CPB Netherlands Bureau for
Economic Policy Analysis

MICSIM 2.0

A behavioural microsimulation model for the analysis of tax-benefit reforms in the Netherlands: an updated version

This background document describes three recent improvements of the Micsim-model: the inclusion of individuals with unemployment benefits (WW), the inclusion of the extensive margin of the self-employed, and an improvement of the estimated preferences.

These improvements in Micsim have not changed the labour supply responses substantially.

CPB Background Document

Henk-Wim de Boer, Roel van Elk, Eugene Verkade

April 2020

MICSIM 2.0

A behavioural microsimulation model for the analysis of tax-benefit reforms in the Netherlands: an updated version

Henk-Wim de Boer* Roel van Elk† Eugene Verkade‡

April 2020

*CPB Netherlands Bureau for Economic Policy Analysis. E-mail: H.W.de.Boer@cpb.nl.

†CPB Netherlands Bureau for Economic Policy Analysis. Corresponding author. CPB, Bezuidenhoutseweg 30, 2594 AV, The Hague, The Netherlands. Phone: +31 6 29037909. E-mail: R.A.van.Elk@cpb.nl.

‡CPB Netherlands Bureau for Economic Policy Analysis. E-mail: E.M.Verkade@cpb.nl.

1 Introduction

In November 2014, the CPB presented the MICSIM model: a behavioural microsimulation model for the analysis of tax-benefit reforms (Jongen et al., 2014). MICSIM has been further developed in recent years, and this background document describes three recent improvements of the model. Section 2 gives a brief description of the Random Utility Model (RUM) in MICSIM, which is the core of the model. Next, second 3 describes the first extension of the model. In 2015, MICSIM was extended by including individuals with unemployment benefits (WW) in the model to capture the response of these individuals to changes in benefits and taxes. Second, in 2016, we incorporated the extensive margin decision of the self-employed into MICSIM, as described in section 4. Finally, we improved the estimation of the preferences and section 5 presents the updated results.

2 The random utility model in MICSIM

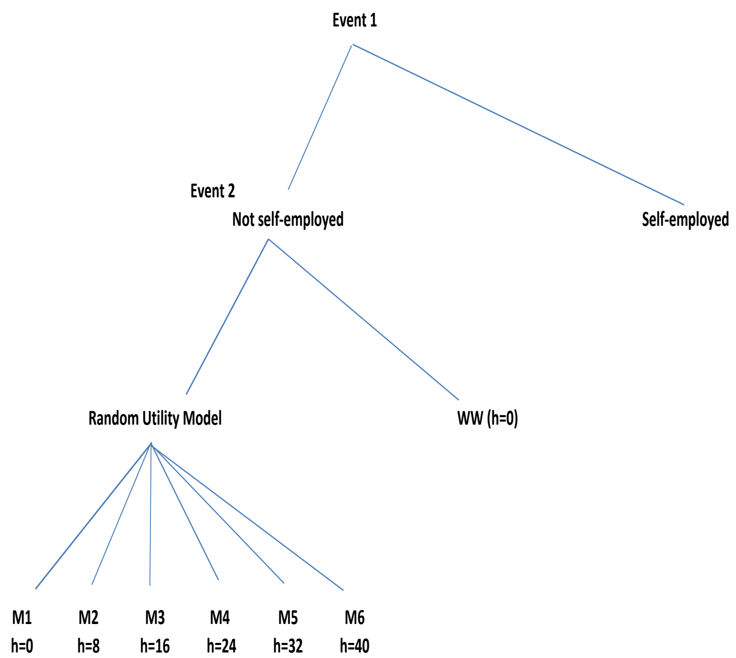
We first introduce the RUM for labour supply which is the core of MICSIM. We presented the MICSIM model in November 2014, and a more elaborate description of the model can be found in De Boer et al. (2015), Jongen et al. (2014) and De Boer et al. (2018). Figure 1 shows a nested decision tree of the complete MICSIM model. The first decision is whether to become self-employed or not (event 1).[1] An individual who becomes self-employed does not enter the next stage. Unfortunately, we do not observe the number of working hours by the self-employed in our data set, and therefore we assign the average number of working hours of employees in MICSIM to the self-employed in MICSIM. In the second event, an individual (not self-employed) faces the 'choice' between the RUM and WW.[2] Individuals who are not self-employed or on WW enter the discrete choice set of the RUM, and we describe the set-up of the RUM below. We discuss the decisions between RUM and WW, and between self-employed or not, in sections 3 and 4 respectively.

Households choose their preferred combination of hours of work from a finite set of alternatives. For singles without childcare, we have 6 alternatives (M1–M6) in the RUM. The first alternative (M1) represents the non-working alternative with

[1] In reality, individuals may not be able to freely choose between self-employment or not.

[2] Only employees are eligible for WW benefits and the job loss cannot be voluntary.

