ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Myck, Michal; Najsztub, Mateusz

Working Paper

Data and model cross-validation to improve accuracy of microsimulation results: Estimates for the Polish household budget survey

DIW Discussion Papers, No. 1368

Provided in Cooperation with:
German Institute for Economic Research (DIW Berlin)

Suggested Citation: Myck, Michal; Najsztub, Mateusz (2014) : Data and model cross-validation to improve accuracy of microsimulation results: Estimates for the Polish household budget survey, DIW Discussion Papers, No. 1368, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:
https://hdl.handle.net/10419/95243

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

DIW BERLIN

1368

Discussion Papers

Deutsches Institut für Wirtschaftsforschung 2014

Data and Model Cross-Validation to Improve Accuracy of Microsimulation Results

Estimates for the Polish Household Budget Survey

Michal Myck and Mateusz Najsztub

Opinions expressed in this paper are those of the author(s) and do not necessarily reflect views of the institute.

IMPRESSUM

© DIW Berlin, 2014

DIW Berlin
German Institute for Economic Research
Mohrenstr. 58
10117 Berlin

Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
http://www.diw.de

ISSN print edition 1433-0210
ISSN electronic edition 1619-4535

Papers can be downloaded free of charge from the DIW Berlin website:
http://www.diw.de/discussionpapers

Discussion Papers of DIW Berlin are indexed in RePEc and SSRN:
http://ideas.repec.org/s/diw/diwwpp.html
http://www.ssrn.com/link/DIW-Berlin-German-Inst-Econ-Res.html

Data and model cross-validation to improve accuracy of microsimulation results: estimates for the Polish Household Budget Survey

Michal Myck, Mateusz Najsztub*

March 26, 2014

Abstract

We conduct detailed analysis of the Polish Household Budget Survey data for the years 2006-2011 with the focus on its representativeness from the point of view of microsimulation analysis. We find important discrepancies between the data weighted with baseline grossing-up weights and official statistics from other sources. A number of re-weighting exercises is examined from the point of view of the accuracy of microsimulation results and we show that using a combination of demographic calibration targets with several economic status variables or tax identifiers from the microsimulation model substantially improves the correspondence of model results and administrative data. While demographic re-weighting is neutral from the point of view of income distribution, calibrating the grossing-up weights to adjust for economic status and tax identifiers significantly increases income inequality. We argue that although data re-weighting can substantially improve the accuracy of microsimulation it should be used with caution.

Keywords: microsimulation, re-weighting, household data analysis

JEL Classification: D31, I38

*Michal Myck (corresponding author) is Director of CenEA - the Centre for Economic Analysis in Szczecin and Research Fellow at DIW-Berlin (e-mail: mmyck@cenea.org.pl); Mateusz Najsztub is Research Economist at CenEA. Michal Myck gratefully acknowledges the support of the Polish National Science Centre (NCN) under project grant number 6752/B/H03/2011/40. Mateusz Najsztub acknowledges the support of the Batory Foundation under the project number 22078. Data used for the analysis have been provided by the Polish Central Statistical Office (GUS), who take no responsibility for the results of the analysis. The paper builds on the initial analysis conducted in the process of development of the SIMPL microsimulation model (version V4S) in cooperation with Adrian Domitrz, Leszek Morawski and Aneta Semeniuk. Their role in the development of the model has been indispensable. The usual disclaimer applies.

1 Introduction

The majority of large scale household surveys conducted by statistical offices or private survey agencies are conducted in such a way as to be "representative" of the population which the respective samples are drawn from. Due to frequent non-random survey participation this representativeness is usually less than perfect and the problems of under-representation of certain groups of the population - for example the very rich or the very poor - have been long recognized. As we demonstrate using the example of the Polish Household Budget Surveys, if this under- or over-representation of certain groups is unaccounted for in the process of generation of population grossing-up weights, the resulting population structure may differ substantially from administrative records. This in turn has significant consequences for the accuracy of tax and benefit simulations using a microsimulation model and thus the reliability of the model for the purpose of policy analysis (Klevmarken (2002)). Because validity of any microsimulation model relies to a large extent on the degree of correspondence between model outcomes and the administrative records, significant deviations in terms of the age or economic activity distribution and the resulting simulation discrepancies might lead to questioning of the models' role for policy purposes. In such cases even if model calculations for each particular household are correct, the grossed-up values are bound to be wrong.

We examine the data of the Polish Household Budget Surveys (PHBS) for years 2006–2011 from the perspective of tax and benefit microsimulation. The exercise presented in this paper serves primarily the purpose of improving consistency of microsimulation results with administrative data on aggregate tax burden and benefit expenditure. We are thus far from either questioning the general approach of the Polish Central Statistical Office (CSO) to the generation of PHBS grossing-up weights or from arguing that our approach should be applied more broadly in other applications of household micro-level data. We show, however, that a relatively simple method of data re-weighting along the lines of Gomulka (1992), Deville and Sarndal (1992) and Creedy and Tuckwell (2004), which has recently been applied widely in various types of micro-data analysis (e.g.: Brewer et al. 2009 and Navicke et al. 2013) can significantly improve the accuracy of tax and benefit microsimulations in many dimensions. We present an approach to calibration of population weights which extends the criteria used by the CSO in several stages. In the first stage the calibration of weights is done with respect to demographic variables.[1] In the second stage we extend this to include economic status variables, while in the third we use a process of "cross-validation" where we calibrate population weights in the data with respect to a set of tax identifiers related directly to microsimulation. Naturally the latter approach relies on the assumption of the reliability of the basic model outcomes used for the calibration.

The second and third stage of this exercise generate significant improvements in terms of the performance of the tax and benefit microsimulation with respect to a chosen set of key tax and benefit parameters. The re-weighting exercise conducted in this paper demonstrates how a careful approach to household survey data may improve the accuracy of microsimulation of taxes and benefits thereby making the model much more applicable for policy analysis. In

[1] For re-weighting using age groups see: Cai, Creedy, and Kalb (2006).

the Appendix we extend our cross-validation approach further and make more extensive use of the outcomes of microsimulation to correct for the under-representation of high income households using simulated values of tax advantages.

The rest of the paper is organised as follows. In Section 2 we briefly describe the Polish Household Budget Survey data used for the analysis including the sampling frame of the survey and the approach of the Central Statistical Office to the computation of population weights provided with the data, referred to as "baseline" weights. Using these weights in Section 3 we present the differences in the age distributions between the baseline PHBS data and other official sources on the demographic structure of the population as well as the correspondence of the economic status information in the data and administrative statistics. The divergence between these distributions forms the principal motivation for the weight calibration in the paper. In Section 3 we show how these underlying differences find their reflection in discrepancies of microsimulation results using a number of key tax and benefit parameters from the microsimulation model SIMPL which is applied to the PHBS data.[2] The method of weight re-calibration, which follows Gomulka (1992) and Creedy and Tuckwell (2004) is briefly discussed in Section 4 including details on different stages of calibration. Results of the process in the form of a comparison of tax and benefit microsimulation outcomes with administrative information is presented in Section 5. It should come as no surprise that the consistency of the grossed-up population structure in the data in terms of demographics and economic activity has a substantial influence on the accuracy of microsimulation. Interestingly weight adjustments for economic activity status and tax identifiers have significant implications for the level and trends in inequality indicators. Conclusions including some words of caution regarding the methods applied in this paper follow in Section 6.

2 Polish Household Budget Survey data

The Polish Household Budget Survey (PHBS) is an annual representative survey covering in recent years over 37,000 Polish households. The first survey was conducted in 1957 and it has been a regular source of information on income, consumption and quality of life of Polish citizens. Through the years it has undergone a number of more and less significant methodological changes, related among other things to the political and economic transition after 1989 and various forms of standardization to international procedures, but the survey continues to collect detailed information on the household structure, income sources and household expenditure. The information covered by the survey includes:

- socio-demographic composition of the household;
- life quality and housing conditions;
- durable goods and housing equipment;
- economic activity of household members;

[2]For more details about the model see: Bargain et al. 2007 and Morawski et al. 2008.

- level and sources of individual and household-level incomes;
- detailed household expenditures.

The PHBS methodology has changed over the years, but fieldwork procedures and the overall methodology since 2006 has seen only minor modifications (Główny Urząd Statystyczny (2011a)). In recent years the sample includes all households with the exception of collective dwellings such as prisons, cloisters, retirement homes or boarding schools. The sampling methodology since 2006 targets 3132 of different dwellings in each month which are selected for the survey giving a total of expected 37584 dwellings each year. Due to the possibility of multiple households in one dwelling and survey interruptions (which are not replaced with reserve households), the actual number of surveyed households may slightly deviate from the target, as summarized in Table 1.

The data from the PHBS has been used in a number of studies of incomes and consumption and has long been the main survey used in the Polish microsimulation model SIMPL.[3] Most of the information collected in the PHBS, and in particular incomes and expenditures, covers the survey period of one month. For every quarter of the year each household being surveyed in that quarter is once again asked to fill a questionnaire regarding durable goods present in the household, as well as rare income and expenditure (e.g. buying or selling property, buying a car, health care services) and other sources of income such as employment fringe benefits.

Table 1 HERE

Table 1 gives a summary of the number of households and individuals in the PHBS in the years covered by the analysis, and the split at the household and individual level by some main characteristics. Over the years we can observe the increasing number of people with higher education and the higher (unweighted) average age of participants in the survey. The average household size fell from over three individuals per household in 2006 to 2.87 in 2011.

2.1 PHBS sampling scheme

PHBS relies on a two-stage random sampling scheme with clustering and rotation (Lysoń (2012)). First the country is divided into around 30,000 Primary Sampling Units (PSUs) consisting of at least 250 dwellings in the urban areas and at least 150 dwellings in rural areas. The PSU's are clustered into 109 layers. 1566 PSUs are selected and divided to two sub-samples containing 783 PSUs each. Each sub-sample is drawn for two subsequent years and is exchanged every year forming two rotation groups. In the second sampling stage 24 dwellings are drawn in each PSU (two for each month of the survey) together with additional 150 reserve households in case of refusal of participation among the primary dwellings. All households in every dwelling are included in the survey.

