespective of form of study are excluded from the eligibility condition.

Validation results for the material needs benefit differ substantially based on the weighting scheme that has been used. When using the original weighting, the income distribution has not been taken into account and low-income earners in input datasets were under-sampled. This translated into undersampling of the number of recipients of MNB, leading to ratio 73% and 81% of the official statistics in 2011 and 2012, respectively (see Table A2 in the Appendix). Overall, this resulted into underweighting of the aggregate amount of this benefit in 2011 (83% of the official statistics). Using the calibrated weights, more weight has been placed on low-income earners who are also the most likely material needs benefit recipients, and finally this led to overestimation of this transfer in terms of the amount of benefits received (19% and 27% in 2011 and 2012, respectively). These results are in line with the evidence documented in the empirical literature suggesting considerable non-take-up of means tested benefits (Wiemers, 2015 or Matsaganis et al., 2008). In Slovak EUROMOD, MNB transfer is simulated differently compared to SIMTASK (see section 4.2.1). Overestimation of several components of MNB leads to even more pronounced overestimation of both the total payments and the number of recipients.

In the simulation of the unemployment benefit, only the length of the eligibility period is simulated but not its allocation to recipients. However, the number of recipients declared in the input dataset is only around 45% of the official statistics (both in 2011 and 2012). Therefore, the total number of recipients as well as aggregate amount of payments of unemployment benefit is substantially underestimated when both weighting schemes are applied.

Aggregate validation of total number of recipients of main simulated benefits leads to comparable results as those presented in previous paragraphs. Detailed results can be found in Table A2 in the Appendix.

The aggregate sum of payroll taxes compared to the official statistics is more precise when using calibrated weights and SIMTASK. Detailed output related to personal income tax and social (SIC) and health (HIC) insurance instruments is depicted in Table 8.

The aggregate sum of tax liabilities (including tax credits and tax allowances) shows almost perfect fit to the official statistics (1.03 and 0.95 in 2011 and 2012, respectively) when using SIMTASK with calibrated weights. A difference in the validation of simulations of SIC for employees and employers and HIC can be observed when the two weighting schemes are compared - using the calibrated weights leads again to the almost perfect fit. SIC paid by self-employed should be

interpreted differently and results documented here are only indicative. The reason is an inconsistency in variables that are equated; profit/loss of self-employed reported in SK-SILC versus the assessment base for SIC in the official SSA database that is based on the performance two years prior to the income reference period.

Table 8 Personal income tax and social insurance contributions: Aggregate amounts

Model Weights	EUROMOD		SIMTASK	
	original	calibrated	original	calibrated
	2011			
Personal income tax	1.17	1.05	1.14	1.03
Social Insurance Contrib. (SIC)				
SIC: Employer	1.24	1.03	1.23	1.02
SIC: Employee	1.25	1.04	1.25	1.04
SIC: Self-employed	1.72	1.77	1.44	1.50
Health Insurance Contrib. (HIC)				
HIC: economic active pop.	1.19	1.01	1.18	1.00
HIC: economic inactive pop.	1.02	1.09	0.94	1.02
	2012			
Personal income tax	0.94	0.97	0.91	0.95
Social Insurance Contrib. (SIC)				
SIC: Employer	1.09	1.01	1.09	1.01
SIC: Employee	1.11	1.02	1.11	1.02
SIC: Self-employed	1.63	1.67	1.38	1.41
Health Insurance Contrib. (HIC)				
HIC: economic active pop.	1.07	1.01	1.06	0.99
HIC: economic inactive pop.	1.03	1.06	0.96	0.99

Source: Authors' calculations using SIMTASK and EUROMOD, official statistics Ministry of Finance (PIT and HIC), SSA (SIC).

Note: Numbers display ratios of aggregate number of recipients computed by EUROMOD and SIMTASK to aggregate number of recipients referred by official statistics.