Figure 1: Full decision tree MICSIM



individuals working zero hours per week ($h = 0$). These individuals are on welfare benefits or without any income resources. Individuals with WW do not enter the RUM and find themselves at node WW in figure 1 with ($h = 0$). The alternatives M2–M6 represent the alternatives for work, running from 8 to 40 hours per week with an interval of 8 hours. Households are assumed to maximize a unitary utility function subject to a budget constraint and a time constraint. We use a flexible specification for preferences: a translog utility function, also used in e.g. Van Soest (1995). The choice of hours of work is the result of a coordinated decision of the two adult household members m and f. Define y as household income and h_m and h_f as the number of hours worked by the respective partners. We also explicitly model the use of formal childcare for households with young children, where c denotes the number of childcare hours per week. The most elaborate specification is then as follows:

$$\begin{aligned} U^d(\nu) &= \nu'\mathbf{A}\nu + \mathbf{b}'\nu + \mathbf{d}'\mathbf{1}[\mu > \mathbf{0}], \\ &\quad \nu = (\log(y), \log(1 - h_m/T), \log(1 - h_f/T), \log(c)), \\ &\quad \mu = (h_m, h_f, c), \end{aligned} \tag{1}$$

where we use the weekly time endowment T to transform the number of working hours into leisure.[3] The vector v consists of the logarithms of disposable household income (y), leisure of the man ($1 - h_m/T$), leisure of the woman ($1 - h_f/T$) and hours of formal childcare (c). The matrix $\mathbf{A}$ is the symmetric matrix of quadratic coefficients, and the vector $\mathbf{b}$ contains the coefficients corresponding to vector v. The vector $\mathbf{d}$ captures fixed costs of work for men and women. These are fixed costs related to working, which are expected to be negative terms for options where the respective person is working. As shown by e.g. Van Soest (1995), fixed costs are necessary to reproduce the low share of individuals that work only few hours per week. Of course, there are sound economic arguments to include them. Fixed costs of work represent disutility from work such as travelling costs, search costs or market frictions and demand side restrictions.[4] They also play a crucial role in

[3]We use total number of hours per week, e.g. 168, as the weekly time endowment. Different values for T hardly affected the results.

[4]Travelling costs are fixed costs of having a job, search and market frictions are costs incurred to find a job, and demand side restrictions imply that the availability of small part time jobs is limited on the labour market.

the distinction between the extensive (participation) and intensive (hours per week) response to changes in financial incentives. We do not include them in income or leisure, but simply include a dummy in utility metric, as in Van Soest (1995). Similarly, we also include fixed costs of using formal childcare.

We allow for preference variation through observed individual and household characteristics $\mathbf{x}_2$, $\mathbf{x}_3$ and $\mathbf{x}_4$ in parameters b_2, b_3 and b_4:

$$\begin{aligned}&\mathbf{b} = (b_1, b_2, b_3, b_4),\\ &b_1 = \beta_1, \quad b_2 = \mathbf{x'}_2\beta_2 + \psi_2, \quad b_3 = \mathbf{x'}_3\beta_3 + \psi_3, \quad b_4 = \mathbf{x'}_4\beta_4 + \psi_4,\end{aligned} \tag{2}$$

which are the linear utility terms in leisure of the male, leisure of the female, and hours of formal childcare, respectively. The same variation is also allowed for the fixed costs parameters $\mathbf{d}$. We further allow for unobserved preference heterogeneity in the preference parameters for leisure (ψ_2 and ψ_3, for the male and female, respectively) and formal childcare (ψ_4).[5] We do not allow for observed and unobserved preference heterogeneity in the coefficient b_1 of income, because it is hard to identify this preference heterogeneity separate from the preference heterogeneity in leisure and childcare.

For workers, we observe gross hourly wages which are used to compute the work-related part of income for each alternative in the choice set.[6] For non-workers, we simulate wages using estimates from a model that accounts for selection (Heckman, 1979)[7], and taking multiple draws from the estimated wage error distribution, see Jongen et al. (2014) and De Boer et al. (2018). Similarly, for households that use formal childcare we use the observed hourly prices of formal childcare, and for non-users we simulate hourly prices using estimates from a model that accounts for selection and taking multiple draws from the estimated gross hourly price error distribution.

For our empirical specification we use a discrete-choice model. Households choose their preferred combination of hours of work and hours of formal childcare from a finite set of alternatives $j \in \{1, ..., J\}$. Next to the deterministic part of household

[5]We use Halton sequences to draw the random terms (Train, 2003). For simplicity, we assume that there is no correlation between these unobserved preference heterogeneity terms.

[6]We use administrative data on hours worked and wages, hence measurement error is less of a concern.

[7]Here we follow e.g. Blundell et al. (2007) and Bargain et al. (2014).

utility $U^d(\nu)$ defined above, utility also contains an individual and option specific random utility term ε_j, necessary to reproduce heterogeneous choices for otherwise similar individuals as observed in the data:

$$U(\nu_j) = U^d(\nu_j) + \varepsilon_j. \tag{3}$$

ε_j is assumed to be identically and independently distributed across individuals and options, according to an Extreme Value Type-I distribution: This results in a convenient multinomial logit specification for the probabilities for observing individuals in particular options (McFadden, 1978).

Random preference heterogeneity, along with the draws from the estimated wage and price equations for non-workers and non-users of formal childcare, respectively, complicate the estimation of the likelihood. We use $R = 50$ (independent) draws from the wage distribution and for non-working men and women, the price distribution for non-users of formal childcare and the random terms for unobserved heterogeneity.[8] We use simulated maximum likelihood, where the likelihood is given by:

$$L = \prod_{i=1}^{N} \frac{1}{R} \sum_{r=1}^{R} \left(\frac{\exp(U_k^i(w_{i,m,r}, w_{i,f,r}, p_{cc,r}, \psi_{i,2,r}, \psi_{i,3,r}, \psi_{i,4,r}))}{\sum\limits_{j=1}^{J} \exp(U_j^i(w_{i,m,r}, w_{i,f,r}, p_{i,cc,r}, \psi_{i,2,r}, \psi_{i,3,r}, \psi_{i,4,r}))} \right)^{D_{ki}}, \tag{4}$$

with D_{ki} being an indicator function taking the value 1 for the observed choice for household i, and zero otherwise.[9]

For some household types the full log-quadratic specification was too flexible, resulting in a significant share (>5%) of households with negative marginal utility of income in the observed choices. This drives down the labour supply elasticities to implausible values. To solve this problem we dropped the interaction terms between income and leisure for these household types. For some households we also obtained an 'inverted' pattern for the marginal utility of income, with a negative (log) linear term and a positive (log) quadratic term. This results in implausible income effects,

[8]The number of draws is kept relatively low to limit the computational complexity of the model. However, increasing the number of draws did not change the predictions of the structural model.

