ORIGINAL PAPER

Analysing tax-benefit reforms in the Netherlands using structural models and natural experiments

Henk-Wim de Boer[1] · Egbert L. W. Jongen[1] (ID)

Abstract

We combine the strengths of structural models and natural experiments in an analysis of tax-benefit reforms in the Netherlands. We first estimate structural discrete-choice models for labour supply. Next, we simulate key past reforms and compare the predictions of the structural model with the outcomes of quasi-experimental studies. The structural model predicts the treatment effects well. The structural model then allows us to conduct counterfactual policy analysis. Policies targeted at working mothers with young children generate the largest labour supply responses but generate little additional government revenue. Introducing a flat tax, basic income or joint taxation is not effective.

Keywords Structural models · Natural experiments · Tax-benefit reform · Netherlands

JEL Classification C25 · C52 · H31 · J22

1 Introduction

The ageing of the population, the consequences of the Great Recession, and more recently the COVID-19 pandemic put the sustainability of public finances at risk. Governments are looking for cost-effective ways to increase labour supply. To this end, policymakers need good empirical information on how financial incentives

---

✉ Egbert L. W. Jongen
  e.l.w.jongen@cpb.nl

  Henk-Wim de Boer
  h.w.de.boer@cpb.nl

[1] CPB Netherlands Bureau for Economic Policy Analysis, P.O. Box 80510, 2508 GM The Hague, Netherlands

affect the labour supply of different groups and the associated changes in public expenditures and revenues. Policymakers receive information from two different sources when considering the effectiveness of tax-benefit reforms. On the one hand, quasi-experimental studies report on the effects of specific reforms ('natural experiments') implemented in the past. Advantages of this approach are that it relies on minimal assumptions and has a transparent source of identification (Angrist and Pischke 2009; Heckman 2010). However, disadvantages include the absence of an underlying economic model, the external validity of the treatment effects, and the limited scope for doing counterfactual policy analysis.[1] On the other hand, policymakers get simulation results from analyses with structural models. The strong points of this approach are the use of an explicit economic model, the possibility to predict treatment effects in external environments, and the wide scope for conducting counterfactual policy analysis (Keane 2010; Heckman 2010). However, structural models rely on more assumptions and the source of identification is not always transparent. Looking at the strengths and weaknesses of both approaches, they can be said to complement each other.

In this study, we combine structural models with quasi-experimental studies to evaluate the effectiveness of tax-benefit reforms in the Netherlands. Specifically, we estimate structural models for labour supply of a large number of subgroups, compare the simulated treatment effects of policy reforms in the structural models with the results from three quasi-experimental studies, and then use the structural model to study counterfactual policy reforms that feature prominently in the political debate in the Netherlands and abroad.

We estimate a structural discrete-choice model for labour supply, building on a large body of literature, e.g. Aaberge et al. (1995, 1999), Van Soest (1995), Keane and Moffitt (1998), Brewer et al. (2006), and Bargain et al. (2014). Discrete-choice models have the advantage of being able to take into account all the complexities in the budget set that result from the tax-benefit system, such as kinks and non-convexities. Furthermore, the estimation of discrete-choice models does not require ex ante imposition of quasi-concavity of preferences.[2] We use an exceptionally large and rich administrative household data set that was constructed specifically for this project by Statistics Netherlands: the Labour Market Panel (*Arbeidsmarktpanel*). Our sample consists of more than 840,000 observations. The size of this data set allows us to precisely estimate preferences over income, leisure, and formal childcare (for parents with young children) for a large number of subgroups.

Next, we use the estimated structural model to simulate a number of key reforms implemented in the past and compare the simulated treatment effects with quasi-

---

[1]Some of these concerns are overcome in the so-called sufficient statistics literature (e.g. Chetty 2009). In the sufficient statistics literature, authors use an explicit economic model to derive, e.g. elasticities that are estimated in the programme evaluation literature. However, this approach can be used only for the analysis of counterfactual small reforms, and cannot be used for the counterfactual analysis of major reforms such as, e.g. the introduction of a flat tax system.

[2]Studies using continuous labour supply choices and piecewise-linear budget constraints need to impose global quasi-concavity of preferences ex ante, which may have led to upward biased estimates of labour supply elasticities in these studies (MaCurdy et al. 1990).

experimental studies on the same reforms. In particular, we compare the simulated treatment effects of the 2005–2009 reform of childcare subsidies and in-work benefits for households with young children with the estimated treatment effects presented in Bettendorf et al. (2015). Furthermore, we compare the simulated treatment effects of the 2002 reform of the in-work benefit for single parents with the estimated treatment effects presented in Bettendorf et al. (2014). Finally, we compare the simulated intensive margin (hours worked per employed person) elasticities of the structural model with the estimated intensive margin elasticities presented in Bosch and van der Klaauw (2012) and Bosch and Jongen (2013), who use the 2001 tax reform that substantially reduced marginal tax rates. Part of the variation used in the quasi-experimental studies is also used in the estimation of the structural model, but part of the variation also comes from outside the data period used to estimate the structural model.[3]

Our main findings are the following. First, we uncover large differences in labour supply elasticities between demographic groups and decision margins. Indeed, there are large differences in labour supply elasticities by age of the youngest child. We also find large differences in the relative importance of the extensive margin (participation) and the intensive margin (hours per employed person), with the intensive margin response still substantial for women with young children but small for most other groups. Second, we find that the structural model gives a good prediction of the treatment effects estimated in quasi-experimental studies on past reforms. Therefore, we feel confident to simulate counterfactual tax-benefit reforms using the structural model. Third, we find that reducing marginal tax rates is not an effective way to promote labour supply. In-work benefits targeted at low-wage earners appear to be more effective, and policies targeted at working mothers with young children generate the largest labour supply response and also cause a noticeable drop in gender inequality in wage income. However, policies targeted at working mothers with young children generate little additional revenue for the government because that group already receives substantial subsidies when working. Using the structural model we also simulate some large counterfactual tax reforms that feature prominently in the political debate. We find that proposals for a move towards a flat tax system, a basic income, or joint taxation are not effective, including a steep rise in gender inequality. Indeed, an efficient tax system accounts for the large heterogeneity in behavioural responses we uncover and hence cannot be too simple.

This study contributes to the existing literature in a number of ways. The size of our data set allows us to estimate preferences separately for subgroups that earlier studies did not consider (previous studies typically only focussed on couples, e.g., Van Soest and Das 2001; Van Soest et al. 2002) or had to pool due to the limited number of observations (previous studies typically pooled households with and without children, e.g., Bargain et al. 2014). We consider a large number of household types, do not pool the data over the various household types, and uncover much more heterogeneity in the behavioural responses than previous studies. We also exploit

---

[3]Hence, our comparison of the structural model with the quasi-experimental studies goes beyond a test of 'goodness-of-fit'.

detailed administrative data on the use and price of formal childcare, information that is not readily available in most other labour supply studies (e.g. Blundell et al. 2000; Van Soest and Das 2001; Blundell and Shephard 2012; Bargain et al. 2014). Furthermore, previous studies had to rely on one cross-section of data or on a few repeated cross-sections from a period when there was hardly any change in the tax system (e.g. Van Soest and Das 2001; Van Soest et al. 2002; Bargain et al. 2014). Hence, identification in those studies comes only from cross-sectional differences in financial incentives due to non-linearities in the tax-benefit system. In contrast, we use several years of data, which includes a major reform of childcare subsidies and in-work benefits for working parents. These reforms strengthen our identification by generating large exogenous variation in budget constraints. Also, we compare the predictions of the structural model with results from a number of quasi-experimental studies on key past reforms and thus contribute to the growing body of literature that evaluates the performance of structural models by comparing simulated policy responses with the results from (quasi-)experimental studies (e.g. Attanasio et al. 2011; Bargain and Doorley 2017; Brewer et al. 2006; Cai et al. 2008; Todd and Wolpin 2006; Geyer et al. 2015; Hansen and Liu 2015; Pronzato 2015; Thoresen and Vatto 2015). In addition, because our structural model is fully integrated with a detailed tax-benefit calculator, we are able to study the effectiveness of various reform proposals, taking into account the budgetary effects of the behavioural responses to the reform. Indeed, the integrated model allows us to go beyond back-of-the-envelope calculations on the effectiveness of different reform proposals (as in, e.g. Blau 2003; Lokshin 2004).

The outline of the paper is as follows. Section 2 describes the Dutch labour market and policy environment. Section 3 outlines the structural model and empirical strategy. Section 4 describes the data set used to estimate the structural model and Section 5 presents the estimation results and corresponding labour supply elasticities. Section 6 then compares the simulated treatment effects of the structural model on a number of key past reforms with the estimated treatment effects from quasi-experimental studies. Next, Section 7 studies the effectiveness of a number of counterfactual tax-benefit reforms. Section 8 discusses our findings and concludes. Supplementary material is given in an Appendix.