[3]For examples of analysis using the PHBS see for example: Brzeziński and Kostro (2010), Brzeziński (2010), Morawski and Myck (2010), Haan and Myck (2012), Myck, Kurowska, and Kundera (2013).

Importantly from the point of view of this analysis the sampling scheme determines the way observation weights are assigned to each household as the inverse of selection probability for every household. These weights are then adjusted by post stratification based on the data from the National Census (2002 census used for years before 2010 and the 2010 census used in later years). Stratification is based on 12 strata. The reference characteristics used to form the strata are the place of residence (rural or urban) and size of the household (single, 2 persons, 3, 4, 5 and 6+ persons).[4] No additional information on sex, age or education is included in the generation of sample weights. As we show below the resulting distribution of even such basic characteristics as age or economic status may as a result significantly differ from the National Census data and other administrative records. While such discrepancies might not matter in many types of analysis they are of crucial importance from the point of view of reliability and policy relevance of results from microsimulation studies which often present grossed-up population values of the elements of the tax and benefit system.

3 Grossed-up PHBS and other data sources on the Polish population

Validation of survey data against other sources is notoriously problematic given various definitional differences and the nature of the specific survey. Thus not only grossing-up weights of the survey data will determine discrepancies between different sources of information. In this Section we present three types of variables from the PHBS which are set against other data sources in a validation exercise using the baseline grossing-up weights provided by the CSO (and derived along the lines outlined above). These three types of variables are:

- demographics: age, education, residence;
- economic status: employment, self-employment, pension and unemployment benefit receipt;
- microsimulation output: aggregated tax and benefit values; the number of tax payers and benefit recipients.

The grossed up values of these variables from the PHBS together with the most appropriate counterpart information from other sources are presented in Tables 2, 3 and 4 respectively. The information used to validate the PHBS data derives principally from CSO's Statistical Yearbooks based on alternative data sources (principally National Census data). Administrative information on taxes, insurance contributions and benefits comes from published and online statistics of the Ministry of Finance, the Ministry of Labour and Social Policy and the Polish Social Security Institution (ZUS).[5]

[4]For discussion of stratified sampling see: Cameron and Trivedi (2005).

[5]Information on age composition for 2006 and 2007 was taken from the CSO Demographic Yearbook (Główny Urząd Statystyczny (2006 - 2007)). For the years 2008 – 2010 from population reports as of 30 June (Główny Urząd Statystyczny (2008 - 2010)) and for 2011 from the National Census report

As we can see from Table 2 the gross-up population of Poland using the CSO baseline weights accounts for about 98-99% of the total population in the official statistics. This small discrepancy is partly driven by lack of survey coverage of collective dwellings, but may also be a result of small definitional differences concerning residence status between survey and census data.[6] More importantly though, a more detailed look at the population distribution by age suggests significant under and over-representation of some of the age groups. Details for the years covered by the analysis are given in Figure 1 in the form of population pyramids by 5-year age groups. The dark-colored bars represent the PHBS population, while the lighter colored ones the census-based official statistics. In all years we examine we find over-representation of children and under-representation of those aged 20-49 in the PHBS data relative to demographic statistics, both among men and women. There is also some under-representation of the oldest groups of the population. Such difference in the demographic structure of the population is to some extent surprising, in particular because in this case there is little room for any definitional discrepancies. As we show below these differences have significant consequences regarding the microsimulation of several important elements of the tax and benefit system. This should not be surprising as many of the elements of the tax and benefit system, such as pensions or family benefits, are related to the demographic composition of households.

FIGURE 1 HERE

Table 2 suggests also that the PHBS under-represents individuals with higher education by about 20% relative to other sources, while in Table 3 we can see that relative to external statistics employees are over-represented in the sample by about 10% and the self-employed are under-represented in the more recent years of the data. In these cases there might be issues of definitional comparability of the different sources of data and of survey measurement error. Some of the over-representation of employees may also arise from the fact that in the data we cannot identify unofficial employment and thus treat all declarations of work at "face value" as official employment.[7] Additionally people may confuse their employee/self-employed status given the popular forms of contracts between firms and single person enterprizes. In these cases, in order to lower labor costs, while people are officially self-employed,

(Główny Urząd Statystyczny (2011)). Data on employment, pensions, education and place of residence can be found in the Polish Statistical Yearbooks (Główny Urząd Statystyczny (2006 - 2011a) and Główny Urząd Statystyczny (2006 - 2011b)). More details on pensions are obtained from the Social Insurance Institution (ZUS) reports (Zakład Ubezpieczeń Społecznych (2006 - 2011)). Information on Personal Income Tax were obtained from the Polish Ministry of Finance reports (Ministerstwo Finansów (2006 - 2011)). Data on Family Allowance was taken from the Polish Ministry of Labour and Social Policy reports (Ministerstwo Pracy i Polityki Społecznej (2006 - 2011)), and data on health insurance and some detailed social security statistics is taken from unpublished sources from ZUS - we are very grateful for making these available.

[6]For details see: Nowak (2013) and Główny Urząd Statystyczny (2011). The official statistics published in the Demographic Yearbooks are based on the 2002 and 2010 Census results which for the remaining years are updated with data from local administration registers.

[7]The number of people in the grey economy in 2010 was estimated to be around 800,000 (Główny Urząd Statystyczny (2011b)), but many of these individuals could combine official and unofficial employment. It is likely that unofficial employment would not be declared in the survey.

they perform their tasks in the same way as an employee would. The PHBS data from years prior to 2010 significantly over-represent farmers which may be caused by definitional problems in survey and administrative data, but probably reflects also the fact that weights prior to 2010 were based on the 2002 Farming Census and the structure of farming in Poland has undergone substantial changes since then. The data beginning with 2010 with weights based on the 2010 Farming and National Census are much closer to other administrative records on the number of farmers in Poland. In Table 3 in addition to employment status comparisons we also present the correspondence of the PHBS data with administrative records with regard to the number of recipients of the main Social Security benefits. The correspondence of these numbers to official statistics differs in different years, but the number are generally relatively close. The main exception are Family Pensions, which are under-represented by up to 25%. The explanation behind this is that these pensions include survivors pensions which in the survey are likely to be declared by the surviving spouses as retirement pensions. This in turn might explain the over-representation of retirement pensions in the data.

Table 2 HERE

Table 3 HERE

Table 4 HERE

The values presented in Table 4 compare direct output from the SIMPL microsimulation model to administrative statistics on the main elements of the tax and benefit system. We present the simulated number of individuals contributing to Social Security (SSC), Health Insurance (HI) and Personal Income Tax (PIT), and the number of recipients of Family Benefits including the principal Family Allowance (FA) and four main supplements (for large families: SLF, for starting school: SSS, for child birth: SCB, and education of disabled children: SEDC). In the case of each year we use the SIMPL microsimulation model to simulate the baseline tax and benefit system which operated in that year. Within the HI and PIT categories we show the total number of contributors and additionally list the numbers by those paying contributions on permanent employment and self-employment income. For PIT we also give the numbers of recipients of the Child Tax Credit, a generous tax credit for families with children introduced in 2007. The simulations of Social Security contributions are relatively close to the official figures, with overestimate of 6.3% in 2011. Similarly to the overall employment status we overestimate Health Insurance contributions for the employees (by between 2% and 20%) and underestimate them for the self-employed (by 6% - 17%). The number of people paying income taxes, on the other hand, overall match relatively well with administrative statistics. However, since we are unable to account for the details of tax deductions among the self employed, and cannot identify clearly the specific ways people file their taxes, the number of tax payers in this category is substantially higher compared to the official statistics. Given the over-representation of children in the data it is not surprising that the model significantly overestimates the number of recipients of the Child Tax Credit as well as the means-tested Family Benefits. In the latter case the basic Family Allowance

is overestimated by about 32% in 2006 and by 17% in 2011, while the supplement to large families by as much as 104% in 2005 and 40% in 2011. These figures suggest that while the data generally over-represent families with children, it might be over-representing the households with a high number of children to a much larger degree than households with one or two children.

4 Weight calibration

In this Section we present the approach we take to calibration of weights applied to the PHBS data. In the exercise the primary source of external data with respect to which the baseline weights are calibrated is information on demographics and employment status. This is then supplemented with information on income sources and finally with a small number of variables simulated in the microsimulaiton model. The weight calibration exercise follows the approach of Vanderhoeft (2001) and Creedy (2004) described also in Deville and Sarndal (1992).

The main principle of the approach is that it assumes validity of the "target" data to which the weights are calibrated. In our case, in the first stage it implies the fact that we trust the external information on the age distribution, while in the second also the data on labor market and income status. In the final stage the assumptions also imply that we trust the procedures applied in the tax and benefit model in that they correctly identify those who pay the "target" taxes.

The calibration procedure does not change the observations themselves. Instead, it changes the household weights in such a way as to represent different aggregated population characteristics in the best possible way, taking into account a "minimum-distance" criterium which minimizes the sum of differences between the old and new weights. The general idea is the following. Having m variables and n observations, we have a vector x_{jk}, where $j = 1, 2, \ldots, m$ and $k = 1, 2, \ldots, n$. We can then define population totals for every variable t_j such that $t_j = \sum_{k=1}^{n} d_k x_{jk}$, where d_k are the initial (baseline) weights.

The goal of the exercise is to minimise the distance $w_k G(\cdot)$, where $G(\cdot)$ is a distance function:

$$\min_{w_k} \sum_{k=1}^{n} w_k G\left(\frac{w_k}{d_k}\right) \tag{1}$$

subjected to m calibration constraints:

$$\sum_{k=1}^{n} w_k x_{jk} = t'_j,\ j = 1, 2, \ldots, m \tag{2}$$

where w_k are the new calibrated weights equal to $w_k = g_k d_k$, with g_k representing the factors by which baseline weights are adjusted, and t'_j are the target population totals, set as targets for the calibration exercise.