[9]Note that for workers and users of formal childcare we take the actual gross hourly wage and actual hourly price, respectively, for each draw r.

and for these households we dropped the (log) quadratic term in income. Finally, for some household types the log-quadratic specification was not flexible enough. In particular, in some cases we do not capture the distribution of hours worked at the top very well, and we introduce a third order term for leisure, which then improves the fit at the top.

Jongen et al. (2014) used latent classes to account for unobserved preference heterogeneity (Train, 2008). They assumed that the analyzed population can be divided into a finite set of latent classes with households having homogeneous preferences within each class but heterogeneous preferences across the classes. This approach is very flexible (e.g. separate set of estimated parameters for each class) but has a strong computational burden and therefore the number of draws from the wage distribution had to be limited at 10. Furthermore, the latent classes approach led to implausible results for some subgroups which were caused by a relatively large share of households with negative marginal utility of income in the observed outcome. For these subgroups, Jongen et al. (2014) estimated a homogeneous model without unobserved heterogeneity. We improve the estimation by using random preferences heterogeneity for all groups. This method is less computationally demanding and allows us to increase the number of draws, from the wage distribution (for non-workers) and price distribution (for non-users of childcare), from 10 to 50. Increasing the number of draws reduces potential simulation bias and gives more precise estimation results (Cameron and Trivedi, 2005).

Households choose their preferred combination of hours of work and childcare from a finite set of alternatives $j \in \{1, ..., J\}$. We experimented with a number of discretizations, an interval of 8 hours (a normal working day in the Netherlands) running from 0 to 40 hours gave a good fit to the data and worked well in the estimations. For singles without young children, we then have 6 discrete options, and for couples without young children we have 6 x 6 = 36 discrete options. The discrete choice set becomes larger for households who potentially use formal childcare. For the use of formal childcare, we allow for 0, 1, 2 or 3 days[10], where the data show that a typical day in a daycare centre equals 10 hours, and a typical day in out-of-school care equals 5 hours. Specifically, we have 6 x 4 = 24 alternatives for lone parents with young children, and 6 x 6 x 4 = 144 alternatives for couples with a young

[10] The data show that using formal child care for more than 3 days per week is rare in the Netherlands.

children.

Disposable income in each discrete option is calculated as:

$$y = w_m h_m + w_f h_f - T(w_m, h_m, w_f, h_f; q) - TC(p_c, c; q) + S(p_c, c, y_t; q), \quad (5)$$

where w_m and w_f represent the gross hourly wage for the man and the woman. For households with young children, who potentially use childcare, we also take the costs of childcare $TC(.)$ and the childcare subsidy S into account. Here, the vector q denotes individual and household characteristics, $TC(.)$ is the total cost of formal childcare, with p_c denoting the price per hour of formal childcare, and $S(.)$ is the childcare subsidy, which depends on the hourly price of formal childcare, hours of formal childcare, taxable income y_t and the age distribution of the children.

To estimate the preferences of the different household types we use the Labour Market Panel (in Dutch: *Arbeidsmarktpanel*) of Statistics Netherlands (2012). The backbone of the Labour Market Panel are the annual observations of the Labour Force Survey (in Dutch: *Enquete Beroepsbevolking*) for the period 1999–2009, which contains the education level of adult members of the household. Statistics Netherlands supplements this data set with three additional data sources. First, administrative data from municipalities for the period 1999–2009 (in Dutch: *Gemeentelijke Basisadministratie*) that contains information on individual and household characteristics like age, ethnicity, ages of the children and area of residence. Second, administrative data from the Social Statistical Panel for the period 1999–2009 (in Dutch: *Sociaal Statistisch Bestand*) on hours worked and gross income. Third, administrative data on formal childcare from the Formal Childcare Database of the Tax Office for the period 2006–2009 (in Dutch: *Wet Kinderopvangtoeslag*). With respect to formal childcare, a distinction is made between daycare (children 0–3 years of age) and out-of-school care (children 4–11 years of age).

We estimate a structural model for the simultaneous choice of labour supply and, if applicable, the use of formal childcare.[11] Because data on childcare in our data set is available from 2006 onwards, we restrict the sample to the period 2006–2009. Furthermore, formal childcare subsidies are available to parents up to the point where the child goes to secondary school. Therefore, we only allow households with

[11] Unfortunately, informal childcare is not in our administrative dataset. However, De Boer et al. (2015) show that including informal childcare, calculated as the overlap in working hours of parents minus the hours of formal childcare, does not affect the results.

a youngest child of 0–11 years of age to choose formal childcare. Before the age of 4, children can go to daycare, whereas older children can go to out-of-school care. For households without children, or with a youngest child of 12 years of age or older, the childcare terms in the utility function drop out. We exclude households with missing information on individual or household characteristics.