2 The Dutch labour market and tax-benefit system

Over the past decades, the Netherlands, like many other developed countries, witnessed a substantial shift in the household composition of the population, see Fig. 1. In particular, the share of couples with children has declined, whereas the share of couples without children has increased. Furthermore, the share of singles, both with and without children, has increased. Hence, empirical knowledge of the behavioural responses of singles and single parents is becoming increasingly relevant.[4]

Another important development has been the rise in the employment rate of women. Figure 2 shows the employment rate for men and women in the Nether-

---

[4]Single parents consist of both lone mothers that decided to raise the child by themselves and of divorced parents.
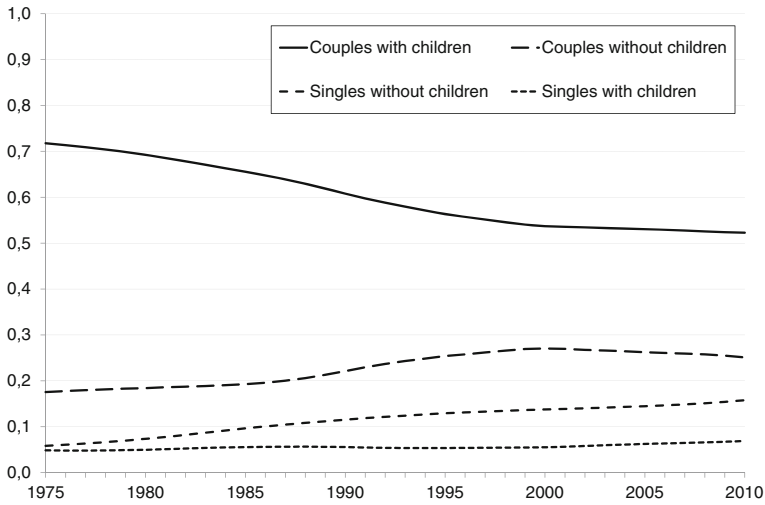
Fig. 1 Shares of individuals per household type in the Netherlands. Source: Statistics Netherlands (statline.cbs.nl). Individuals 15–64 years of age

lands over time, along with the development in the employment rate for a number of other countries. The employment rate for men in the Netherlands has increased somewhat over the 1975–2014 period, but with a noticeable dip in the 1980s due to the recession and the increased use of early retirement and disability benefits. Since the 1980s, early retirement benefits have been cut down substantially and have become more actuarially fair, and access to disability schemes has become more difficult. As a result, the employment rate of men has rebounded to pre-1980 levels. In 2014, the employment rate of men in the Netherlands was one of the highest of all OECD countries. The rise in the employment rate of women has been nothing short of spectacular, from just 30% in 1975 to 70% at the start of the Great Recession. Indeed, by 2014, the Netherlands had one of the highest employment rates for women in the OECD. An age-period-cohort analysis by Euwals et al. (2011) indicates that cohort effects are an important driver behind this increase, in particular for women with
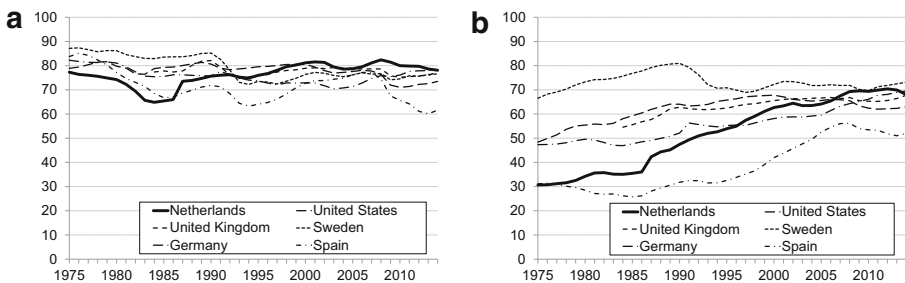

Fig. 2 Employment rates in selected countries. Source: OECD (2016). Individuals 15–64 years of age
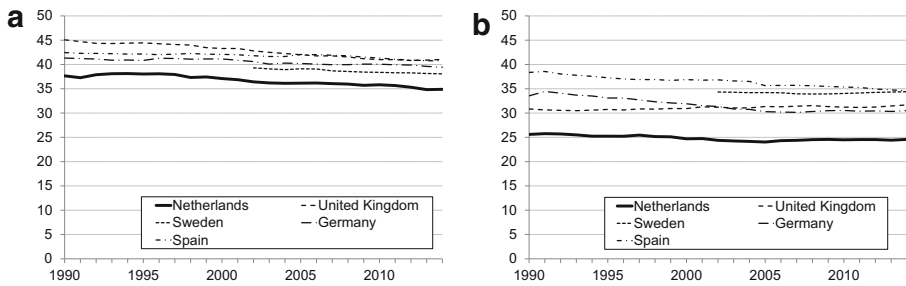
Fig. 3 Average usual weekly hours worked on the main job in selected countries. Source: OECD (2016). Individuals 15–64 years of age. **a** Men. **b** Women

young children. The steep rise in the employment rate of women has important implications for the behavioural responses to changes in financial incentives. We estimate much lower labour supply elasticities for women in couples than previous studies for the Netherlands (e.g. Van Soest 1995; Van Soest and Das 2001) that used data of the 20th century.[5]

Whereas the rise in the employment rate of women in the Netherlands has been spectacular, their hours worked remained remarkably stable over the past decades, see Fig. 3. If anything, hours worked per week by employed women decreased slightly. However, a noticeable gap of 5 to 10 h per week remains with their peers in other European countries. The changes in hours worked per week by employed men were also limited over the past decades, with a slightly downward trend. However, also for men, the Dutch, on average, work a few hours per week less than their peers in other European countries. Below, we study how changes in tax-benefit policies affect both the participation and the hours-worked-per-week decisions.

Considering the Dutch tax-benefit system, the Netherlands has an individual tax system, with specific tax credits and subsidies targeted at certain groups. However, income-dependent income support is based on household rather than individual income. The financial incentives implicit in the tax-benefit system are illustrated in the appendix to De Boer and Jongen (2020), where we present so-called effective marginal tax rates (EMTRs) and participation tax rates (PTRs) for different subgroups.[6] The analysis shows that EMTRs and PTRs are particularly high for singles and single parents, and particularly low for secondary earners. Below, we consider how changes in the EMTRs and PTRs following tax-benefit reforms affect the participation and hours-worked-per-week decisions of different groups on the labour market.

---

[5]This is in line with the findings for the USA in Blau and Kahn (2007) and Heim (2007), they find that the female labour supply elasticity has declined substantially over the past decades, along with the rise in female employment rates. Furthermore, using a cross-section of countries, Bargain et al. (2014) show that the female labour supply elasticity is lower for countries that have a higher female employment rate.

[6]Following, e.g. Brewer et al. (2010) and OECD (2016).

3 Structural model

We develop a structural model, where households are assumed to maximise a unitary household utility function. The most elaborate specification, for couples with young children, is outlined below. In this household, both partners choose their labour supply as well as their hours of formal childcare.[7,8] The utility functions for the other household types (defined below) are a special case of this utility function.

The systematic part of household utility, U^s, depends on disposable income y, hours worked by the male h_m, hours worked by the female h_f, and hours of formal childcare cc. We ignore saving and borrowing, and hence consumption equals disposable income. For the functional form of U^s, we use the translog specification:

$$U^s(v) = v'\mathbf{A}v + \mathbf{b}'v + \mathbf{d}'\mathbf{1}[\mu > 0],$$
$$v = (\log(y), \log(1 - h_m/T), \log(1 - h_f/T), \log(cc)),$$
$$\mu = (h_m, h_f, cc), \tag{1}$$

with $\mathbf{A}$ being a symmetric matrix of quadratic coefficients and $\mathbf{b}$ being a vector of linear coefficients corresponding to the vector of the variables contained in v. Note that we allow childcare to enter directly in the utility function (next to entering via disposable income). The hours worked variables h_m and h_f in the vector v have been transformed into indicators of leisure utilisation, representing the fraction of weekly time endowment T which is spent on activities unrelated to work. The vector $\mathbf{d}$ captures fixed costs of work, for the male and the female separately, and fixed costs of using formal childcare, for the household as a whole.

For some household types, the full translog specification resulted in a significant share ($>5\%$) of households with negative marginal utility of income in the observed choices. This is not consistent with utility maximisation and drives down the labour supply elasticities to implausible values. For these household types, we dropped the interaction terms between income and leisure, which resulted in a low share of households with negative marginal utility of income ($<5\%$). For some households, we also obtained an 'inverted' pattern for the marginal utility of income, with a negative (log)

---

[7]Unfortunately, we do not have data on the use of paid leave by mothers or fathers, which may interact with the use of childcare and labour supply. However, the duration of paid maternity leave and paid paternity leave was and is rather short in the Netherlands, 16 weeks and 0.4 weeks for mothers and fathers (or partners for same-sex couples) respectively, compared to an average of 18.1 and 1.4 weeks respectively in the OECD (OECD Family database). Furthermore, there is no legislated paid parental leave for mothers or fathers following maternity or paternity leave, compared to 35.8 and 6.7 weeks on average for mothers and fathers respectively in the OECD. However, 16 out of the 99 largest (in terms of employees covered) collective labour agreements have paid parental leave. In 2018, 32% of mothers and 17% of fathers with a child younger than 8 years of age used parental leave (Statistics Netherlands 2018).

[8]Another concern may be that rationing on the formal childcare market prohibits parents from realizing their childcare demand, which may then restrict their labour supply. However, the available data on waiting lists suggest that these are rather small during our data period, and that the change in waiting lists was much smaller than the change in filled childcare places. For example, the survey data reported in Van Rens and Smit (2011) suggest that the waiting list for daycare (out-of-school care) dropped from 10% (11%) of filled places in 2007 (the first year of the survey) to 7% (6%) of filled places in 2009. The drop in waiting lists is much smaller than the increase in the number of children going to daycare and out-of-school care, which increased by 49% (19%) and 139% (55%) respectively between 2005 and 2009 (2007 and 2009).

linear term and a positive (log) quadratic term. This results in implausible (positive) income effects, and for these households, we dropped the quadratic term in income. Finally, for certain other household types, the translog specification was still 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 (log) leisure, which then improves the fit at the top.[9]

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:

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

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).[10] 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.[11]

Disposable household income is given by:

$$y = w_m h_m + w_f h_f - T(w_m, h_m, w_f, h_f; q) - TC(p_{cc}, cc; q) + S(p_{cc}, cc, y_t; q),$$
$$(3)$$

where w_m and w_f denote the gross hourly wage for the male and the female,[12] $T(.)$ denotes taxes and employees' premiums, q denotes individual and household characteristics, $TC(.)$ is the total cost of formal childcare, with p_{cc} denoting its price per hour, and $S(.)$ is the childcare subsidy, which depends on the hourly price of formal childcare, the hours of formal childcare, taxable income y_t, and household characteristics (like the ages of the children[13]).

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.[14] For non-workers, we simulate wages using estimates from a model that accounts for selection (Heckman

---

[9]Given the large number of observations we have, we found it less useful to use the Akaike information criterion or the Bayesian information criterion for model selection. They will favor ever large models when the sample size is very big, since the relative efficiency loss of allowing for an extra coefficient is getting small. Instead, we opted for model selection based on whether the labour supply elasticities were consistent with utility maximization and gave a good fit to the data.

[10]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.

