The effectiveness of fiscal stimuli for working parents

Boer, H.W. de; Jongen, E.L.W.; Kabatek, J.

Bron: Boer, H.-W. de, Jongen, E. and J. Kabátek, 2022, The effectiveness of fiscal stimuli for working parents, *Labour Economics*, forthcoming

The Effectiveness of Fiscal Stimuli for Working Parents*

Henk-Wim de Boer† Egbert L.W. Jongen‡ Jan Kabatek§

March 2022

Abstract

Child care subsidies and in-work benefits are prominent policies used to promote the labor participation of parents. We study these policies in a structural model of labor supply and child care use for couples in the Netherlands. Major reforms in family policies benefit the identification. We use differences-in-differences to assess the reliability of the model predictions. In-work benefits for secondary earners that increase with income are shown to be the most cost-effective tool for stimulating parental labor supply. Child care subsidies are less effective, as substitution of informal for formal care drives up public expenditures. We also relate our findings to related policies in the UK and US.

JEL classification codes: C21, C25, C52, H31, J22
Keywords: Household labor supply, structural models, differences-in-differences, work and care policies

*We have benefitted from comments and suggestions by the editor Wilbert van der Klaauw, two anonymous referees, Rolf Aaberge, Hans Bloemen, Tom Kornstad, Magne Mogstad, Andrew Shephard, Arthur van Soest, Thor Thoresen, Trine Vatto and participants at numerous conferences and seminars. The usual disclaimer applies.

†CPB Netherlands Bureau for Economic Policy Analysis. Corresponding author. CPB, P.O. Box 80510, 2508 GM The Hague. Phone: +31-6-41603330. Email: h.w.de.boer@cpb.nl.

‡CPB Netherlands Bureau for Economic Policy Analysis, Leiden University and IZA. Email: e.l.w.jongen@cpb.nl.

§Melbourne Institute of Applied Economic and Social Research, University of Melbourne, LCC, IZA and Netspar. Email: j.kabatek@unimelb.edu.au.

1 Introduction

One of the central goals of family tax systems around the world is to encourage and support parental employment. This is evidenced by large endowments of fiscal instruments that are designed to do so. However, given the prominence of these instruments in national fiscal systems, it is interesting to note that there are large cross-country differences in the actual mix of fiscal policies targeted at working parents. For example, Scandinavian countries direct much of their public support to child care subsidies (Kleven, 2014; OECD, 2015b), whereas Canada and the US rely more on in-work benefits (Immervoll and Pearson, 2009).[1] And although these policies may differ to some extent in their objectives, a common goal is to stimulate employment of parents with young children.[2]

A large body of literature studies the employment effects of child care subsidies (and related programs like pre-kindergarten and pre-school)[3], and a similarly large body of literature studies the employment effects of in-work benefits for families with children.[4] These two policy instruments are typically shown to have a positive effect on parental employment, However little is known about their relative effectiveness in terms of additional employment per dollar or euro spent. Cross-country comparisons of these policies are inevitably obfuscated by differences in the respective institutional and cultural settings, and within-country comparisons can be distorted by differences in timing of the reforms. It is therefore difficult to infer which policy works best for stimulating parental employment. This exercise is complicated further by the targeting of these policies, which is also highly country-specific.[5] Indeed, the policy responses may be heavily influenced by targeting,

[1] Of note, the Canadian province of Quebec introduced '5 dollar per day' daycare for young children at the end of the 1990s/early 2000s (see e.g. Baker et al., 2008; Lefebvre and Merrigan, 2008), and the Biden administration is planning to substantially reduce the parental fee for child care for low- and middle-income families in the US, see e.g. `https://www.whitehouse.gov/briefing-room/statements-releases/2021/04/28/fact-sheet-the-american-families-plan`.

[2] Another goal of child care subsidies may be to promote skill formation among disadvantaged children, whereas another goal for in-work benefits may be to provide income support for low-income families.

[3] See Blau (2003) for an overview of the earlier literature, and e.g. Lokshin (2004), Tekin (2007), Baker et al. (2008), Lefebvre and Merrigan (2008), Cascio (2009), Havnes and Mogstad (2011) and Fitzpatrick (2012) for some more recent analyses. Bettendorf et al. (2015) also provides an overview of more recent studies.

[4] Two major in-work benefit programs that have received much attention in the literature are the Earned Income Tax Credit (EITC) in the US and the Working Families' Tax Credit (WFTC) in the UK. See e.g. Brewer and Browne (2006), Meyer (2010) and Brewer and Hoynes (2019) and the references therein for empirical studies on the impact of the EITC and WFTC on employment, respectively. In Section 4 we relate our findings on the reforms we consider for the Netherlands to the setup of and empirical findings for the EITC and the WFTC.

[5] For example, in-work benefits in the US and the UK are targeted primarily at low-income families (Brewer et al., 2009), whereas in-work benefits in the Netherlands are targeted more at middle- and high-income families (see below).

and the optimal targeting strategy is likely to depend on the relative importance of labor supply responses on the extensive margin (participation) and intensive margin (hours worked per employed) (Saez, 2002).

In this paper, we offer a systematic analysis of child care subsidies and in-work benefits for couples with children, assessing their effectiveness in stimulating couples' labor supply.[6] Leveraging a structural model and large reforms of the Dutch family tax system, we quantify the public spending required per additional parent employed for several considered policy instruments. In doing so, we take into account parents' behavioral responses to the policy instruments. Furthermore, we consider to what extent the targeting of these instruments at different income groups matters for their effectiveness. In this way, we study the equity-efficiency trade-off of family tax policies. We also consider how the effects of in-work credits based on individual earnings differ from those based on joint household earnings.

Our analysis rests on a structural discrete choice model of household labor supply and child care use in the Netherlands.[7] We use a large and rich administrative household data set, the Labor Market Panel of Statistics Netherlands (2012), for the period 2006–2009 to estimate the couples' work and child care preferences. We distinguish between the couples with the youngest child aged 0–3 (pre-primary school age) and couples with the youngest child aged 4–11 (primary school age). The model allows for the simultaneous choice of labor supply by the mother and the father, and the use of formal child care.[8] In our preferred specification, we model unobserved heterogeneity using the random preference approach of Van Soest (1995). As a robustness check, we also estimate a more flexible specification of our model using the latent-class framework (Train, 2008; Pacifico, 2013;

[6]The Dutch tax and benefit system contains other benefits and in-work credits that are targeted at lone parents. For an analysis of the effects of these policies on the labor supply of lone parents, see De Boer and Jongen (2017).

[7]Building on the work by Van Soest (1995), discrete choice models have become a popular tool for the structural modelling of labor supply, see e.g. Keane and Moffitt (1998), Blundell et al. (2000), Gong and Van Soest (2002), Blundell and Shephard (2012) and Bargain et al. (2014).