Individuals who adjust their labour supply in our discrete choice model are employed, on welfare benefits or without any income resources. We do not model and effectively ignore the labour supply of the following types of individuals: students and retired or disabled individuals. Below we will refer to these individuals as having 'inflexible' labour supply. Also, self-employed are not included in the estimation of the preferences since we do not observe their number of working hours.[12] We also drop individuals with unemployment benefits (WW) when estimating the parameters of the utility function, implicitly assuming that they are constrained in their labour supply choice (Bargain et al., 2014). Furthermore, we also drop same sex households.[13] Finally, we drop individuals under 18 years of age, and individuals over 63 years of age.[14]

For the empirical analysis, we distinguish between 1-flex households and 2-flex households. Couples are 2-flex households when both partners are able to adjust their labour supply, and 1-flex households if only one partner has a flexible labour supply. For example, one partner who is employee and the other partner receives disability benefits. However, we account for the inflexible partner income when calculating the budget constraint of the flexible partner. In the estimations we distinguish 15 household types: childless singles (1), single parents with a youngest child aged 0–3, 4–11, 12–17 or 18 years of age or older (2–5), adult children living with their parent(s) (6), couples without children with both partner flexible (7), couples without children where only the man is able to adjust his labour supply (8),

[12]This also holds for owners of small corporations (*in Dutch: Directeur Grootaandeelhouders*), who are not included.

[13]Here we follow the standard approach in the literature to estimate separate elasticities for men and women within couples (see, for example, Bargain et al. (2014). About 2% of the population are same sex couples. The possible effect of this group on the total labor supply effects is therefore limited.

[14]For individuals older than 63 years of age, the early retirement decision is more likely to be important too. This calls for a dynamic modeling approach which is beyond the scope of this study.

couples without children where only the woman is able to adjust her labour supply (9), couples where both partners are flexible and with a youngest child aged 0–3, 4–11, 12–17 or 18 years of age or older (10–13), couples with children where only the man can adjust his labour supply (14), and couples with children where only the woman can adjust her labour supply (15).

3 Incorporating unemployment benefits

In 2015 we included individuals with unemployment benefits (WW) in the model to capture the response in individuals on WW to changes in taxes and benefits. We only add individuals with WW as their primary income source to the model. Ericson and Flood (2012) presents a shortcut for incorporating WW responses in a microsimulation model for the analysis of tax-benefit changes on labour participation.[15] Essentially they estimate a (quasi-fixed effects) logit model for the outcome of whether or not an individual is in the state of WW. The probability of being in the state of unemployment benefits depends on individual and household characteristics and the replacement rate. A higher replacement rate lowers the incentive to search and accept a new job, thereby increasing the probability of being in the state of WW. Taking draws from the error term distribution we can predict whether or not an individual is on WW (0 or 1). Individuals that are in WW do not enter the discrete choice model. Individuals that are not in WW do enter the discrete choice model.

To simulate the effect of a change in WW benefits or taxes on the probability of WW, we need to calculate the replacement rate. This is typically not observed in the data. We either observe wages or we observe WW benefits. Here we can follow Ericson and Flood (2012) again. For everybody we have an hourly wage rate, either observed or predicted. We can use this to predict gross labour income when working 40 hours per week for the whole year, and predict gross unemployment benefit income when unemployed for 40 hours per week (taking into account the maximum income over which unemployment benefit applies). The replacement rate is then defined as disposable household income ('NBI') when the individual is on

[15] A more elaborate description of the method can be found in the working paper version Ericson et al. (2009).

unemployment benefits over disposable household income when the individuals is employed, taking the income and employment status of the partner as given.

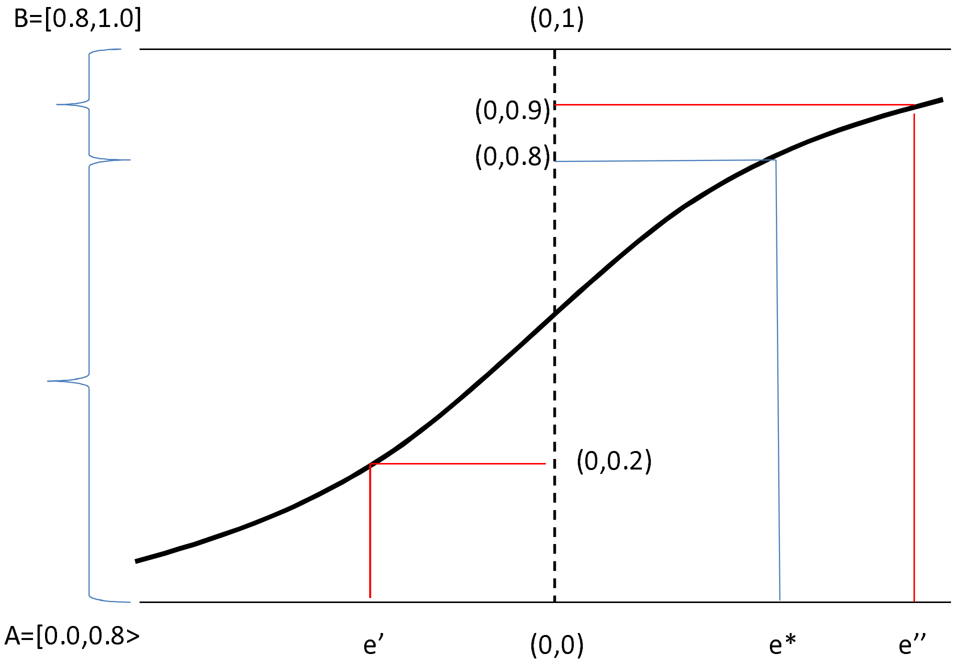
We estimated logit models for the probability of being in WW, using data for the period 2006–2009 (in line with the rest of MICSIM). The individual and household characteristics have the expected sign. However, the replacement rate always has the 'wrong' sign, where a higher replacement rate lowers the probability of being in WW. This is probably due to insufficient exogenous variation in replacement rates for WW in our data. However, we can still use the methodology of Ericson and Flood (2012) to incorporate WW responses in MICSIM, calibrating the response to changes in the replacement rate on the findings of empirical studies. Below we outline the method.