[11]As a robustness check, we allowed for unobserved preference heterogeneity by using a latent classes approach, the results are similar (available on request).

[12]For simplicity, we assume that the gross hourly wage does not depend on the hours worked.

[13]We focus on the youngest age of the child; we do not allow for preferences to depend on the number of children.

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

1979)[15], and taking multiple draws from the estimated wage error distribution, see the appendix to De Boer and Jongen (2020). 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, see the appendix to De Boer and Jongen (2020).

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 systematic part $U^s(v_j)$, the utility function contains alternative-specific stochastic terms ε_j:

$$U(v_j) = U^s(v_j) + \varepsilon_j. \tag{4}$$

These stochastic terms are assumed to be independent and identically distributed across alternatives, and to be drawn from a type 1 extreme value distribution. This leads to a multinomial logit specification (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, complicates the estimation of the likelihood. We use $R = 50$ (independent) draws from the wage distribution for non-working men and women, the price distribution for non-users of formal childcare, and the random terms for unobserved heterogeneity. We use simulated maximum likelihood, where the likelihood is given by:

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

$$\tag{5}$$

with D_{ki} being an indicator function taking the value 1 for the observed choice for household i, and zero otherwise.[16]

4 Data structural model

We use data from the Labour Market Panel (LMP) of Statistics Netherlands (2012). This data set was constructed specifically for the empirical analysis presented here. The LMP is a large administrative household panel data set over the period 1999–2009, containing a rich set of individual and household characteristics, including gender, month and year of birth, the level of education and ethnicity of all adult household members, the ages of the children, and place of residence. In addition, the LMP also contains administrative data on gross income from different sources (e.g.

---

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

[16]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.

wages, profits, benefits) and on hours worked. Finally, the LMP contains administrative data on the use and gross hourly price of formal childcare for each child in formal childcare.[17] Because data on childcare is only available from 2006 onwards, we restrict the sample to the 2006–2009 period.

In the empirical analysis, we model the labour supply decision for employed people, those on welfare benefits, and those without personal income.[18] We make a number of additional selections. We exclude people under 18 years of age (most of them are in education), and those over 63 years of age (we do not model the retirement decision). Furthermore, we do not model the labour supply decision of students, people on retirement or disability benefits, and the self-employed. We do not model their labour supply decision because reliable information is not available on their hours worked or because we are unable to determine their budget constraint. Furthermore, same-sex households are also excluded, as are households for which characteristics about individual members or the household are missing.[19] In the end, we use 840,348 observations in the estimations.

We estimate structural discrete-choice models for the following 15 household types: childless singles; single parents with a youngest child aged 0–3, 4–11, 12–17, or 18 years of age or older, respectively; adult children living with their parent(s);[20] couples without children where we model the labour supply decision of both partners; couples without children where we only model the labour supply decision of the man (because the woman is a student, on disability or retirement benefits or self-employed, see above); couples without children where we only model the labour supply of the woman (because the man is a student, on disability or retirement benefits or self-employed); couples where we model the labour supply of both partners and that have a youngest child aged 0–3, 4–11, 12–17, or 18 years of age or older, respectively; couples with children where we only model the labour supply of the man (because the woman is a student, on disability or retirement benefits or self-employed); and couples with children where we only model the labour supply of the woman (because the man is a student, on disability or retirement benefits or self-employed).

We discretise the data for the discrete-choice model in the following way. Adults can choose from 6 labour supply options: working 0, 1, 2, 3, 4, or 5 days per week,

---

[17]Unfortunately, informal childcare is not included in our administrative data set. However, De Boer et al. (2015) estimate preferences using the overlap in working hours of parents minus the hours of formal childcare as a proxy for informal childcare. The resulting labour supply and formal childcare price elasticities are very similar to the model without the proxy for informal childcare.

[18]We remove people on unemployment benefits from the structural model, implicitly assuming that they are constrained in their labour supply choice. In the simulation model, we add a reduced-form model on the probability of people being on unemployment benefits, following Ericson and Flood (2012), see the appendix to De Boer and Jongen (2020).

[19]For couples with a youngest child of between 0 and 3 and between 4 and 11 we use a 50% subsample. These groups have the largest discrete-choice set, and using the full sample was not possible due to memory restrictions in Stata.

[20]We model adult children living with their parents as a separate household category that is not entitled to welfare benefits, in accordance with Dutch legislation.

each day equalling 8 h.[21] For childcare, we allow for 0, 1, 2, and 3 days[22], with data showing a typical childcare day to equal 10 h,[23] and a typical out-of-school-care day to equal 5 h.[24] Couples with a youngest child aged 0 to 3 or 4 to 11 have the largest choice set: $6 \cdot 6 \cdot 4 = 144$ alternatives.

To determine disposable household income in each discrete option, we use the advanced tax-benefit calculator MIMOSI (Koot et al. 2016).[25] MIMOSI is the official tax-benefit calculator of the Dutch Government for the (non-behavioural) analysis of the impact of reform proposals on the disposable income distribution and the government budget. MIMOSI takes into account all (national[26]) taxes, social security premiums, and income-independent subsidies and tax credits. Furthermore, MIMOSI also calculates the childcare subsidy in each option.[27] Furthermore, in accordance with the law, we ensure that household disposable income (excluding childcare costs and childcare subsidies) cannot drop below the welfare level.[28] For each discrete option, we also calculate the net transfer from the household to the government (positive or negative). This allows for an accurate calculation of the net budgetary costs of the reforms we simulate.

5 Estimation results structural model

For each household type, we allow preferences for leisure to depend on age and fixed costs of work to depend on the level of education (in three classes) and ethnicity (in three classes). Furthermore, for households with a youngest child aged 0–3 or 4–11,

---

[21] The discretization of the number of working hours per week is as follows: 0–4 h = not working (j =1), 5–12 h = 1 day (j=2), 13–20 h = 2 days (j=3), 21–28 h = 3 days (j=4), 29–36 h = 4 days (j=5), and 37 h or more = 5 days (j=6).

[22] The data show that using formal childcare for more than 3 days per week is rare in the Netherlands. The remaining childcare needs are usually met by informal care or parents themselves.

[23] Classified as follows: 0 h = no child care (j=0), 0–14 h = 1 day (j=1), 15–24 h = 2 days (j=2): 25 h or more = 3 days (j=3). Opening hours are typically from 7:00 until 18:00 or later, enough to cover a full working day.

[24] Classified as follows: 0 h = no child care (j=0), 0–7 h = 1 day (j=1), 8–12 h = 2 days (j=2): 13 h or more = 3 days (j=3). School days typically end around 14:30 or 15:00, 12:00 on Wednesdays, and out-of-school care then is open until 18:00 or later.

[25] Disposable incomes in the estimations and simulations are in 2006 prices. We use the CPI to convert nomimal values in later years into 2006 prices.

[26] Local taxes account for only a small portion of total taxes in the Netherlands (3.3% in 2007, European Union, 2014).

[27] Only working parents are entitled to receive the childcare subsidy, with the subsidy level depending on the gross hourly price of childcare per type of childcare (daycare or out-of-school care), and only up to a maximum price beyond which parents receive no additional subsidy, household income (subsidies are lower for higher incomes), and number of children (subsidies are higher for second, third and subsequent children in formal childcare).

[28] We ignore potential nontake-up of welfare benefits, which may affect labour supply choices (Hoynes 1996; Keane and Moffitt 1998). However, non take-up of welfare benefits in the Netherlands is small. Indeed, in our data set only 0.2% of households have no income. These households may be living from their wealth. Unfortunately, we do not have information on wealth in our data set.

we allow preferences for the use of formal childcare to depend on the level of education and ethnicity, and whether or not their residential location was in a large city (>150,000 inhabitants). The preference parameters are not discussed here, because it is often a combination of preference parameters that drives behavioural responses (the estimated preferences are available on request).[29] Furthermore, there is no analytical solution for the labour supply elasticity in discrete-choice models. Therefore, following the literature (e.g. Bargain et al., 2014), we simulate these elasticities by increasing gross wages by 10%. We present the total elasticity (% change in total hours worked over the % change in the gross wage rate), and the decomposition into the extensive margin elasticity (% change in the participation rate over the % change in the gross wage rate) and the intensive margin elasticity (% change in hours worked per employed person over the % change in the gross wage rate).

Figure 4 gives the simulated labour supply elasticities for couples where both partners can adjust their labour supply. We find small, positive labour suppy elasticities for men. Labour supply elasticities are higher for women, on the extensive margin but also on the intensive margin. Furthermore, labour supply elasticities for women in couples are particularly high when they have a young child.[30]

Figure 5 shows that the labour supply elasticity is relatively low for childless singles without children.[31] The labour supply elasticity is much higher for single parents with a youngest child of 0–3 years of age. The labour supply elasticity of single parents whose youngest child is over the age of 3 is lower, though still higher than that of childless singles. Also note that the differences among single parents are primarily driven by differences in the extensive margin elasticity. The intensive margin response for single parents is quite small.[32]

Figure 5 also gives the labour supply elasticities for couples where only one partner can choose his or her labour supply (because the other partner is a student, on disability or retirement benefits or self-employed). For these groups, we pool couples with children of all ages. Most men in these couples are working, and typically also full-time (descriptive statistics are available on request). Hence, there is little upward potential in terms of total hours worked, and they have a relatively low labour supply elasticity. For women in these couples, there is more upward potential in total hours worked, they have a higher labour supply elasticity. Finally, adult children living with their parents generally have a high participation rate, resulting in a relatively low labour supply elasticity.[33]

---

[29] The models generate a good fit of the hours worked distribution and the distribution of the use of child care (available on request).

[30] The cross-elasticities, e.g. the % change in total hours worked by one partner over the % change in the gross wage rate of the other partner, are close to zero for men, but are negative and non-negligible for women (available on request).

[31] We also estimated preferences separately for single men and single women, but they were almost identical, so we pooled these groups in the estimations.

[32] Their budget constraint plays an important role here, as working extra hours or days may generate little additional net income due to their relatively high effective marginal tax rate, see the appendix to (De Boer and Jongen 2020).

