

Ericson, Peter; Flood, Lennart; Wahlberg, Roger

Working Paper

SWETaxben: a Swedish tax/benefit micro simulation model and an evaluation of a Swedish tax reform

IZA Discussion Papers, No. 4106

Provided in Cooperation with:

IZA – Institute of Labor Economics

Suggested Citation: Ericson, Peter; Flood, Lennart; Wahlberg, Roger (2009) : SWETaxben: a Swedish tax/benefit micro simulation model and an evaluation of a Swedish tax reform, IZA Discussion Papers, No. 4106, Institute for the Study of Labor (IZA), Bonn, <https://nbn-resolving.de/urn:nbn:de:101:1-2009051329>

This Version is available at:

<https://hdl.handle.net/10419/35720>

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.

IZA DP No. 4106

**SWETaxben:
A Swedish Tax/Benefit Micro Simulation Model
and an Evaluation of a Swedish Tax Reform**

Peter Ericson
Lennart Flood
Roger Wahlberg

April 2009

SWEtaxben: A Swedish Tax/Benefit Micro Simulation Model and an Evaluation of a Swedish Tax Reform

Peter Ericson

Sim Solution

Lennart Flood

*University of Gothenburg
and IZA*

Roger Wahlberg

*University of Gothenburg
and IZA*

Discussion Paper No. 4106

April 2009

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

SWEtaxben*:** **A Swedish Tax/Benefit Micro Simulation Model** **and an Evaluation of a Swedish Tax Reform

The purpose of *SWEtaxben* is to evaluate the impact of changes in the tax/benefit systems on households as well as the central governmental budget. Relating to the micro simulation literature this model can be labeled a static micro simulation model with behavioral changes. This behavioral change takes two different forms and use two different types of models; first binary models that describe mobility in/out from non-work states such as old age pension, disability, unemployment, long term sickness and second models that describe change in working hours and welfare participation. Thus, apart from the choice to work or not to work, working hours conditional on working as well as welfare participation are treated as endogenous variables. As an application the model is used to evaluate the recent Swedish "make work pay" reform, effective from 2007 and further reinforced in 2008 and 2009. The key characteristic of this reform is an in-work tax credit and decreased state tax rate. Simulations performed by *SWEtaxben* show increased working hours at both the intensive as well as extensive margin. The tax decrease together with dynamic changes results in a strong increase in household's incomes but also a reduction in income inequality. However, even considering the increase in hours of work, the reform is far from being self-financed.

JEL Classification: C8, D31, H24

Keywords: micro simulation, tax-benefit system, in-work tax credit reform

Corresponding author:

Lennart Flood
School of Business, Economics and Law
University of Gothenburg
Department of Economics
P.O. Box 640
SE 405 30 Göteborg
Sweden
E-mail: Lennart.Flood@economics.gu.se

* Financial support from the Jan Wallander and Tom Hedelius Foundation for Research in Economics and from the Confederation of Swedish Enterprise is gratefully acknowledged.

1. Introduction

The purpose of *SWEtaxben* is to evaluate the impact of changes in the tax/benefit systems on households as well as the central governmental budget. For the household the main effects are changes in disposable income (income after tax and transfers), labor supply and welfare participation. Since the method is based on micro data, the effects are evaluated for each individual/household, and aggregated measures such as total revenues/expenditure can be obtained and hence the impact of a policy change on central governmental budget can be calculated. Furthermore, due to the micro approach, distributional effects can be produced, thus Gini- or similar measures on income can be presented for each evaluated reform. A part of the evaluation also include a measure of change in social welfare, a welfare function has been estimated and changes in welfare as well as in income and working hours is included in the reform evaluation. The sample used are based on 2006 year LINDA, the sample size correspond to almost 8% of the Swedish population, thus all the output is given with a high precision and since the sampling weights are known aggregate population measures can be produced.

Relating to the micro simulation literature *SWEtaxben* can be labeled a *static micro simulation model with behavioral changes*. Thus, different from a pure static model this model allows for a change in individual behavior as a response to a change in economic incentives. This behavioral change takes two different forms and use two different types of models; first binary models that describe mobility in/out from non-work states such as old age pension, disability, unemployment, long term sickness and second models that describe change in working hours and welfare participation. Thus, apart from the choice to work or not to work, working hours conditional on working as well as welfare participation are treated as endogenous variables.

Compared to most other static micro simulation models *SWEtaxben* stands out in at least three dimensions; the quality of the data, the detailed description of tax/benefit rules and the sophistication of the included statistical models.² The model consists of two major parts, one is a tax/benefit program, which gives a detailed description of the tax and benefit rules, and the other is the prediction program that calls the tax/benefit program (repeatedly) to calculate disposable income and replacement rates needed for prediction of variables such as occupational status, working hours, income, taxes etc.

² Another advantage relates to the maintenance of the changes in tax/benefit rules. The fact that the tax/benefit rules in *SWEtaxben* is a subset of a tax/benefit model supported by Statistical Sweden, makes it is easy to include yearly updates.

In order to present the model we start by describing the data needed for simulation, then the structure of the tax/benefit program followed by the structure of the full model. After that the estimated behavioral models will be presented and a special attention is given to the social welfare function. Finally, as an illustration an evaluation of a proposed tax reform is also provided.

2. Data

The data used for simulation comes from the Swedish register-based LINDA³. LINDA consists of a large panel of individuals and their household members; the sample used in this study comes from the 2006 wave of LINDA. Since the whole LINDA survey of 785 341 individuals is included no selections are used⁴. In order to include all individuals the values for some individuals on some of the variables needed for the simulation have been imputed, most notably hourly wage rate, W , and yearly working hours, H .

Hourly wage data was collected from the official statistics by Statistics Sweden, based on employers' reports. Employers report monthly earnings, expressed in full-time equivalents, thus giving the amount the individual would have earned if employed on a full-time contract. To obtain hourly wage rates, the monthly earnings are divided by 165. Yearly hours of work, H , is then defined as total labor earnings divided by the hourly wage rate. The hourly wage rate used here is different from that obtained by dividing observed earnings by observed hours, which has a tendency to include measurement errors. The data used here is not subject to the same problem since these measures come from employer's record and not from interviews; as a result no top- or bottom coding is needed. However, there are missing values from the employers reported earnings and as usual; a remaining problem is that wage rates are missing for non-working individuals. Wage equation, for males and females separately, are estimated and based on the estimates predicted values are used for those that miss information.

Even if LINDA includes an unusual rich set of information some information is still missing, for instance there is no information about cost of housing, needed as an input in the calculation of housing allowance. Therefore, the cost of housing was imputed using information from an alternative data source, the Swedish Household Income Survey 2006, also supplied by Statistics Sweden. In the imputation we used the method of minimum-distance, using age,

³ For a description of LINDA see Edin and Fredriksson (2000).

⁴ The only requirement is that the individual is included in the population register.

number of children, earnings, place of residence, and citizenship as classification variables. Another variable with no information is cost of child care. However, the childcare fees could still be calculated since the rules are known and since the take-up ratio is high we assume that everyone utilize municipal child-care.

The *SWEtaxben* model includes several statistical models used for predictions. The variables used in these models are: age, education, age and number of children, Swedish born, years of working experience, marital status and region. Apart from these variables also disposable income, or replacement rates, is included; this is calculated in the tax/benefit program as will be explained below.

An important feature of the model is that it allows mobility between different occupational states. For this purpose the sample is divided into different categories based on age or occupation, here the following nine different groups are used:

1. Child, 0-15 years of age
2. Old age pensioner, from age 61-
3. Student
4. Disability pensioner, 18-64 and old age pensioner after 64
5. Parental leave
6. Unemployed, 18-64 and old age pensioner after 64
7. Other (no income from states 2-6, 8, 9 but can have income from social assistance)
8. Long term sick, 18-64 and old age pensioner after 64
9. Working, 18-70 and old age pensioner after 70

This classification relates to full time status during the base year (2006) and is primarily based on the main income source. If an individual gets the largest part from old age pension then he is classified as a pensioner, if it comes from disability pension then he is classified as disabled and so on. There are also some age related criteria that overrules the income source. Thus all individuals less than 16 are classified as a child and all individuals above 70 as an old age pensioner. An individual can only be classified as disabled, unemployed or long term sick up to 64 year of age, above that he is classified as an old age pensioner.

In order to have some point of reference we have adopted the income limits so that the numbers of individuals in the different categories are reasonable similar to statistics based on the labor force survey (LFS). Table 1 compares our classification based on LINDA for year 2006

data with the corresponding LFS statistics. As follows from the table the correspondence is reasonably close.

Table 2 is based on a more detailed classification. Note that it is not possible to distinguish long term sick from working, so working includes both these two groups and also parental leave is not included in LFS.

Again, apart from the group classified as other, the correspondence between LFS and LINDA is reasonable close. However, it is important to remember that the simulation is based on predicted status and not observed, thus predicted before and after a reform is what matters. Since predicted status differ from the statistics in the tables above it is important not to pay too much attention to absolute numbers but rather focus on percentage changes.

Even if all individuals in LINDA are included in the simulations, not all are included in the stochastic models. Individuals classified as children, students or on parental leave as well as older children living with their parents are only included indirectly. All their available information of income is used in the calculation of household disposable income but this information is calculated in the initial start data and does not change in the simulation. Since *SWEtaxben* is a static model people do not age, thus a child is a child before and after a reform (it is also true that a pensioner is a pensioner above age 70). For students we have not tried to model the entry/exit probability. This would be complicated in a static model, but in a future development of the model it is in principle possible to allow also for students to change status due to a policy change. Individuals on parental leave have little incentives to leave this state and this is also a rather small group to be of an interest in a policy simulation (unless the policy simulation is related to a change in parental leave). Again a model that addresses the possibility that economic incentives affect the probability of parental leave can be included in a future version. Older children could in principle be allowed to change behavior; however that requires a model of household labor supply that allows for more than two decision makers per household. We are not aware of any such model, but alternatively they could be treated as singles, however neglecting that they are part of a household is not very realistic. This is also an area for a future version. However, again individuals not included in stochastic models are still included in calculation of household income and in the simulation but they keep their observed information from the start data set.

3. Mini_FASIT a tax/benefit program

The tax/benefit part of *SWEtaxben* which we refer to as *MINI_FASIT* is primarily a tool for calculation of household budget sets.⁵ For the two earners household the budget (disposable income or net income after tax and transfers) evaluated at observed working hours is given as⁶:

$$(1) \quad C = I_m + I_f + B_s + B_h - B_c \quad \text{where } I_i = W_i H_i + Y_i + V_i - t(X_i), \quad i = m \text{ (male), } f \text{ (female)}$$

Apart, from hourly wages, W_i and yearly working hours, H_i , Y_i represent non-earned taxable income (e.g. capital income, old age pension and benefits from unemployment, disability and long term sickness) and V_i non-earned non-taxable income (e.g. child allowance), t is a tax function defined on taxable income, X_i , ($X_i = W_i H_i + Y_i - D_i$, where D_i is deductions for work related expenses or part of premium for private pension savings). The three means-tested (that is dependent on H_i) transfers considered are social assistance (B_s), housing allowance (B_h) and cost of child care (B_c), finally old-age pension (B_p) is also included.