[8]For an overview of structural models that explicitly include child care see Blau (2003). Recent applications include Lokshin (2004), Kornstad and Thoresen (2006, 2007), Tekin (2007), Blundell and Shephard (2012), Gong and Breunig (2017), Apps et al. (2016) and Thoresen and Vatto (2019). Closest to our specification is the model of Thoresen and Vatto (2019) who estimate a joint model of parental labor supply and formal child care choice using Norwegian data. While we adopt a similar modeling strategy, the two papers differ substantively in their focus: Thoresen and Vatto (2019) explore the active role of fathers as informal caregivers, advocating for joint modeling of parental labor supply decisions; in accord with the authors, we embrace this modeling framework and use it to evaluate the effects of various fiscal stimuli for working parents. Our data is superior in this regard, spanning a period of large-scale family tax reforms which aid identification of the model parameters. We also extend the analysis in Kurowska et al. (2017) who consider potential reforms to support for couples in Poland, including a 'double earner' premium. Specifically, since we have high quality data on child care use and hours worked, we can also study child care reform and hours worked in reforms targeted at dual earners.

Apps et al., 2016). The identification of the structural parameters benefits from a major reform of child care subsidies and in-work benefits for working parents that took place in the period of observation, generating large exogenous variation in the couples' budget sets. We also compare the predictions for the reform effects to the results of a differences-in-differences analysis. Furthermore, we consider an out-of-sample prediction, withholding the data for 2009 from the estimation of the structural parameters, and predicting the corresponding labor supply and childcare responses using our model. Finally, we also consider the contribution of the reform to the overall increase in the participation rate and hours worked of mothers and fathers in couples with young children.

Our main findings are as follows. First, we find that the most effective fiscal stimulus for working parents is an in-work benefit targeted at secondary earners that rises with their individual income. This policy provides incentives both on the extensive and the intensive margins of labor supply, and it is targeted at workers who are relatively elastic in their employment choices (mostly women with young children). Second, we find that child care subsidies are less effective than in-work benefits for secondary earners. The underlying reason is that childcare subsidies incentivize parents to substitute away from other, informal types of child care arrangements. This substitution drives up public expenditures, since the parents who switch child care types are often already employed. However, child care subsidies are still much more effective than in-work benefits for both primary and secondary earners, because the labor supply of primary earners is rather unresponsive to financial incentives. Third, we find that the effect of child care subsidies is not much lower when targeted at low-income households, compared to the policy targeted at middle- and high-income households. However, the knock-on effects (i.e., changes in public expenditures and receipts due to workers' behavioral changes) are more favorable for the policy targeting the middle and higher income families, which is attributable to their higher tax burdens. So in the end, there is still a trade-off between equity and efficiency when it comes to the targeting of fiscal stimuli for working parents. We also show that implementing an in-work credit based on household as opposed to individual income would be detrimental to labor supply in the Netherlands, in particular for women with young children.

We make several contributions to the existing literature. We have a large and rich data set, combining high-quality administrative data, including working hours, hours of formal child care use, with relevant survey data on e.g. education and job search, for several years in which there was a large reform of child care subsidies an in-work benefits for families with young children. Hence, we can split the sample into subgroups, and estimate the model for families with pre-school children, and families whose children are in primary school, and allow for a rich set of interactions of preferences. This allows our model to capture differences by many demographic dimensions, like age of the youngest child, education and migration background. The reform also provides us with credible exogenous variation in budget sets, in contrast to previous structural analyses of labor supply and child care that mostly relied on cross-sectional variation. Furthermore, we can also relate the simulated reform effects of the structural model with to a differences-in-differences analysis (Bettendorf et al., 2015).[9] Also, the structural model enables us to decompose the labor participation effect of the reform package into the effects caused by changes to the child care subsidies and the effects caused by changes to the in-work benefits. In addition, because our structural model is fully integrated with a detailed tax-benefit calculator, we are also able to study the effectiveness of fiscal stimuli for working parents in terms of additional employment generated per additional public euro spent. The integrated model allows us to go beyond back-of-the-envelope calculations of policy effectiveness (Blau, 2003; Lokshin, 2004). Although we study reforms implemented in the Netherlands, we argue that our findings are relevant for other developed OECD countries as well. Indeed, the participation rate of mothers and fathers in the Netherlands, as well as public spending on formal child care and pre-primary education, takes an intermediate position between Scandinavian and Anglo-Saxon countries (OECD, 2015b). We show how a reform akin to the EITC in the US and WFTC in the UK, and the related CTC reforms, could impact labor market outcomes for couples in the Netherlands.

The paper is organized as follows. Section 2 describes the labor market and policy environment in the Netherlands. Section 3 outlines the structural model, the empirical methodology and data used to estimate the preferences, and it also presents the empirical results in terms of simulated labor supply and price of child care elasticities. We also compare the predictions of the structural model for the policy reforms in the data period with a differences-in-differences analysis. In Section 4 we study the effectiveness of different fiscal stimuli for working parents, and also relate these policy reforms and estimated effects to the policies that are in place in the UK and US (like the WFTC and EITC, respectively). Section 5 discusses the significance our findings and concludes. An online appendix contains supplementary material.

[9] We contribute to a small but growing literature that evaluates the performance of structural models by comparing simulated policy responses with the results from (quasi-)experimental studies (Todd and Wolpin, 2006; Geyer et al., 2015; Thoresen and Vatto, 2015; Hansen and Liu, 2015; Thoresen and Vatto, 2019).

2 Labor market and policy environment

The labor participation rate of women in the Netherlands has changed markedly over the last 50 years. In the mid 1970s it was close to 30% (OECD, 2015a), which was relatively low by international standards. However, following the economic crisis in the early 1980s it started to rise, with the rise being particularly strong among mothers with young children (Euwals et al., 2011). By 2004, the Netherlands reached a female participation rate of 70%, taking an intermediate position between the somewhat higher participation rates in Scandinavian countries, and the somewhat lower participation rates in the US and the UK.[10] The participation rate of men in the Netherlands was falling between mid 1970s and mid 1980s. In the face of adverse labor market conditions, many men were sent into early retirement and disability programs. However, the generosity of these schemes was cut back in the 1990s and 2000s, and the male participation rate returned to levels comparable to other developed OECD countries.

To further promote the labor participation of parents with children (in persons but also in hours worked per week), the Dutch government implemented a series of reforms over the period 2005–2009. Below we give a short chronological overview of these policy changes, preceded by a brief introduction to the child care market in the Netherlands.