5.2 Validation of the predictive accuracy

SIMTASK is designed so that it can be used also for ex-ante evaluation of the proposed legislative reforms of Slovak tax and social system. In order to test for the predictive accuracy of SIMTASK we have performed a simulation exercise. We show that simulation results match the official statistics adequately and the observed discrepancies are qualitatively not different from those reported in our previous ex-post simulations.

As it has been already outlined above, we proceed in two steps. First, selected income variables in the input SK-SILC dataset (income reference year 2012) were updated with the corresponding growth factors to refer to 2013 and 2014, respectively. In the next step, new weights in the updated

datasets were calibrated to match the population totals in 2013 and 2014 using the selected socio-demographic groups, groups defined based on economic activity and labour income distribution defined in terms of calibration factors.

Table 9 Simulated benefits (calibrated weights)

Aggregate amounts (I)					
	2013	2014		2013	2014
Unemp. benefit	0.41	0.48	Child birth grant	1.03	1.21
Parental allowance	0.94	0.92	Material needs benefit	1.32	1.15
Child benefit	1.15	1.15			
Recipients (II)					
	2013	2014		2013	2014
Unemp. benefit	0.48	0.54	Material needs benefit	1.07	1.05
Parental allowance	0.97	0.97	Housing allowance	1.10	1.03
Child benefit	1.08	1.09	Activation allowance	1.59	1.31
Child birth grant	0.98	1.21	Protection allowance	1.07	0.86

Source: Authors' calculations using SIMTASK, official stat. SSA (unemployment benefit), COLSAF (other benefits).

Note: Numbers display ratios of aggregate amount of payments to individuals (I) and number of recipients (II) computed by SIMTASK to official statistics.

Aggregate validations of simulation of transfers, tax and social security contributions are summarized in Table 9 and Table 10. Overall picture is comparable to validation statistics of simulations for 2012 when calibrated weights have been used. This is not a surprise since the same underlying input dataset has been used, although weights were calibrated differently using the updated external statistics. To sum up, observed departures from the official statistics (either under- or over-sampling) are similar both in direction and magnitude to those reported for 2012.

Table 10 Personal income tax, social and health insurance contributions (calibrated weights)

	2013	2014		2013	2014
Personal income tax	1.00	0.99	SIC: Employer	0.99	1.00
HIC: economic active pop.	1.04	1.06	SIC: Employee	1.00	1.00
HIC: economic inactive pop.	0.96	1.02	SIC: Self-employed	1.42	1.71

Source: Authors' calculations using SIMTASK, official stat. Ministry of Finance (PIT and HIC), SSA (SIC).

Note: Numbers display ratios of aggregate amount of payments computed by SIMTASK to aggregate amounts computed by using official statistics. SIC stands for social insurance contributions and HIC for health insurance contributions.

5.3 Impact on income distribution, inequality and poverty measures

In the following part we present a comparison of indicators of income distribution, inequality and poverty reported by Eurostat and estimated by EUROMOD and SIMTASK. Results published by Eurostat are reported for reference, as they are not directly comparable to the estimates by EUROMOD and SIMTASK. Several reasons may explain differences between computed results. In particular, although Eurostat results are also based on SILC data and use the original weighting scheme, equalised households' disposable income definition includes different components compared to definitions used by EUROMOD and SIMTASK (that are comparable). On top of that, some income sources that enter to the computation of disposable income may have different values due to the fact that they are simulated in EUROMOD and SIMTASK.

We can, however, provide a meaningful comparison of inequality measures calculated on data generated by different simulation tools. This should give the reader a flavour of the impact of the weights calibration and of the closer match of the model with legislation on simulated inequality measures. Indicators in EUROMOD and SIMTASK were computed using the same methodology like Eurostat. In particular, results were calculated on the basis of the total equalised disposable income attributed equally to each member of the household. Disposable income is defined as a sum of all monetary income sources of all household members net of paid payroll taxes. Household members are equalised by weighting using the modified OECD equivalence scale that assigns a value 1 to the household head, 0.5 to each additional adult member and 0.3 to each child under 14.