Different distance functions can be used for the calibration procedure. The approach used here follows the Deville-Sarndal distance function described in Deville and Sarndal (1992) that eliminates negative weights and constrains the new weights not to exceed a

specified lower and upper bound, relative to the old weights. The optimization problem (1) constrained by (2) is solved numerically using the Lagrange multipliers method. More details on finding the solution and properties of different distance functions can be found in Deville and Sarndal (1992), Vanderhoeft (2001) and Creedy (2004). Calibration procedure according to the above methodology is available in Stata in the REWEIGHT package of Pacifico (2010). The Deville and Sarndal (1992) distance function allows setting the minimum and maximum factors by which new weights may differ relative to the old ones, and the package permits automatic adjustment of these values once the initial factors prove too restrictive for the iterative algorithm.

4.1 Three stages of calibration

There is clearly an endless number of ways in which weight calibrations could be conducted, conditional on the choice and number of target variables as well as specific methods of calibration. In this paper we conduct the calibration exercise in three stages in each case using the same set of target variables and the same calibration method for every year of data. The target variables used at each stage of calibration are given in Table 5.

Table 5 HERE

In each of the three calibration stages we target two principal variables which underlie the generation of baseline weights at the CSO in order for the calibration exercise not to diverge from these basic criteria. Thus we target household size (6 categories) and the place of residence (2 groups). These three target variables are generated from the data using the baseline weights so that they remain unchanged in the calibration exercise. The additional criterion on which weights are calibrated in stage 1 (S1) are demographic targets related to the age distribution. In this case the target variables are the official population statistics from Demographic Yearbooks by age group.[8] Stage two (S2) extends these targets by adding seven types of basic income sources as declared in the PHBS, while in stage three (S3) we use two indicators for income receipt (any social security pension and unemployment benefit) and supplement this by a number of outcome variables from the SIMPL microsimulation model. These variables relate to the number of individuals paying Personal Income Tax and employee and self-employed Health Insurance. The starting weights for the calibration exercise are the baseline weights as provided by the CSO. Results generated using these weights are labelled as S(0).

A note of caution is necessary before we discuss the results. As we noted above there is an endless number of weight calibrations one could conduct and there is no obvious or objective criterion one should target. The calibrations presented in this paper aim primarily at finding a practical solution to the problem of under or over-representation of certain groups of the population with respect to data which form a point reference for the purpose

[8]These need to be proportionally rescaled in order to keep the grossed-up population in the PHBS data unchanged. For the full set of references to external statistics see footnote 5.

of microsimulation. We must also remember that there is a limit to the accuracy of calibration and to the number of target variables one can choose, as with too large a number of targets the algorithm may not converge. We thus start from the most basic correction of the discrepancies between the grossed-up survey data and other official statistics, namely the age distribution. As we shall see already this change brings in a significant improvement in the accuracy of a number of simulation outcomes. The next two stages of calibrations aim more specifically at adjusting the distribution with respect to the representativeness of households by the type and level of taxes and benefits they pay. In stage two of the calibration process (S2) the target variables are economic activity categories defined by the source of external income - i.e. income which is not simulated in the microsimulation. This covers market incomes from employment and self-employment and social security benefits such as different types of pensions and unemployment benefit. Because it is unclear if the administrative statistics on employment and self employment correspond to the economic status variables in the survey, in the third calibration stage we replace this by the number of those who pay income tax and health insurance while keeping the total number of pensioners and recipients of unemployment benefit as target variables. The principal reason behind this is to ensure consistency of definitions between administrative and survey data since we no longer need to distinguish between the source of income as long as it is subject to income tax. One argument against this type of approach might be that we may rely too much on the correct nature of the identification of the tax contributors and if these are wrong introduce an erroneous target variable for the calibration procedures. In our view because we target only the number of those paying taxes and health insurance contributions and not its values, the scope for introducing such error is very small and as we shall see in the case of several outcome variables the advantages in terms of the accuracy of simulations are substantial.

5 Results

The effect of weight calibration under the above three scenarios is presented in two separate categories. First we show how the calibrations affect the correspondence of economic status and social security benefit receipt relative to external data, and secondly we present aggregate outcomes of the microsimulation model as compared to administrative statistics. In this way one can immediately see the effect on the distribution of economic status and on the accuracy of simulations relative to administrative data. In a similar way to the detailed results presented in Table 4 for each of the years considered in the analysis we apply the baseline tax and benefit system for the given year. The differences in the grossed-up number of individuals in specific economic status and SSC benefit receipt category as well as with regard to the simulation outputs between S(0) and S(1)-S(3) result purely from changes in the values of grossing-up weights. The results in the form of ratios of PHBS based figures and external sources for the economic status and SSC benefit receipt are presented in Table 6. The headcount figures for the simulated contributions and tax outcomes are shown in Table 7, while for Family Benefits outcomes in Table 8. For these parameters we also show the ratios between the simulated and administrative information for headcount and aggregate amounts

in the form of radar charts in Figures 2 and 3.[9] The Tables and Figures include the same tax and benefit outcomes as those chosen for the baseline validation presented in Table 4. The closer the relative values are to 1, the closer are the simulated values to their administrative counterparts. The list of the tax and benefit parameters and their labels is given in Table 9.

Table 6 HERE

Table 7 HERE

Table 8 HERE

Table 9 HERE

FIGURE 2 HERE

FIGURE 3 HERE

As shown in Tables 7 and 8 and summarized in Figures 2 and 3, weight calibration generally leads to an improved correspondence in the results for most of the selected simulated parameters. The most significant improvements apply to the Family Allowance and its supplements. The biggest relative deviation from the administrative data can be observed in the Supplement for Large Families (SLF) which is oversimulated by nearly 100% in 2006 in terms of the headcount measure and by 43% in terms of amount when using the baseline weights. For all calibration targets in the given period the number of recipients and the aggregate value of the SLF drops substantially and gets closer to the administrative records. The headcount values are still oversimulated but by much less compared to the baseline weights. In terms of total spending most of the simulations generate results closely matching the administrative values. A similar picture can be seen for the Supplement for Starting School (SSS) in years 2010 and 2011.

The results on the contributions side are not as straightforward, and there are important differences between the accuracy of results by headcount and aggregate amounts. Figures 2 and 3 show that, as we would expect given the target variables in S2 and S3, the second and third stage of the calibration substantially improve the respective number of contributors to Social Security and Health Insurance and taxes. This is the result of calibrating the number of recipients of the main types of incomes in S2 and of selected types of contributions in S3. The improvements in terms of aggregate amount of contributions and taxes, however, are less clear cut. The details are presented in Table 7 and we can see that generally while there are improvements in terms of the numbers of contributors to Social Security, both S2 and S3 calibrations result in higher deviations in terms of Health Insurance amounts taxes, although these deteriorations are usually small in magnitude. The reason behind this is that targeting the working population through either income sources or contributions, lowers the weight on market incomes and in particular assigns lower weights to top income recipients in the data

[9] These are generated using radar graphs for Stata (Mander (2007)).

who generate a significant proportion of tax and contributions incomes. In the Appendix we propose an alternative calibration approach which directly uses microsimulation output to re-weight high income households and as a result produces a further correction to the overall values.

The accuracy of calibrations is summarized through a simple indicator covering the correspondence of the selected tax and benefit parameters. This is computed as the total squared relative deviation of the simulated values from administrative data:

$$d = \frac{1}{l} \sum_{i=0}^{l} (1 - s_i)^2 \tag{3}$$

where l is the number of tax and benefit outcomes included in the analysis (as presented in Table 9), and s_i is the ratio of simulated to administrative values for that particular outcome. These indicators are computed separately for the headcount measure and for aggregate amounts of taxes and benefits. The summary of the comparison for the chosen set of parameters is presented in Figure 4. The summary results reflect our detailed discussion above and overall for the second and third calibration stages show improvements in the accuracy of simulations - even despite the less precise results on insurance and income taxes in some cases. The only case when the indicator with calibrated weights is higher compared to the baseline is in the case of amounts for S1 calibrations in 2008 and 2009. These results relate to a substantial undersimulation of a number of elements of the Family Benefits system in these two years. It is interesting to note that the summary indicators for the second and third stage of calibrations are very close to each other for all of the analyzed years in terms of both the headcount and aggregate amounts.

Finally, on top of the comparison of the headcount and aggregate amounts we also analyse the extent to which changes in grossing-up weights get reflected in inequality statistics. One could expect that changes in the composition of the population generated by the alternative weights would result in changed distributions of income, and it is interesting to examine whether the approach to weight calibration affects the trends in income inequality under different calibration scenarios. Inequality statistics - the Gini coefficients and the 9/1 decile ratios, together with their 95% confidence intervals, are shown in Figure 5.[10] Differences in inequality measures when calibrating only to age groups are insignificant when compared to the system with baseline weights. This is not the case with the two extended systems. In both cases inequality indicators are substantially higher and the differences with respect to the measures using baseline weights are statistically significant.

FIGURE 4 HERE

FIGURE 5 HERE

[10] To compute inequality statistics we use Inqdeco by Jenkins (1999). Confidence intervals are generated using bootstrapping with 1000 draws.

6 Conclusions and notes of caution

As we saw in Section 5 calibration of grossing-up weights by demographic characteristics (S1) and indicators related to the economic status (S2) and tax identifiers (S3) results in substantial improvements in the accuracy of microsimulation results on the side of taxes as well as benefits. Interestingly major gains in the performance of the microsimulation models appear already after correcting for the distribution of age, which is probably the least controversial and arbitrary adjustment, and these changes are neutral from the point of view of the overall income distribution. The latter two stages of calibration significantly change the distribution of income as reflected in inequality measures, although both of these forms of re-weighting bring additional benefits in terms of accuracy of microsimulation.