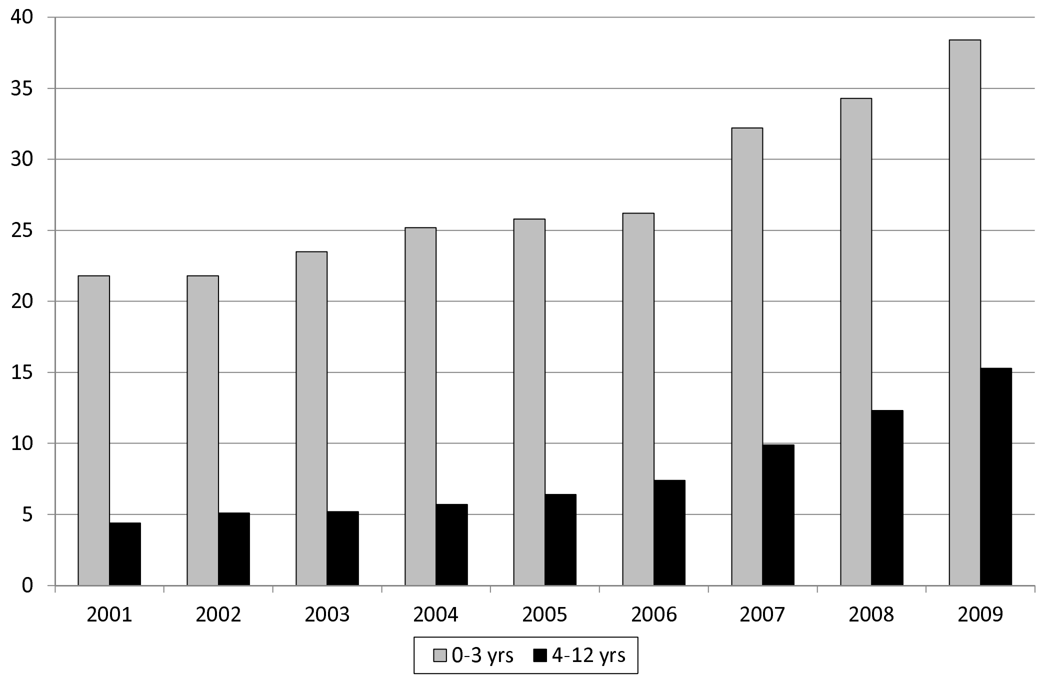
Children in the Netherlands enter primary school when they turn 4 years of age, and most children are 12 years old when they enter secondary school. Before the age of 4, children can attend centre-based daycare, community-based playgroups (*peuterspeelzalen*) and informal care. Before the introduction of the Law on Child Care (*Wet kinderopvang*) in 2005, centre-based daycare was subsidized at varying rates.[11] The enrollment rate of children 0–3 years of age in centre-based care was 25% in 2004, see Figure 1. Primary school-aged children (4–12 years of age) can attend centre-based after school care, also subsidized at varying rates before the introduction of the Law on Child Care, or informal care. The enrollment rate of children 4–12 years of age in centre-based care was 6% in 2004, also shown in Figure 1.

The series of reforms we consider started with the introduction of the Law on Child Care in 2005. This law unified the subsidies for formal child care services. As a result, parents became eligible for the same subsidy irrespective of their choice of formal child care provider. The subsidy was awarded conditional on the labor market participation

[10] Whereas the participation rate of women in the Netherlands has converged to other well-developed OECD countries, there remains a sizeable and stable gap in hours worked by employed women (OECD, 2015a). In 2004, employed women in the Netherlands worked on average approximately 24 hours per week, while their counterparts in other OECD countries worked 5 to 10 hours per week more.

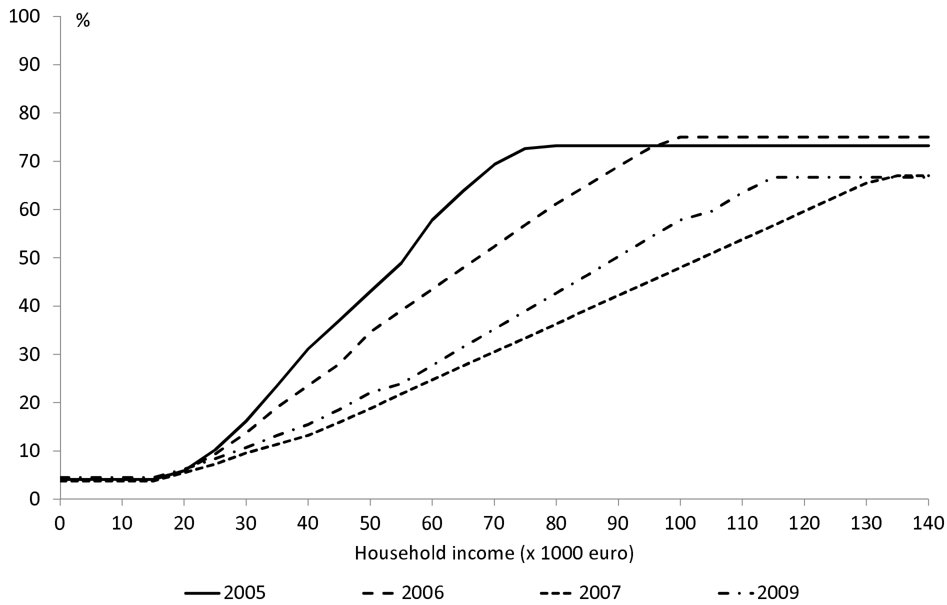
[11] All the data on the use of formal child care in this section are retrieved from Statistics Netherlands (http://statline.cbs.nl).

Figure 1: Share of children in formal child care (in %)



Source: Statistics Netherlands (http://statline.cbs.nl).

Figure 2: Parental contribution rate for the first child



Source: Own calculations using publicly available subsidy tables.

Table 1: Public spending on child care and in-work benefits for parents (millions of euro)

Year	2002	2003	2004	2005	2006	2007	2008	2009
Child care subsidies	725	755	1,028	1,001	1,343	2,058	2,825	3,034
In-work benefits for parents	410	460	738	830	871	984	971	1,290
– *Combination benefit*[a]	410	460	479	484	314	324	247	0
– *Income-dependent combination benefit*[b]	0	0	259	346	557	660	724	1,290

Source: Ministry of Finance (2010) and own calculations (imputation of employers' contribution for child care up to 2007 with data from the Ministry of Social Affairs and Employment (personal communication) and split of the in-work benefits for parents in its two components using the MIMOSI model of CPB). [a]The Combination benefit applies to primary earners, secondary earners and working single parents with the youngest child up to 12 years of age. [b]The Income-dependent combination benefit applies to secondary earners and working single parents with the youngest child up to 12 years of age.

of both parents (or one parent in the case of a single-parent households). The subsidy amount was specified as an income-dependent fraction of the total child care costs, leaving parents to pay a pre-specified contribution rate.

The income-dependent component of the subsidy is illustrated in Figure 2 which plots the effective parental contribution rate for the 'first child' against the household income, and documents its changes over the reform period.[12] The contribution rate in 2005 was the lowest among parents with annual income below 15,000 euro, who were required to pay 4% of the child care costs. The contribution rate was set to increase relatively linearly with additional income before levelling off at 74% for parents with annual income above 80,000 euro. Even though the 2005 reform unified the childcare subsidies, it did not lead to an immediate decrease of parents' own contribution rates (because it effectively replaced the pre-existing subsidy programs). Indeed, the unification of the subsidies had only a minor effect on public spending on formal child care, with the public spending falling slightly from 2004 to 2005, as reported in Table 1.