The distribution of equalised disposable income by deciles is reported in Table 11. Results show that shares of disposable income are very similar when estimated by EUROMOD or by SIMTASK and when using the same weighting scheme. Differences can be noticed when outcomes using original and calibrated weights are contrasted. To sum up, differences between results from Eurostat, EUROMOD and SIMTASK are small, whereas estimates based on calibrated weighting scheme appear to be closer to Eurostat results.

Table 11 Shares of equalised disposable income by deciles

Weights	EUROSTAT	EUROMOD		SIMTASK	
		original	calibrated	original	calibrated
	2011				
Decile 1	3.5	3.9	3.7	3.8	3.5
Decile 2	5.7	6.1	5.7	6.0	5.6
Decile 3	6.9	7.5	7.1	7.1	6.8
Decile 4	7.9	8.2	8.1	8.0	7.7
Decile 5	8.6	8.9	8.8	8.9	8.6
Decile 6	9.5	9.8	9.4	9.8	9.5
Decile 7	10.7	10.6	10.5	10.9	10.7
Decile 8	12.1	11.9	11.8	12.1	12.1
Decile 9	14.1	14.0	14.4	14.0	14.6
Decile 10	21.0	19.1	20.5	19.3	20.8
	2012				
Decile 1	3.6	3.9	3.5	3.9	4.0
Decile 2	5.7	6.2	5.9	6.1	5.4
Decile 3	6.8	7.4	7.1	7.3	7.0
Decile 4	7.7	8.2	8.0	8.2	7.9
Decile 5	8.7	9.0	8.7	9.0	8.7
Decile 6	9.7	9.8	9.6	9.9	9.7
Decile 7	10.9	10.8	10.7	10.9	10.8
Decile 8	12.3	11.9	12.0	12.1	12.3
Decile 9	14.3	13.9	14.3	14.0	14.4
Decile 10	20.3	18.8	20.1	18.7	19.9

Source: EUROSTAT and authors' calculations using SIMTASK and EUROMOD.

Note: The ratio of disposable income in the corresponding decile to the population.

Computed for individuals based on household disposable income and equalised by the modified OECD equivalence scale.

In Table 12, some income inequality and poverty measures are presented. Differences in estimated indices by EUROMOD and SIMTASK that can be attributed to refined simulations are small (when EUROMOD and SIMTASK with the same weights are compared). Disparities are larger when the results based on two weighting schemes are compared. In accordance with disposable income distribution presented in Table 11, when calibrated weights are used, results are closer to official figures reported by Eurostat.

Table 12 Income inequality and poverty rates

	EUROSTAT	EUROMOD		SIMTASK	
		original	calibrated	original	calibrated
Weights					
		2011			
GINI	25.7	23.3	25.6	23.5	26.0
S80S20 ratio*	3.8	3.3	3.7	3.4	3.9
At risk of poverty rate**					
Total population	13.0	12.1	13.4	12.1	13.5
Females	12.8	12.0	13.5	12.0	13.6
Males	13.1	12.3	13.3	12.2	13.3
		2012			
GINI	25.3	22.4	24.4	22.6	24.7
S80S20 ratio*	3.7	3.2	3.6	3.3	3.6
At risk of poverty rate**					
Total population	13.2	11.5	12.9	11.7	13.1
Females	13.2	11.3	12.4	11.5	12.8
Males	13.3	11.6	13.3	11.9	13.4

Source: EUROSTAT and authors' calculations using SIMTASK and EUROMOD.

*The ratio of total income received by the top 20 % of population to that received by the bottom 20 %.

**Percentage of population below 60% of median equalised income.

Measures are computed for individuals based on disposable household income and equalised by the modified OECD equivalence scale.