There are several reasons why one should approach re-weighting of data with caution, and they apply particularly strongly to the more sophisticated calibrations which are necessarily based on a number of subjective and to some extent arbitrary choices made in the process. When considering the re-weighting procedures one first of all has to bear in mind the reasons for the discrepancies between the grossed-up population on the basis of the survey data and external statistics. In the least controversial cases, as in our case the distribution of age, we can count on reliable and directly comparable external data. The most likely reason for discrepancies in such cases is the selective survey non-response with respect to these characteristics and the fact that the baseline weights do not account for them sufficiently. Adjusting the grossing-up weights in such cases, in particular if it also accounts for categories taken into consideration in generating the baseline weights, seems justified, and as we saw in Section 5 it might bring substantial gains in the accuracy of simulation results.

Discrepancies between survey and external data might however result also for other reasons. These include survey reporting errors (e.g.: Giles and McCrae (1995) and O'Donoghue, Sutherland, and Utili (1999)), definitional differences between survey and external data, as well as in the case of comparisons of tax and benefit values, the issues of computational accuracy of the microsimulation model. For the latter reason one ought to approach the use of microsimulation output data for the purpose of re-weighting with a lot of caution. In the case presented in this paper in the third stage of calibration we rely purely on identification of individuals contributing to taxes and health insurance rather than calibrating specific benefits or values of tax amounts.[11] A clear example of a survey reporting error which we could see in the above analysis is the likely confusion of retirement and survivor pensions, which in the Polish data leads to the seeming under-representation of recipients of Family Pensions (which is how survivor pensions are classified in the Polish system). In this case targeting the number of recipients of Family Pensions in the calibration procedures (which we do in S2) may not bring the desired improvements in overall microsimulation results. Similar arguments apply in the case of potential confusion of contractual employment and self-employment. Re-weighting will by definition give us the correct number of individuals in the target categories but it will not correct the reporting errors and if these are systematic

[11] A more extensive use of microsimulation for the purpose of re-weighting is presented in the Appendix where we re-weight the top end of the income distribution by calibrating the value of a particular tax advantage.

in some dimension it might actually further distort the accuracy of microsimulation rather than improve it. For example, if Family Pensions other than survivor pensions are lower in value (as they in fact are in reality), then calibrating the overall number of recipients of these pensions will not correct the overall total value of these pensions. Similarly, corrections of grossing-up weights using targets in the case of definitional differences might also lead to distortions of the overall picture. A clear example of such problems is the definition of contractual employment, which is one of the main reasons why we use a more reliable target of contributors to Health Insurance and taxes in the third stage of the simulations. From this point of view the third stage of re-weighting might be seen as a trade-off between using a more directly comparable data on contributions at the cost of having to take advantage of microsimulation information in the form of tax identifiers in the survey data. As we saw in Section 5 in our case this trade-off pays off in terms of higher accuracy in many but not all aspects of the simulations.

The final note of caution relates to the issue of the effect of re-weighting on the so-called residual groups, i.e. groups of the population which are largely missing from the specified calibration targets. The best example of such groups in our case are farmers. The potential definitional discrepancies and lack of accuracy of external data which we discussed in Section 3 is the main reason why we leave them out of the calibration process, but the comparison of stage 2 and 3 of re-weighting very clearly reflects the effects of calibrating a high number of targets (in stage 2) on this residual group. We can see for example in Table 6 that the second stage of the calibration has a very substantial effect on the number of farmers in the data, with much less pronounced implications of using the less restrictive set of calibration targets under stage 3.

There are thus a number of reasons why re-weighting ought to be used with caution and why it should take into account a number of issues relating to the way survey data is collected and the nature of external statistics to which we compare the grossed-up values. As the exercise in this paper shows, however, a careful approach to data re-weighting may significantly improve accuracy of and thus make the analysis much better suited for the purpose of policy analysis. As we demonstrated various approaches to re-weighting are possible and some of them will have substantial consequences for the resulting distribution of income in the survey data and the implied levels and trends in inequality. Such differences in the distribution of income in our view deserve further careful analysis. Although they might just be artifacts of the calibrations, there are also reasons to believe that the different levels of inequality in re-weighted data may in fact be more accurate reflections of reality.

Appendix: calibrating the top end of the income distribution using income tax microsimulation

A further possibility of using output from a microsimulation model for sample re-weighting relies on using more specific information generated in the model. In the example presented here we use the information from the model to re-weight the top end of the income distri-

bution in such a way as to match administrative statistics on tax records, and specifically on the tax advantages from joint personal income taxation. The under-representation of high income households is a well known phenomenon in income surveys and has important implications for the simulated values of taxes and contributions (e.g.: O'Donoghue, Sutherland, and Utili (1999)). In this example we use the re-weighting approach to correct this under-representation.[12]

In the Polish tax code it is possible for couples to file taxes jointly, in which case the tax schedule is applied to half of the total taxable income of the couple and the resulting taxes are then multiplied by two. Given the relatively low degree of progressivity in the Polish tax system, this treatment of couples brings highest benefits to high income couples with large differences in incomes between the partners, and specifically when one of the partners has incomes exceeding the higher rate threshold (85,528 PLN or about 20,500 euro per year) and the other one does not.

The Polish Ministry of Finance regularly published the costs of the major tax advantages, including joint taxation which in 2011 cost the budget 2.98bn PLN (0.7bn euro).[13] The simulated costs of these advantages in the SIMPL microsimulation model using the baseline weights is only 2.20bn PLN. The re-weighting conducted here adjusts the household weights in such a way so that the simulated cost of joint taxation matches closely the officially published figure. While the re-weighting targets the value of the tax advantage, the parameter included in the process specifies the number of high income earners in the population, whom we define as individuals with earnings equal to five times the official average gross wage, which in 2011 was equal to 3,399.52 PLN or about 820 euro. Thus the re-weighting implies increasing the number of high earners sufficiently to match the total cost of joint taxation.

We thus first define the cost of joint taxation (JTC) as a function of the number of people (n_{tr}) exceeding the specified income threshold (tr):

$$JTC = JTC_{data}\left(n_{tr}, tr\right) \tag{4}$$

with JTC_0 being the administrative cost of joint taxation we next define our optimization problem as:

$$\min_{n_{tr}} \left(JTC_0 - JTC\right)^2 = \min_{n_{tr}} \left(JTC_0 - JTC_{data}\left(n_{tr}, tr\right)\right)^2. \tag{5}$$

The solution we applied involved using the Nelder-Mead method (Nelder and Mead (1965)), as for each set of weights we need to compare two sets of outputs from the model to specify the cost of joint taxation (one set with and one without the tax advantages). The exercise is presented only for 2011 and in this stage of calibration we use the information on joint taxation on top of the other targets used in stage 3 for the 2011 re-weighting (S3+JTax).

[12] Another approach used in the literature to correct for under-representation of the top incomes relies fitting of the tail of income distribution to Pareto distribution. See for example Brzeziński and Kostro (2010) for an application to PHBS data.

[13] See Ministerstwo Finansów (2012).

The additional criterium results in improved simulation results, as shown in Figure 6 where we present the baseline results, together with results from stages 3 and 4 of reweighting. Compared with other calibration systems we see improvement in simulating the overall amount of PIT. The underestimation of PIT falls from 15% using the baseline weights, to 18% with S3 weights (Table 7) and to 6% if we additionally target the total cost of joint taxation. Naturally, since we increase the grossed-up number of high income households the inequality measures grow further above those obtained with S3 weights. The Gini coefficient goes up significantly from 0.315 to 0.349 and the p90/p10 ratio increases from 3.725 to 4.059.

References

Bargain, O., L. Morawski, M. Myck, and M. Socha (2007): "As SIMPL As That: Introducing a Tax-Benefit Microsimulation Model for Poland," IZA Discussion Papers 2988, Institute for the Study of Labor (IZA).

Brewer, M., J. Browne, R. Joyce, and H. Sutherland (2009): "Micro-simulating Child Poverty in Great Britain in 2010 and 2020," Working Paper 06-31, National Poverty Center.

Brzeziński, M. (2010): "Income Affluence in Poland," *Social Indicators Research*, 99(2), 285–299.

Brzeziński, M., and K. Kostro (2010): "Income and consumption inequality in Poland, 1998–2008," *Bank i Kredyt*, 41(4), 45–72.

Cai, L., J. Creedy, and G. Kalb (2006): "Accounting For Population Ageing In Tax Microsimulation Modelling By Survey Reweighting ," *Australian Economic Papers*, 45(1), 18–37.

Cameron, A., and P. Trivedi (2005): *Microeconometrics: Methods and Applications*chap. Stratified and Clustered Samples, pp. 813 – 857. Cambridge University Press.

Creedy, J. (2004): "Survey Reweighting for Tax Microsimultaion Modeling," *Research on Economic Inequality*, 12, 229–249.

Creedy, J., and I. Tuckwell (2004): "Reweighting Household Surveys for Tax Microsimulation Modelling: An Application to the New Zealand Household Economic Survey," *Australian Journal of Labour Economics*, 7(1), 71–88.

Deville, J.-C., and C.-E. Sarndal (1992): "Calibration Estimators in Survey Sampling," *Journal of the American Statistical Association*, 87(418), 376–382.

Giles, C., and J. McCrae (1995): "TAXBEN: the IFS microsimulation tax and benefit model," IFS Working Papers W95/19, Institute for Fiscal Studies (IFS).

Gomulka, J. (1992): "Grossing-up revisited," in *Microsimulation Models for Public Policy Analysis: New Frontiers*, ed. by R. Hancock, and H. Sutherland, pp. 121–132. Suntory-Toyota International Centre for Economics and Related Disciplines, London School of Economics and Political Science, London.

Główny Urząd Statystyczny (2006 - 2007): *Rocznik demograficzny*, Statystyka Polski. Zakład Wydawnictw Statystycznych, Warszawa.

——— (2006 - 2011a): *Mały rocznik statystyczny Polski*. Zakład Wydawnictw Statystycznych, Warszawa.