Again, we estimate a logit model for the probability of being in the state of WW, with as explanatory variables individual and household characteristics. The logit model bounds the probability of being in the state of WW to the $[0, 1]$ interval. The logit model produces a probability of being in the state of WW. To prevent messing up the behavioural responses in rest of the model, it seems useful to reproduce the observed outcome in the data, e.g. whether or not an individual is on WW. We can use the inversion method. This method is illustrated in Figure 2.

For each individual we have a prediction for the probability that he or she has unemployment benefits. The logit specification gives the following expression for the probability: $P(WW = 1) = \frac{e^{X\beta}}{1+e^{X\beta}}$. Suppose that given the characteristics of the individual this value is 0.20. This means that starting from minus infinity, for the first 80 percent of the random draws for the error term the individual is predicted not to be in WW, and for the remaining 20 percent the individual is predicted to be in WW. The critical value is indicated by e^*. For individuals with these characteristics that are observed (not) to be in WW we take a random draw from the CDF of the logistic distribution above (below) e^*. Hence, for an individual observed to be in WW we take a random uniform draw of the interval [0.8,1.0], for example 0.9, and find the inverse of the CDF for this value (e''). For an individual observed not to be in WW we take a random uniform draw of the interval [0.0,0.8>, for example 0.2, and find the inverse of the CDF for this value (e').

The inverse of the logistic distribution (e.g. finding x belonging to y) is $-\log(1/y-1)$. Denote the predicted probability for an individual by $P_i(X\beta)$ (using predict). Hence, for an individual with characteristics $X\beta$ observed to be in WW we generate

Figure 2: Inversion method



a random draw from the logistic distribution using

$$e'' = -\log(1/((1 - P_i(X\beta)) + P_i(X\beta) * runiform) - 1), \tag{6}$$

and for an individual with characteristics $X\beta$ observed not to be in WW we generate a random draw from the logistic distribution using

$$e' = -\log(1/((1 - P_i(X\beta)) * runiform) - 1). \tag{7}$$

We include the replacement rate of WW for all individuals in the logit model to calibrate the elasticity of the probability of being in WW with respect to the replacement rate. Empirical studies on unemployment benefits, see CPB (2015) for an overview, suggest a value of 0.5 for the elasticity of the probability of having unemployment benefits with respect to the benefit level.[16] Specifically, in the logit model we include a term $\beta_{rr}(rr_{simulation} - rr_{base})$, and by trial-and-error look for

[16]CPB (2020b) (forthcoming) will present an update of the empirical literature on unemployment benefits. Based on the new empirical insights the current elasticity of 0.5 may be adjusted in MICSIM.

the value of β_{rr} that generates the desired average elasticity over all groups. We take the difference of the replacement rate from a policy simulation and the baseline scenario. Hence, if a policy simulation does not affect the replacement rate, the difference equals 0 and the number of individuals with WW does not change. A tax-benefit reform that increases net income from work leads to a lower replacement rate and therefore to a lower probability of being in WW (compared to the baseline scenario).[17]

4 Extensive margin of the self-employed

In 2016, we have used the same short cut provided by Ericson and Flood (2012) to incorporate the extensive margin decision of the self-employed. Applying this method to self-employed means we want to estimate the probability of being self-employed. The probability depends on individual and household characteristics and on the level of self-employed income in relation to wage income. Hence, we need to calculate the difference in income from self-employment and wage earnings. However, this is typically not observed in the data. We either observe wages or we observe the level of profits. For each individual we need to simulate income in the alternative situation, where we use profit as an approximation for labour costs (and vice versa).[18] Taking draws from the error term distribution we can predict whether or not an individual is self-employed (0 or 1). Individuals that are self-employed do not enter the RUM. We include all self-employed with profit as their primary income source, both self-employed without employees and the self-employed who employ workers. However, we drop owners of small corporations (*in Dutch: Directeur-Grootaandeelhouder*) because their personal liability differs and their tax treatment is not comparable to the tax treatment of the majority of the self-employed. These owners have high administrative positions within private or

[17] We do not use MICSIM to determine the effects of changes in the level of the WW benefit or the maximum benefit duration. For such scenarios we can simply use the suggested elasticity of 0.5 by the empirical studies (CPB, 2015). With respect to the WW, the added value of MICSIM is that we can simulate the effects of changes in the tax-benefit system (e.g. changes in the effective tax rates) on the replacement rates.

[18] An alternative is to use an hourly wage rate, either observed or predicted. We can use this to predict gross labour income when working 40 hours per week for the whole year. However, we do not observe the number of working hours by the self-employed.

public limited liability companies (*In Dutch: BV/NV*) and hold a large proportion of the shares.

Following Ericson and Flood (2012) we estimate logit models for the probability of being self-employed. The individual and household characteristics have the expected sign. On average, the self-employed are older, more often men and living in couples. However, we do not have enough identifying variation to determine the effect of the relative earnings differential (e.g. income from self-employment and wage income) on the probability of self-employment (Bosch and De Boer, 2019). We would expect that a reform that increases income from self-employment relative to wage income, results in a higher probability of being self-employed. Although the binary regression produces the 'wrong' sign, we can still use the methodology of Ericson and Flood (2012) and calibrate the response to changes in (relative) income using elasticities of empirical studies. Again, we use an elasticity of 0.5 based on a literature overview of empirical studies on the occupational choice of the self-employed, see Bosch and De Boer (2019). To prevent messing up the behavioural responses in the rest of the model, it seems useful to reproduce the observed outcome in the data, e.g. whether or not an individual is self-employed. We again use the inversion method described above.