More important were the changes that followed in 2006 and 2007. In these years the child care subsidy rates were increased substantially, in particular in 2007. These changes were particularly prominent among middle income households, for whom the subsidy rates went up by 20 to 40 percentage points. For the lowest income households (whose subsidies amount to 96% of the full child care price), the subsidy rate hardly changed. On average, the effective price of child care dropped from 37% of the full price in 2005 to 18% of the

[12]The Tax Office defines the first child as the child for which the parents have the highest child care expenditures. For most households the first child is the youngest child since more hours are needed for daycare than for after-school care.

Annual in-work benefit parents with children

Figure 3: Primary and secondary earners

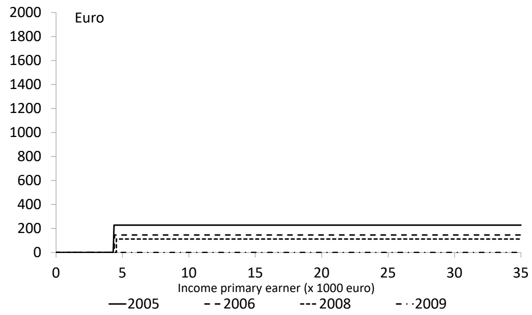
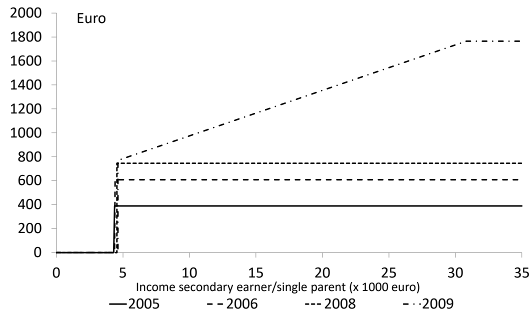


Figure 4: Secondary earners



Source: Tax Office.

full price in 2007.[13][14] In 2008 there were virtually no changes in child care subsidies, and in 2009 there was a small reversal of the initial increase in child care subsidies, see again Figure 2.

Over the period 2005–2009, public spending on formal child care in the Netherlands went from 1 to 3 billion euro. By 2009, with public spending on child care and pre-primary education representing 0.5% of GDP, the Netherlands took an intermediate position between Sweden and Norway that spent respectively 1.4 and 1.2% of GDP, and the US and Canada that spent respectively 0.4 and 0.2% of GDP (OECD, 2015b). The increase of public spending was fueled by increasing utilization of formal child care, see Figure 1.

Next to changes in child care subsidies, the period 2004–2009 also witnessed a number of changes to in-work benefits for working parents. Initially, these were available in the form of a fixed 'Combination Benefit' (*Combinatiekorting*), with the name referring to the combination of work and care. All working parents with the youngest child less than 12 years of age qualified for this in-work benefit.[15] Over the reform period, the Combination Benefit was gradually phased out, see Figure 3.

[13]Source: Tax Office data provided by the Ministry of Social Affairs and Employment (personal communication).

[14]Despite the steep increase in the subsidy rate, the average prices of formal child care places grew more or less in line with the CPI. Next to the drop in parental fees, from 2007 onwards schools were obliged to act as an intermediary for parents and child care institutions to arrange after school care.

[15]The award rate of the Combination Benefit was fixed, and the benefit was available for all parents whose earned income was above a minimum-income threshold (approximately 25% of the annual gross minimum wage).

The phase-out of the Combination Benefit was accompanied by phase-in of a new in-work benefit, called Income-Dependent Combination Benefit (*Inkomensafhankelijke Combinatiekorting*). This benefit was introduced in 2004 and its award rates were gradually increasing throughout the reform period (see Figure 4). The new benefit was awarded only to the secondary earners (and single parents), primary earners being excluded. Up to 2008, the award rate structure of the new benefit resembled the old one, with a fixed payment awarded to parents who were above the minimum-income threshold. In 2009, the benefit became income-dependent, increasing in the income of the secondary earner, with a phase-in rate of 3.8% up to a maximum of 1,765 euro.[16]

Over the reform period, public expenditures on the Combination Benefit dropped from 484 million euro in 2005 to 0 in 2009, whereas public expenditures on the Income-Dependent Combination Benefit rose from 346 million euro in 2005 to 724 million euro in 2008, and then to 1,290 million euro in 2009, see Table 1.

3 Structural model

Below we outline the functional forms of the structural model, the empirical methodology, the data set we use to estimate the model, and the empirical results.

3.1 Preferences and empirical methodology

Households with children are assumed to maximize a static unitary household utility function.[17] The systematic part of household utility U^s depends on three choice variables: weekly hours worked by the male h_m, weekly hours worked by the female h_f, and weekly hours of formal child care used c. These choice variables influence the household utility directly through the preference channel, and indirectly through their effects on household disposable income, y. For the functional form of U^s we choose the translog specification,

[16]This maximum is reached at 30,803 euro, approximately 160% of the minimum wage.

[17]Some studies use collective household labor supply models where both partners have their own preferences. In these studies, individual labor supply is determined alongside a bargaining process between household members. A priori, it is not clear whether the behavioral responses would change significantly if we would use a collective household model instead of a unitary model. Vermeulen (2005) compares the behavioral responses in a collective model with those in a unitary model, using data from the Netherlands 1995–2003, and finds comparable labor supply elasticities. Similarly, the collective model of Blundell et al. (2007) yields elasticities that are comparable to the elasticities in unitary models.

following Mastrogiacomo et al. (2017):

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

The key component of this functional form is the vector ν, which contains logarithmic transformations of the aforementioned choice variables and household disposable income.[18] Following the convention in the literature, labor supply is expressed in terms of leisure utilization (i.e., the share of weekly time endowment $T = 168$ left after subtracting the hours worked).

A is a symmetric matrix of quadratic preference parameters and **b** is a vector of linear preference parameters, both corresponding to the vector ν. **d** is a vector of preference parameters capturing the fixed costs of work by the male and by the female and fixed costs of using formal child care. These fixed costs are imposed on households with non-zero hours of work or formal child care, and they represent a composite of several distinct cost factors such as intrinsic disutility from work, commuting costs, market frictions and other costs associated with work and child care use.

Household disposable income is defined as follows:

$$y = w_m h_m + w_f h_f - T(w_m h_m, w_f h_f; q) - TC(p_c, c) + S(p_c, c, h_m, h_f; q), \tag{2}$$

where w_m and w_f denote the gross hourly wages for the male and the female, $T(.)$ represents taxes and employees' premiums, which depend on spousal earnings $(w_m h_m, w_f h_f)$ and other factors q that are relevant for the fiscal system, including household characteristics and the policy regime in the year of observation. $TC(.)$ is the total cost of formal child care, which depends on the gross hourly price of formal child care p_c and the number of child care hours used. $S(.)$ is the child care subsidy, which depends on the hourly price of formal child care, number of child care hours used, parental labor supply, and other relevant factors, such as the ages of the children and the policy regime in the year of observation.