——— (2006 - 2011b): *Rocznik Statystyczny Rzeczpospolityj Polskiej*. Zakład Wydawnictw Statystycznych, Warszawa.

——— (2008 - 2010): *Ludność. Stan i struktura w przekroju terytorialnym.*, Informacje i opracowania statystyczne. GUS, Departament Badań Demograficznych, Warszawa.

——— (2011a): *Metodologia badania budżetów gospodarstw domowych*. Główny Urząd Statystyczny, Departament Warunków Życia, Warszawa.

——— (2011b): *Praca nierejestrowana w Polsce w 2010 roku*, Informacje i opracowania statystyczne. Główny Urząd Statystyczny, Departament Pracy, Warszawa.

Główny Urząd Statystyczny (2011): *Wyniki wstępne Narodowego Spisu Powszechnego Ludności i Mieszkań 2011*. Zakład Wydawnictw Statystycznych, Warszawa.

Haan, P., and M. Myck (2012): "Multi-family households in a labour supply model: a calibration method with application to Poland," *Applied Economics*, 44(22), 2907–2919.

Jenkins, S. P. (1999): "INEQDECO: Stata module to calculate inequality indices with decomposition by subgroup," Statistical Software Components, Boston College Department of Economics.

Klevmarken, N. A. (2002): "Statistical inference in micro-simulation models: incorporating external information," *Mathematics and Computers in Simulation*, 59(1–3), 255 – 265, Selected Papers of the MSSANZ/IMACS 13th Biennial Conference on Modelling and Simulation.

Lysoń, P. (2012): *Budżety gospodarstw domowych w 2011 r.*, Informacje i opracowania statystyczne. Główny Urząd Statystyczny, Warszawa.

Mander, A. (2007): "RADAR: Stata module to draw radar (spider) plots," Statistical Software Components, Boston College Department of Economics.

Ministerstwo Finansów (2006 - 2011): *Informacja dotycząca rozliczenia podatku dochodowego od osób fizycznych*. Departament Podatków Dochodowych, Warszawa.

——— (2012): "Preferencje podatkowe w Polsce," Preferencje podatkowe w Polsce 3, Ministerstwo Finansów (MF).

Ministerstwo Pracy i Polityki Społecznej (2006 - 2011): *Informacja o realizacji świadczeń rodzinnych*. Departament Polityki Rodzinnej, Warszawa.

Morawski, L., O. Bargain, M. Myck, and M. Socha (2008): "Model podatkowo-świadczeniowy dla Polski - SIMPL2003," *Wiadomości Statystyczne*, (4), 30–38.

Morawski, L., and M. Myck (2010): "'Klin'-ing up: Effects of Polish tax reforms on those in and on those out," *Labour Economics*, 17(3), 556–566.

Myck, M., A. Kurowska, and M. Kundera (2013): "Financial support for families with children and its trade-offs: balancing redistribution and parental work incentives," *Baltic Journal of Economics*, 13(2), 59–83.

Navicke, J., O. Rastrigina, and H. Sutherland (2013): "Nowcasting Indicators of Poverty Risk in the European Union: A Microsimulation Approach," EUROMOD Working Papers EM11/13, EUROMOD at the Institute for Social and Economic Research.

Nelder, J. A., and R. Mead (1965): "A Simplex Method for Function Minimization," *The Computer Journal*, 7(4), 308–313.

Nowak, L. (2013): *Ludność: stan i struktura demograficzno-społeczna : Narodowy Spis Powszechny Ludności i Mieszkań 2011*. Zakład Wydawnictw Statystycznych, Warszawa.

O'Donoghue, C., H. Sutherland, and F. Utili (1999): "Integrating output in Euromod: an assessment of the sensitivity of multi country microsimulation results," EUROMOD Working Papers EM1/99, EUROMOD at the Institute for Social and Economic Research.

Pacifico, D. (2010): "REWEIGHT: The Stata command for survey reweighting," Center for the Analysis of Public Policies (CAPP) 0079, Universita di Modena e Reggio Emilia, Dipartimento di Economia Politica.

Vanderhoeft, C. (2001): "Generalised Calibration at Statistics Belgium: SPSS Module G-CALIB-S and Current Practices," Working Paper 3, Statistics Belgium, Bruxelles.

Zakład Ubezpieczeń Społecznych (2006 - 2011): *Ważniejsze informacje z zakresu ubezpieczeń społecznych*. Departament Statystyki i Prognoz Aktuarialnych, Warszawa.

Table 1: PHBS sample summary for years 2005–2011.

Year	2006	2007	2008	2009	2010	2011
Number of HH	37508	37366	37358	37302	37412	37375
Number of individuals	114311	111992	109819	108038	107967	107239
Place of residence:						
- town over 500k	11187	11262	10890	10427	10633	10787
- town over 200k	9097	8721	8670	8704	8471	8284
- town over 100k	7949	7820	7101	6700	6617	6918
- town over 20k	19137	18022	17930	17493	16917	17004
- town up to 20k	11675	11957	11401	11568	12516	12076
- village	55266	54210	53827	53146	52813	52170
Gender:						
- adult male	40749	40006	39552	39359	39614	39491
- adult female	46258	45665	45160	44817	44943	44825
Children (<18 y.)	27304	26321	25107	23862	23410	22923
Labour market status:						
- Is employed	44625	45903	46074	44734	44144	43719
- Is self-employed	13017	12476	11833	11607	10771	10274
Education:						
- Higher	10311	10832	11488	12267	13778	14529
- Secondary	29980	29882	29518	29279	28347	28182
- Primary	56402	54215	52509	50765	50108	48873
Mean age (sample)	36.49	36.88	37.54	38.03	38.22	38.55
Mean HH size (sample)	3.05	3.00	2.94	2.90	2.89	2.87

Source: PHBS data 2006-2011, unweighted sample statistics.

Table 2: PHBS and external statistics: socio-demographics for years 2006-2011 using baseline CSO weights.

Variable	2006	2007	2008	2009	2010	2011
Population [M]	37.703	37.708	37.721	37.717	37.726	37.723
- external [M]	38.132	38.116	38.116	38.154	38.187	38.512
- relative	0.989	0.989	0.990	0.989	0.988	0.980
Mean:						
Age	37.101	37.160	37.363	37.486	37.577	37.735
HH size	2.828	2.828	2.829	2.829	2.830	2.829
Residence (in mn):						
T > 200k	8.102	8.117	8.140	8.041	7.933	7.903
- external [M]	7.972	7.946	7.923	7.912	7.902	7.890
- relative	1.016	1.022	1.027	1.016	1.004	1.002
T < 200k	3.063	3.078	2.913	2.911	2.928	3.035
- external [M]	3.059	3.049	3.044	3.040	3.039	3.006
- relative	1.001	1.009	0.957	0.958	0.964	1.010
T < 100k	7.474	7.209	7.384	7.311	7.050	7.194
- external [M]	7.408	7.389	7.403	7.395	7.406	7.514
- relative	1.009	0.976	0.997	0.989	0.952	0.957
T < 20k	4.595	4.834	4.807	4.978	5.336	5.124
- external [M]	4.931	4.933	4.917	4.931	4.918	4.976
- relative	0.932	0.980	0.977	1.010	1.085	1.030
Rural	14.469	14.469	14.477	14.476	14.479	14.468
- external [M]	14.757	14.799	14.848	14.890	14.936	15.153
- relative	0.981	0.978	0.975	0.972	0.969	0.955
Education (in mn):						
Primary	17.940	17.579	17.331	17.028	16.795	16.516
- external	16.574	16.253	15.947	15.653	15.365	15.079
- relative	1.082	1.082	1.087	1.088	1.093	1.095
Secondary	10.351	10.450	10.462	10.454	10.083	10.005
- external	10.908	10.964	11.091	11.147	11.183	11.176
- relative	0.949	0.953	0.943	0.938	0.902	0.895
Higher	3.770	3.989	4.232	4.522	5.073	5.334
- external	4.699	5.075	5.467	5.769	6.075	6.408
- relative	0.802	0.786	0.774	0.784	0.835	0.832

Source: SIMPL model based on PHBS data 2006-2011 and external statistics (see: footnote 5), weighted with baseline weights.

Table 3: PHBS and external statistics: income data for years 2006-2011 using baseline CSO weights.

Variable	2006	2007	2008	2009	2010	2011
Employment type and farmers (in million):						
Employed	14.492	15.216	15.606	15.369	15.224	15.141
- *relative to external:*	1.126	1.154	1.138	1.116	1.101	1.096
Self-employed	3.787	3.710	3.570	3.571	3.357	3.223
- *relative to external:*	0.986	0.973	0.908	0.915	0.861	0.830
Farmers	2.415	2.277	2.140	2.067	1.818	1.715
- *relative to external:*	1.314	1.239	1.164	1.125	0.989	0.914
Temporary employment	0.682	0.594	0.564	0.521	0.463	0.412
- *relative to external*	1.921	1.621	1.491	1.299	1.041	0.810
SSC benefit recipients (in million):						
Retirement pension	6.096	6.151	6.457	6.634	6.545	6.501
- *relative to external*	1.038	1.023	1.043	1.040	1.030	1.033
Disability pension	1.980	1.824	1.690	1.545	1.450	1.398
- *relative to external*	1.048	1.021	1.009	0.986	0.967	0.972
Family pension	1.198	1.149	1.060	1.047	1.039	1.023
- *relative to external*	0.862	0.826	0.763	0.754	0.747	0.735
Pre-retirement pension	0.452	0.370	0.234	0.147	0.157	0.146
- *relative to external*	0.990	1.005	0.914	0.891	1.030	0.981
Unemployment benefit	0.287	0.234	0.187	0.304	0.295	0.245
- *relative to external*	0.867	0.884	0.830	0.909	0.855	0.793

Source: SIMPL model based on PHBS data 2006-2011 and external statistics (see: footnote 5), weighted with baseline weights.