As indicated by (1), disposable income of a household consists of both earned and unearned income. The income components included in V (such as the national child allowance) are not dependent on H and therefore calculated in the start data and kept constant during the simulation. The income components included in Y can be either dependent or independent on H . Capital gains for instance are independent of H whereas unemployment benefits and other transfers that are income dependent are calculated in *MINI_FASIT*. The simple principle (which however is not always followed) is that an income component that is not dependent on H is calculated in start data and kept constant in the simulation, whereas all other incomes are calculated in *MINI_FASIT*. The reason for this principle is that *MINI_FASIT* is called repeatedly for every individual and to speed up the evaluations only the minimum amount of calculation have been included. As mentioned earlier calculations for part of the sample, such as older children living in parent's household, are also done once in the start data and then kept constant.

Given information of individual and household characteristics as well as income components that are independent of H , *MINI_FASIT* includes imputation of all relevant sources of income and taxes needed to calculate individual and household disposable income. First part

⁵ The reason it is called MINI is because it is a simplified version of FASIT, which is a tax/benefit program developed by Statistics Sweden and the Swedish Ministry of Finance. *MINI_FASIT* is a subset of FASIT concerning tax/benefit rules but an extension concerning rules that depends on hypothetical choices. For instance, pensions in FASIT are based on observed pensions from data, in *MINI_FASIT* pensions can be predicted at different retirement ages.

⁶ The one-earner household is of course a trivial simplification.

calculates the individual component in (1), $I_i = W_i H_i + Y_i + V_i - t(X_i)$. The sequence in the individual part is as follows;

First, for individuals classified as disabled, income from disability pension is calculated. Next, we calculate income for those on long-term sicknesses and after that for those who are unemployed. In *MINI_FASIT* income from these sources are basically dependent on potential income from work given that the individual work full time (1 800 hours/year). For instance an individual that is unemployed has a replacement approximately 80 percent of potential income from work⁷. Finally for old age pensioner income from pension is calculated. This step requires detailed information from the start data set as both pensions from the old as well as the new system are considered⁸. For this purpose income records back to 1960 has been used to construct current pension rights. Apart from public pension, income from occupational pension is also calculated, however, due to the complexity in these systems a simplification has been used and all pensioners are compensated according to a system similar to the pension agreement for central governmental employee. Finally income from private pension is also included. This is possible since we have calculated accumulated savings and hence know current private pension wealth for each individual. The pension part of *MINI_FASIT* can be used to calculate the old age pension for any individual above age 61.

After this taxes on income as well as capital gain and properties are calculated and hence at this stage the individual component of disposable income is known. Next step include calculation of the household specific components of three means-tested programs; social assistance, housing allowance, and cost of childcare.

Social assistance is determined by nationwide rules and depends on household composition. To be entitled to social assistance, a household must have an income below the maximum benefit-level. Housing allowance is also determined by nationwide rules. The amount of housing allowance a household is entitled is determined by household total income, rent, the number of children and the age of the parents. The maximum-childcare fee-reform, which was implemented in 2002, is based on household income, but only up to a rather low ceiling above which the fee is constant. For the first child the fee is 3 percent, for the second child 2 percent, and for the third child 1 percent of gross household income. No fees are charged for further children. The ceiling is set fairly low, and as a result most households paid the monthly

⁷ According to the rules for the income year 2006 the replacement is 80 % first 200 days and 70% from 200-300 and 65% thereafter.

⁸ Those born before 1938 are only included in the old system and those born after 1938 are on the old and new system and those born after 1952 are only on the new. The old system is based on income during the 15 best years and the new system on income during whole life.

maximum amount SEK 1,260, 840, and 420 for the first, second, and third child in child care.

Given all information above, the individual part of disposable income as well as the household specific components, household disposable income can be calculated. The design of the model is such that a change in working hours produces a different measure of disposable income and by repeatedly calling *MINI-FASIT* with different hours the household's budget set can be obtained. This budget set is the key variable in the prediction on households' behavior of a policy change. In the models used for prediction of entry/exit from the non-work states (pension/disability/unemployment/sickness) the replacement rate is the economically crucial variable. This replacement rate is defined as disposable income in one of these four states, divided by income from full time work. For the model that predict working hours as well as welfare participation the economic variable is disposable income evaluated at seven different working classes, ranging from zero to long time. This will be described closer below when the sequential structure of a simulation is explained.

4. The structure of *SWETaxben*

The main sequential steps are given in Figure 1. As a first step the model is applied on the tax/benefit rules before the reform, the rules before can of course either be existing rules, for example those existed in 2006, or hypothetical rules. As described above these rules are implemented in *MINI_FASIT*. Next step involves changing *MINI_FASIT* to include new rules, for instance including the changes in the in-work tax credit and governmental tax rate that will be implemented in 2009⁹. Note that if the simulation is a comparison over time, for instance 2006 and 2009, than changes in prices and income has to be considered. This is considered by different price- and income base amounts but earnings have also been indexed and of course also other income sources can be indexed. After the simulation it is possible to deflate the results to prices before or after the reform. For instance the comparison can be made between predicted base scenario in 2006 and predicted outcome of a policy change in 2006 year prices. The steps described below will be the same both for the simulation before as after, the only difference is changes in *MINI_FASIT*.

First step (see Figure 1) involves the definition of replacement rate for disability pension. The population at risk is individuals age 18-64 (but not older children) with status disabled/unemployed, long term sick or working. For couples at least one of the spouses should

⁹ See the Central Governmental budget proposal 2008.

belong to the population at risk. For each individual in this population *MINI_FASIT* is called first to calculate disposable income assuming that everyone is classified as on full time disability. Next, for same individuals, income is calculated assuming full time work ($H=1800$). The ratio, disposable income from disability divided by disposable income from work, is the replacement rate. For instance a replacement rate of 0.7 means that an individual who receives full time compensation from disability insurance, receive 70 percent of the disposable income he would have as a full time worker. A change in a tax/benefit that has an effect on the replacement rate will also have an effect on the probability of entry to, staying in or exit from disability. Given the replacement rate, as well as all other variables included in the model, the probability of disability is calculated. In the calculation of this probability two stochastic terms enter, first a random draw from a normal distribution (with an estimated mean and variance) representing individual heterogeneity and second a Monte Carlo experiment. If the simulated probability is less than a random draw from a uniform (0-1) distribution, then the event takes place; that is, the individual is classified as disabled. Note that the random errors for each individual are the same before and after a reform. Individuals not classified as disabled gets the temporary status (10) and enter the next stochastic model in the sequence.

Next step involves unemployment and the population at risk is unemployed, long term sick, working and those belonging to the temporary status. The steps involved are the same as for disability, thus after this step the individuals in the risk population are either classified as unemployed or in the temporary state.

After this follows the long term sick, the population at risk is now long term sick, working and in the temporary state. Again same procedure and as a result of this module individuals are in the status long term sick or temporary.

The final binary model is old age pension; the population at risk is old age pensioner, other, working or temporary status at age 61-70. An individual below 61 is not eligible for an old age pension and all individuals above 70 are by default old age pensioners. Again after this step individuals are classified as old age pensioners or in the temporary state.

After these binary models a simple imputation follows, all individuals with the temporary status who before the reform belonged to one of the binary states, that is individuals who have exit from one of the binary states, are imputed as entering the working state and is given a working hour equal to 1 800. Note that we are not saying that an individual that for instance belong to disability (before the reform) and predicted to exit disability is given $H=1\ 800$. However, if this individual will pass the remaining three binary states and still be predicted as an exit, then he will get $H=1\ 800$. This concludes the first part of the model where the binary

models are used. Next we will explain imputation of working hours and participation in social assistance.

Every individual in the risk population (status other or working) are considered as working or voluntarily non-working. Thus this is typically the risk population in standard labor supply studies. For every individual in this population *MINI-FASIT* is called repeatedly in order to evaluate the budget set. For individuals classified as singles this requires 14 calls (7 working classes with and without social assistance) and for couples the creation of the budget set requires 98 calls ($7*7*2$)¹⁰. Note that for the couples at least one of the spouses should belong to the population at risk. Given the budget set and all other variables included in the labor supply models, working hours as well as the probability of social assistance is predicted. The stochastic experiment for those models involves draws from an extreme value distribution. Also note that different models have been estimated depending on family type. At this stage of the simulation every individual has a predicted status as well as working hours and welfare participation.

Finally, at predicted status and working hours, *MINI_FASIT* is called a last time to calculate predicted disposable income. Thus, this is the predicted disposable income for the individuals/households that are the results of the tax/benefit rules in *MINI_FASIT*. By changing the rules in *MINI_FASIT* and repeating the simulation we have the possibility of comparing disposable income before and after a reform. Of course the same is also true for any variable of interest that might have changed due to a reform, such as taxes, working hours, labor force participation, social assistance etc. The procedure for a policy evaluation is then to use the saved data sets (at least two) and run a program that creates a large amount of summary statistics that facilitates the policy evaluation. This output includes results from two evaluations; one without any behavioral responses (a traditional static analysis with no change in H), and one with potential behavioral responses where individuals are allowed to adjust hours of work and enter or leave different “out of work” statuses. An example of such an output is displayed at the end of this document.

At this stage a natural question is related to the sequence of binary models. Naturally an interesting and less restrictive approach would be based on a simultaneous model. However this would be much more complicated to implement and the main argument for the sequential structure is simplicity. However, the sequential approach also has the advantage of allowing models that differ in specification, thus it allows a higher flexibility in the specifications. The argument for the specific order in this sequence is based on observing the mobility using real

¹⁰ Of course in practice *MINI_FASIT* is evaluated 7 times for single and 49 for spouses and each time disposable income with and without social assistance is calculated.

data over a few years. We are not arguing that it is a clear pattern that supports the order implemented here but there is some support for it. Also the population at risk involved in the different models is partly based on observing the data and partly on common sense. For instance few individuals move from other (no main source of income except social assistance) to disability, unemployment or long term sick. This is expected since there is an implicit work requirement in the compensation from those states. Thus an individual has to have a work history in order to get unemployment insurance and should not be allowed to move from other (7) to unemployed (6). However a huge complication is that we construct states that are intended to reflect whole year states, in reality an individual can be working a few months become unemployed during a period and perhaps after that long term sick and so on. This makes it much more difficult to exclude some patterns of mobility based on inspection of data. A future improvement of the model includes the possibility of part time states.

Of course the simulation results are dependent on the econometric models and for this reason a detailed presentation follows next.

5. Binary models of occupational status

As mentioned above four econometric models are included in the simulation for the probability of disability, unemployment, long term sickness and old age pension. All of these models have been estimated as dynamic random-effects logit models. The data used for the estimation is a balanced LINDA panel from 2000-2006.