Equation 2 is used to compute household-specific values of disposable income for each possible combination of parental work and child care choices that is considered by our model. Here, we face the standard problem of partial observability: gross hourly wages are observed only among parents who work, and gross child care prices are observed only

[18] As a sensitivity check, we have also estimated a quadratic specification (without logs). The results of this model were in line with the results corresponding to the translog specification.

among parents who use child care. For parents with missing wages and child care prices, we use Heckman selection models to impute this information.[19] Detailed description of both imputation protocols can be found in the online appendix (sections A and B).[20]

The structural model is set in a discrete choice framework. Parents choose their preferred combination of hours of work and hours of formal child care from a finite set of alternatives $j \in \{1, ..., J\}$, yielding a utility valuation $U_j(\nu)$. This valuation has a systematic part $U_j^s(\nu)$, and an alternative-specific stochastic term ε_j,

$$U_j(\nu) = U_j^s(\nu) + \varepsilon_j. \tag{3}$$

The stochastic terms are assumed to be i.i.d. across alternatives, and drawn from the Type 1 Extreme Value distribution. This leads to the multinomial logit specification of the discrete choice model (McFadden, 1978), which is commonly used in the structural labor supply literature (Van Soest, 1995; Blundell and Shephard, 2012; Bargain et al., 2014).[21]

Our model allows for both observed an unobserved preference heterogeneity. Observed preference heterogeneity is captured by the linear parameters b_2, b_3 and b_4, which are interacted with a vector of observed individual and household characteristics $\mathbf{x}_2$, $\mathbf{x}_3$ and

[19] Disposable incomes of parents who are observed to work and/or use child care are computed using their observed wages and/or prices. An alternative would be to use imputed information for all parents, however it is not clear whether this approach is preferable. Further, Mastrogiacomo et al. (2017) find that using simulated wages for workers yields labor supply elasticities does not yield substantively different elasticities.

[20] The imputation of wages and child care prices is done outside of the structural model, which contrasts with the simultaneous approach advocated by Blundell and Shephard (2012). While the simultaneous approach is better suited for capturing wage heterogeneity, our use of a sophisticated external tax-benefit calculator renders it practically infeasible. Hence, we follow the majority of the structural labor supply literature and use the two-step estimation procedure (see Loeffler et al., 2014, for an overview). The wage and price heterogeneity is then approximated by taking multiple draws from the corresponding error distributions, adding them to the imputed wages and prices, and integrating them out in the likelihood function.

[21] We do not require parents to provide a minimal amount of child care to their children, meaning that both parents can work full-time without using any formal child care. This modeling choice reflects the fact that parents often rely on informal child care providers (such as grandparents, siblings or friends), whose services are unobserved. Supplementary analyses of LISS survey data confirm that 40% of parents of pre-school children use informal child care (which is overwhelmingly provided by grandparents), with 29% combining both formal and informal sources of child care, and 11% relying on informal sources exclusively. One may be also concerned that waiting lists are preventing parents from realizing their formal child care demand, but this does not appear to be the case. Waiting lists in the Netherlands are relatively uncommon (less than 10% of filled places) and their prevalence decreased slightly over the period of observation, see Bettendorf et al. (2015). In a robustness check we include a proxy for the use of informal child care as an additional argument in the utility function. The resulting labor supply and child care elasticities are similar to the baseline specification (see below).

$\mathbf{x}_4$.[22]

$$b_2 = \mathbf{x_2'}\beta_2, \quad b_3 = \mathbf{x_3'}\beta_3, \quad b_4 = \mathbf{x_4'}\beta_4 \tag{4}$$

The same interactions are also imposed on the fixed cost parameters $\mathbf{d}$. For a full list of individual and household characteristics used, see the online appendix Section E.

To account for the possibility that observationally equivalent families have different tastes for leisure and formal child care, we consider two ways of modelling unobserved preference heterogeneity. In our baseline model specification, we include random preference heterogeneity in the linear utility parameters:

$$b_2 = \mathbf{x_2'}\beta_2 + \gamma_2, \quad b_3 = \mathbf{x_3'}\beta_3 + \gamma_3, \quad b_4 = \mathbf{x_4'}\beta_4 + \gamma_4, \tag{5}$$

where γ_2, γ_3 and γ_4 are assumed to be normally distributed with mean zero and variance σ_2, σ_3 and σ_4, respectively. To keep the numerical optimization of the likelihood practically feasible, we do not parameterize quadratic coefficients in matrix $\mathbf{A}$.[23] The corresponding likelihood function is given in appendix C.

The second way we model unobserved heterogeneity is using the latent class framework. In this case, we assume that there is a finite number of types of households K (called classes) that are allowed to differ in all preference parameters.[24] The details of the latent class specification are given in appendix D.

3.2 Data

To estimate preference parameters of the structural model we use the Labor Market Panel (LMP) of Statistics Netherlands (2012). The LMP is a large administrative household panel data set with annual data for the period 1999–2009. The LMP contains a rich set of individual and household characteristics, including gender, month and year of birth, the highest completed level of education and migration background for all adult members of the household, the ages of the children and the area of residence. The LMP also contains administrative data on hours worked and gross income from different sources (wages, profits, benefits etc.). Furthermore, the LMP contains administrative data on the use and gross hourly price of formal child care for each child participating in formal child care.[25]

[22]We leave the linear parameter b_1 without any interactions, which helps to reduce the computational complexity of the problem. Given that the utility function is identified up to a monotonic transformation only, this does not seem overly restrictive.

[23]For a detailed discussion of this specification, see Van Soest et al. (2002) and Kabátek et al. (2014).

[24]The latent class model is more flexible in its treatment of unobserved heterogeneity, relying on fewer functional form assumptions than our baseline model specification. As such, it can help us assess whether or not the assumptions underlying our baseline model specification are likely to hold.

[25]The LMP does not contain data on the quality of child care.

The data on formal child care are available only for a subset of years (2006–2009), and hence we restrict the estimation sample to this period.[26]

We impose the following sample selection criteria. We restrict the sample to couples with the youngest child being less than 12 years old, since older children do not qualify for childcare subsidies. We exclude couples in which at least one parent is either self-employed (8% of observations), or has multiple sources of income (7% of the remaining observations), because in both cases we cannot determine the effective budget constraint. We also exclude couples in which at least one of the partners is on disability or unemployment benefits (3% of the remaining observations), assuming that they are constrained in their labor supply choice. After these selections are made, we further drop couples with missing information on individual or household characteristics (7% of the remaining observations), which leaves us with 61,220 observations (couples times years observed).

We discretize the data for the discrete choice model. Both parents can choose from 6 labor supply options: working 0, 1, 2, 3, 4 or 5 days per week, where each day equals 8 hours. For child care, we allow for 0, 1, 2 and 3 days of care.[27] The full choice set for each household contains $6 \cdot 6 \cdot 4 = 144$ alternatives. The observed choice is defined as the alternative which is closest to the hours of work and child care hours recorded for the given family.