[33] Comprehensive surveys of the labour supply elasticity in a large number of countries can be found in Blundell and MaCurdy (1999), Bargain and Peichl (2016), and the recent estimates presented in Bargain
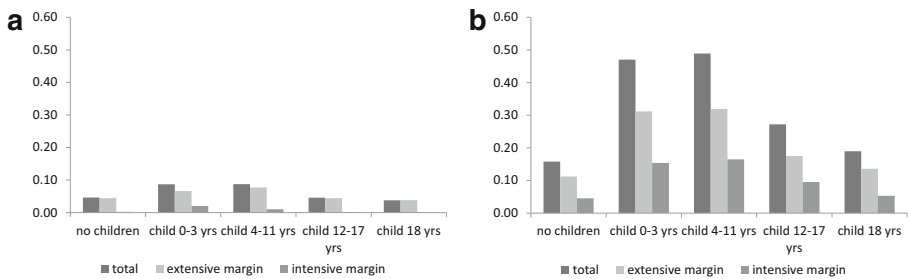
Fig. 4 Couples where both partners have a labour supply choice. **a** Men. **b** Women

6 Comparison predictions structural model with results from quasi-experimental studies

Next, we consider whether the structural model can predict the effects of past reforms. There is a growing body of literature that compares the simulated policy responses in structural models with the results from (quasi-)experimental studies.[34] We present three such comparisons: (i) one for couples with a youngest child of 0–3 and 4–11 years of age, (ii) one for single parents with a youngest child of 12–15 years of age, and (iii) one for the intensive margin responses for a number of household types. For the comparison, we have re-estimated the structural model using data from 2006 only, so that the parameter estimates are only estimated on data in the absence of the policy change of the first reform we consider.[35]

Table 1 shows our structural model results for couples with a youngest child of 0–3 and 4–11 years of age. Bettendorf et al. (2015) use differences-in-differences (DD)

---

et al. (2014). We compare our results with the recent estimates for Europe and the USA in Bargain et al. (2014). Bargain et al. (2014) find that for women in couples the total hours elasticity ranges from 0.1 to 0.6 among countries (with a mean of 0.27). Our estimates for women in couples with young children fall within this range. The estimates for women in couples with older children or no children are somewhat lower. However, the participation rate of women in the Netherlands is relatively high from an international point of view. For men in couples, the total hours elasticity in Bargain et al. (2014) ranges from 0.05 to 0.15 among countries (with a mean of 0.10). Our estimates for men in couples without and with children are on the lower end of that range. For single men, Bargain et al. (2014) find a total hours elasticity ranging from 0.0 to 0.4 (and some even higher). For single women, they find an elasticity ranging from 0.1 to 0.5 (and again some even higher). Our estimates for singles are on the lower end of that range, and our results for single parents are more in the middle and upper part. Bargain et al. (2014) find that the extensive margin elasticity is typically (much) more important than the intensive margin elasticity and that cross-elasticities for women in couples are non-negligible and are close to zero for men in couples. These are also our findings.

[34] See, e.g. Attanasio et al., (2011); Bargain and Doorley, (2017); Brewer et al. (2006); Cai et al. (2008); Todd and Wolpin (2006); Geyer et al. (2015); Hansen and Liu (2015); Pronzato (2015); Thoresen and Vatto (2015).

[35] However, using the structural model estimated on data for the period 2006–2009 instead results in simulated policy responses that are very similar to those using data from 2006 only (available on request). This shows that the estimation of the preferences also depends very much on the within-year variation across households.
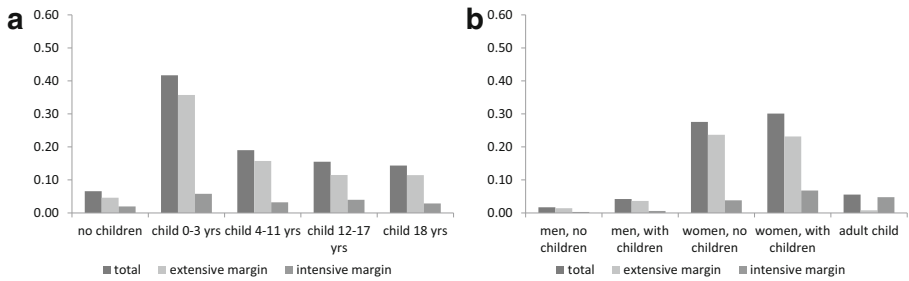
Fig. 5 Households where only one person has a labour supply choice, and adult children. **a** Singles and single parents. **b** Couples where only one person has a labour supply choice, and adult children living at home

Table 1 Comparison with results quasi-experimental study: couples with young children

	Structural model[a]				DD analysis[b]
	Childcare subsidies	Combi. credit	Income-depend. combi. credit	Total	Coefficient
Absolute changes					
Youngest child 0–3 yrs					
Participation rate women	0.016	−0.007	0.026	0.035	0.020
	(0.000)	(0.000)	(0.000)	(0.000)	(0.007)
Hours worked per week women	0.616	–0.124	0.724	1.216	1.099
	(0.003)	(0.001)	(0.003)	(0.006)	(0.215)
Participation rate men	0.002	–0.002	0.004	0.004	0.006
	(0.000)	(0.000)	(0.000)	(0.000)	(0.004)
Hours worked per week men	0.044	–0.027	0.043	0.060	–0.364
	(0.001)	(0.001)	(0.001)	(0.000)	(0.229)
Youngest child 4–11 yrs					
Participation rate women	0.002	–0.007	0.022	0.017	0.022
	(0.000)	(0.000)	(0.000)	(0.000)	(0.007)
Hours worked per week women	0.070	–0.112	0.581	0.539	0.741
	(0.001)	(0.001)	(0.003)	(0.003)	(0.212)
Participation rate men	0.000	–0.001	0.002	0.002	0.003
	(0.000)	(0.000)	(0.000)	(0.000)	(0.004)
Hours worked per week men	0.008	–0.005	0.011	0.013	–0.141
	(0.001)	(0.001)	(0.001)	(0.000)	(0.227)

[a]Standard errors in parentheses. The standard errors from the structural model are based on 200 draws of the estimated preference parameters. [b]Additional estimates using the same sample as (Bettendorf et al. 2015), full regression results available on request

Table 2 Comparison with results quasi-experimental study: single parents

	Structural model[a]	DD analysis[a] Coefficient	RD analysis[b] Coefficient
	Absolute changes		
Participation rate			
Youngest child of 12–15 years	0.000	0.004	0.004
	(0.000)	(0.006)	(0.011)
Youngest child of 4–11 years	–0.006		
	(0.000)		
Youngest child of 0–3 years	–0.011		
	(0.000)		

[a] Standard errors in parentheses. The standard errors from the structural model are based on 200 draws of the estimated preference parameters. [b] For details of the DD and RD results, see Bettendorf et al. (2014)

to analyse the employment effects of a combination of reforms during the period 2005–2009 targeted at households with children of 0–11 years of age. These reforms contained three elements: (1) an increase in childcare subsidies (column *Childcare subsidies*), (2) an increase in the in-work benefit for secondary earners with children of 0–11 years of age (column *Income-depend.combi.credit*), and (3) the in-work benefit for both primary and secondary earners with children of 0–11 years of age was abolished (column *Combi.credit*).[36] Bettendorf et al. (2015) use data for the 1995–2009 period from the Labour Force Survey. They present estimation results for women in couples with a youngest child of 0–11 years of age. To make the comparison with the DD as clean as possible, we use the same sample as Bettendorf et al. (2015) to estimate the treatment effects by the subgroups with a youngest child of 0–3 or 4–11 years of age, respectively. The results are given in Table 1, along with the simulation results from the structural model. The results for the structural model are consistent with the results from the DD analysis for women with children. The estimated effects on the participation rate of men with children are also in line with the predictions from the structural model. The only coefficient of the DD analysis which differs somewhat from the prediction of the structural model is the intensive margin response by men, for which the DD analysis suggests a larger, negative though not statistically significant response than the structural model.

Table 2 shows the evaluation of the predictions made by our structural model for single parents with a youngest child of 12–15 years of age. Bettendorf et al. (2014) use DD and regression discontinuity (RD) to study the impact of an in-work benefit targeted at single parents. In 2002, the children's age of eligibility was increased, and the target group of the in-work benefit was extended from single parents with a youngest child of 0–11 years of age to single parents with a youngest child of 0–15 years of age. The outcome of the analysis is that they find a small effect on labour supply, in both the DD and RD analyses, not significantly different from zero.

---

[36] See Bettendorf et al. (2015) for a detailed description of the reforms.

Table 2 shows the effect of abolishing the in-work benefit targeted at single parents. We simulate the effect on the participation rate of single parents with a youngest child of 0–3, 4–11, and 12–15 years of age.[37] We find the effect on the participation rate with –1.1 percentage points to be quite sizeable for single parents with a youngest child of 0–3 years of age, whereas it drops to –0.6 percentage points for single parents with a youngest child of 4–11 years of age, and to 0.0 percentage points for single parents with a youngest child of 12–15 years of age. Indeed, single parents with a youngest child 12–15 years of age are already less responsive to financial incentives than single parents with a youngest child 0–3 years of age. More importantly, again, the structural model predicts treatment effects that are in line with the results from the quasi-experimental study.

Finally, we compare the intensive-margin responses in the structural model with a quasi-experimental study on intensive-margin responses. Figures 4 and 5 show that intensive-margin responses are typically rather small, and much smaller than extensive-margin responses. We compare the intensive-margin elasticities with results from the DD analysis in Bosch and Jongen (2013). They use the 2001 tax reform, which generated large heterogeneous variation in marginal tax rates. For men in couples, they find very low intensive-margin elasticities with a point estimate of 0.00 (s.e. 0.01), in line with the structural model. For women in couples, they find larger intensive-margin elasticities, with a point estimate of 0.15 (s.e. 0.06).[38] This is in line with the structural-model results on the response at the intensive margin for women with young children, and somewhat higher than for the other groups of women in couples. For singles and single parents, Bosch and Jongen (2013) also find somewhat higher intensive-margin elasticities than the structural model, 0.15 to 0.20, respectively. We should note though that the comparison is somewhat complicated because we compare gross wage elasticities of the structural model[39] to the net wage elasticities of the DD. The latter are typically higher (Bargain et al. 2014).