Table 4: PHBS and external statistics: SIMPL output macrovalidation for years 2006–2011 with baseline CSO weights.

Variable	2006	2007	2008	2009	2010	2011
Contributions and taxes: headcount in million:						
Retirement and disability SSC	12.343	13.165	13.641	13.592	13.691	13.682
- *relative to external*	1.036	1.100	1.105	1.076	1.069	1.063
Health Insurance:	20.041	20.644	21.206	21.146	21.158	21.121
- *relative to external*	0.978	0.996	1.005	0.985	0.950	0.946
- permanent employment	10.108	10.980	11.569	11.382	11.539	11.639
- *relative to external*	1.095	1.177	1.208	1.182	1.199	1.199
- self-employment	1.351	1.414	1.420	1.497	1.536	1.503
- *relative to external*	0.926	0.904	0.861	0.885	0.881	0.833
Personal income tax (PIT)	21.791	22.312	22.657	22.569	22.495	22.433
- *relative to external*	0.974	0.974	0.978	0.974	0.966	0.972
- permanent employment	8.015	8.618	9.115	8.767	8.460	6.881
- *relative to external*	1.324	1.344	1.357	1.315	1.304	1.019
- self-employment	0.728	0.769	0.821	0.804	0.793	0.452
- *relative to external*	2.931	3.141	3.398	2.780	2.441	1.310
Child Tax Credit	-	6.869	7.357	7.299	7.461	7.352
- *relative to external*	-	1.142	1.157	1.163	1.187	1.168
Benefit recipients: headcount in million:						
Family Allowance (FA)	6.047	4.896	4.021	3.674	3.319	3.238
- *relative to external*	1.316	1.148	1.067	1.108	1.105	1.170
FA supplements:						
- large families (SLF)	1.582	1.374	1.222	1.092	0.964	0.645
- *relative to external*	2.041	1.944	1.941	1.953	1.910	1.396
- starting school (SSS)	4.156	3.340	2.629	2.348	2.978	2.918
- *relative to external*	1.300	1.107	1.089	0.980	1.395	1.440
- child birth (SCB)	0.232	0.219	0.176	0.158	0.156	0.154
- *relative to external*	0.906	0.983	0.822	0.810	0.857	0.960
- education of disabled child (SEDC)	0.230	0.196	0.176	0.162	0.151	0.163
- *relative to external*	0.999	0.897	0.866	0.871	0.874	1.006

Source: SIMPL model based on PHBS data 2006-2011 and external statistics (see: footnote 5), weighted with baseline weights.

Table 5: Summary of calibration targets.

System	Target variables	Description
S0:	—	Baseline weights
S1:	Household size	6 groups by household size (1, 2, 3, 4, 5, 6+);
	Place of residence	2 groups: rural or urban;
	Age	16 groups by 5 year threshold;
S2:	S1 + recipients of 8 income sources (as declared in PHBS)	employee: permanent and temporary;
		self employment;
		pensions: pre-retirement, retirement, disability and family pensions;
		unemployment benefit;
S3:	S1 + recipients of 2 income sources : (as declared in PHBS)	all pensions;
		unemployment benefit;
	+ SIMPL output:	number of contributors:
		Personal Income Tax;
		Health Insurance on permanent employment;
		Health Insurance on self-employment;

Table 6: PHBS and external statistics: income sources (headcount) by weight calibration for 2006–2011.

Income source/Weights	2006	2007	2008	2009	2010	2011
Permanent employment						
S0	1.159	1.196	1.194	1.183	1.200	1.198
S1	1.222	1.262	1.266	1.264	1.273	1.268
S2	1.000	1.000	1.000	1.000	1.000	1.000
S3	1.087	1.062	1.027	1.034	1.041	1.030
Temporary employment						
S0	1.921	1.621	1.491	1.299	1.041	0.810
S1	1.970	1.634	1.488	1.270	1.007	0.805
S2	1.000	1.000	1.000	1.000	1.000	1.000
S3	3.868	4.110	3.513	3.261	3.011	2.073
Self employment						
S0	0.702	0.729	0.693	0.739	0.758	0.746
S1	0.706	0.737	0.701	0.749	0.770	0.765
S2	1.000	1.000	1.000	1.000	1.000	1.000
S3	0.780	0.843	0.829	0.858	0.892	0.922
Farmer						
S0	1.314	1.239	1.164	1.125	0.989	0.914
S1	1.389	1.322	1.249	1.218	1.070	0.995
S2	1.646	1.700	1.664	1.530	1.438	1.308
S3	1.477	1.478	1.449	1.446	1.260	1.159
Retirement pension						
S0	1.038	1.023	1.043	1.040	1.030	1.033
S1	1.071	1.078	1.092	1.085	1.086	1.077
S2	1.000	1.000	1.000	1.000	1.000	1.000
S3	1.051	1.055	1.071	1.067	1.059	1.068
Disability pension						
S0	1.048	1.021	1.009	0.986	0.967	0.972
S1	1.050	1.027	1.027	1.020	0.995	1.016
S2	1.000	1.000	1.000	1.000	1.000	1.000
S3	1.002	1.002	1.012	1.008	1.016	1.006
Family pension						
S0	0.862	0.826	0.763	0.754	0.747	0.735
S1	0.850	0.808	0.731	0.723	0.730	0.734
S2	1.000	1.000	1.000	1.000	1.000	1.000
S3	0.831	0.789	0.700	0.712	0.723	0.700
Pre-retirement pension						
S0	0.990	1.005	0.914	0.891	1.030	0.981
S1	0.937	0.966	0.906	0.880	1.045	1.011
S2	1.000	1.000	1.000	1.000	1.000	1.000
S3	0.877	0.905	0.899	0.832	0.992	0.949
Unemployment benefit						
S0	0.867	0.884	0.830	0.909	0.855	0.793
S1	0.880	0.909	0.852	0.955	0.887	0.815
S2	1.000	1.000	1.000	1.000	1.000	1.000
S3	1.000	1.000	1.000	1.000	1.000	1.000

Source: PHBS data 2006-2011 and external statistics (see: footnote 5), weighted with baseline (S0) and calibrated (S1-S3) weights.

Table 7: PHBS and external statistics: contributions and taxes in SIMPL by weight calibration for 2006–2011.

Output and weights	2006	2007	2008	2009	2010	2011
Social security contributions (SSC)						
S0	1.041	1.042	1.049	1.061	1.114	1.092
S1	1.090	1.086	1.102	1.122	1.167	1.140
S2	0.962	0.932	0.940	0.965	0.990	0.976
S3	1.006	0.946	0.957	0.971	0.994	0.990
Health insurance contributions (HI)						
S0	0.954	0.937	0.945	0.940	0.937	0.913
S1	0.986	0.972	0.985	0.985	0.977	0.948
S2	0.901	0.866	0.871	0.874	0.859	0.840
S3	0.922	0.870	0.878	0.880	0.862	0.850
Personal Income Tax (PIT)						
S0	0.917	0.901	0.843	0.812	0.882	0.849
S1	0.921	0.921	0.868	0.854	0.922	0.890
S2	0.916	0.920	0.844	0.819	0.871	0.842
S3	0.877	0.862	0.809	0.780	0.842	0.817
Child Tax Credit (in PIT)						
S0	–	1.214	1.241	1.262	1.299	1.278
S1	–	1.132	1.131	1.142	1.171	1.129
S2	–	1.061	1.066	1.062	1.088	1.076
S3	–	1.016	1.011	1.023	1.023	1.046

Source: PHBS data 2006-2011 and external statistics (see: footnote 5), weighted with baseline (S0) and calibrated (S1-S3) weights.

Table 8: PHBS and external statistics: Family Benefits in SIMPL by weight calibration for 2006–2011.