To determine disposable household income in each discrete option we use the advanced tax-benefit calculator MIMOSI (Romijn et al., 2008). MIMOSI allows for a very accurate calculation of the budget constraints, taking into account all national taxes[28], social security premiums, income independent subsidies and tax credits. MIMOSI also calculates the child care subsidy applicable for each household in each option.[29]

We estimate the preferences separately for couples with the youngest child aged 0–3, and for couples with the youngest child aged 4–11.[30] This is to acknowledge that there can

[26] During this period, the economic conditions in the Netherlands were stable. The labor market had not yet been affected by the large labor demand shocks of the Great Recession, with the employment rate increasing slightly from 74.4% in 2006 to 77.0% in 2009. At the same time, the unemployment rate fell from 3.9% to 3.2% (OECD, 2015a).

[27] The data show that using formal child care for more than 3 days per week is rare in the Netherlands. The remaining child care needs are usually accommodated by informal carers or by parents themselves. To map the reported child care hours to modeled days of care, we use the following metric: One day of care equals to 10 reported hours of daycare per child, or 5 reported hours of after school care per child. This metric is again guided by our data, which shows spikes at the respective reported hours.

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

[29] 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, both before and after we account for the behavioral responses.

[30] Table E.1 in the online appendix gives descriptive statistics of our principal estimation sample. The full set of results for our preferred specification can be found in Table F.1 in the online appendix. Section

be non-trivial differences in child care requirements and labor supply incentives faced by parents with younger and older children (Bernal, 2008). The preferences are estimated in a pooled cross-section framework, using a 15% random sample of our selected household-year observations.[31]

3.3 Results

Rather than interpreting the individual estimates of preference parameters, we focus on elasticities derived from these parameters. First, we consider the Marshallian (uncompensated) labor supply elasticities (Table 2). We simulate the elasticities by increasing gross wages, which is common in the literature.[32] We find substantial own-wage elasticities for mothers: 0.43 for mothers with the youngest child 0–3 years of age, and 0.41 for mothers with the youngest child 4–11 years of age. About two-thirds of the response is on the extensive margin, corresponding to the change of the participation rate, and about one-third is on the intensive margin, corresponding to the change of working hours among employed. The response on the intensive margin is relatively high, which is in line with Bargain et al. (2014), who find comparable intensive margin responses among Dutch women.[33] Apart from mother's own labor supply, her wage influences other household decisions as well. We find negative but small cross-wage elasticities for fathers, and substantial wage elasticities of the use of formal child care. Fathers' own-wage elasticities are much smaller, 0.08 for both groups of fathers, and most of the response is on the extensive margin. We also find a sizeable negative cross-wage elasticity for total hours worked by mothers, and a modest elasticity of the use of formal child care with respect to the gross hourly wage of fathers. Results for the child care price elasticities are given in appendix H.

3.4 Robustness Checks

We present an extensive list of robustness checks for the structural model in the online appendix.

G in the online appendix lists figures documenting the goodness of fit of our model. The model gives a good fit of the distribution of working days and days of formal child care.

[31] This is to further limit the computational burden of our models. We have tested different sample sizes, and found that 15% sample was sufficient for the stability of the preferences and the elasticities. To this end, we have drawn another 15% subsample from the remaining 85% of the data and re-estimated our model. Both samples yielded very similar estimates, as shown in Table J.3 of the robustness checks.

[32] Bargain et al. (2014) point out that high tax countries have smaller net wage increases and this may explain smaller gross wage elasticities for those countries.

[33] We follow Bargain et al. (2014) who decompose the total number of hours (H) as $H = ph$, where p is the extensive margin (e.g. the fraction of the period of the reference period when the individual is employed) and h represents the intensive margin which is defined as the total number of hours H divided by the fraction of the reference period in employment.

Table 2: Gross wage elasticities

	Hourly wage women +1%				Hourly wage men +1%			
	0–3 yrs		4–11 yrs		0–3 yrs		4–11 yrs	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Labor supply women	0.425	0.002	0.411	0.003	–0.161	0.002	–0.121	0.002
– Extensive margin	0.276	0.001	0.269	0.002	–0.110	0.001	–0.074	0.001
– Intensive margin	0.144	0.001	0.138	0.001	–0.052	0.001	–0.047	0.001
Labor supply men	–0.036	0.001	–0.026	0.001	0.082	0.001	0.080	0.001
– Extensive margin	–0.009	0.000	–0.013	0.000	0.071	0.001	0.069	0.001
– Intensive margin	–0.027	0.000	–0.014	0.000	0.010	0.000	0.011	0.000
Formal child care	0.449	0.003	0.485	0.005	0.087	0.004	0.256	0.008

Notes: Bootstrapped standard errors based on 200 draws of the estimated preference parameters.

First, we perform an auxiliary 'out-of-sample' fit test by re-estimating the model using data only for the period 2006–2008, and using these estimates to predict the distribution of labor supply and child care use in 2009. Figures presented in Appendix I show that the model fits the out-of-sample distribution well.

Next, we assess the importance of exogenous variation in the budget constraints by comparing our results to the result we get if we ignore this variation and perform a cross-sectional estimation for each year separately. Tables J.1 and J.2 show that there is quite some variation in the wage elasticities of women and the price elasticity of child care over the years. By using only cross-sectional data, it becomes clear the responses move further away from the responses found in the DID-analysis (discussed below).

As a further robustness check we estimate a model in which we add time dummies to the parameter vector **d**. This is to capture potential preference shifts towards work or the use of child care over the period of observation. Table J.4 shows that this model arrives at similar elasticities as our baseline specification, which suggests that the preferences for work and child care in this period were stable.

We further estimate a model without child care to assess the importance of this model feature for the determination of wage elasticities of parental labor supply. Table J.5 shows that the wage elasticities for couples with a youngest child aged 0–3 are very similar in both models. However, for women with a youngest child aged 4–11 we find a substantial difference, which highlights the importance of child care modeling for this age group.

In a similar exercise, we impose zero parameter values on the cross effects between leisure and child care in the matrix **A**. Table J.6 shows that a structural model without cross effects yields higher child care elasticities and labor supply elasticities for women with a youngest child 0–3 years of age. This result highlights potential complementarities between women's leisure and child care: if women have an intrinsic preference for combining leisure and child care (e.g., as a means for relaxation), then shutting down this channel will make women's labor supply more elastic with regard to their wages.

Table J.7 shows that the results for the model without unobserved heterogeneity exhibit only minor differences compared to the results for the model with random preference heterogeneity. J.8 and J.9 give the elasticities for the models with one (no unobserved heterogeneity), two and three latent classes. The results are similar for the latent class specifications as for the base model with random preference heterogeneity. Only in the 2-class specification for mothers with the youngest child aged 0–3, the elasticities are somewhat lower. This is due to a lower extensive margin response for mothers.