7 Simulating tax-benefit reforms

Overall, the structural model predicts the treatment effects of past reforms rather well. We then exploit the strength of the structural model by simulating counterfactual tax-benefit reforms. We first consider changes in a selected set of policy parameters, motivated by recent reforms in the Netherlands, to illustrate which policies are more or less cost-effective in terms of stimulating labour supply. Subsequently, we consider a number of major tax reforms that feature prominently in the policy debate in the Netherlands, for example in the election proposals of Dutch political parties, and

---

[37](Bettendorf et al. 2014) do not consider the effect of the reform on hours worked.

[38]Bosch and van der Klaauw (2012) apply the estimator of Blundell et al. (1998) to estimate the intensive-margin responses for women in couples, using data from the Labour Force Survey. They find a coefficient insignificantly different from zero.

[39]Which is common in the literature on discrete-choice labour supply, see for instance Bargain et al. (2014).

abroad. Also for the policy simulations, we use the estimated preferences using data from 2006 only.[40],[41]

7.1 Changes in selected tax-benefit policies

Marginal tax rates We first consider the effects of changes in marginal tax rates. More specifically, we consider the effects of decreasing the tax rate of the first, second, third, and fourth (open) tax bracket, so that, in each simulation, tax revenues decrease by 1.5 billion euros before behavioural responses. Table 3 gives the results in columns (1)–(4), respectively.[42]

We report the effects on hours worked per week and on the participation rate. Hours worked per week includes the zeros for the non-employed. The participation rate is the number of persons employed over the total number of employed and not employed.[43] We also calculate the effect on labour productivity per hour worked, which is obtained by subtracting the change in hours worked from the change in labour costs, where the change in labour costs, in turn, is an approximation for the change in output. Furthermore, we also calculate the so-called knock-on effects for the government budget. These are the net budgetary savings due to behavioural responses, expressed as a percentage of the ex-ante (before behavioural responses) budgetary 'shock'.[44] This is particularly relevant for simulations that increase the participation of (potential) secondary earners, who pay little in taxes and typically get substantial subsidies when they start working. An alternative strategy would be to simulate policy reforms that are budgetary neutral after taking into

---

[40]Before we can simulate these reforms, we have to prepare the model for policy simulation. We estimated the preferences of the different household types using data from the LMP. However, CPB uses data from the Income Panel of Statistics Netherlands to calculate the effects of tax-benefit reforms on the income distribution and the government budget (see Koot et al. (2016)). To have a single model that will generate all relevant output, we integrate the discrete-choice model for labour supply and childcare with the tax-benefit calculator using the Income Panel data. The appendix to De Boer and Jongen (2020) describes this process and compares the resulting labour supply elasticities.

[41]The simulation results are very similar when we use the preferences estimated on data for the period 2006–2009 instead (available on request).

[42]To keep the table to a manageable size, we aggregate the results to some broader categories. Specifically, we use the following aggregates: (1) 'Men in couples youngest child 0–17' and 'Women in couples youngest child 0–17' which are respectively men and women in couples for which we model both the labour supply decision with a youngest child of 0–17 years of age, (2) 'Men in other couples' and 'Women in other couples' are men and women in couples without children (couples with a youngest child of 18 years of age or older are also classified as couples without (dependent) children), or where one of the partners in the couple is a student, on disability or retirement benefits or self-employed (see above), (3) 'Single parents youngest child 0–17', and (4) 'Singles' consist of singles without (dependent) children and adult children living with their parents (this group also includes single parents with a youngest child 18 years of age or older). Furthermore, the total results over all groups are for people whose labour supply is determined within the model only, so excluding the potential labour supply of the partners that are a student, on disability or retirement benefits or self-employed.

[43]Excluding the partners that are a student, on disability or retirement benefits or self-employed.

[44]We do not consider the change in the system costs of running the tax-benefit system. However, since we are essentially only changing parameters of existing tax and benefit policies, these should be rather small compared to the budgetary effects of the policy changes.

Table 3 Changes in selected tax-benefit policies

Simulation	Base	(1) First bracket[a]	(2) Second bracket[b]	(3) Third bracket[c]	(4) Fourth bracket[d]	(5) Welfare benefits[e]	(6) In-work tax credit[f]	(7) Combination credit[g]	(8) Childcare subsidies[h]	(9) Child benefit[i]
Ex-ante government spending (in billion euros)		1.5	1.5	1.5	1.5	−0.5	1.5	0.5	0.5	−0.5
	Levels[j]	Changes in percentages								
Hours worked per week	27.4	0.00	0.10	0.08	0.03	0.66	0.19	0.18	0.20	0.29
– Men in couples youngest child 0–17	35.8	−0.08	0.05	0.13	0.09	0.78	−0.04	0.04	0.07	0.53
– Women in couples youngest child 0–17	17.4	0.16	0.21	−0.03	−0.12	0.84	0.78	1.35	1.75	1.16
– Men in other couples	34.6	−0.07	0.02	0.08	0.07	0.55	−0.05	0.00	0.00	0.06
– Women in other couples	19.8	0.06	0.17	0.11	0.02	0.82	0.38	0.06	0.00	0.20
– Single parents youngest child 0–17	20.2	−0.04	0.21	0.15	0.05	2.07	0.37	0.70	0.28	0.56
– Singles	30.6	0.03	0.10	0.08	0.02	0.37	0.22	0.00	0.00	0.00
Participation rate	0.84	0.02	0.04	0.00	−0.02	0.64	0.26	0.17	0.16	0.26
– Men in couples youngest child 0–17	0.95	−0.04	0.06	0.04	0.00	0.80	0.14	0.11	0.09	0.46
– Women in couples youngest child 0–17	0.78	0.15	0.00	−0.17	−0.14	0.73	0.63	0.81	0.94	0.86
– Men in other couples	0.93	−0.04	0.02	0.04	0.03	0.52	0.04	0.00	0.00	0.05
– Women in other couples	0.72	0.07	0.07	0.01	−0.02	0.77	0.41	0.06	0.00	0.16

Table 3 (continued)

Simulation	Base	(1) First bracket[a]	(2) Second bracket[b]	(3) Third bracket[c]	(4) Fourth bracket[d]	(5) Welfare benefits[e]	(6) In-work tax credit[f]	(7) Combination credit[g]	(8) Childcare subsidies[h]	(9) Child benefit[i]
Ex-ante government spending (in billion euros)		1.5	1.5	1.5	1.5	−0.5	1.5	0.5	0.5	−0.5
– Single parents youngest child 0–17	0.70	−0.02	0.11	0.04	0.01	2.01	0.47	0.62	0.21	0.40
– Singles	0.88	0.01	0.05	0.04	0.01	0.33	0.17	0.00	0.00	0.00
Labour productivity per hour		−0.02	−0.01	0.03	0.06	−0.11	−0.11	−0.03	−0.02	−0.04
Knock-on effects[j]		−4	8	11	13	119	0	13	−41	52
Gini coefficient[k]	0.24	−0.22	0.30	0.87	1.33	0.78	−0.35	−0.01	0.01	0.43

[a]Decrease in bracket rate of −0.8%. [b]Decrease in bracket rate of −2.1%. [c]Decrease in bracket rate of −3.5%. [d]Decrease in bracket rate of −14%. [f]An increase in the in-work tax credit of 440 euros, by increasing the phase-in rate from 19.7 to 23.9%. The higher in-work tax credit is phased out from annual incomes of 34,000 euros onwards at 4%. [g]Increase in the combination credit for working secondary earners and working single parents with a youngest child of up to 12 years of age by 1,109 euros, by increasing the phase-in rate from 4 to 8%. [h]An increase in childcare subsidies for working parents, the change in childcare subsidies is set in such a way that there is a proportional decline of 38% in the parental contribution rate. [i]A decrease in the income-dependent child benefit for parents with a youngest child of up to 18 years of age. The subsidy is phased out from 19,676 euros at a rate of 6.75%. We decrease the subsidy by 45%, and keep the phase-out rate the same. [j]Knock-on effect as a percentage of the initial additional government expenditures, for people in the labour supply sample. [k]Gini coefficient of disposable household income, using equivalence scales. The Gini coefficient is calculated over the full Dutch adult population with a gross income of over 66% of the annual minimum wage. [j]In order to limit the size of the table, we do not present the standard errors of the policy simulations here. These standard errors are very small, ranging from 0.00 to 0.01

account behavioural responses. However, this is rather time consuming and does not change the relative effectiveness of the different policies. Finally, we report the effect on household income inequality (before behavioural changes), using the Gini coefficient.[45]

Column (1) gives the results for the decrease in the tax rate in the first tax bracket. Overall, hours worked and the participation rate hardly change. However, this is the net result of some groups that decrease and some groups that increase their labour supply. In particular, there is a modest decrease in hours worked by men in couples, due to an income effect (for them, the first tax bracket is typically inframarginal), and a modest increase in hours worked by women in couples, for them the substitution effect dominates the income effect. Income inequality, as measured by the Gini coefficient, decreases.

Column (2) gives the effect of lowering the tax rate in the second tax bracket. The effect on overall labour supply is positive but modest. Men in couples now also work a bit more due to the substitution effect, while the effect on women in couples is larger than under reform (1). When comparing the effects on total hours worked per week with those on the participation rate, most of the response comes from the intensive margin rather than the extensive margin. Income inequality increases somewhat, as the lowest incomes do not benefit from a lower second tax bracket rate.

Column (3) then considers the effects of a decrease in the third tax bracket rate. The increase in overall labour supply in hours is somewhat smaller than in reform (2), because of the smaller effect on hours worked by women. Indeed, although for some of these women the third tax bracket is the relevant marginal tax bracket, their own income effect and the income effect from a higher income of their partner dominates.[46] For single parents and singles, we find a positive effect on labour supply in hours, they do not have an income effect from a partner and the substitution effect of the lower marginal tax rate dominates. Income inequality increases more than under reform (2).