Now, to simulate the effect of changes in taxes and benefits on the probability of self-employment, we need to calculate the difference in income from self-employment and wage earnings. In the logit model we include a term $\beta_{rr}(rr_{simulation} - rr_{base})$, and by trial-and-error look for the value of β_{rr} that generates the desired average elasticity over all groups. Here, the term rr is disposable household income from self-employment divided by disposable household income from wage earnings. When taxes and benefits change, the status of an individual of being in or out of self-employment may change. Suppose a policy reform lowers the tax exemption of the self-employed. Consequently, disposable income of the self-employed falls whereas disposable income from wage earnings is unaffected. Stated differently, the reform lowers the term $rr_{simulation}$ and therefore the probability of self-employment compared to the baseline scenario. Individuals who are self-employed do not enter RUM. Furthermore, individuals who change from being employee to self-employed must leave the RUM.

5 Updated labour supply elasticities

The previous sections discussed the three most important improvements of the model: incorporation the extensive margin decision of the self-employed, including individuals with WW and improving the estimation of the parameters of the utility function. This sections presents the updated labour supply elasticities.

The estimated preferences can be found in De Boer et al. (2018). In discrete choice models, we do not have an analytical solution for the labour supply elasticity. This has to be simulated. We simulate these elasticities by increasing gross wages by 10%. Here, we present the total elasticity (the percentage change in total hours worked over the percentage change in the gross wage rate), and the decomposition of this total elasticity into the extensive margin elasticity (the percentage change in the participation rate over the percentage change in the gross wage rate) and the intensive margin (the percentage change in hours worked by the employed over the percentage change in the gross wage rate).

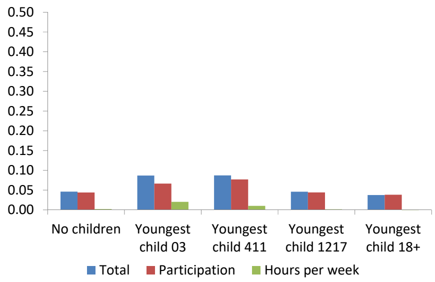
Figure 3 gives the simulated labour supply elasticities for couples in which both partners can choose whether or not to work and for how many days per week. We estimate this for several subgroups, where subgroups are defined by the age of the youngest child, including a category for flexible couples without children. We find small, positive labour suppy elasticities for men, see panel (a). The labour supply elasticities are much higher for women, on the extensive margin but also on the intensive margin, see panel (b). Furthermore, the labour supply elasticities for women in couples are particularly high when the youngest child is 0–3 years of age (pre primary school age) or 4–11 years of age (primary school age).

Figure 4 gives the so-called cross-elasticities, e.g. the percentage change in total hours worked by one partner over the percentage change in the gross wage rate of the other partner. Panel (a) shows that cross-elasticities are negative but close to zero for men. But for women, cross elasticities are non-negligible.

Next, Figure 5 panel (a) shows that the labour supply elasticity is relatively low for singles without children. The labour supply elasticity is much higher for single parents with young children. The labour supply elasticity of single parents whose youngest child is no longer in primary school is much lower, though still higher than for singles without children. Also note that the differences across single parents are primarily driven by differences in the extensive margin elasticity. The intensive

Figure 3: Households with two flexible persons

(a) Men



(b) Women

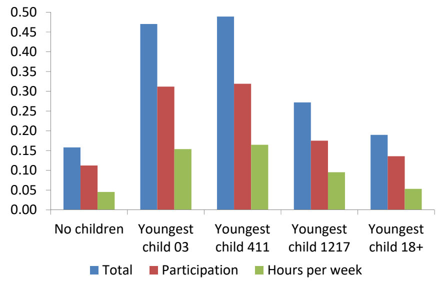
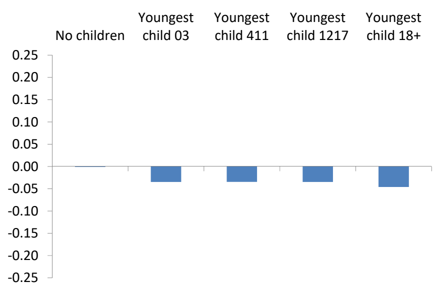


Figure 4: Cross elasticities in households with two flexible persons

(a) Men



(b) Women

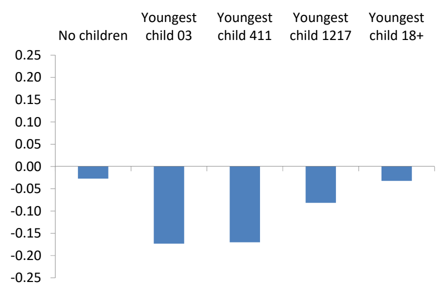
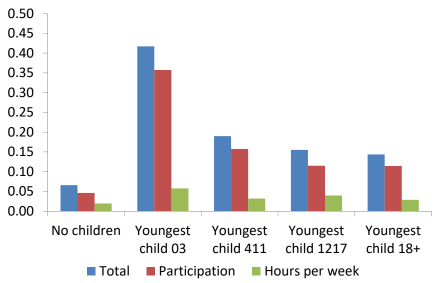
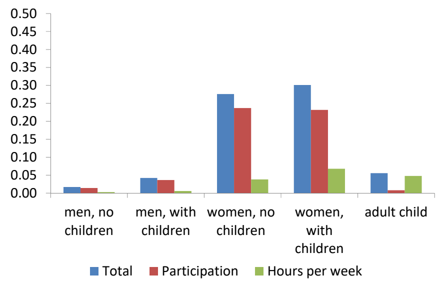


Figure 5: Households with one flexible person, and adult children

(a) Singles and single parents



(b) Individuals with an inflexible partner, and adult children living at home



margin response for single parents is quite small.