Table J.10 considers the labor supply elasticities when we include a proxy for informal child care in the utility function.[34] The results with the informal care proxy are quite similar to the base model without the informal care proxy.

Finally, we extend the model by allowing for involuntary unemployment (e.g. Bargain et al., 2010). Appendix K outlines this extended model and shows that allowing for the possibility of involuntary unemployed lowers the labor supply responses somewhat, though the impact on behavioral responses is small. This conforms with the fact that the share of involuntarily unemployed within our data period is rather low (only 3%).

3.5 Comparison of structural model with diff-in-diff analysis

Finally, we compare the predictions of the structural model with the results of a differences-in-differences (DD) analysis. The DD analysis builds on Bettendorf et al. (2015), using data from the Dutch Labor Force Survey for 1995-2009 (which also allows for pre-reform placebo estimates). The details of the DD analysis are given in appendix L. The main difference with Bettendorf et al. (2015) is that we restrict the sample to couples with children, and hence exclude single parents.

We simulate the effects of the 2005–2009 reform using our structural model, and compare the simulated effects to the medium to long run effects of the DD analysis. For the simulation exercise we use the structural parameters estimated using the 2006–2009 LMP data. We subject the LMP households to the fiscal regimes applicable in years 2005 and 2009, and we predict their work and childcare choices under the two regimes. The differ-

[34] We assume that parents coordinate their hours worked so as to minimize the use of child care. Total demand for child care then equals $c_{tot} = \max((h_m + h_f - \overline{h}), 0)$, where $\overline{h}$ equals hours per week in a fulltime job. The demand for informal child care is then given by $c_{inf} = \max((c_{tot} - c), 0)$. For simplicity we do not distinguish between informal care and hours spent at school for households with the youngest child 4–11 years of age.

Table 3: Comparison of structural model with DD analysis: policy reforms 2005–2009

	Structural model				DD analysis[a]	
	(1) Child care subsidies	(2) Combination Credit	(3) Income-depend. Combi. Credit	**(4) Total effect**	**(5) Coefficient**	SE
Youngest child 0-3 yrs			Changes in levels			
Participation rate women	0.019	–0.006	0.020	**0.033**	**0.020**	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.007)	
Hours worked per week women	0.752	–0.106	0.630	**1.275**	**1.099**	
	(0.006)	(0.008)	(0.005)	(0.009)	(0.215)	
Participation rate men	0.003	–0.002	0.004	**0.005**	**0.006**	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.004)	
Hours worked per week men	0.067	–0.033	0.046	**0.080**	**–0.364**	
	(0.003)	(0.001)	(0.002)	(0.003)	(0.229)	
Youngest child 4-11 yrs						
Participation rate women	0.003	–0.005	0.016	**0.014**	**0.022**	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.007)	
Hours worked per week women	0.105	–0.086	0.453	**0.471**	**0.741**	
	(0.002)	(0.001)	(0.007)	(0.008)	(0.212)	
Participation rate men	0.000	–0.001	0.003	**0.002**	**0.003**	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.004)	
Hours worked per week men	0.018	–0.024	0.051	**0.045**	**–0.141**	
	(0.000)	(0.000)	(0.001)	(0.001)	(0.227)	

[a]DD treatment effect for 2008–2009, see Table L.2.

ence between the predictions constitutes the simulated reform effect. Table 3 presents the comparison of the two effects, with the simulated reform effects presented in column (4) and the DD effects presented in column (5).[35]

Table 3 shows that the simulated effects of the structural model with random preference heterogeneity for mothers are in line with the results of the DD analysis in terms of the participation rate and hours worked. For fathers, the simulated effects for the participation rate are again in line with the DD analysis. For hours worked per week, the structural model predicts an effect close to zero, whereas the DD analysis suggests a small negative effect. However, the DD estimates for hours worked per week by fathers are not statistically

[35]In columns (1)-(3) we split the total simulated effect into partial effects corresponding to the separate policy instruments. Column (4) is the sum of the three partial effects, including potential interactions between the reforms.

significantly different from zero.[36,37]

4 Relative effectiveness of fiscal stimuli

Now that we have shown that the structural model not only gives a good fit of the data, but also generates behavioral responses consistent with a quasi-experimental analysis of the recent reforms, we move to the analysis of counterfactual scenarios. Namely, we leverage the flexibility of the structural model to study the effectiveness of different fiscal stimuli for working couples with children. The policy reforms we consider are motivated by the actual reforms that have occurred in the Netherlands over the past decade (an extensive overview of these reforms can be found in De Boer et al., 2018). However, we also relate our reforms and findings to related reforms in the UK and US that have received much attention in the empirical literature.

First, we analyze three types of fiscal stimuli for working parents: i) an increase of child care subsidy rate, ii) the introduction of an (additional) in-work benefit for secondary earners with children 0–11 years of age, and iii) the introduction of an in-work benefit for both primary and secondary earners with children 0–11 years of age. For each of these fiscal stimuli, we consider two types of reforms: a) a subsidy/benefit that does not depend on income, and b) a subsidy/benefit that rises with income. By comparing the responses to the income-dependent and income-independent policies, we can study to what extent there is an equity-efficiency trade-off between these two types of fiscal stimuli. Next, we analyze reforms that are closer to the reforms in the UK and US.

For the baseline prediction, we subject the LMP households to the fiscal regime of the year 2009. To make the stimuli comparable, in all simulations we consider the effects of a reform that costs 100 million euro given the initial distribution of labor supply and child care choices, *e.g.*, without taking behavioral effects into account. Then, we calculate the

[36] Another relevant question is how much of the change (2005–2009) in the participation rate and hours worked is explained by the structural model. Comparing the effects of the reforms to the change in the LFS for the same sample, we find that for women with a youngest child 0–3 years of age we explain 35% and 38% of the increase in the participation rate and hours worked per week, respectively. For women with a youngest child 4–11 years of age we explain 19% and 17% of the increase in the participation rate and hours worked per week, respectively. For men we predict hardly and change in the participation rate and hours worked per week. Part of other drivers are captured by our structural model, like the change in wages. However, part of these changes may be the result of changes in household characteristics, like the level of education of the partners, preference shifts in the willingness to work and use formal child care, or changes in labor demand conditions (the overall unemployment rate went down from 5.9% in 2005 to 4.4% in 2009). However, in this respect it is reassuring to see that the behavioral responses are quite similar when we include time dummies in the fixed costs of work and the use of child care (appendix J.4) or involuntary unemployment (appendix K).