Following Wooldridge (2002, 2005) we define the model,

$$y_{it} = 1 \{ \rho y_{it-1} + \beta_1 x_{1it} + \beta_2 x_{2it} + \dots + \beta_k x_{kit} + c_i + \varepsilon_{it} \}$$

where

$$(2) \quad c_i = \alpha_0 + \alpha_1 y_{i0} + \gamma_1 \bar{x}_{1i} + \gamma_2 \bar{x}_{2i} + \dots + \gamma_m \bar{x}_{mi} + \mu_i$$

with

$$\varepsilon_{it} \sim N(0,1) \text{ and } \mu_i \sim \text{logistic}$$

where y_{it} denote the occupational status of individual i year t (1 if individual is disabled, unemployed and so on), y_{it-1} is the occupational status the previous year and the x -variables are measure of observed characteristics, like age, education and so on. Unobserved determinants are either time-invariant, c_i , or time-variant ε_{it} . The time-invariant unobserved component c_i is

allowed to be correlated with the occupational status of the initial period and the time-average of the explanatory variables. The purpose is to address the initial conditions problem and the possible endogeneity of explanatory variables with respect to time-invariant characteristics. The \bar{x} variables denote the average value for the time invariant variables for each individual.

The variables included in the different models are:

Year effect:

Year=2000, Yes=1, No=0,

Year=2001, Yes=1, No=0,

.

.

If year=2006, Yes=1, No=0,

Age effect:

If age 20-24, Yes=1, No=0,

If age 25-29, Yes=1, No=0,

.

.

if age 60-, Yes=1, No=0,

Educational effect:

If highest education is primary school, Yes=1, No=0,

If highest education is secondary school, Yes=1, No=0,

If highest education is a university degree, Yes=1, No=0,

Regional effect:

If Big city, Yes=1, No=0,

If Medium city, Yes=1, No=0,

If Countryside, Yes=1, No=0,

Other variables:

If born in Sweden, Yes=1, No=0,

Gender, Male=1, Female=0,

Single, Yes=1, No=0,

If income above ceiling for pension rights in the old age pension (7.5 income base amount),

Yes=1, No=0,

Replacement rate, disposable income given full time benefit divided by disposable income given full time market work ($h=1800$),

Occupational status of individual i at $t-1$, Yes=1, No=0,

Occupational status year 2000 (the initial value), Yes=1, No=0,

The results for disability pension are presented in Table 3. The importance of the initial condition as well as the status year $t-1$ signals that disability pension to a large extent is an absorbing state, that is of all individuals that enters this state only a few are rehabilitated and able to exit. The year dummies show an increase over time and the age effect show that older have a much higher probability than younger. Other variables that imply a higher incidence of disability is low education, living in the countryside, being single, born outside Sweden and finally being a male. The replacement rate has a positive effect on the probability of disability. Thus reforms such as an in-work credit tax reform, which lower the replacement rate, reduce the probability of being disabled. However, due to the large lag and initial effect replacement rate plays only a minor role. The estimated ratio of the two error terms, individual over time specific (ρ), shows that over 80% is due to the individual (time-invariant) error term, again, this shows the limited mobility over time.

The estimates for the unemployment model are given in Table 4. As expected the lag and initial condition are relatively less important in unemployment compared to disability. The level of unemployment is increasing up to year 2004 and then decreasing. The age effect is u-shaped, high level for younger low for mid-age and again higher for older. The risk for unemployment is much higher for low educated and also for singles and those born outside Sweden. The replacement rate is positive and the point estimate is smaller compared to the corresponding rate for disability, but the smaller estimates for the initial condition and the lag value indicate that a change in replacement rate potentially has a larger effect on unemployment. This is also evident by the (much) smaller estimate of the error ratio.

Long term sickness (Table 5) is also characterized by smaller inertia, smaller lag and initial condition effect. There is no time trend but a strong age effect implying that older have a much higher likelihood of long term sickness. Individual characteristics have similar effects as for disability and unemployment, thus a low education, born outside Sweden and single all signal a higher probability of long term sickness. In contrast to unemployment, males have a higher probability. The replacement rate indicates a relatively high sensitiveness. Thus, the low error

ratio, only 25% due to individual heterogeneity, together with a rather weak lag and initial condition effects imply that exit/entry for long term sickness is affected by changes in tax/benefit rules.

Finally, the results for old age pension are presented in Table 6. The econometric specification for old age pension differs from the other binary models. Old age pension is almost completely an absorbed state, almost no individual in the sample, returns to work after retirement and further there is a very strong age impact. These characteristics made it necessary to drop the lag effect; however the initial condition is still included. The time effect is strong and increasing and not surprisingly the age dummies indicate a strong increase in the probability of old age pension after age 64. In common with the other binary models an increase in the probability is found for low education and born outside Sweden but males have also a higher incidence. Replacement ratio matters and since there is a ceiling in the public pension system an indicator is included that measures if the individual has an income above this threshold. The estimation shows that there is a higher probability of old age pension if income is above the cap. This result may be explained by the importance of other pillars in the pension system, such as occupational or private pension.

Next section presents the labor supply model estimated and simulated on the population at risk defined as status 7 and 9 (other and working).

6. The labor supply model

We model labor supply as a discrete choice, following previous work by van Soest (1995); the household model is described in Flood *et al.* (2004) and the model for the single headed household in Flood *et al.* (2007). The discrete choice model allow us to include as many details as needed regarding the budget set and it extends naturally into a household model, where husbands and wives jointly determine their labor supply. Specifically, we assume that each individual can choose among seven alternatives in the choice set of income-leisure combinations and hence the choice set for a household contains $7*7$ different hours of work combinations.

We assume that family utility depends not only on consumption and leisure, but also on participation in social assistance. We further assume that the utility function is increasing in income and leisure and decreasing in welfare participation. The disutility from participation in a social assistance is assumed to reflect the non-monetary costs, such as fixed costs or “stigma”,

and is included to account for nonparticipation among eligible families.¹¹

Following van Soest (1995), we use a trans-log specification of the direct utility function, and for any specific household we have:

$$(3) \quad \begin{aligned} U(C, T - h_j) = & \alpha_1 \log C + \alpha_{11} (\log C)^2 + \alpha_2 (\log(T - H_H)) + \alpha_{22} (\log(T - H_H))^2 \\ & + \alpha_3 (\log(T - H_W)) + \alpha_{33} (\log(T - H_W))^2 + \alpha_{12} \log C * (\log(T - H_H)) \\ & + \alpha_{13} \log C * (\log(T - H_W)) + \alpha_{23} (\log(T - H_H)) * (\log(T - H_W)) - \\ & \alpha_4 P_{SA} - b_{fch} D_H - b_{fcw} D_W \end{aligned}$$

and for single headed household as:

$$(4) \quad \begin{aligned} U(C, T - h_j) = & \alpha_1 \log C + \alpha_{11} (\log C)^2 + \alpha_2 (\log(T - H_H)) + \alpha_{22} (\log(T - H_H))^2 \\ & + \alpha_{12} \log C * (\log(T - H_H)) - \alpha_4 P_{SA} - b_{fc} D_w \end{aligned}$$

where C is household disposable income described above and (T-H_j) is leisure (j= M, (male) or F (female) and T is an upper limit (4000) P_{SA} is a binary variable, one if the household is a receiver of social assistance else zero. D_j is also a binary variable, one if working hours is above zero; this reflects the importance of “fixed cost” of working.

In order to implement the model, we also have to specify the nature of heterogeneity in household preferences and the stochastic disturbances. For the household model heterogeneity in preferences for leisure is introduced as,

$$(5) \quad \alpha_2 = \sum_{k=1}^{K_x} \alpha_{2,k} z_{2,k} + \phi_H$$

$$(6) \quad \alpha_3 = \sum_{k=1}^{K_x} \alpha_{3,k} z_{3,k} + \phi_W$$

and in the specification of welfare participation as,

$$(7) \quad \alpha_4 = \sum_{j=1}^J \alpha_{4,k} z_{4k} + \phi_{SA}$$

The corresponding specifications for the single person is given by,

$$(8) \quad \alpha_2 = \sum_{k=1}^{K_x} \alpha_{2,k} z_{2,k} + \phi_H$$

¹¹ What may appear as “stigma” or disutility from welfare participation may also result from the inability of the econometrician to measure true welfare eligibility. Moreover, imperfect information regarding benefit eligibility on behalf of the household is also included in this non-monetary cost.

$$(9) \quad \alpha_4 = \sum_{j=1}^J \alpha_{4,k} z_{4k} + \phi_{SA}$$

The z-vector includes measurable individual and household characteristics and the ϕ 's represents unobserved variables that affect preferences for leisure. As usual it is assumed that an important source for population heterogeneity in terms of preferences for leisure is unobserved. In order to account for this, we formulate a finite mixture model, which allows for unobserved heterogeneity in a very flexible way without imposing a parametric structure. To make the model estimable, additional random disturbances are added to the utilities of all choice opportunities (for details see Flood *et al.* (2004) and Flood *et al.* (2007)).

In the estimation seven different classes or intervals of working hours per year have been used; 0, 1-500, 501-1000, 1001-1500, 1501-2000, 2001-2500, and above 2500.

The estimated parameters are presented in Table 7-8 and goodness of fit as well as elasticities in Tables 9-11. Since the parameters in these highly non-linear models do not have a simple interpretation, Table 9 below presents wage elasticities. These are within the bounds typically presented in the literature, higher for females and a negative male-female cross elasticity.

Table 10 and 11 present the predicted hour distribution. In most cases the fit is good and especially the frequency of zero hours is picked up quite well. This is important since many reforms are designed to increase hours at the extensive margin and the effects of such reforms can be under- or overestimated if this ratio is not reasonable correct.

Even if the elasticities seem reasonable it is important to remember that this is only a measure that summarizes the sensitivity of all individuals in the sample. It is not easy to draw any conclusion about a likely response to a tax/benefit reform based on these estimates. This is the reason we turn to a policy simulation. However, before the results are presented a short description of the welfare function is presented.

7. The social welfare function¹²

So far we have not discussed a suitable measure that can be used to characterize the welfare benefits of a reform. Even if it is relevant to measure the change in level or distribution of disposable income this is not an ideal measure for evaluating whether a reform should be preferred or not. For instance, a reform such as the in-work tax credit and decreased state tax that

¹² This section is largely based on Aaberge and Flood (2008) and Aaberge and Colombino (2008) and for further details we refer to these two sources.

is evaluated below is found to both increase the level of income as well as reduce the inequality of the income distribution. Even if this can be judged as an improvement it does not necessarily imply increased welfare for every individual. The reason is that some households have increased their income due to increased working hours whereas others have obtained higher income without increased hours. Since we assume that working hours is a disutility, that is leisure is a normal good, the increase in income and decrease in leisure might very well result in a decreased welfare. For this reason a social welfare function is estimated and will be used as one part of the evaluation of a reform.

We introduce the idea of a social planner who wants to implement a tax/benefit design to optimize welfare. The well known problem of interpersonal comparability is solved by assuming the existence of a common individual welfare function which is assumed to increase in income and leisure. The formal definition of the individual welfare function (Ψ) determined by the social planner is given by

$$(10) \quad \Psi_i = f(L_i, C_i)$$

where L is leisure and C is disposable income. In this application a quadratic utility function has been used and the estimation is described below.