Panel (b) gives the labour supply elasticities for men and women in couples whose partner's labour supply is inflexible (because this person is e.g. disabled or retired). For these groups we pool couples with children of all ages. Most men with an inflexible partner work, and typically also fulltime. Hence, there is little upward potential in terms of total hours worked, and they have a relatively low labour supply elasticity. For women there is more upward potential in total hours worked, both in terms of the participation rate and in terms of hours worked per employed. Women with an inflexible partner have a higher labour supply elasticity, in particular on the extensive margin. Panel (b) also gives the labour supply elasticity for adult children living at the home of their parents. They have a very high participation rate (when they are not disabled etc.), resulting in a very low labour supply elasticity.

6 Discussion

This background document has described three recent improvements of the MICSIM model since its introduction in 2014. Next, we have provided a detailed overview of the heterogeneous labour supply responses to changes in financial incentives in the Netherlands. We find substantial differences between men and women in couples, in particular when children are present. Furthermore, the age of the youngest child seems to play an important role in labour supply responses, mothers with young children being particularly responsive. We have also shown that the decision whether or not to participate is more responsive to financial incentives than the hours per week decision. Comparing the updated results in this paper with the results found in Jongen et al. (2014), we can conclude that the recent improvements in MICSIM have not changed the results substantially. Indeed, Van Elk et al. (2020) compare the differences in labour supply elasticities and results of policy simulations, and conclude that both methods produce similar results. In addition, De Boer et al. (2018) presents a comparison of predictions by the structural model with the findings from three recent quasi-experimental studies, and shows that the results of the structural model are still very much in line with the results found by the quasi-experimental studies.

The past decade, MICSIM has been used extensively to simulate the labour supply effects of many policy scenarios. We do not present the results here but they can be found in De Boer et al. (2018), CPB (2020a) and Van Elk et al. (2020).

The focus of the model is a detailed modelling of labour supply responses. However, various mechanisms are not present in the model which are potentially relevant for tax-benefit reforms. We discuss these mechanism below.

We assume that individuals are free to choose whether or not to participate, and how many hours or days to work per week. However, individuals can be involuntary unemployed, or they may not be able to work the number of hours or days per week that they would prefer (given the budget constraint). Previous studies have shown that accounting for involuntary unemployment, or the difference between preferred and actual working hours, can make quite a difference in terms of employment responses (Euwals and Van Soest, 1999; Bargain et al., 2010). We have put considerable effort in investigating the issue of involuntary unemployment, estimating a double-hurdle model (Cragg, 1971). However, for all household types we find that accounting for involuntary unemployment makes little difference to the employment responses to changes in financial incentives (De Boer, 2018). Very few individuals in the data are classified as involuntary unemployed. Note however that we use data for the period 2006–2009. Since then, unemployment has gone up considerably in the Netherlands and accounting for involuntary unemployment is likely to have a larger effect at this point in time. However, recall that we are simulating the structural effects of tax-benefit reform. The structural level of (involuntary) unemployment is probably not that different from the period 2006–2009. Regarding the difference between actual and preferred hours of work, we do not have data on preferred hours of work in our dataset. However, this seems to be a much smaller problem in the Netherlands than in many other OECD countries. For example, OECD (2013) reports that just 5% of part-time working women would like to work more hours, compared to e.g. 13% in Germany, 28% in France and 55% in Spain.

In modelling the decision between self-employment or wage employment, we do not take into account the differences in the costs (and benefits) of pension build-up and insurance for unemployment or disability. These differences are likely to affect the decision to become self-employed or not as well.

In the model we focus on the labour supply responses of changes in the tax-benefit system. Part of the modern public finance literature looks at a broader range of behavioural responses, by considering the so-called elasticity of taxable income, see Saez et al. (2012) for an excellent overview. The elasticity of taxable income also captures e.g. changes in effort more generally, occupational choice, and tax avoidance.[19] For the majority of workers, changes in taxable income mainly reflect changes in labour supply. Indeed, in a recent study for the Netherlands, Jongen and Stoel (2019) find that for the average worker the elasticity of taxable (labour) income is not that different from the labour supply elasticity. However, for high incomes they find that the labour supply elasticity is lower, whereas the elasticity of taxable income is higher, consistent with the literature, see again Saez et al. (2012). Hence, for high incomes, the labour supply response only captures part of the response in the tax base. Therefore, to determine e.g. the budgetary consequences of an increase in the top tax rate, one needs to consider the other behavioural responses next to the labour supply response.

We ignore general equilibrium effects on prices and wages. However, this may not be a bad approximation for the long run with a perfectly elastic labour demand (Aaberge and Colombino, 2014), which seems particularly relevant for a small open economy like the Netherlands.[20] This is in line with the stylized fact for the development of labour supply and employment in the Netherlands over a longer period. Data of Statistics Netherlands[21] show there has been a strong increase in labour supply, of mostly women and older workers, in the Netherlands in the period 1975–2010. This increase was accompanied by a strong increase in employment without a substantial increase in structural unemployment.

Perhaps more problematic is that we ignore the lifecycle. A number of studies have shown that accounting for lifecycle effects can be important for the analysis of tax-benefit reform, see e.g. Imai and Keane (2004) and Keane (2011). However, we do not have the data (e.g. on consumption or savings) to model lifecycle responses to tax-benefit reforms. Furthermore, there is often a trade-off in modelling different parts of economic behaviour, due to the numerical complexities that arise.