6 CONCLUSION

This paper provides a summary on the construction of the Slovak tax and transfers microsimulation model SIMTASK. This model has been built up due to the Slovak Council for Budget Responsibility's (CBR) need to have a model being able to assess the static effects of policy changes as well as the long-run consequences of tax and benefit reform strategies. Therefore, a microsimulation model, which works in the common environment and thus can be easily incorporated as a part of more complex models used within CBR was developed.

A number of challenges were addressed during the process of development. First, we considered issues related to the simulation of social structures themselves, i.e. we identified possible improvements (compared to Slovak component of EUROMOD) such that the national tax and benefit system can be replicated as closely as possible. At this point, a major task was to precisely replicate the valid legislation in the corresponding years. At the same time, we inspected the used micro dataset in great detail and we compared it with appropriate administrative statistics. We re-weighted the input data sample such that the new calibrated weights replicate, among other factors, also the earned income distribution and selected age cohorts directly. Hence, the validity of

simulated output was interpreted further in light of differences between simulations using original and new weighted survey data on one side and the official statistics on the other side.

We conclude that weight calibration considerably improves the fit of the model with respect to important income tax and social security contributions categories. However, some distortions when using calibrated weights result too. These involve mainly non-simulated transfers with a lower number of recipients (and in lower total volumes) such as benefits for orphans, disabled or maternity benefits. Weights calibration helps SIMTASK to become a more convincing tool to simulate and evaluate ex-post and ex-ante the impact of selected tax and transfer system policies. However, we showed that re-weighting is not a panacea and the focus of further analysis that user is interested in is of relevance.

REFERENCES

Callan T, Keane C, Walsh J R and Lane, M (2010) 'From Data to Policy Analysis: Tax-Benefit Modelling using SILC 2008', Papers WP359, Economic and Social Research Institute (ESRI).

Creedy J and Tuckwell, I (2003) 'Reweighting the New Zealand Household Economic Survey for Tax Microsimulation Modelling', Treasury Working Paper Series 03/33, New Zealand Treasury.

Creedy J (2004) 'Survey Reweighting for Tax Microsimulation Modelling'. In *Research in Economic Inequality*, Volume 12 (ed. by Y. Amiel and J. Bishop), pp. 229-249. New York: AI Press.

Figari F, Paulus A and Sutherland H (2015) 'Micro-simulation and Policy Analysis', in A.B. Atkinson and F.Bourguignon (Eds.) *Handbook of Income Distribution*, Vol. 2B Amsterdam: Elsevier, pp. 2141–2221.

Frankovic B (2013) 'Calibration of weights of statistical surveys in R language', *Forum Statisticum Slovacum*, 19-37.

Glaser-Opitzova H, Ivancikova L and Frankovic B (2015) 'The Calibration of Weights Using Calmar2 and Calif in the Practice of the Statistical Office of the Slovak Republic', European conference on quality in official statistics Q2014, Vienna.

Horvath M, Senaj M, Siebertova Z and Svarda N (2015) 'The End of the Flat Tax Experiment in Slovakia', CBR Working paper 4/2015, Council for Budget Responsibility.

Immervoll H, Lindström K, Mustonen E, Riihelä M and Viitamäki H (2005) 'Static data ageing techniques. Accounting for population changes in tax-benefit microsimulation', EUROMOD Working paper, EM7/05, Institute for Social & Economic Research, University of Essex.

Inglic R, Seljak R, Intihar S (2013) 'Final Quality Report, EU-SILC-2010, Slovenia', Statistical Office of the Republic of Slovenia.

Kump N and Navicke J (2014) 'Re-weighting EUROMOD for demographic change: an application on Slovenian and Lithuanian data', EUROMOD Working Paper, EM13/14, Institute for Social & Economic Research, University of Essex.

LeGuennec J and Sautory O (2002) 'CALMAR2: une nouvelle version de la macro CALMAR de redressement d'échantillon par calage', *Actes des Journées de Méthodologies*, INSEE, Paris.

Matsaganis M, Paulus A and Sutherland H (2008) 'The take up of social benefits', Research Note 5/2008, Social Situation Observatory, European Commission.