The simplest measure of aggregate welfare is simply to sum all individuals' measure of welfare. However, since this implies equal welfare weights to the individuals, independent of the welfare level, this specific welfare function ignores distributive considerations. In order to address distributive justice individuals with a low welfare should be assigned larger welfare weights than those that are better off. This is described by the following family of rank-dependent welfare functions,

$$(11) \quad W_k = \frac{1}{n} \sum_{i=1}^n p_k\left(\frac{1}{n}\right) \Psi_i, \quad k=1,2,\dots,$$

where $\Psi_1 \leq \Psi_2 \leq \dots \leq \Psi_n$ is the ordered individual welfare levels Ψ , and $p_k(t)$ is a weight function defined by

$$(12) \quad p_k(t) = \begin{cases} -\log t, & k=1 \\ \frac{k}{k-1}(1-t^{k-1}), & k=2,3,\dots \end{cases}$$

The implication of (11) is that the weights given to low-welfare, individuals decrease with increasing k. As $k \rightarrow \infty$, W_k approaches inequality neutrality and coincides with the linear additive welfare function defined by

$$(13) \quad W_{\infty} = \frac{1}{n} \sum_{i=1}^n V_i .$$

To provide a simple guide to understand the inequality aversion profiles exhibited by W_1 , W_2 , W_3 and W_{∞} Table 12 provides ratios of the corresponding weights – as defined by (11) – of the median individual and the one percent poorest, the five percent poorest, the thirty percent poorest and the five percent richest individual for different social welfare criteria.

7.1 Estimation of the Social welfare function

The welfare function can be estimated either on household or on individuals. In either case the problem of comparing single- and non-single household must be solved. In Aaberge et. al. (2008) the welfare measure is based on individuals and in order to make individuals in couple and single household comparable the disposable income of the couple is divided by the square root of two. In this study we choose the household as the unit for estimation and welfare evaluation. We do this by transforming a couple-household into one representative household by using the average of disposable income (C) and working hours (H). Further, in order to compensate for number of children we divide household disposable income by the square root of number of children. A discrete choice translog model,

$$(14) \quad \Psi = \beta_l \ln(T - H) + \beta_{ll} (\ln(T - H))^2 + \beta_c \ln(C) + \beta_{cc} (\ln(C))^2$$

is estimated on a sample of households younger than 70 (mean age for couples). Thus, this is a simplified version of model (4) described above and as in (4) T=4 and hours H are divided by 1 000 and income C by 100 000. The estimated parameters are presented in Table 13. These parameters produce a welfare function increasing in leisure and income. The indifference curves presented in Figure 2 shows that the welfare function is more sensitive for a change in income rather than leisure.

An application of these welfare measures is provided in the next section when an evaluation of a Swedish tax reform is discussed.

8. Evaluation of the proposed 2009 tax reform

The Swedish income tax system consists of a flat municipal tax and a progressive national tax regime for earnings as well as for non-labor income. The individual is the taxation-unit and income taxes are independent of marital status. The flat municipal tax rate varies across

municipalities; the average municipal tax-rate in 2007 was 31.55 percent, the lowest 28.89 and the highest 34.24.

The evaluated tax reform consists of two parts; an in-work tax credit and a reduced national tax level. The in-work tax credit, includes the level introduced in 2007, 2008 as well as the proposed change 2009. Thus, what is referred to as the reform “after” corresponds to the level of in-work tax credit and state tax level suggested in the budget proposition for year 2009, the “before” reform is obtained by removing the in-work tax credit completely and also increasing the state tax to its 2008 year level.

Figure 3 displays the marginal and average tax rates for the income year 2009, before and after the reform. The marginal taxes before the reform have an irregular shape up to about SEK 350,000, the break point for governmental tax, and the reform imply a reduced tax rate by increasing this level. The irregular shape, before the reform, to the left of the break point is explained by the phase-in and phase-out of a basic tax deduction. This basic tax deduction remains unchanged after the reform but the tax credit is designed such that it smooths the irregularities created by the basic deduction. The result is an increasing step-wise marginal tax rate with the following tax rates and kink points (for an average municipal tax rate). Zero tax rate up to a taxable income of SEK 40,000, next about 24% up to SEK 116,000, about 29% up to SEK 300,000, almost 32% up to SEK 364,000, 52% up to SEK 507,100 and finally 57%. It should also be mentioned that the in-work tax credit used in Figure 3 applies for age younger than 65 and for an average municipal tax rate. For individual above 65 years of age and still working the tax credit is much more generous. This is considered in the simulation but, again, not in Figure 3.

The distribution of taxable income (in 2009 prices) shows that most individuals after the reform face a marginal tax rate lower or close to the municipal tax rate. About 17 percent has a higher income than the first (after reform) breakpoint for governmental tax and very few pay the highest rate. Most individuals face lower marginal tax rates due to the reform, the only exceptions are those above the breakpoint (after the reform). Hence, the average tax rate declines for everyone with a positive labor income, but much more for low than for high incomes. However, since the tax credit only applies for income from work, the tax reform increases the incentives for transitions to job participation for those who did not work before the reform. The incentive effects for those who were working before the reform are mixed. Below the after reform breakpoint the marginal tax rates have been reduced but for high income earners located above the breakpoint, marginal tax rates are unchanged but the average taxes have declined. Thus, a negative income effect might result in reduced working hours for high income earners.

Finally, note that the reform results in a substantial tax cut, especially for low income earners. For an income of SEK 150,000 the average tax rate is less than 20 percent and at 300,000 less than 25 percent.

To avoid choosing low hours of work most in-work tax credits programs has a phase in region and to avoid an income support to wealthy households there is also typically a phase out region. However, the Swedish design, implemented in 2007 and further extended in 2008 and 2009, differs from the mainstream design; it has no phase out region and the credit applies for all individuals with income from work. Moreover, unlike the US and UK reform the credit is not refundable.

Apart from the in-work tax credit there is also a basic tax deduction, see Figure 4. The basic deduction, which all individuals can claim, reduces taxable income (the income base for municipal and governmental tax). The basic deduction have three levels depending on the level of taxable income; SEK 18,000, 33,000 and 12,500. The in-work tax credit, which is deducted from the tax rate, increases with taxable income and reach a maximum value at SEK 18,000.

The tax credit reform applies for all individuals with an income from work. In the governmental budget proposal for 2008 the tax relief of this reform is estimated to about SEK 50 billion and the increased tax credit together with the reduced state tax rate suggested in the budget proposal for 2009 is estimated to an additional SEK 15 billion, assuming no behavioral effects from the reform. Also note that the in-work tax credit for individuals older than 65 is much more generous, this profile is not displayed in Figure 4 but included in the simulations.

9. Results

Table 14 presents the aggregated results for the central governmental budget. The strong reduction in taxes results in an increase in disposable income (almost 6 percent). Earnings have increased much less which of course implies a reduction in tax revenues (SEK 70 billion or close to 12 percent). This is almost completely due to the in-work tax credit. Even if revenues from taxes have reduced the increase in household's income also implies a reduction in governmental expenditures for transfer payments. The strongest effect is the reduction in sickness benefit whereas there is a much smaller decrease in unemployment- as well as disability benefits. As expected, due to the increase in disposable income, there is a sharp reduction in means tested benefits like housing allowance and social assistance. Apart from these effects we also consider change in VAT and payroll taxes. Assuming an average VAT of 17.6 percent on the increase in

disposable income adds more than SEK 14 billion to the governmental revenue. Finally a payroll tax of 32.42 percent on the increase of earnings adds another SEK 5.4 billion. If all entries are summed together the net result is a deficit of SEK 47 billion.

Of course the change in working hours are the driving force behind the dynamic effects reported in Table 14. These changes are summarized in Table 15 and it follows that working hours has increased 1.5% and number of working individuals by 1.1%.

Next, the change in the distribution of working hours is presented. As explained above the labor supply models are based on a classification of working hours in to seven discrete classes and Table 16 displays the before-after mobility between these classes. All entries on the main diagonal show the share of individuals within each class with same hours before and after. As expected the typical outcome is no change, but as indicated by all positive entries at the right hand side of the main diagonal, for those that change it is predominantly from lower to higher hours. The expected results from the in-work tax credit reform is an increase in work-participation, however, Table 16 might not give the impression that the most important change is from zero to positive hours, but even if the percentage change is not that big the number of individual is. For those that change their hours it is typically to longer hours, 1 500 or more. The decrease in state tax makes it more economically rewarding to work longer hours. Apart from working hours an important set of information relates to the mobility between the work/non-work states before and after the reform. The mobility tables are presented in Table 17. As follows by the entries on the main diagonal for several of the groups the change is small (not counting children, students and parental leave who should be unchanged), for instance only 0.6 percent of those individuals classified as old age pensioner before the reform have left this state and most of those have moved to the working state after the reform and a similar result is also found for disability. These results follows from the estimated binary models presented above, both old age and disability pension is too much of an absorbing state. Even a generous in-work tax credit for older people that is part of the reform has a strong influence.

Unemployed show a stronger response and 1.5 percent have moved to the work state. Long term sickness a much stronger change in fact 6.7 percent have left sickness and almost all have started to work. As expected the strongest effect is reported for the group classified as other and 9 percent have changed to work.

The change from non work to work is explained by the change in replacement rate primarily caused by the in-work tax credit. Table 18 show average replacement rates before and after the reform as well as the percentage change. For instance the average rate for unemployed before the reform was 69.7 percent, i.e. as an unemployed you have 69.7 percent of disposable

income compared to if you worked full time. Due to the in-work tax credit this rate decrease to 64.8 percent, thus a decrease of seven percent.

Next we turn to the change in distribution of disposable income caused by the reform. Figure 5 is constructed by classifying individuals into deciles based on the income before the reform. Then for each decile the change in income is calculated. The distribution shows that individuals with a medium or higher income have gained most. However, this graph includes all individuals, children as well as retired and as an alternative Table 19 presents the change in disposable income for different family types as well as the overall Gini-coefficient. Single mother increase their disposable income from SEK 141,000 to 149,000, a change of SEK 8,000 or 6%. Single men and couples gain more in levels but about the same percentage change. The Gini-measure indicates a somewhat smaller income inequality.

The information discussed above can be misleading since it was based on all individuals, regardless of whether they are allowed to change working hours or not. In order to focus on the part of population that are most likely to be affected by the reform the information above is repeated but now excluding individuals who are children (older as well as younger), students, on parental leave, and older than 70. As expected as follows from Figure 6 low income earners are now better off. The increase in the lowest decile is of course due to the increase in labor force participation.

The results in Table 20 are similar in percentage changes but since children and older are not included, the level is higher. The Gini-measure indicates a stronger decrease in inequality. The main explanation is the decrease in non-workers, by decreasing the ratio of non-workers had a positive effect on the income distribution.

Finally we turn to the effects on social welfare, Table 21. Using the social welfare function discussed above and applying the four different weights for inequality aversion, we find that social welfare have increased due to the reform. However, the profile show that the increase is larger if a larger weight is given for those with a low welfare ($W_1 > W_2 > W_3 > W_4$). Thus, even if individuals with a low level of welfare before the reform often had a low income due to not working, to start working implies higher welfare. This demonstrates that the disutility from working is dominated by the increased welfare from consumption. The evaluation of the social welfare effect adds an important dimension the tax reform evaluation. The challenge is to design a reform not only to maximize after tax income, but rather to produce the highest social welfare.