Finally, column (4) gives the effects of lowering the tax rate in the fourth (open) tax bracket. This has only a small positive effect on overall hours worked, and the effect on labour supply in persons is negative (due to the 'added worker effect'). But, where the increase in hours worked is much smaller under reform (4) than reform (3), labour productivity increases more due to a composition effect, workers with income in the fourth tax bracket are more productive. Also, because high-income individuals pay a relatively large amount of taxes, the knock-on effect for the government budget is higher than for reforms (1)–(3). Lowering the top rate leads to the biggest increase in inequality of reforms (1)–(4).

Participation tax rates Next, we consider policy reforms targeted more at the 'participation tax rate', the effective tax on the transition from non-employment to

---

[45]Where household income is weighted by the equivalence scales of Statistics Netherlands.

[46]The cross-effect of a higher income of men in couples on the labour supply of women in couples is an illustration of the 'added worker effect' (Lundberg 1985).

employment. Specifically, we consider lowering the participation tax rate through a 'carrot' or a 'stick'.

Column (5) gives the simulated effects of a reduction in welfare benefits by 14% (the stick), for a total amount of 500 million euros.[47] This leads to a substantial increase in overall labour supply, both in total hours worked and in persons, of +0.7% and +0.6% respectively. The effects are much larger than the reforms considered before, because welfare benefits operate on the extensive margin. The response is particularly large for single parents; 32% of single parents are on welfare benefits in the base.[48] The knock-on effects for the government are very high, because there is a sizeable reduction in the expenditures on welfare benefits due to behavioural responses. On the downside, this simulation causes a steep rise in income inequality.

In column (6) we use the 'carrot' instead, and consider an increase in the (general) in-work tax credit for a total amount of 1.5 billion euros, targeting the increase at low-income workers.[49] This also has a larger effect on total hours worked than reducing tax bracket rates because it is targeted at the extensive margin. Also on the upside, income inequality decreases, as the reform targets low income workers. Indeed, in this simulation, there is an increase in hours worked as well as a decrease in income inequality (see also Saez, 2002 for the potential welfare gains from in-work tax credits for low-income workers).[50] Furthermore, as most of the increase in participation is by women, this reform also causes a noticeable reduction in gender inequality. Indeed, Table 5 in the Appendix gives the change in average hours worked by men minus the average hours worked by women in couples for all the reforms in Table 3.[51] We find that the inequality in hours drops, because women are more responsive to the change in the in-work tax credit. However, on the downside, there is a sizable drop in average productivity, due to a change in the workforce composition, and the knock-on effect is also close to zero.

---

[47]In 2015, the year of the simulations, the working age population in the Netherlands was 12.6 million persons, of which 8.2 million were working and 436 thousand were on welfare benefits.

[48]Fortin et al. (2004) also finds substantial effects of the benefit level in welfare on the duration of welfare spells for singles in Canada.

[49]Specifically, we increase the maximum in-work tax credit by 441 euros, and start the phase-out at a rate of −4% as in the base, but from an annual income of 34,000 euros as opposed to 49,770 euros in the base.

[50]Several studies for the US have also shown that an EITC may stimulate labour supply of single parents (Hoynes and Patel 2018), although it may deter labour supply of secondary earners that are in the phase-out range of the EITC (Eissa and Hoynes 2004). The latter is because the EITC in the US depends on family income. The general in-work credit in the Netherlands depends on individual income, and hence stimulates participation of secondary earners. Annabi et al. (2013) study the introduction of the Working Income Tax Benefit (WITB) in 2007 in Canada. The WITB substantially reduces effective marginal tax rates for low incomes. Similar to our study they find that this stimulates employment, in particular of single parents, and reduces income inequality as measured by the Gini-coefficient.

[51]We do not distinguish between single men and single women. We first estimated preferences for single men and single women, and found that their preferences were very similar, and subsequently grouped them together in the estimation of their preferences. Furthermore, we do not distinguish between single fathers and single mothers, because the group of single fathers is too small to reliably estimate their preferences. Hence, we look at inequality in hours worked in couples.

Subsidies for households with young children Finally, we consider a number of reforms targeted at households with young children.[52] This group is of particular interest because there are many policies targeted at this group, and because mothers with young children appear particularly responsive to changes in financial incentives. In these simulations, we use a smaller increase in government expenditures than before, because these reforms target only a subgroup of the working age population.

In column (7), we increase the income-dependent part of the in-work tax credit for secondary earners and single parents with a youngest child of up to 12 years of age.[53] This leads to a substantial increase in the number of hours worked given the budgetary impulse, because it targets the groups with the highest labour supply elasticity. Since the response is mostly by women, this reform also causes a substantial decline in gender inequality, see also Table 5 in the Appendix. However, the knock-on effects are still limited, as secondary earners and working single parents with a young child already receive large subsidies in the base.

In column (8), we increase childcare subsidies. In particular, we consider a proportional decrease of 38% in the parental contribution that results after deducting the subsidy from the full hourly price. Again, there is a substantial increase in hours worked. The effect on total hours worked is somewhat larger than for reform (7), and also reduces gender inequality in hours worked more (again, see Table 5 in the Appendix). However, the childcare reform also leads to substantial substitution of informal care for formal care. As a result, the knock-on effect for the government budget is negative, making this reform less cost-effective than reform (7).

Moving from carrots to sticks, in column (9), we decrease the income-dependent child benefit for parents with young children.[54] This also leads to a substantial increase in hours worked and in the participation rate, and a decrease in gender inequality, in particular for couples with young children (Table 5 in the Appendix).[55] However, again, the downside of this 'stick' reform is that it increases income inequality, as we take benefits away from low-income households.

7.2 Major tax reforms

After considering changes in single policy instruments, we now consider a number of major tax reforms that feature prominently in the current policy debate, in the Nether-

---

[52]In 2015, the year of the simulations, there were 1.3 million households with a child younger than 12 years old, the potential target group for the combination credit and child care subsidies, and 1.9 million households with a child younger than 18 years old, the potential target group for child benefits.

[53] 'Inkomensafhankelijke Combinatiekorting' in Dutch, which depends on income, with a phase-in rate of 4% for income above the threshold of 4857 euros, reaching its maximum of 1119 euros at a personal income level of 32,832 euros. We increase the maximum credit level by 1100 euros, via an increase in the phase-in rate of 4 percentage points.

[54]'Kindgebonden Budget' in Dutch, this benefit is a fixed amount, which however depends on the number and ages of the children, up to a gross annual household income of 19,676 euros, after which it is phased out at a rate of 6.75% in the base. We decrease the maximum amount for all families by 45% and keep the phase-out rate the same.

[55]Note that there is also a small effect on men and women in other couples, these are the men and women in couples with a partner whose labour supply is fixed, but who have a dependent child.

lands and abroad. Specifically, we consider the introduction of a flat tax system, a basic income system, and a move towards joint rather than individual taxation.[56] The simulation results are given in Table 4.

In column (1), we change the four tax brackets rates in the baseline, 36.5%, 42%, 42%, and 52%, respectively, to a flat tax rate of 39.7%. This scenario is budgetary neutral before behavioural changes. We see that this flat tax increases overall labour supply. Men in couples and singles without children increase their labour supply, because the flat tax rate is lower than their initial marginal tax bracket rate. For women in couples, there are two opposing effects. First, women paying taxes in the first tax bracket now face a higher marginal tax rate and some women withdraw from the labour force. Second, women paying taxes in the second and higher tax bracket now face a lower marginal tax rate, triggering a positive response at the intensive margin. For women in couples, with children aged 0–17, the first effect dominates, whereas for women in other couples the second effect dominates. Finally, single parents increase their labour supply, which is caused by the sharp drop in the net welfare benefit due to the increase in the first bracket rate. The downside of this flat tax proposal is a substantial increase in income inequality. Furthermore, gender inequality in hours worked in couples increases, see Table 6 in the Appendix.

In column (2), we consider a flat tax system that is budgetary neutral *and* 'Gini neutral'. Specifically, we introduce a lump-sum subsidy for all adults of 1.950 euros and finance this with a flat tax rate of 45.3%. We find that to arrive at the same income inequality as in the baseline, the flat tax reduces labour supply.[57] Indeed, the flat tax increases marginal and participation tax rates at the lower end of the income distribution, which is more responsive to tax changes in terms of hours worked than the upper end of the income distribution.[58] Gender inequality in hours worked in couples increases further (Table 6).

Column (3) shows that generic income support via the introduction of an unconditional basic income has a strong adverse effect on labour force participation. We simulate a basic income of 50% of the net welfare benefit level. All adults qualify for this basic income. For adults receiving social benefits (e.g. welfare, unemployment, disability or retirement benefits) we reduce the benefit level, so that together with the basic income their disposable income does not change. We finance the basic income scenario by abolishing the general in-work tax credit for all workers, which is in line with the idea that this type of income support is unconditional, and a flat tax rate of 56.6%.[59] The flat tax rate of 56.6% implies a considerable increase in

---

[56] CPB used the simulation model MICSIM to evaluate the government's Tax Plan for 2016 (Ministry of Finance 2015). The aim of the Tax Plan was to create more employment by reducing the tax burden on labour with 5 billion euros. Important policy measures in the tax plan concerned a higher tax credit for all workers, a higher tax credit for working parents and a higher child care subsidy. There has been much debate among politicians and in the media about the employment effects of the new Tax Plan, which came into effect on 1 January 2016. Our model predicts that labor supply increases with 35,000 full-time equivalents (FTEs).

[57] A point made earlier by Jacobs et al. (2009).

[58] Simulation studies for other countries find that a flat tax may stimulate labour supply, but only when marginal tax rates and basic allowances are set relatively low (Aaberge et al. 2000; Fuest et al. 2008).

[59] When the general in-work tax credit is not abolished, we still obtain a strong negative effect on labour supply.