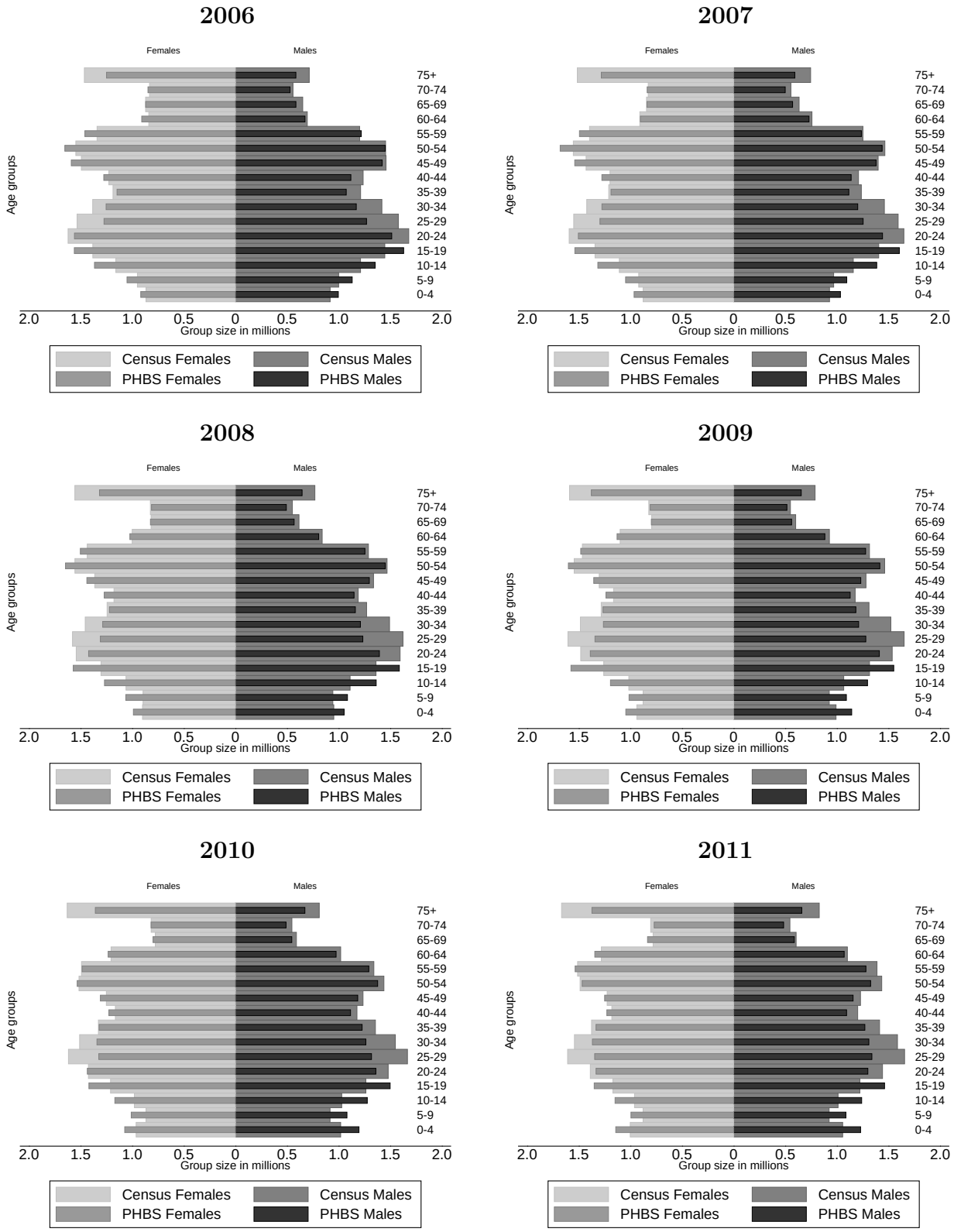
Output and weights	2006	2007	2008	2009	2010	2011
Family Allowance (FA)						
S0	1.319	1.154	1.072	1.124	1.112	1.174
S1	1.147	0.942	0.848	0.860	0.839	0.938
S2	1.181	1.047	0.954	1.047	1.016	1.071
S3	1.177	1.041	0.983	1.005	1.031	1.065
Family Allowance supplements:						
- large families (SLF)						
S0	1.425	1.344	1.344	1.359	1.346	1.397
S1	1.152	0.887	0.875	0.796	0.764	1.015
S2	1.014	0.941	0.909	0.917	0.846	0.893
S3	1.124	1.058	0.933	0.964	1.044	0.927
- starting school (SSS)						
S0	1.216	1.107	1.089	0.980	1.395	1.441
S1	1.043	0.887	0.845	0.735	0.707	0.762
S2	1.076	0.982	0.947	0.902	0.853	0.871
S3	1.068	0.974	0.981	0.858	0.862	0.870
- child birth (SCB)						
S0	1.024	0.984	0.822	0.810	0.857	0.960
S1	0.927	0.838	0.692	0.646	0.674	0.781
S2	0.964	0.913	0.732	0.742	0.772	0.861
S3	1.014	0.957	0.776	0.779	0.784	0.811
- education of disabled child (SEDC)						
S0	1.063	0.904	0.862	0.874	0.866	1.009
S1	0.921	0.719	0.648	0.673	0.667	0.786
S2	1.026	0.858	0.783	0.828	0.833	0.900
S3	1.031	0.874	0.894	0.853	0.854	0.984
Child Birth Allowance (CBA):						
S0	1.119	1.067	1.037	1.069	1.070	1.112
S1	0.888	0.942	0.962	0.952	0.932	0.952
S2	0.989	0.969	0.955	0.960	0.941	0.969
S3	0.943	0.991	1.076	1.051	1.061	1.036

Source: PHBS data 2006-2011 and external statistics (see: footnote 5), weighted with baseline (S0) and calibrated (S1-S3) weights.

Table 9: Elements of tax and benefit used as performance measures

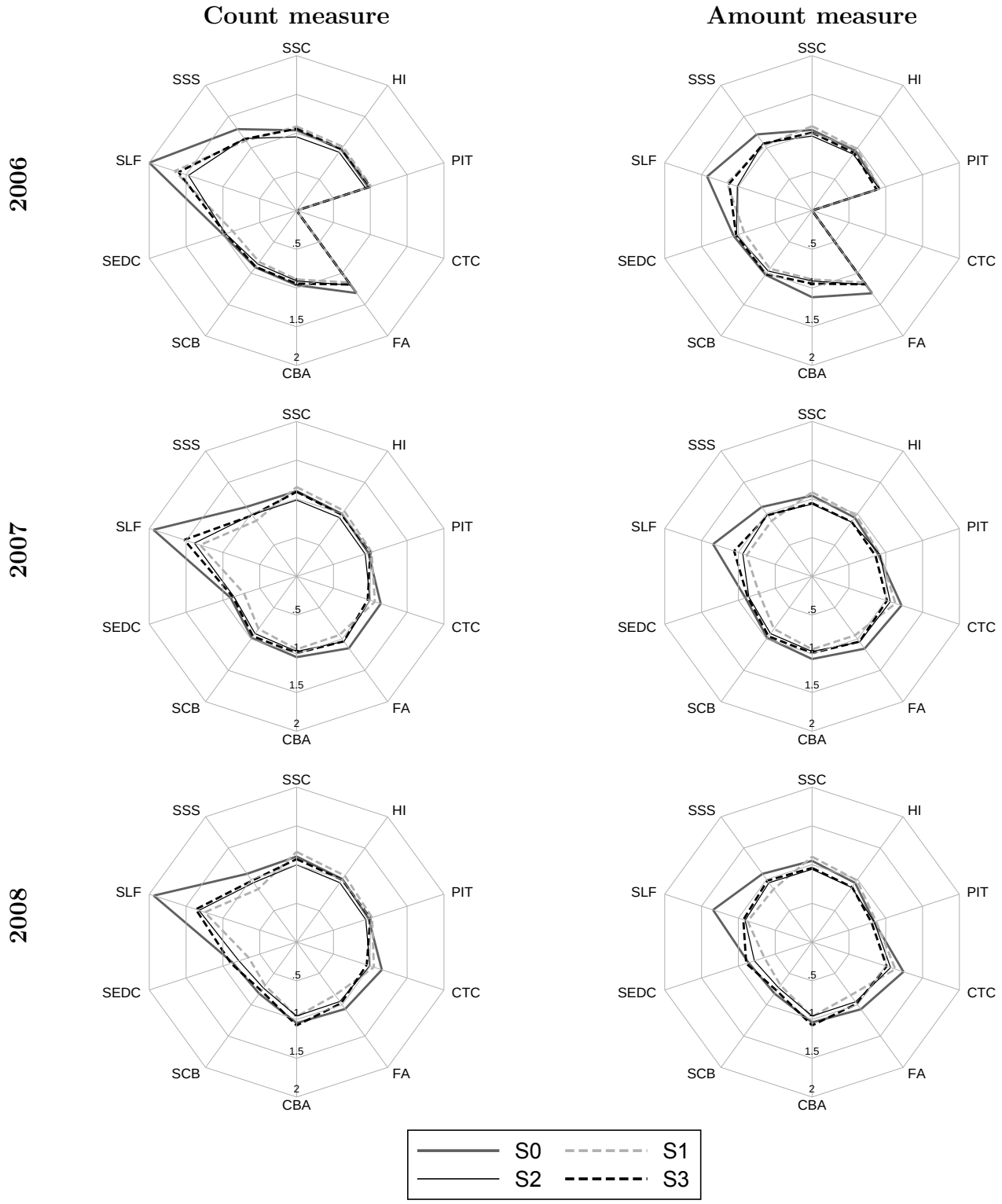
Abbreviation	Full name
Taxes and contributions	
SSC	Social Security Contributions
HI	Health Insurance contributions
PIT	Personal Income Tax
CTC	Child Tax Credit (within PIT)
Family Benefits:	
FA	Family Allowance
SCB	FA Supplement for Child Birth
SEDC	FA Supplement for Education and Rehabilitation of Disabled Child
SLF	FA Supplement for Large Families
SSS	FA School Starting Supplement
CBA	Child Birth Allowance

Figure 1: Population age structure in baseline PHBS and demographic CSO statistics: 2006-2011.