10. Conclusion and future directions

To include behavioral changes in a static micro simulation model has many advantages but also several challenges. From our perspective the most important reason for including behavioral changes is the demand for a framework to evaluate the impact of in-work tax credits. Owens (2005) discusses the importance of the “make work pay” policies and presents the design in several OECD countries. These policies always have a work requirement and the target group is disadvantaged groups with a low or zero employment rates.

SWETaxben is a model that allows an evaluation of “make work pay” reforms on working hours, income and welfare of individuals “inside” or “outside” the labor market. The Swedish version of “make work pay” offers an informative illustration to the challenges that has to be addressed in order to provide a realistic evaluation. The in-work tax credit highlights the importance to analyze changes at the extensive margin. This implies a need to distinguish between different groups of “outsiders”. Individuals that are early retired or have disability pension have different compensation schemes than those who are unemployed or long term sick. The compensation schemes across the different groups have to be included in the benefit calculations and since the calculation of replacement rates also requires detailed information of other benefits such as social assistance and housing allowance also have to be included.

The Swedish reform has a much higher in-work tax credit for older and this raises the extremely important question of the retirement decision or alternatively the decision to move from retirement back to work decision. This requires inclusion of the pension system in the tax/benefit calculations and also a statistical model for the retirement decision.

Since welfare, or social assistance, is closely related to the working decision the rules for social assistance and also other welfare programs have to be included but also the take up problem have to be considered. Interestingly the take up problem cannot be avoided even in a purely static model. If we do not correct for the fact that a (large) portion of households that are entitled to social assistance do not apply for it, we will overestimate the cost of social assistance. In SWETaxben the take up problem is part of the labor supply model.

Another characteristic of the Swedish reform is that lower replacement rates are not only obtained by in-work tax credit but also by lower unemployment benefits. Again, this is an argument for distinguishing the different groups of “outsiders”.

To assess the impact of an in-work tax reform it is necessary to distinguish the heterogeneity in the “outsiders” as well as the “insiders”. The high degree of non-linearity

created by a careful treatment of these differences is a strong argument for the micro-simulation approach. To add the reasonable assumption that the “make work pay” reform affect behavior is a further argument for a simulation model that allows individuals to adapt to economic incentives.

Even if the present version of *SWETaxben* allows for an evaluation of a “make work pay” reform (and obviously to almost any alternative tax/benefit design) there is still room for many improvements in future versions. To include those individuals that are not affected by a policy change, such as students and older children, is one reasonable simple extension. To allow for part time occupations, for instance part time retirement and part time work, is a much more complicated issue. However, in terms of policy relevance this change has a high priority.

Another more ambitious change is to reduce the restriction implicit in the sequential structure in *SWETaxben*. To allow for the simultaneous decision of working hours as well as occupational status would be much less restrictive, but also much more complicated. An interesting possibility is to extend the approach suggested in Aaberge and Flood (2008).¹³ This paper analyzes the effect on in-work tax credit on lone mothers including individuals on disability, unemployment, long term sickness but not old age pension. In principle this approach can be extended to old age pension, a household model and part time occupations. The method allows for a large number of combinations of working hours, part time occupations and welfare participations and also allows for restrictions in the choice set.

Apart from the suggested extensions there is also a need to have a better assessment of the long run effects of a policy change. To correctly measure the effects of an increase in labor force participation it is clearly not enough to just measure the immediate (one-year effects). The effects on the individual’s life cycle should be incorporated. Even if this is an argument for a dynamic microsimulation model we like to make the contrast between the need to evaluate a specific reform, with all the details, and the need to evaluate general changes. To develop a dynamic model that allows for a detailed evaluation of complicated tax/benefit rules is quite a challenge. The approach suggested here is instead based on a static model, but allowing for a dynamic interpretation. For instance applying the direct effect on earnings profiles can give a crude measure of the effects of a tax change on life time income as well as on life time taxes.

An advantage with *SWETaxben* is that it uses the same data as a dynamic simulation model (SESIM) developed at the Swedish Ministry of Finance.¹⁴ An exciting future development is to integrate *SWETaxben* and SESIM. The same individuals included in *SWETaxben* can be

¹³ For recent applications, see Aaberge *et al.* (1995) and Aaberge *et al.* 1999.

¹⁴ See Flood (2008)

simulated over time in SESIM. The effects of a policy change on life time income, taxes, benefits can be recorded. This would produce a much better understanding of the long run effects of a reform but still allowing for the detailed level required in order to keep it at a realistic level.

References

- Aaberge, R., J.K. Dagsvik, and S. Strøm. (1995). "Labor Supply Responses and Welfare Effects of Tax Reforms." *Scandinavian Journal of Economics* 97(4): 635-659.
- Aaberge, R., U. Colombino, and S. Strøm. (1999). "Labor Supply in Italy: An Empirical Analysis of Joint Household Decisions, with Taxes and Quantity Constraints." *Journal of Applied Econometrics* 14(4):403-422.
- Aaberge, R., and U. Colombino. (2008). "Designing Optimal Taxes with a Microeconomic Model of Household Labour Supply." *CHILD Working Paper* no. 06/08.
- Aaberge, R., and L. Flood. (2008). "Evaluation of an In-Work Tax Credit Reform in Sweden: Effects on Labor Supply and Welfare Participation of Single Mothers." *IZA Discussion Paper No. 3736*.
- Flood, L. (2008). "SESIM: A Swedish Micro-Simulation Model." Simulating an Ageing Population A microsimulation approach applied to Sweden. Ed. by Anders Klevmarken and Bjorn Lindgren. *Emerald Group Publishing Limited, Howard House UK, pp. 55-83, ISBN 978-0-444-53253-4*
- Flood, L., J. Hansen, and R. Wahlberg. (2004). "Household Labor Supply and Welfare Participation in Sweden." *Journal of Human Resources*, 39(4): 1008 -1032.
- Flood, L., R. Wahlberg, and E. Pylkänen. (2007). "From Welfare to Work: Evaluating a Proposed Tax and Benefit Reform Targeted at Single Mothers in Sweden." *LABOUR*, 21(3): 443-471.
- Governmental budget proposition 2009. 2008/09:1
- Moffitt, R. (1983). "An economic model of welfare stigma." *American Economic Review*, 73(5):1023-1035.
- Owens, J. (2005). "Fundamental Tax Reform: The Experience of OECD Countries." *Background Paper Number 47, Tax Foundation*.

Van Soest, A. (1995). "Structural Models of Family Labor Supply." *Journal of Human Resources* 30(1), 63-88.

Wooldridge, J. (2002). *Econometric Analysis of Cross Section and Panel Data*, MIT Press.

Wooldridge, J. (2005). "Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity." *Journal of Applied Econometrics* 20, 39 - 54.

Table 1. Population age 16-64 in and outside the labor force. LFS and LINDA 2006 (in thousands).

	<i>Labor force</i>		<i>Not in labor force</i>	Population 16-64
	Working	Unemployed		
LINDA	4 241	229	1 384	5 854
LFS	4 340	245	1 239	5 825

Source: LFS <http://www.scb.se/statistik/AM/AM0401/2007M01/AKU2006.pdf>
LINDA own calculations

Table 2. Population age 16-64 according to occupational status. LFS and Linda 2006 (in thousands).

	<i>Children</i>	<i>Pensioner</i>	<i>Student</i>	<i>Disability</i>	<i>Parental leave</i>	<i>Unemployed</i>	<i>Other</i>	<i>Working including Long term sick</i>	Total
LINDA	0	74	535	333	76	229	365	4 241	5 854
LFS	0	80	525	395		245	238	4 340	5 825

Source: LFS Table 16 <http://www.scb.se/statistik/AM/AM0401/2007M01/AKU2006.pdf>
LINDA own calculations

Note: Other includes home worker, job seekers, military service and other, see Table 16.

Table 3. Random-effects logistic regression of disability pension.

<i>Variable</i>	<i>Coefficient</i>	<i>Standard error</i>
Initial condition	13.290***	0.201
Occupational status <i>t-1</i>	4.743***	0.063
Intercept	-10.407***	0.247
Year = 2002	0.605***	0.054
Year = 2003	1.014***	0.056
Year = 2004	1.334***	0.058
Year = 2005	1.658***	0.060
Year = 2006	1.620***	0.063
Age 20 – 24	-3.012***	0.484
Age 25 – 29	-2.881***	0.191
Age 30 – 34	-2.435***	0.118
Age 35 – 39	-1.998***	0.090
Age 40 – 44	-1.674***	0.080
Age 45 – 49	-1.409***	0.077
Age 50 – 54	-0.922***	0.072
Age 55 – 59	-0.329***	0.062
Primary school	2.664***	0.087
Secondary school	1.360***	0.075
Big city	-0.495***	0.062
Medium city	-0.106*	0.055
Single	0.925***	0.051
Born in Sweden	-2.041***	0.066
Gender	1.310***	0.051
Replacement rate	0.877***	0.289
$\hat{\sigma}_c^2$	3.879***	0.043
$\hat{\sigma}_\varepsilon^2$	0.821***	0.003
Log Likelihood value		-41,031.73
Number of observations		895,938

Note: *, **, *** denote significance at the 10, 5, and 1 percent levels, respectively.

Table 4. Random-effects logistic regression of unemployment.

<i>Variable</i>	<i>Coefficient</i>	<i>Standard error</i>
Initial condition	2.435***	0.053
Occupational status <i>t-1</i>	2.587***	0.033
Intercept	-5.810***	0.104
Year = 2002	0.170***	0.034
Year = 2003	0.328***	0.034
Year = 2004	0.572***	0.033
Year = 2005	0.420***	0.034
Year = 2006	0.382***	0.034
Age 20 – 24	0.222	0.143
Age 25 – 29	0.097	0.066
Age 30 – 34	-0.089*	0.051
Age 35 – 39	-0.350***	0.047
Age 40 – 44	-0.490***	0.045
Age 45 – 49	-0.620***	0.046
Age 50 – 54	-0.642***	0.047
Age 55 – 59	-0.516***	0.044
Primary school	0.760***	0.043
Secondary school	0.455***	0.035
Big city	-0.042	0.031
Medium city	0.028	0.028
Single	0.250***	0.026
Born in Sweden	-0.635***	0.035
Gender	-0.091***	0.025
Replacement rate	0.534***	0.109
$\hat{\sigma}_c^2$	1.638***	0.025
$\hat{\sigma}_\varepsilon^2$	0.449***	0.008
Log Likelihood value		-63,954.20
Number of observations		822,288

Note: *, **, *** denote significance at the 10, 5, and 1 percent levels, respectively.