Table 4 Major tax reforms[a]

	Base	(1) Flat tax[b]	(2) Flat tax, Gini neutral[c]	(3) Basic income[d]	(4) Joint taxation[e]
Simulation					
Ex-ante government spending (in billion euros)		0.0	0.0	0.0	0.0
	Levels[f]	Changes in percentages			
Hours worked per week	27.4	0.24	−0.90	−5.60	−2.23
– Men in couples youngest child 0–17	35.8	0.67	−0.27	−3.30	0.14
–Women in couples youngest child 0–17	17.4	−0.71	−4.79	−21.96	−16.65
– Men in other couples	34.6	0.48	−0.02	−2.59	0.00
– Women in other couples	19.8	0.11	−1.76	−8.70	−3.19
– Single parents youngest child 0–17	20.2	0.62	−0.08	−2.25	−0.76
– Singles	30.6	0.12	−0.39	−2.71	−0.32
Participation rate	0.84	−0.07	−0.61	−4.67	−2.78
– Men in couples youngest child 0–17	0.95	0.25	−1.47	−2.94	−0.48
–Women in couples youngest child 0–17	0.78	−1.06	−0.37	−16.65	−14.18
– Men in other couples	0.93	0.26	−0.62	−2.04	−0.10
– Women in other couples	0.72	−0.22	−0.64	−6.70	−3.00
– Single parents youngest child 0–17	0.70	0.24	−0.84	0.24	−0.60
– Singles	0.88	0.09	−0.29	−0.57	−0.19
Labour productivity per hour		0.19	0.25	0.62	0.37
Gini coefficient[g]	0.24	5.18	0.00	−7.95	0.76

[a]Knock-on effects are not presented for budgetary neutral scenarios, as these effects are expressed as a percentage of the initial budgetary impulse which equals zero in the first four scenarios. [b]A flat tax rate of 39.7%. [c] A flat tax rate of 45.3% and a lump-sum benefit of 1.950 euros for all adults. [d]An introduction of an unconditional basic income of 50% of the net welfare benefit (8248 euros) and a flat tax rate of 56.5%. [e]Joint taxation, financed through an increase of 1.9 percentage points in the tax rates in the four tax brackets. [f]In order to limit the size of the table, we do not present the standard errors of the policy simulations here. These standard errors are very small, ranging from 0.00 to 0.01. [g]Gini coefficient of disposable household income, using equivalence scales. The Gini coefficient is calculated over the full Dutch adult population with a gross income of over 66% of the annual minimum wage

all marginal tax rates, in particular for adults in the lower tax brackets. Furthermore, the introduction of a basic income increases income for non-working partners. Labour supply decreases both on the extensive and intensive margin, and in total by

−5.6%.[60] There is a dramatic increase in gender inequality in terms of hours worked (Table 6). However, on the upside, income inequality decreases, by almost 8%.

Finally, column (4) simulates a move from individual to joint taxation, as in the tax systems of, e.g. France, Germany, and the USA.[61] We simulate joint taxation by taking the sum of taxable income from both partners and then assign half of the total taxable household income to both partners. We finance this scenario by increasing marginal tax rates in all four tax brackets by 1.9 percentage points. Total labour supply decreases by −2.2%. Most women in couples are secondary earners and face a relatively low marginal tax rate under the tax system in the baseline. Joint taxation means that the marginal tax rate increases for secondary earners and they reduce their labour supply. The effective marginal tax rate for primary earners decreases, and they increase their labour supply, but to a much lesser extent. Furthermore, income inequality increases, and there is again a rather dramatic increase in gender inequality in terms of hours worked (Table 6). Hence, this scenario scores unfavourably in terms of hours worked, income inequality and gender inequality.[62]

8 Discussion and conclusion

In this paper, we used both structural models and quasi-experimental studies to study the effectiveness of tax-benefit reforms. Using a very large and rich data set, we estimate structural discrete-choice models for a large number of household types. We uncover large differences in the labour supply responses between various demographic groups, mostly related to the age of the youngest child. We also find that the decision of whether or not to participate is more responsive to financial incentives than the hours-per-week decision, although the hours-per-week decision is still non-negligible for women in couples with children. We used the structural model to simulate a number of key reforms from the past, and compared the predictions of the structural model with the outcomes of quasi-experimental studies on the same reforms. We find that the structural model predicts the estimated treatment effects from the quasi-experimental studies rather well.

We then conduct a counterfactual policy analysis with the structural model, and study the effectiveness of potential tax-benefit reforms in stimulating labour supply. We find that reducing marginal tax rates is not an effective way to promote labour

---

[60]Clavet et al. (2013) and Horstschraer et al. (2010) also use a structural labour supply model to study the labour supply effects of a basic income, for Quebec (Canada) and Germany, respectively. Both studies find a strong decrease in total hours worked. Our results are also consistent with the findings from the basic income experiments in the US and Canada. Hum and Simpson (1993, Table 2) gives an overview. For the United States, hours worked by men in couples, by women in couples, and by single women declined by −5 to −7%, −17 to −21%, and −13 to −17%, respectively. For Canada, hours worked by men in couples, by women in couples, and by single women declined by a respective −1%, −3%, and −7%. For a recent assessment of the potential of universal basic incomes in the USA and other advanced countries, see Hoynes and Rothstein (2019).

[61]In the Netherlands, the Reformed Political Party (SGP), in their election manifesto, proposed a tax system which would allow incomes to be split.

[62]Kabatek et al. (2014) estimate a discrete-choice labour supply model for French couples to evaluate a shift from joint taxation to individual taxation. They find that a system with joint taxation discourages female labour supply. Callan et al. (2009) estimate a discrete-choice model for couples, using Irish data. They find that a shift from joint taxation to individual taxation stimulates women's labour supply, also consistent with our findings.

supply. In-work benefits targeted at low-wage earners are more effective. Policies targeted at working mothers with young children generate the largest labour supply response, reduce gender inequality, but generate little additional revenue for the government. With the structural model we also simulate some major tax reforms that feature prominently in the current policy debate. We find that proposals for a move to a flat tax system, a basic income system or a system with joint taxation are not effective in stimulating labour supply, and cause a steep rise in gender inequality. Indeed, an efficient tax system accounts for the large heterogeneity in responses, between different demographic groups (e.g. primary vs. secondary earners and with vs. without young children) and different decision margins (e.g. extensive vs. intensive margin), and therefore cannot be too simple.

Although we believe that our analysis makes a number of improvements over previous studies on the effectiveness of tax-benefit reforms, it still has a number of limitations. We ignore involuntary unemployment (and a potential difference between preferred and actual working hours). However, estimating a double-hurdle model (Cragg 1971), we find that accounting for involuntary unemployment makes little difference in the employment responses to changes in financial incentives (De Boer 2018).[63] Furthermore, we ignore responses to marginal (and participation) tax rates other than labour supply. Part of the modern literature on public finance looks at a broader range of behavioural responses, by considering the so-called elasticity of taxable income, see Saez et al. (2012) for an overview. Indeed, a recent study by Jongen and Stoel (2019) for the Netherlands shows that the elasticity-of-taxable-income may be higher than the labour supply elasticity, suggesting larger distortions from tax rates than by looking solely at labour supply. We further ignore the life cycle. A number of studies have shown that accounting for life-cycle effects can be important for the analysis of tax-benefit reform (e.g. Imai and Keane 2004, 2011; Blundell et al. 2016). This would be an interesting direction for future research. However, the data set we used does not include data on, for example, consumption or savings, which makes it difficult to estimate a life-cycle model, and it should be noted that there is often a trade-off in modelling different parts of economic behaviour, due to the numerical complexities that arise.[64] Finally, we assume that all people are fully aware of their budget constraint. However, recent work by Chetty et al. (2009) shows that information, or the lack thereof, can play an important role in the behavioural responses to financial incentives. This too seems an interesting direction for future research.[65]

Appendix: Effects reforms on gender gaps in hours worked

---

[63] Regarding the difference between actual and preferred hours of work, we do not have data on preferred hours of work in our data set. 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 women working part-time in the Netherlands would like to work more hours, compared to, for example, 13% in Germany, 28% in France, and 55% in Spain.

[64] Estimating the preferences for all subgroups in our static model already took several weeks.

[65] See, e.g. Bosch et al. (2019) for a bunching analysis of high effective marginal tax rates resulting from income support in the Netherlands. They find rather limited bunching, which may be due to inattention or frictions.

Table 5 Changes in selected tax-benefit policies: gender gap in hours worked couples[a]

Simulation	Base	(1) First bracket	(2) Second bracket	(3) Third bracket	(4) Fourth bracket	(5) Welfare benefits	(6) In-work tax credit	(7) Combination credit	(8) Childcare subsidies	(9) Child benefit
Ex-ante government spending (in billion euros)		1.5	1.5	1.5	1.5	−0.5	1.5	0.5	0.5	−0.5
	Levels	Changes in percentages								
Couples youngest child 0–3	18.7	−0.31	−0.05	0.34	0.25	0.76	−0.90	−1.72	−3.82	−0.34
Couples youngest child 4–11	19.5	−0.30	−0.07	0.28	0.32	0.84	−0.82	−1.34	−0.54	−0.02
Couples youngest child 12–17	16.4	−0.29	−0.20	0.18	0.30	0.44	−0.71	−0.21	0.00	0.20
Couples youngest child 18+	16.7	−0.38	−0.29	−0.03	0.07	0.57	−0.86	0.00	0.00	0.03
Couples without children	19.8	−0.34	−0.25	0.11	0.25	0.82	−0.63	0.00	0.00	0.00

[a] Average hours worked by the men minus average hours worked by the women

Table 6 Major tax reforms: gender gap in hours worked couples[a]

Simulation	Base	(1) Flat tax	(2) Flat tax, Gini neutral	(3) Basic income	(4) Joint taxation
Ex-ante government spending (in billion euros)		0.0	0.0	0.0	0.0
	Levels	Changes in percentages			
Couples youngest child 0–3	18.7	1.94	2.44	21.10	20.98
Couples youngest child 4–11	19.5	2.09	2.31	12.67	16.12
Couples youngest child 12–17	16.4	1.87	3.09	7.83	9.00
Couples youngest child 18+	16.7	1.30	3.88	3.73	5.14
Couples without children	19.8	1.49	3.06	6.03	5.58

[a] Average hours worked by the men minus average hours worked by the women

Declarations

References

Aaberge R, Colombino U, Strom S (1999) Labour supply in Italy: an empirical analysis of joint household decisions, with taxes and quantity constraints. J Appl Econ 14(4):403–422