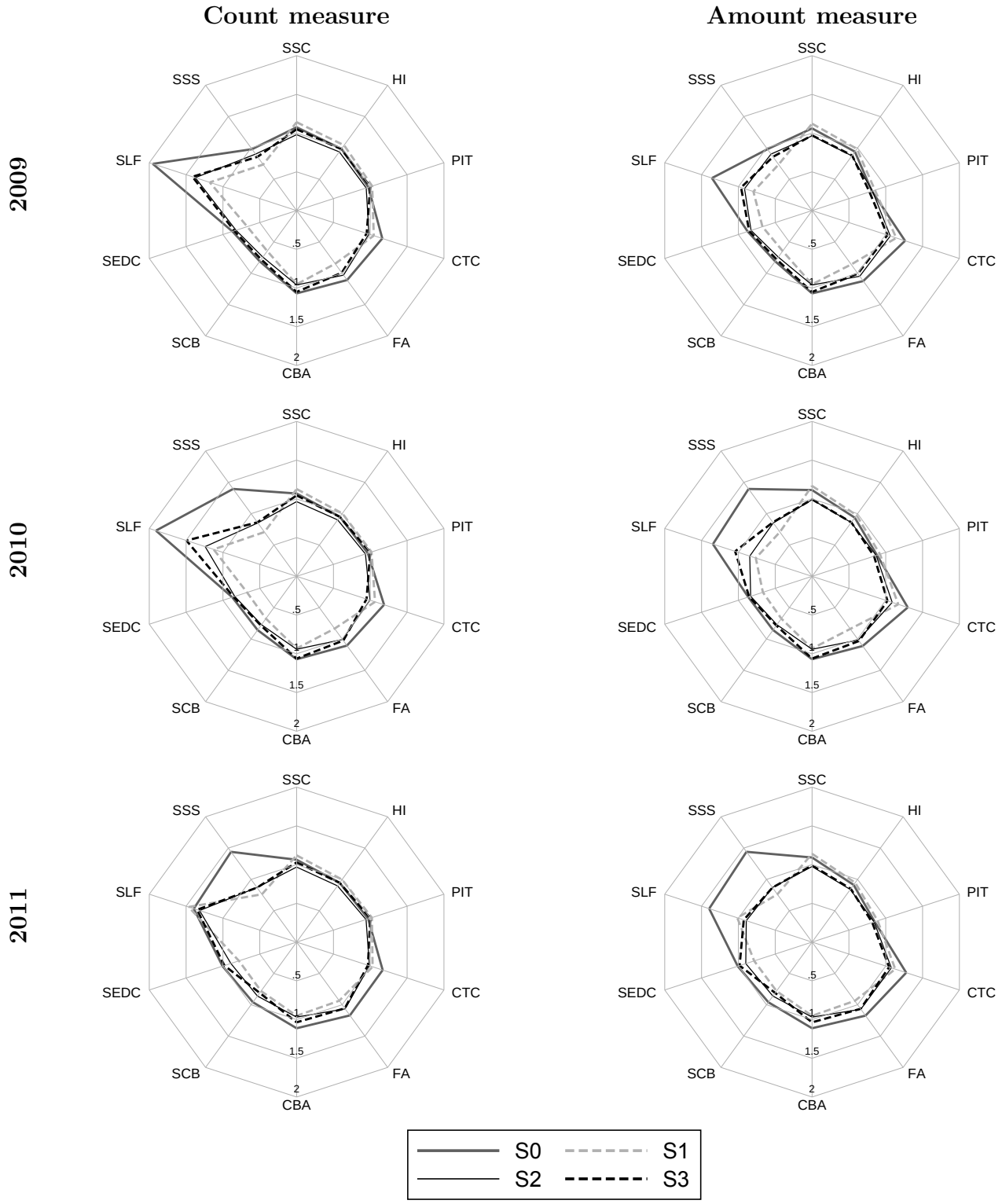
Source: Baseline PHBS 2006-2011 and external statistics (see: footnote 5).

Figure 2: Macrovalidation results: selected tax and benefit outcomes using different weights (S0, S1, S2, S3): 2006, 2007 and 2008 relative to administrative data.



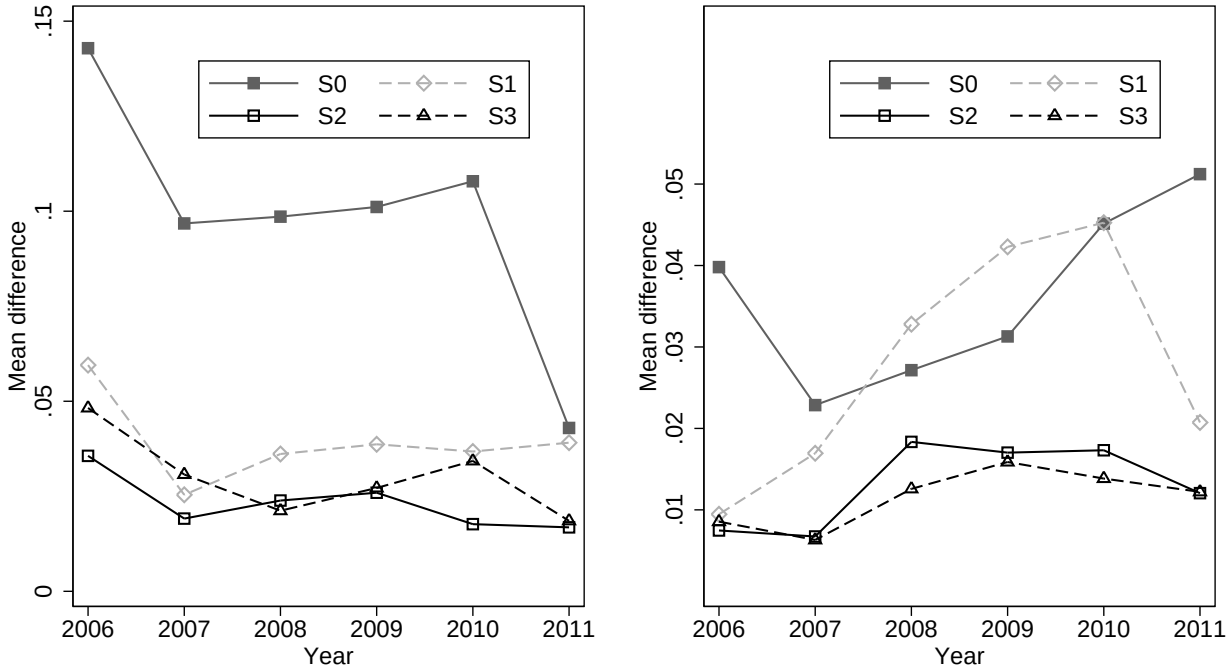
Source: SIMPL model on PHBS 2006-2008 and external statistics (see: footnote 5).

Figure 3: Macrovalidation results: selected tax and benefit outcomes using different weights (S0, S1, S2, S3): 2009, 2010 and 2011 relative to administrative data.



Source: SIMPL model on PHBS 2009-2011 and external statistics (see: footnote 5).

Figure 4: Mean square relative distance measure using different weights (S0, S1, S2, S3).

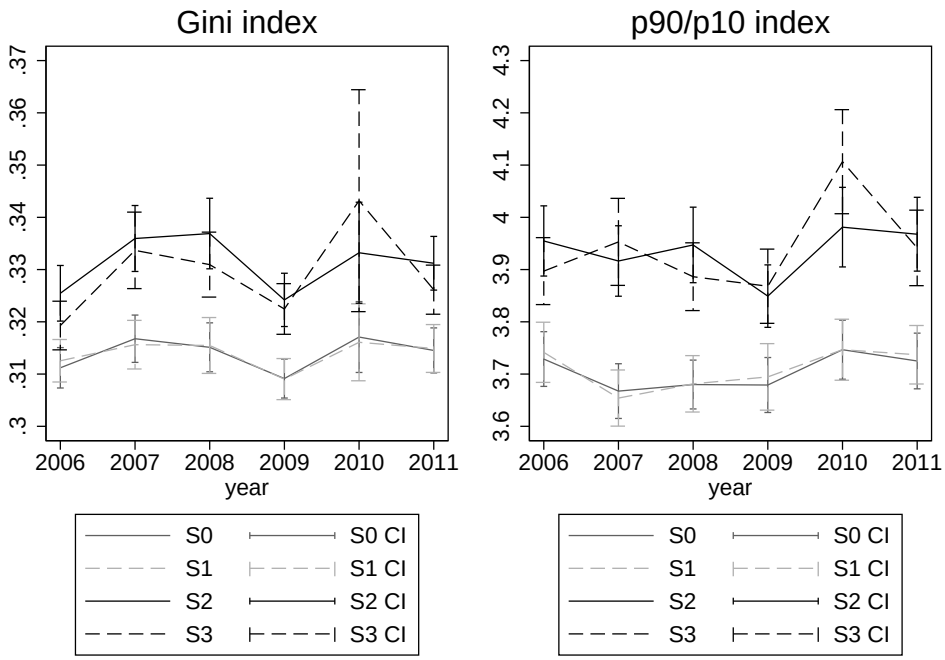


(a) By headcount

(b) By amount

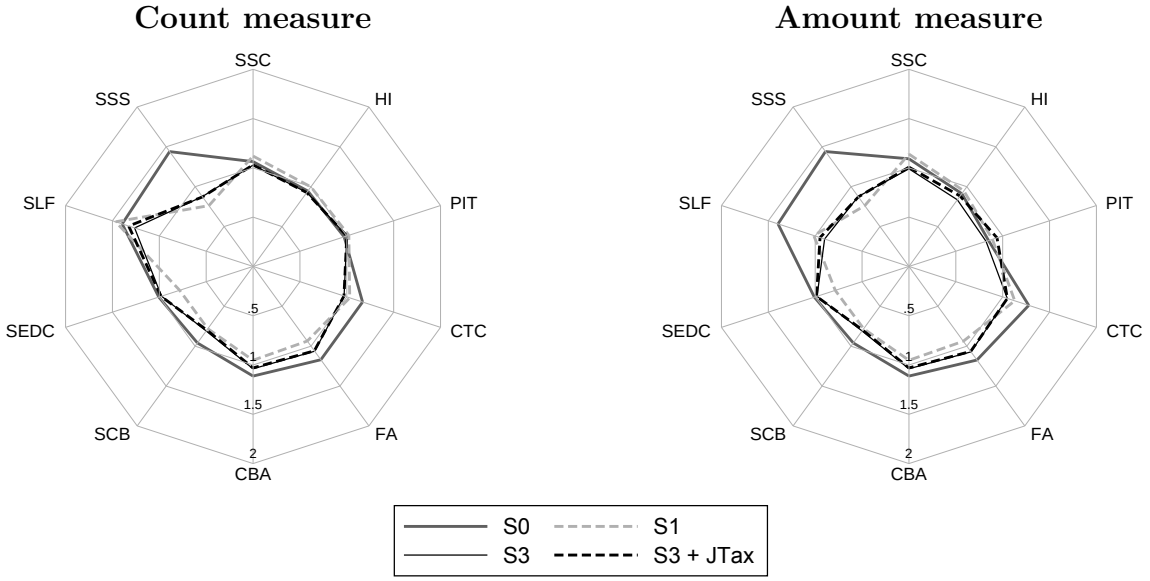
Source: SIMPL model on PHBS 2006-2011 and external statistics (see: footnote 5).

Figure 5: Inequality levels using different weights (S0, S1, S2, S3)



Source: SIMPL model on PHBS 2006-2011.

Figure 6: Macrovalidation results: selected tax and benefit outcomes using different weights (S0, S1, S3, S3+JTax): 2011 relative to administrative data.



Source: SIMPL model on PHBS 2011 and external statistics (see: footnote 5).