Table 5. Random-effects logistic regression of long term sickness

<i>Variable</i>	<i>Coefficient</i>	<i>Standard error</i>
Initial condition	1.116***	0.049
Occupational status <i>t-1</i>	3.197***	0.034
Intercept	-5.782***	0.147
Year = 2002	0.100***	0.028
Year = 2003	-0.002	0.029
Year = 2004	-0.023	0.029
Year = 2005	0.029	0.029
Year = 2006	0.013	0.029
Age 20 – 24	-1.150***	0.205
Age 25 – 29	-0.778***	0.071
Age 30 – 34	-0.577***	0.046
Age 35 – 39	-0.473***	0.039
Age 40 – 44	-0.463***	0.037
Age 45 – 49	-0.430***	0.037
Age 50 – 54	-0.328***	0.036
Age 55 – 59	-0.190***	0.035
Primary school	0.485***	0.030
Secondary school	0.239***	0.024
Big city	-0.051**	0.024
Medium city	0.040*	0.022
Single	0.160***	0.021
Born in Sweden	-0.390***	0.028
Gender	0.448***	0.019
Replacement rate	1.770***	0.205
$\hat{\sigma}_c^2$	1.055***	0.025
$\hat{\sigma}_\varepsilon^2$	0.253***	0.009
Log Likelihood value		-77,431.30
Number of observations		750,846

Note: *, **, *** denote significance at the 10, 5, and 1 percent levels, respectively.

Table 6. Random-effects logistic regression of old age pension.

<i>Variable</i>	<i>Coefficient</i>	<i>Standard error</i>
Initial condition	9.264***	0.128
Income above ceiling	0.365***	0.053
Intercept	-4.483***	0.164
Year = 2002	1.667***	0.057
Year = 2003	2.702***	0.065
Year = 2004	3.375***	0.073
Year = 2005	3.886***	0.082
Year = 2006	4.652***	0.094
Age = 61	-1.330***	0.114
Age = 62	-1.572***	0.094
Age = 63	-1.727***	0.081
Age = 64	-1.554***	0.069
Age = 65	-0.844***	0.060
Age = 66	0.585***	0.052
Primary school	0.824***	0.092
Secondary school	0.505***	0.081
Big city	-0.204**	0.085
Medium city	0.134*	0.080
Born in Sweden	-0.376***	0.103
Gender	0.876***	0.071
Replacement rate	1.656***	0.109
$\hat{\sigma}_c^2$	2.736***	0.041
$\hat{\sigma}_\varepsilon^2$	0.695***	0.006
Log Likelihood value		-22,198.51
Number of observations		77,490

Note: *, **, *** denote significance at the 10, 5, and 1 percent levels, respectively.

Table 7. Labor supply estimates of the Household model

<i>Variable</i>	<i>Coefficient</i>	<i>Standard error</i>
<i>Husband's leisure:</i>		
Intercept	34.277***	1.002
Big City	0.369***	0.069
Primary school	-0.794***	0.114
Secondary school	-1.235***	0.077
Age	-0.296***	0.034
Age ² /100	0.332***	0.037
<i>Wife's leisure:</i>		
Intercept	36.664***	1.194
Big City	-0.277***	0.059
Primary school	0.851***	0.102
Secondary school	0.185***	0.052
Age	-0.348***	0.028
Age ² /100	0.362***	0.031
<i>Utility parameters:</i>		
α_1	8.893***	0.291
α_{11}	2.257***	0.058
α_{22}	-12.594***	0.142
α_{33}	-12.503***	0.154
α_{12}	-0.420***	0.057
α_{13}	-0.630***	0.061
α_{23}	0.857***	0.079
<i>Fixed costs:</i>		
b_{fch}	10.542***	0.125
b_{fcw}	6.617***	0.082
<i>Welfare:</i>		
Intercept	13.826***	0.677
Big City	-0.094	0.150
Swedish born, Husband	2.890***	0.302
Swedish born, Wife	3.363***	0.356
Primary school, Husband	-1.134***	0.213
Primary school, Wife	-1.702***	0.225
Age 18-29, Husband	6.005***	0.404
Age 30-34, Husband	2.536***	0.506
Age 18-29, Wife	-1.857***	0.419
Age 30-34, Wife	-0.647***	0.290

Note: *, **, *** denote significance at the 10, 5, and 1 percent levels, respectively.

Continued...

Table 7 Continued:

<i>Unobserved heterogeneity:</i>		
<i>Husband's leisure:</i>		
ϕ_{H_1}	7.198***	0.004
ϕ_{H_2}	-10.940***	0.603
ϕ_{H_3}	3.742***	0.795
<i>Wife's leisure:</i>		
ϕ_{W_1}	14.578***	0.004
ϕ_{W_2}	-7.448***	0.963
ϕ_{W_3}	-7.130***	0.954
<i>Welfare:</i>		
ϕ_{SA_1}	21.970***	0.004
ϕ_{SA_2}	-9.399***	0.779
ϕ_{SA_3}	-12.571***	0.736
<i>Type probabilities:</i>		
π_1		0.014*** (0.001)
π_2		0.966*** (0.001)
Log-likelihood value:		-58,565.97
Number of observations		27,679

Note: *, **, *** denote significance at the 10, 5, and 1 percent levels, respectively.

Table 8. Labor supply estimates of single females, single mothers, and single males.

<i>Variable</i>	<i>Single Females Coefficient</i>	<i>Single Mothers Coefficient</i>	<i>Single Males Coefficient</i>
<i>Leisure:</i>			
Intercept	31.501*** (1.472)	14.951*** (3.803)	19.094*** (0.623)
Big City	0.039 (0.081)	0.163 (0.482)	0.684*** (0.080)
Primary school	0.797*** (0.177)	2.195*** (0.216)	1.470*** (0.152)
Secondary school	-0.077 (0.126)	0.552*** (0.196)	-0.083 (0.135)
Age	-0.251*** (0.031)	-0.424 (0.922)	-0.185*** (0.024)
Age ² /100	0.295*** (0.037)	0.478 (1.139)	0.211*** (0.030)
<i>Utility parameters:</i>			
α_1	10.798*** (0.607)	2.452 (5.613)	1.209*** (0.289)
α_{11}	1.375*** (0.152)	1.400*** (0.162)	0.364*** (0.045)
α_{22}	-9.922*** (0.420)	-14.781*** (1.854)	-13.097*** (0.187)
α_{12}	-2.103*** (0.182)	0.423 (1.758)	0.024 (0.096)
<i>Fixed costs:</i>			
b_{fc}	4.831*** (0.557)	8.507*** (0.206)	9.759*** (0.187)

Note: Standard errors appear in parentheses. *, **, *** denote significance at the 10, 5, and 1 percent levels, respectively.

Continued...

Table 8 Continued:

<i>Welfare:</i>			
Intercept	8.005*** (0.270)	5.639*** (0.378)	8.964*** (0.275)
Big City	0.148 (0.179)	0.193 (0.507)	-0.044 (0.139)
Swedish born	1.022*** (0.207)	2.821*** (0.368)	0.076 (0.133)
Primary school	-2.541*** (0.253)	-4.455 (8.050)	-5.833*** (0.204)
Age 18-29	-1.329*** (0.228)	-1.900*** (0.520)	-0.300 (0.198)
Age 30-34	0.362 (0.383)	-1.170*** (0.420)	0.154 (0.262)
<i>Unobserved heterogeneity:</i>			
<i>Leisure:</i>			
ϕ_{L_1}	-6.347*** (0.009)	-29.662*** (0.012)	-2.751*** (0.009)
ϕ_{L_2}	-4.858*** (0.554)	13.088* (7.045)	-2.230*** (0.175)
ϕ_{L_3}	11.205*** (0.770)	16.574*** (6.535)	4.981*** (0.210)
<i>Welfare:</i>			
ϕ_{SA_1}	6.992*** (0.009)	4.869*** (0.012)	6.279*** (0.009)
ϕ_{SA_2}	-2.708*** (0.353)	0.518 (7.651)	1.377** (0.129)
ϕ_{SA_3}	-4.283*** (0.278)	-5.388*** (0.785)	-7.657*** (0.257)
<i>Type probabilities:</i>			
π_1	0.541*** (0.016)	0.006*** (0.001)	0.766*** (0.002)
π_2	0.409*** (0.023)	0.859 (1.300)	0.134*** (0.002)
Log-likelihood value:	-14,212.00	-9,284.92	-16,359.95
Number of observations	11,130	7,158	12,826

Note: Standard errors appear in parentheses. *, **, *** denote significance at the 10, 5, and 1 percent levels, respectively.

Table 9. Uncompensated wage elasticities, by family type

	<i>Household</i>		<i>Single mothers</i>		<i>Single</i>
	Male	Female	Female	Female	Male
Wage increase by 1%:					
Male	0.10	-0.07			0.05
Female	0	0.16	0.21	0.38	

Table 10. Observed and predicted working hours and social assistance for households.

Hours interval	<i>Males</i>		<i>Females</i>	
	Observed	Predicted	Observed	Predicted
	%	%	%	%
0	4.41	4.55	4.69	4.46
1-500	0.03	0.00	0.18	0.10
501-1000	0.45	0.10	1.76	1.51
1001-1500	1.03	2.70	12.64	14.15
1501-2000	31.52	29.65	53.81	51.23
2001-2500	59.69	60.44	26.49	28.35
2500-	2.86	2.55	0.44	0.20
	Observed		Predicted	
	%		%	
Social assistance	1.1		0.7	

Table 11. Observed and predicted working hours and social assistance for singles.

Hours interval	<i>Single mothers</i>		<i>Single females</i>		<i>Single males</i>	
	Observed	Predicted	Observed	Predicted	Observed	Predicted
	%	%	%	%	%	%
0	9.85	9.42	7.59	7.63	13.96	13.85
1-500	0.31	0.01	0.73	0.12	0.23	0.00
501-1000	1.16	0.84	2.17	1.34	0.82	0.25
1001-1500	8.52	9.84	5.44	9.68	1.49	3.43
1501-2000	48.41	46.86	42.51	39.78	27.65	26.59
2001-2500	30.96	32.30	40.64	40.19	53.06	53.00
2500-	0.80	0.74	0.92	1.28	2.80	2.87
Social assistance	7.2	6.6	2.6	2.6	5.2	4.7

Table 12. Distributional weight profiles of four different social welfare functions

	W_1 (<i>Bonferroni</i>)	W_2 (<i>Gini</i>)	W_3	W_∞ (<i>Utilitarian</i>)
p(.01)/p(.5)	6.64	1.98	1,33	1
p(.05)/p(.5)	4.32	1.90	1.33	1
p(.30)/p(.5)	1.74	1.40	1.21	1
p(.95)/p(.5)	0.07	0.10	0.13	1

Table 13. Estimated parameters of the Social Welfare Function

	Estimates	Standard errors
Log leisure, β_I	3.0985	0.0215
Log leisure squared, β_{II}	0.3620	0.0151
Log income, β_C	4.7262	0.0224
Log income squared, β_{CC}	-0.0616	0.0089

Table 14. Reform effect on central governmental budget (billions of SEK)

	<i>Before</i>	<i>After</i>	<i>Change</i>	<i>%</i>
Household income/tax				
Disposable income	1,398.4	1,480.1	81.7	5.8%
Earnings	1,390.6	1,407.3	16.7	1.2%
Old age pension	101.5	99.8	-1.7	-1.6%
Total tax	604.9	534.9	-70.0	-11.6%
Transfers				
Unemployment benefit	22.9	22.5	-0.4	-1.5%
Sickness benefit	22.1	20.7	-1.4	-6.4%
Disability pension	47.1	46.8	-0.3	-0.7%
Child allowance	23.2	23.2	0.0	0.0%
Housing allowance	2.9	2.7	-0.2	-7.4%
Social assistance	15.8	14.9	-0.9	-5.7%
Child care fees	10.4	10.5	0.1	0.8%
VAT			14.4	
Payroll taxes			5.4	
Total budget effect			-47.0	

Table 15. Change in working hours and participation.