Porubsky M, Machlica G and Strizencova K (2013) 'EUROMOD Country Report, Slovak Republic (2009-2012)', Institute for Social & Economic Research, University of Essex.

O'Donoghue C and Loughrey J (2014) 'Nowcasting in Microsimulation Models: A Methodological Survey', *Journal of Artificial Societies and Social Simulation*, vol. 17(4), p. 1-12.

Siebertova Z, Svarda N and Valachyova J (2014) 'A Microsimulation model of the Slovak Tax-Benefit System', CBR Discussion paper 4/2014, Council for Budget Responsibility.

Siebertova Z, Senaj M, Svarda N and Valachyova J (2015) 'To Work or Not to Work? Updated Estimates of Labour Supply Elasticities', CBR Working paper 3/2015, Council for Budget Responsibility.

Statistical Office of the Slovak Republic (2014), EU SILC 2012 SR UDB version 20/01/2014.

Statistical Office of the Slovak Republic (2014), EU SILC 2013 SR UDB version 23/07/2014.

Strizencova K and Hagara E (2014) 'EUROMOD Country Report, Slovak Republic (2009-2013)', Institute for Social & Economic Research, University of Essex.

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

Vlacuha R and Frankovic B (2015) 'The Calibration of Weights by Calif Tool in the Practice of the Statistical Office of the Slovak Republic', *Romanian Statistical Review* 2.

Vork A and Paulus A (2014), 'EUROMOD Country Report – Estonia', Institute for Social & Economic Research, University of Essex.

Wiemers J (2015) 'Endogenizing take-up of social assistance in a microsimulation model A case study for Germany', *International Journal of Microsimulation* 8(2), 4-27.

APPENDIX

Table A1 Summary statistics SK-SILC 2013 (income reference period 2012)

Weights	Mean (Std.Dev)			Number of individuals		
	SK SILC original	SK SILC calibrated	Admin. data	SK SILC original	SK SILC calibrated	Admin. data
Labour income (monthly)						
Gross wage employment	692.97 (369.69)	722.59 (530.82)	805.00 * 775.60 **	1,983,176	1,788,927	1,881,598
Income from self-employment	722.89 (558.06)	709.50 (540.66)	382.43 **	347,397	361,935	277,125
Other payments made by employers	17.24 (20.29)	20.79 (29.89)		315,322	278,861	
Income from agreements	70.16 (154.41)	68.96 (93.76)	143.94 **	364,458	341,729	955,330
Non-labour income (monthly)						
Unemployment benefit	271.20 (164.21)	259.96 (148.52)	312.00	58,064	60,399	143,896
Maternity benefit	431.87 (179.48)	451.57 (168.34)	443.00	35,203	71,317	24,221
Child birth grant	60.46 (22.24)	62.95 (23.17)	153.67	38,020	70,298	56,994
Child benefit (incl additional child benefit)	23.49 (12.27)	23.34 (11.19)	37.78	779,739	738,748	688,344
Parental Allowance	200.71 (78.99)	200.43 (71.97)	195.87	140,111	164,278	142,274
Material needs benefit	132.63 (120.49)	122.99 (103.3)	132.62	102,402	118,566	183,091
Nursing allowance	150.05 (100.47)	140.28 (89.04)	142.22	32,411	41,248	58,700
Sickness and nursing benefits	78.84 (81.78)	101.03 (96.78)	209.70	107,431	139,511	119,092
Disability pension	272.53 (112.53)	261.98 (98.64)	260.90	201,729	257,721	227,801
Old-age pension	375.44 (112.07)	374.58 (108.71)	367.00	1,072,056	1,062,519	980,863
Early retirement pension	348.33 (107.57)	325.15 (104.03)	374.00	16,765	16,862	24,404
Widow's pension	153.16 (97.51)	156.99 (100.6)	237.00	339,647	346,049	336,877
Orphan's pension	159.09 (65.77)	146.41 (55.18)	126.00	19,479	16,972	26,923
Sample size	15,426	15,426				

Source: Authors' calculations using SK SILC, official statistics Statistical Office, SSA and COLSAF.