Finally, we assume that all people are fully aware of their full budget constraint. However, the recent work by Chetty (2009) shows that information or the lack thereof can play an important role in the behavioural responses to financial incen-

[19]Furthermore, for top incomes, the contractual hours that we observe in our dataset may not be a good indicator of actual hours.

[20]Interesting exercises with finite demand elasticities can be found in Peichl and Siegloch (2012) and Colombino (2013).

[21]available online at *https://opendata.cbs.nl/statline/*.

tives. This is an important new research area. However, note that we are using policies and changes therein in the past to estimate preferences. So, to the extent that informational frictions play a role in behavioural responses to changes in financial incentives, our estimated preferences implicitly incorporate these informational frictions (if existed) in this period. Also, note that the model does rather well in predicting behavioural responses to past reforms (De Boer et al., 2018).

To conclude, we believe that we have made a big step in modelling the heterogeneous responses to tax-benefit reform. However, interesting topics for future research remain.

References

Aaberge, R. and Colombino, U. (2014). Labour supply models. In O'Donoghue, C., editor, *Handbook of Microsimulation Modelling*, pages 167–222. Emerald.

Bargain, O., Caliendo, M., Haan, P., and Orsini, K. (2010). 'Making work pay' in a rationed labour market. *Journal of Population Economics*, 23(1):323–351.

Bargain, O., Orsini, K., and Peichl, A. (2014). Comparing labor supply elasticities in Europe and the United States: New results. *Journal of Human Resources*, 49(3):723–838.

Blundell, R., Chiappori, P.-A., and Meghir, C. (2007). Collective labour supply: Heterogeneity and non-participation. *Review of Economic Studies*, 74:417–455.

Bosch, N. and De Boer, H. (2019). Income and occupational choice responses of the self-employed to tax rate changes: Heterogeneity across reforms and income. *Labour Economics*, 58:1–20.

Cameron, A. and Trivedi, P. (2005). *Microeconometrics: methods and applications.* Cambridge University Press.

Chetty, R. (2009). Is the taxable income elasticity sufficient to calculate deadweight loss? The implications of evasion and avoidance. *American Economic Journal: Economic Policy*, 1(2):31–52.

Colombino, U. (2013). A new equilibrium simulation procedure with discrete choice models. *International Journal of Microsimulation*, 6(3):25–49.

CPB (2015). *Kansrijk Arbeidsmarktbeleid.*

CPB (2020a). *Kansrijk Belastingbeleid.*

CPB (2020b). *Kansrijk Sociale Zekerheidsbeleid.*

Cragg, J. (1971). Some statistical models for limited dependent variables with applications to the demand for durable goods. *Econometrica*, 39:829–844.

De Boer, H.-W. (2018). A structural analysis of labour supply and involuntary unemployment in the Netherlands. *De Economist*, 166(3):285–308.

De Boer, H.-W., Jongen, E., and Kabatek, J. (2015). The effectiveness of fiscal stimuli for working parents. IZA Discussion Paper 9298, Bonn.

De Boer, H.-W., Jongen, E., and Mastrogiacomo, M. (2018). *Structural Reforms: Moving the Economy Forward*, chapter Tax-Benefit Reforms and Structural Models for Labour Supply, pages 239–264. Springer.

Ericson, P. and Flood, L. (2012). A microsimulation approach to an optimal Swedish income tax. *International Journal of Microsimulation*, 5(2):2–21.

Ericson, P., Flood, L., and Wahlberg, R. (2009). SWEtaxben. IZA Discussion Paper 4106, Bonn.

Euwals, R. and Van Soest, A. (1999). Desired and actual labour supply of married men and women in the Netherlands. *Labour Economics*, 6:95–118.

Heckman, J. (1979). Sample selection bias as a specifcation error. *Econometrica*, 47(1):153–161.

Imai, S. and Keane, M. (2004). Intertemporal labor supply and human capital accumulation. *International Economic Review*, 45(2):601–641.

Jongen, E., De Boer, H.-W., and Dekker, P. (2014). Micsim: A behavioural microsimulation model for the analysis of tax-benefit reform in the netherlands. CPB background document, The Hague.

Jongen, E. and Stoel, M. (2019). Estimating the elasticity of taxable labour income in the Netherlands. *De Economist*, 167(4):359–386.

Keane, M. (2011). Labor supply and taxes: A survey. *Journal of Economic Literature*, 49(4):961–1075.

McFadden, D. (1978). Modeling the choice of residential location. In Karlqvist, A., Lundqvist, L., Snickars, F., and Weibull, J., editors, *Spatial Interaction Theory and Planning Models*, pages 75–96. North-Holland, Amsterdam.

OECD (2013). *OECD Labour Force Statistics 2013*. OECD, Paris.

Peichl, A. and Siegloch, S. (2012). Accounting for labor demand effects in structural labor supply models. *Labour Economics*, 19(1):129–138.

Saez, E., Slemrod, J., and Giertz, S. (2012). The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of Economic Literature 2012*, 50:1:350.

Statistics Netherlands (2012). Documentatierapport Arbeidsmarktpanel 1999-2009V1. Statistics Netherlands, Leidschenveen.

Train, K. (2003). *Discrete Choice Methods with Simulation*. Cambridge University Press, Cambridge.

Train, K. (2008). EM algorithms for nonparametric estimation of mixing distributions. *Journal of Choice Modelling*, 1:40–69.

Van Elk, R., De Boer, H.-W., and Verkade, E. (2020). Arbeidsaanbodelasticiteiten in micsim. *CPB Background document*.

Van Soest, A. (1995). Structural models of family labor supply: A discrete choice approach. *Journal of Human Resources*, 30(1):63–88.