	<i>Before reform</i>	<i>After reform</i>	<i>Change</i>	<i>%</i>
Sum of worked hours (billion)	8.0	8.1	0.1	1.5
Number of workers (thousand)	4,131.7	4,178.8	47.0	1.1

Table 16. Percentage change in working hours before/after reform, individuals age 18-65.

<i>Before\After</i>	<i>0</i>	<i>1-500</i>	<i>501-1000</i>	<i>1001-1500</i>	<i>1501-2000</i>	<i>2001-2500</i>	<i>2501-</i>
0	97.56	0.01	0.05	0.29	1.21	0.86	0.02
1-500	0.00	84.95	0.00	0.00	13.49	1.56	0.00
501-1000	0.06	0.00	86.73	0.98	5.85	6.24	0.14
1001-1500	0.00	0.00	0.02	94.26	2.68	2.97	0.06
1501-2000	0.00	0.00	0.01	0.04	99.15	0.78	0.02
2001-2500	0.00	0.00	0.00	0.08	0.66	99.24	0.03
2501-	0.00	0.00	0.01	0.03	1.04	1.58	97.35

Table 17. Change of occupational status before and after the reform (%)

<i>Before</i>	<i>After</i>									Total
	<i>Child Pension</i>	<i>Student Disability</i>	<i>Parental leave</i>	<i>Unemployed</i>	<i>Other</i>	<i>Sickness</i>	<i>Work</i>			
Child	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	18.4
Pension	0.0	99.4	0.0	0.0	0.0	0.0	0.0	0.0	0.5	17.0
Parental leave	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	5.9
Disability	0.0	0.0	0.0	99.2	0.0	0.1	0.0	0.0	0.6	3.1
Parental leave	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.8
Unemployed	0.0	0.0	0.0	0.0	0.0	98.3	0.2	0.0	1.5	1.5
Other	0.0	0.0	0.0	0.0	0.0	0.0	91.0	0.0	9.0	3.5
Sickness	0.0	0.0	0.0	0.0	0.0	0.0	0.3	93.3	6.5	1.1
Work	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	48.7
Total	18.4	16.9	5.9	3.1	0.8	1.5	3.2	1.0	49.2	100.0

Table 18. Level and change in replacement rates before and after the reform.

	<i>Occupational status</i>			
	<i>Disability</i>	<i>Unemployed</i>	<i>Long term sick</i>	<i>Pension</i>
Before	64.7%	69.7%	82.8%	72.2%
After	60.4%	64.8%	77.0%	66.5%
Change	-6.6%	-7.0%	-7.0%	-7.8%

Table 19. Disposable income per family type and the Gini-coefficient

<i>Family type</i>	<i>Before (SEK 1000)</i>	<i>After (SEK 1000)</i>	<i>Change (SEK 1000)</i>	<i>Change %</i>
Single mother	141	149	8	6
Single female	152	162	10	7
Single men	187	201	14	7
Couple	208	223	15	6
Gini-coefficient	0.258	0.256	-0.002	-0.7

Table 20. Disposable income per family type and the Gini-coefficient

<i>Family type</i>	<i>Before (SEK 1000)</i>	<i>After (SEK 1000)</i>	<i>Change (SEK 1000)</i>	<i>Change %</i>
Single mother	210	225	15	7%
Single female	187	204	17	9%
Single men	206	222	16	8%
Couple	213	231	18	8%
Gini-coefficient	0.275	0.269	-0.006	-2%

Table 21. Change in Social Welfare for different weights of inequality aversion

	High	k=2	k=3	Low
	k=1			k=∞
Change	6.53%	4.93%	4.50%	3.38%

Figure 1. Structure of SWEtaxben.