Note: * Mean monthly wage reported by the Statistical Office, ** means computed by using data from SSA.

Table A2 Simulated benefits: Ratios of aggregate number of recipients

Model	EUROMOD		SIMTASK	
	original	calibrated	original	calibrated
2011				
Unemp. Benefit	0.50	0.44	0.51	0.45
Parental allowance	0.68	1.03	0.67	1.03
Child benefit*	0.88	1.00	1.08	1.15
Child birth grant	0.59	1.07	0.59	1.07
Material needs benefit	1.00	1.36	0.73	1.01
Housing allowance	1.92	2.62	0.79	1.15
Activation allowance	4.68	6.47	1.00	1.50
Protection allowance	1.41	1.87	0.58	0.68
2012				
Unemp. Benefit	0.38	0.41	0.40	0.43
Parental allowance	0.88	0.87	0.87	0.99
Child benefit*	0.96	0.94	1.15	1.08
Child birth grant	0.75	1.02	0.75	1.02
Material needs benefit	1.07	1.43	0.81	1.04
Housing allowance	2.12	2.78	0.82	1.10
Activation allowance	5.25	6.83	1.25	1.56
Protection allowance	2.06	2.80	0.73	0.92

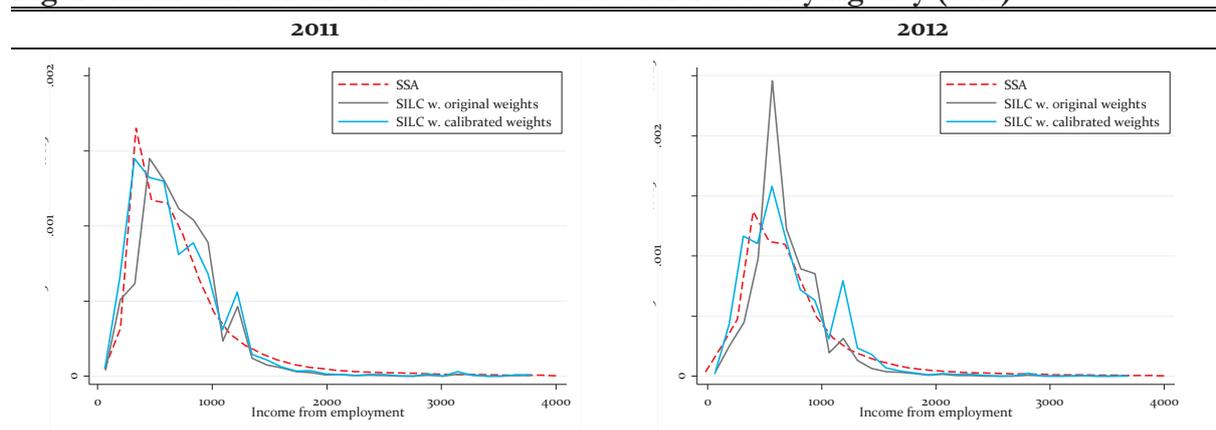
Source: Authors' calculations using SIMTASK and EUROMOD, official statistics SSA (unemployment benefit), COLSAF (other benefits).

* Official statistics on child benefit recipients is taken as the average of monthly data over the year.

Official statistics on other benefits is the total number of individual recipients (incidence).

Note: Numbers display ratios of aggregate number of recipients computed by EUROMOD and SIMTASK to aggregate number of recipients referred by official statistics.

Figure A1 Income distribution SK-SILC and Social Security Agency (SSA)



Source: Authors' calculations using SK-SILC and SSA.

ⁱ We would like to thank the editor of this journal and two anonymous referees for their helpful comments. We are also grateful to Michal Horvath from University of York for his help and support.



CELSI