Aaberge R, Colombino U, Strom S (2000) Labor supply responses and welfare effects from replacing current tax rules by a flat tax: empirical evidence from Italy, Norway and Sweden. J Popul Econ 13(4):595–621

Aaberge R, Dagsvik J, Strom S (1995) Labor supply responses and welfare effects of tax reforms. Scand J Econ 97(4):635–659

Angrist J, Pischke J.-S. (2009) Mostly harmless econometrics: an empiricist's companion. Princeton University Press, Princeton

Annabi N, Boudribila Y, Harvey S (2013) Labour supply and income distribution effects of the working income tax benefit: a general equilibrium microsimulation analysis. IZA Journal of Labor Policy 2(19)

Attanasio O, Meghir C, Santiago A (2011) Education choices in Mexico: using a structural model and a randomized experiment to evaluate PROGRESA. Rev Econ Stud 79(1):37–66

Bargain O, Doorley K (2017) The effect of social benefits on youth unemployment: combining regression discontinuity and a behavioral model. J Hum Resour 52(4):1032–1059

Bargain O, Orsini K, Peichl A (2014) Comparing labor supply elasticities in Europe and the United States: new results. J Hum Resour 49(3):723–838

Bargain O, Peichl A (2016) Steady-state labor supply elasticities: an international comparison. IZA J Labor Econ 5:1–31

Bettendorf L, Folmer K, Jongen E (2014) The dog that did not bark: the EITC for single mothers in the Netherlands. J Public Econ 119:49–60

Bettendorf L, Jongen E, Muller P (2015) Childcare subsidies and labour supply - evidence from a large Dutch reform. Labour Econ 36:112–123

Blau D (2003) Child care subsidy programs. In: Moffitt R (ed) Means-tested transfer programs in the United States. NBER, Cambridge, pp 443–516

Blau D, Kahn L (2007) Changes in the labor supply behavior of married women: 1980-2000. J Labor Econ 25(3):393–438

Blundell R, Chiappori P.-A., Meghir C (2007) Collective labour supply: heterogeneity and non-participation. Rev Econ Stud 74:417–455

Blundell R, Costa-Dias M, Meghir C, Shaw J (2016) Female labour supply, human capital and welfare reform. Econometrica 84(5):1705–1763

Blundell R, Duncan A, McCrae J, Meghir C (2000) The labour market impact of the Working Families' Tax Credit. Fisc Stud 21(1):75–104

Blundell R, Duncan A, Meghir C (1998) Estimating labor supply responses using tax reforms. Econometrica 66:827–861

Blundell R, MaCurdy T (1999) Labor supply: a review of alternative approaches. In: Ashenfelter O, Card D (eds) Handbook of Labor Economics, vol 3. Elsevier, Amsterdam, pp 1559–1695

Blundell R, Shephard A (2012) Employment, hours of work and the optimal taxation of low income families. Rev Econ Stud 79(2):481–510

Bosch N, Jongen E (2013) Intensive margin responses when workers are free the choose: evidence from a Dutch tax reform. Paper presented at EALE

Bosch N, Jongen E, Leenders W, Mohlmann J (2019) Non-bunching at kinks and notches in cash tranfers in the Netherlands. Int Tax Public Financ 26(6):1329–1352

Bosch N, van der Klaauw B (2012) Analyzing female labor supply - evidence from a Dutch tax reform. Labour Econ 19:271–280

Brewer M, Duncan A, Shephard A, Suarez M (2006) Did the Working Families' Tax Credit work? The impact of in-work support on labour supply in Great Britain. Labour Econ 13:699–720

Brewer M, Saez E, Shephard A (2010) Means-testing and tax rates on earnings. In: Mirrlees J. A., Adam S., Besley T. J., Blundell R, Bond S, Chote R, Gammie M, Johnson P, Myles GD, Poterba JM (eds) The mirrlees review – dimensions of tax design, chapter 3. Oxford University Press, Oxford, pp 202–274

Cai L, Kalb G, Tensg Y.-P., Vu H (2008) The effect of financial incentives on labour supply: Evidence for lone parents from microsimulation and quasi-experimental evaluation. Fisc Stud 29(2):285–325

Callan T, Van Soest A, Walsh J (2009) Tax structure and female labour supply: evidence from Ireland. Labour 23(1):1–35

Chetty R (2009) Sufficient statistics for welfare analysis: a bridge between structural and reduced-form methods. Ann Rev Econ 1:451–488

Chetty R, Looney A, Kroft K (2009) Salience and taxation: theory and evidence. Am Econ Rev 99(4):1145–1177

Clavet N, Duclos J, Lacroix G (2013) Fighting poverty: Assessing the effect of guaranteed minimum income proposals in Quebec. Canadian Public Policy 39(4):491–516

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 (2020) Analysing tax-benefit reforms in the Netherlands using structural models and natural experiments. IZA Discussion Paper 12892, Bonn

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

Eissa N, Hoynes H (2004) Taxes and the labor market participation of married couples: the Earned Income Tax Credit. J Public Econ 88(9-10):1931–1958

Ericson P, Flood L (2012) A microsimulation approach to an optimal Swedish income tax. Int J Microsimulation 5(2):2–21

European Union (2014) Taxation trends in the European union. European Union, Brussels

Euwals R, Knoef M, van Vuuren D (2011) The trend in female labour force participation: What can be expected for the future? Empir Econ 40:729–753

Fortin B, Lacroix G, Drolet S (2004) Welfare benefits and the duration of welfare spells: evidence from a natural experiment in Canada. J Public Econ 88(7-8):1495–14520

Fuest C, Peichl A, Schaefer T (2008) Is a flat tax reform feasible in a grown-up democracy of Western Europe? A simulation study for Germany. Int Tax Public Financ 15:620–636

Geyer J, Haan P, Wrohlich K (2015) The effects of family policy on mothers' labor supply: combining evidence from a structural model and a quasi-experimental approach. Labour Econ 36(C):84–98

Hansen J, Liu X (2015) Estimating labor supply responses and welfare participation: using a natural experiment to validate a structural model. Can J Econ 48(5):1831–1854

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

Heckman J (2010) Building bridges between structural and program evaluation approaches to evaluating policy. J Econ Lit 48(2):356–398

Heim B (2007) The incredible shrinking elasticities: married female labor supply, 1978-2002. J Hum Resour 42(4):881–918

Horstschraer J, Claus M, Schnabel R (2010) An unconditional basic income in the family context: labor supply and distributional effects. ZEW Discussion Paper No. 10-091, Munich

Hoynes H (1996) Welfare transfers in two-parent families: labor supply and welfare participation under AFDC-UP. Econometrica (64):295–332

Hoynes H, Patel A (2018) Effective policy for reducing poverty and inequality? The Earned Income Tax Credit and the distribution of income. J Hum Resour (53):859–890

Hoynes H, Rothstein J (2019) Universal basic income in the United States and advanced countries. Annual Rev Econ (1): 929–985

Hum D, Simpson W (1993) Economic response to a guaranteed annual income: experience from Canada and the United States. J Labor Econ 11(1):S263–S296

Imai S, Keane M (2004) Intertemporal labor supply and human capital accumulation. Int Econ Rev 45(2):601–641

Jacobs B, De Mooij R, Folmer K (2009) Analyzing a flat tax in the Netherlands. Appl Econ 42(25):3209–3320

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

Kabatek J, Van Soest A, Stancanelli E (2014) Income taxation, labour supply, and housework: a discrete choice model for French couples. Labour Econ 27:30–43

Keane M (2010) Structural vs. atheoretic approaches to econometrics. J Econ 156:3–20

Keane M (2011) Labor supply and taxes: a survey. J Econ Lit 49(4):961–1075

Keane M, Moffitt R (1998) A structural model of multiple welfare program participation and labor supply. Int Econ Rev 39(3):553–589

Koot P, Vlekke M, Berkhout E, Euwals R (2016) MIMOSI: Microsimulatiemodel voor belastingen, sociale zekerheid loonkosten en koopkracht. CPB Background Document, The Hague

Lokshin M (2004) Household childcare choices and women's work behavior in Russia. J Hum Resour 39(4):1094–1115

Lundberg S (1985) The added worker effect. J Labor Econ 3(1):11–37

MaCurdy T, Green P, Paarsch H (1990) Assessing empirical approaches for analyzing taxes and labor supply. J Hum Resour 25:415–490

McFadden D (1978) Modeling the choice of residential location. In: Karlqvist A, Lundqvist L, Snickars F, Weibull J (eds) Spatial interaction theory and planning models. North-Holland, Amsterdam, pp 75–96

Ministry of Finance (2015) Belastingherziening. Ministry of Finance, The Hague

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

OECD (2016) OECD Labour force statistics 2016. OECD, Paris

Pronzato C (2015) Fighting lone mothers' poverty through in-work benefits: methodological issues and policy suggestions. CESifo Economic Studies 61(1):95–122

Saez E (2002) Optimal income transfer programs: intensive versus extensive labor supply responses. Q J Econ 117(3):1039–1073

Saez E, Slemrod J, Giertz S (2012) The elasticity of taxable income with respect to marginal tax rates: a critical review. J Econ Lit 2012 50:1:3–50

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

Statistics Netherlands (2018) Monitor arbeid, zorg en kinderopvang. The Hague

Thoresen T, Vatto T (2015) Validation of the discrete choice labor supply model by methods of the new tax responsiveness literature. Labour Econ 37(C):38–53

Todd P, Wolpin K (2006) Assessing the impact of a school subsidy program in Mexico: using a social experiment to validate a dynamic behavioral model of child schooling and fertility. Am Econ Rev 96(5):1384–1417

Train K (2003) Discrete choice methods with simulation. Cambridge University Press, Cambridge

Van Rens C, Smit F (2011) Wachtlijsten en wachttijden kinderdagverblijven en buitenschoolse opvang - 6e meting ITS. Radboud University, Nijmegen

Van Soest A (1995) Structural models of family labor supply: a discrete choice approach. J Hum Resour 30(1):63–88

Van Soest A, Das M (2001) Family labour supply and proposed tax reforms in the Netherlands. De Economist 149:191–218

Van Soest A, Das M, Gong X (2002) A structural labor supply model with flexible preferences. J Econometrics 107(1-2):345–374