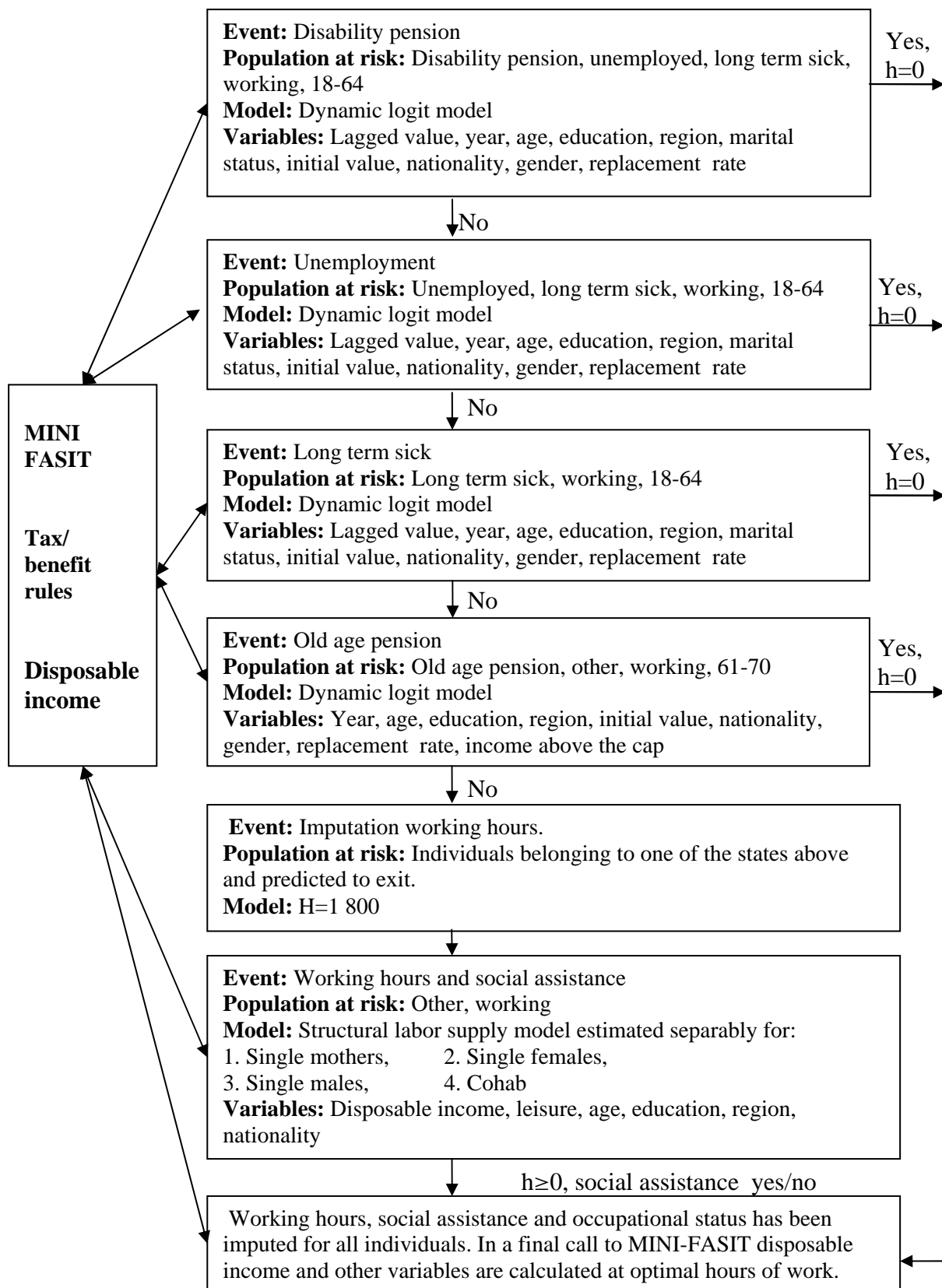


Figure 2. Indifference curves for the Social Welfare Function

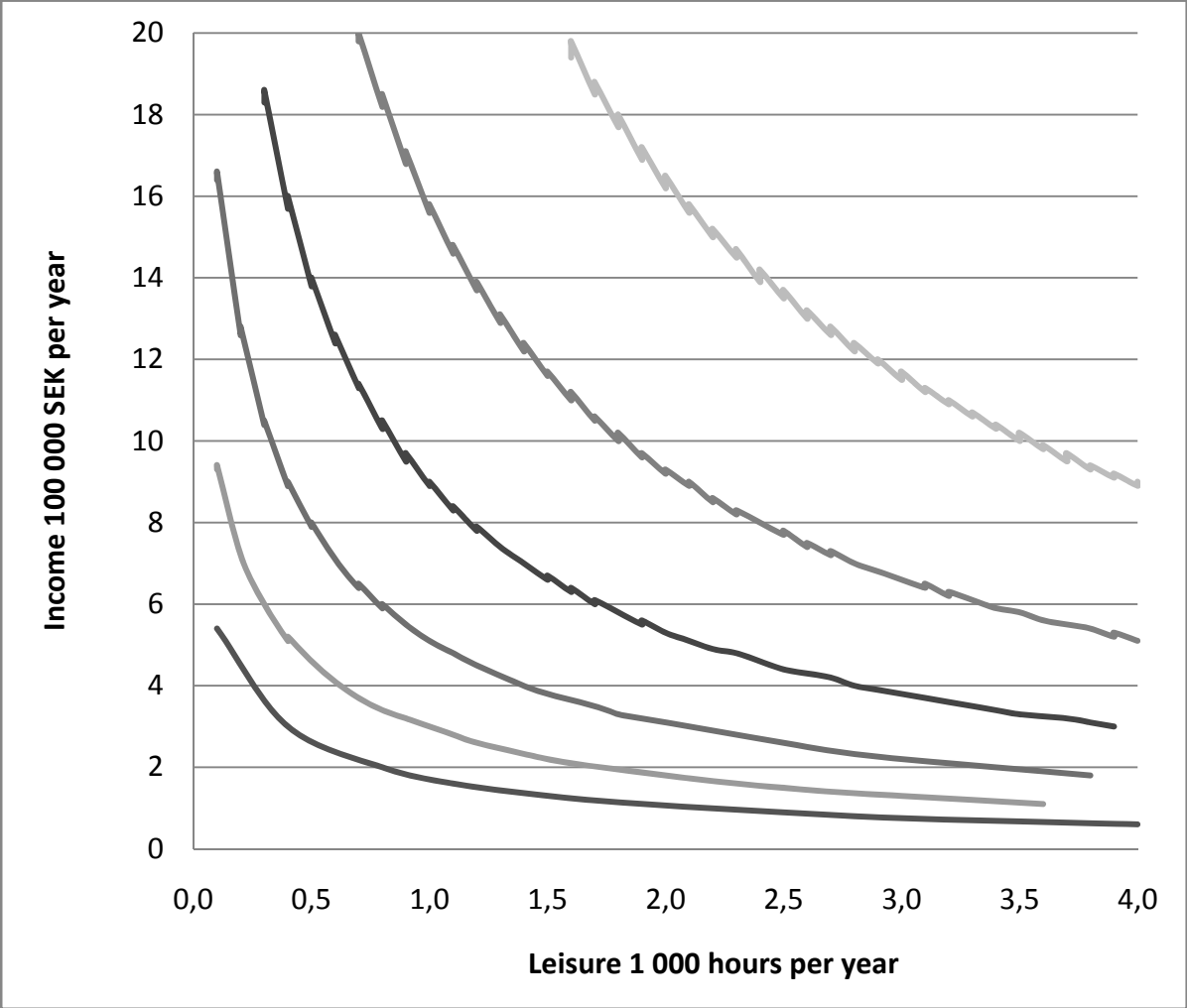
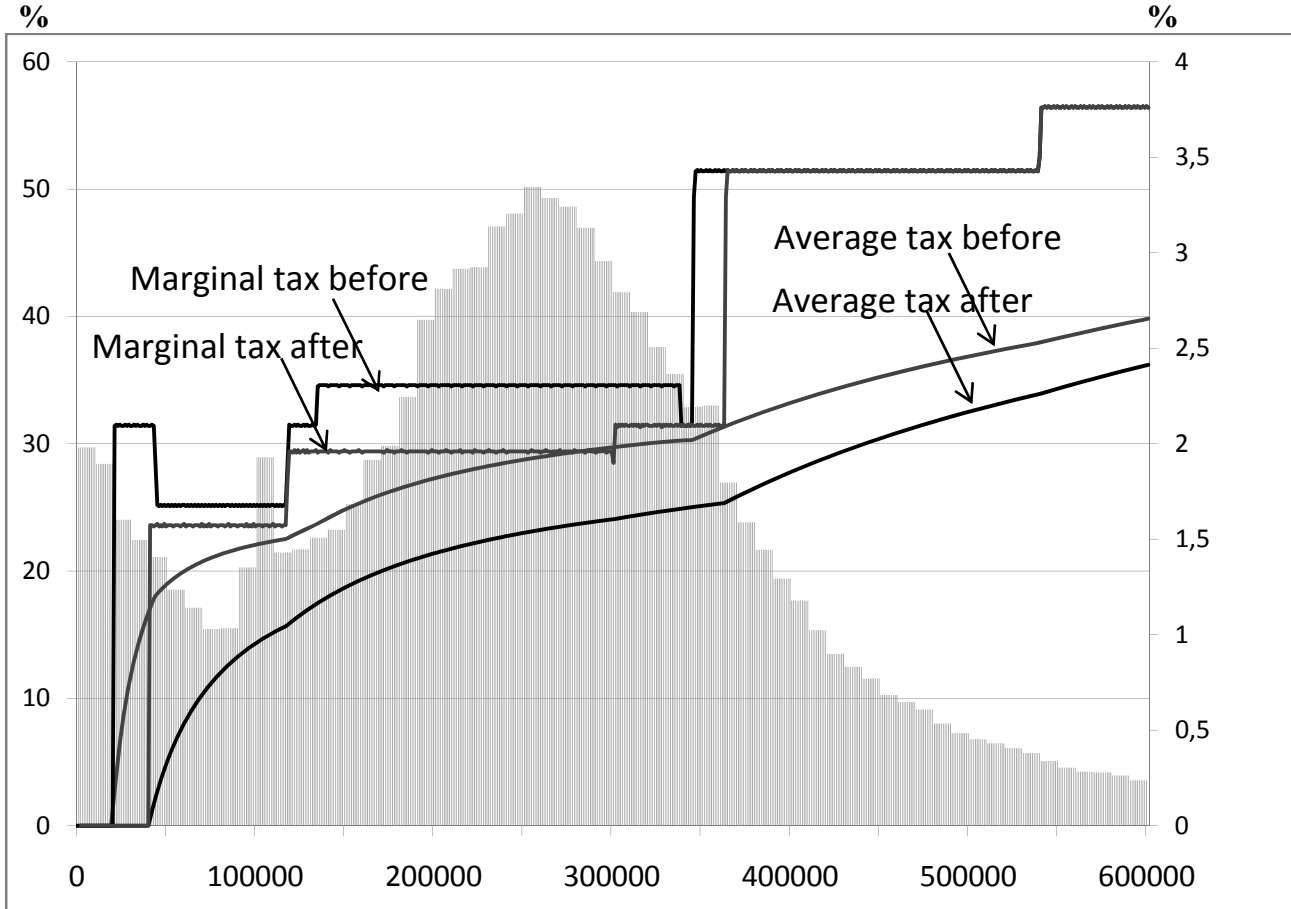


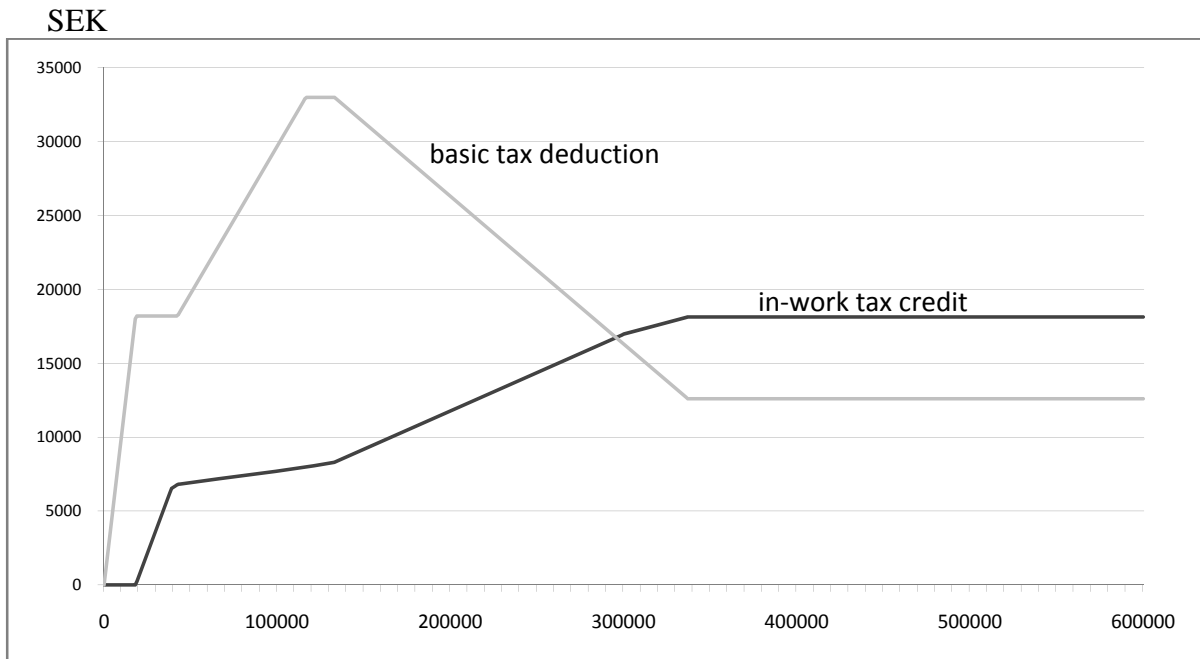
Figure 3. Marginal and average tax rates 2006 and 2009.



Source: Own calculations.

Note. Tax rates calculated at a municipal tax rate of 31.44% for an individual below 66 years of age. The income distribution is based on taxable income for all individuals between age 18-65. Use left hand scale for taxes and right hand scale for the income distribution.

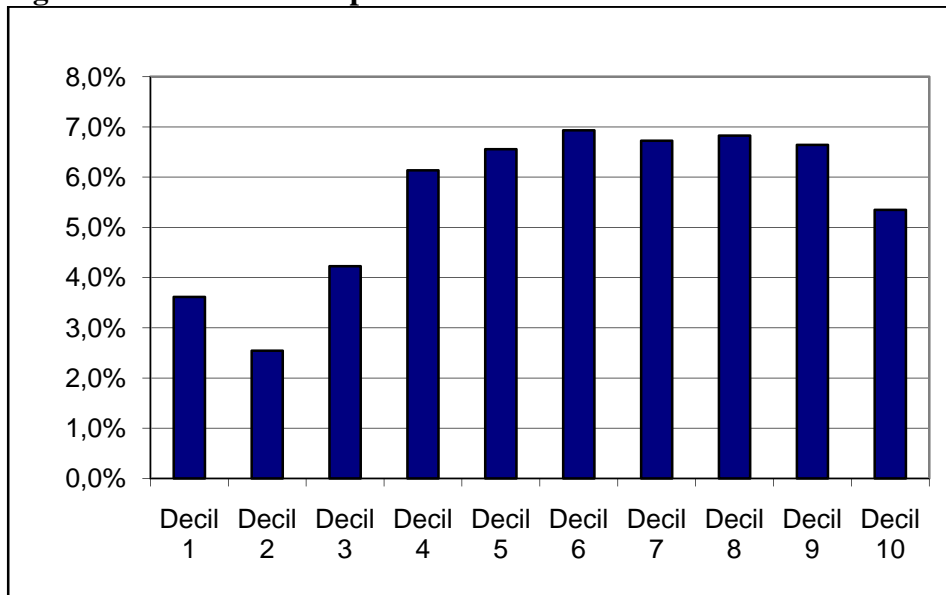
Figure 4. Basic tax deduction and in-work tax credit.



Source: Own calculations

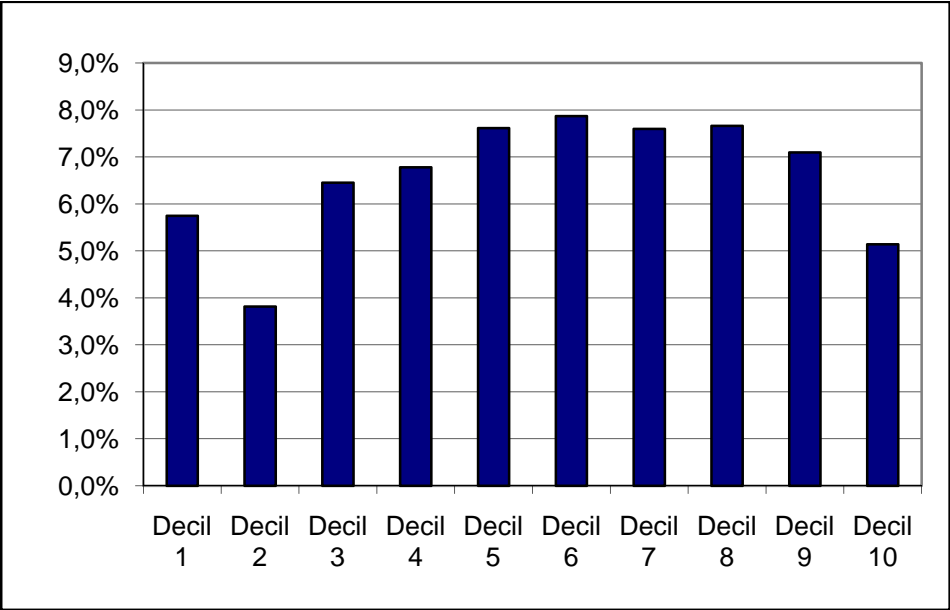
Note: At a municipal tax rate of 31.44% for an age below 65.

Figure 5. Distribution disposable income all individuals.



Note: Household disposable income adjusted for household size.

Figure 6. Distribution disposable income for the active population.



Note: Household disposable income adjusted for household size.