[37] Table L.4 shows that the model with three latent classes and the model without unobserved heterogeneity also generate predictions that are in line with the DD estimates.

effectiveness of the different policies by comparing the increase in total hours worked to the increase in public expenditures, but this time we take into account expenditures induced by workers' behavioral responses. We show that these induced expenditures (so-called knock-on effects) play a particularly important role for assessing the relative effectiveness of the candidate policies.[38]

4.1 Income independent stimuli for working parents

First we consider the results for the following three reforms that are income-independent:

(1) An across-the-board increase in the child care subsidy per hour of child care.[39]

(2) A fixed annual in-work benefit for secondary earners.[40]

(3) A fixed annual in-work benefit for primary and secondary earners.[41]

For the child care reform, although the increase in the subsidy is not income-dependent, it is implicitly targeted more at middle and higher incomes because they use more formal child care than lower income families. Accordingly, there is a positive effect of the reform on the Gini-coefficient for disposable household income (before behavioral changes), as reported in column (1) in Table 4. Next we consider the effects on disposable income for the in-work benefit for secondary earners. As low income households benefit more in percentage terms, this reform leads to a decline in the Gini-coefficient, see column (2) of Table 4. Finally, we consider the effects on disposable income of the in-work benefit for both primary and secondary earners. This reform is targeted even more at lower incomes, in percentage terms, and the Gini-coefficient falls the most in this scenario, see column (3) in Table 4. Figures M.1, M.3 and M.5 in appendix M show the redistributional effects of these reforms, respectively. In these figures we show on the horizontal axis initial disposable household income, and on the vertical axis the percentage change in disposable household income before behavioral responses.[42]

Table 4 presents the effects on labor participation, formal child care and public finances. Column (1) gives the results for the increase in the child care subsidy. First, consider the

[38] Alternatively, we could consider impulses that are comparable in terms of budgetary expenditures ex post, e.g. after taking into account the behavioral responses. However, this is computationally much more demanding, and it generates the same ordering as the procedure we follow.

[39] We increase the child care subsidy by 10.3 percentage points of the hourly price, up to a maximum total subsidy of 100% of the hourly price.

[40] An in-work benefit of 290 euro.

[41] An in-work benefit of 126 euro.

[42] We prefer to use the effects on disposable income before behavioral changes. Using changes in disposable income after behavioral changes ignores the potential offsetting utility effects (the envelope theorem suggests that these are approximately equal to the income gains or losses from changes in behavior).

Table 4: Effectiveness of fiscal stimuli of 100 million euro[a] [b]

	Not targeted			Targeted more at higher incomes		
	(1) Child care subsidy	(2) In-work benefit second. earners	(3) In-work ben. all parents	(4) Child care subsidy	(5) In-work benefit second. earners	(6) In-work ben. all parents
Inequality	Percentage changes					
Gini-coefficient	0.35	-0.10	-0.34	0.93	0.53	-0.14
Labor supply						
Labor supply total	0.54	0.30	0.09	0.53	0.46	0.11
Labor supply youngest child 0–3						
– Women	2.69	0.86	0.26	2.49	1.48	0.29
— Extensive margin	1.69	1.02	0.39	1.30	0.82	0.14
— Intensive margin	0.99	–0.16	–0.13	1.17	0.65	0.15
– Men	0.07	0.06	–0.01	0.12	0.04	0.04
— Extensive margin	0.27	0.15	0.08	0.18	0.13	0.08
— Intensive margin	–0.19	–0.09	–0.07	–0.06	–0.09	–0.04
Labor supply youngest child 4–11						
– Women	0.46	0.79	0.28	0.48	1.22	0.48
— Extensive margin	0.24	0.90	0.37	0.22	0.70	0.22
— Intensive margin	0.22	–0.11	–0.09	0.25	0.51	0.25
– Men	0.03	0.06	–0.01	0.03	0.05	0.03
— Extensive margin	0.03	0.11	0.03	0.03	0.09	0.03
— Intensive margin	0.00	–0.05	–0.04	0.01	–0.04	0.01
Formal child care						
Formal child care total	12.3	1.26	0.64	9.58	1.97	0.72
Formal child care youngest child 0–3	13.2	1.29	0.66	10.1	1.98	0.72
Formal child care youngest child 4–11	7.05	1.08	0.58	6.46	1.91	0.74
Public finances	Millions of euro					
Additional public exp. ex ante	100.0	100.0	100.0	100.0	100.0	100.0
Knock-on effect child care subsidies	119.3	10.9	5.5	90.2	15.9	5.7
Knock-on effect taxes and benefits	–72.6	–31.3	–8.4	–74.1	–52.7	–15.1
Additional public exp. ex post	146.7	79.6	97.0	116.0	63.2	90.7
Effectiveness in terms of hours worked	Euro					
Ex ante spending per FTE	28,986	52,637	180,617	29,775	34,159	139,093
Including effect on formal child care use	63,566	58,374	190,479	56,618	39,597	147,072
Including effect on taxes and benefits	42,514	41,923	175,285	34,543	21,584	126,133

[a]In all simulations we consider the effects of a reform that costs 100 million euro given the initial distribution of labor supply and child care choices. In column (1) we increase the child care subsidy by 10.3 percentage points of the hourly price. In column (2) we introduce an annual in-work benefit for secondary earners of 290 euro. In column (3) we introduce an annual in-work benefit for primary and secondary earners of 126 euro. In column (4) we increase the child care subsidy so that the parental fee falls by 41% for all incomes. In column (5) we introduce an in-work benefit for secondary earners starting at zero at an annual gross labor income of 4,000 euro, and then rising with 2.2% per euro of income up to a maximum of 581 euro at an income of 30,000 euro. In column (6) we introduce an in-work benefit for primary and secondary earners in couples, starting at zero at an annual gross labor income of 4,000 euro, and then rising with 0.6% of income up to a maximum of 168 euro per year at an income of 30,000 euro. See the main text for a full description of the reforms.

[b]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.

effects on labor participation of couples with the youngest child aged 0–3. For women, there is a substantial positive effect on both the extensive and the intensive margin.[43] The overall effect on hours worked by men is small. We observe similar though somewhat smaller labor supply effects for couples with the youngest child 4–11 years of age. Overall, total hours worked by couples with the youngest child 0–11 years of age increase by 0.54%, which corresponds to approximately 0.5 hours per week. In terms of child care, couples with the youngest child aged 0–3 increase their demand by 13%, whereas couples with older children demand 7% more child care. Overall, this corresponds to an increase in hours of child care per week of approximately 1.3 hours. Evidently, the rise in formal child care is much bigger than the rise in total hours worked, which underscores the role of child care type substitution.[44]

Next, we consider the effect of the child care stimulus on public finances, excluding and including the knock-on effects. The increase in hours worked increases tax receipts and reduces benefit expenditures by 73 million euro. However, this effect is dominated by the increase in child care subsidy expenditures due to behavioral changes of 119 million euro.[45] Indeed, many couples switch to heavily-subsidized formal child care, which makes the child care reform rather costly to the government. In the last three rows we approximate the costs of the child care stimulus by calculating the public spending required per additional fulltime equivalent (fte) worker employed. Ignoring the knock-on effects, additional public spending per additional fte is 29 thousand euro. However, taking into account the increase in formal child care spending raises the costs to 64 thousand euro. Finally, taking into account additional tax receipts and savings on benefits, we arrive at the final figure of 43 thousand euro per additional fte employed.

Column (2) in Table 4 gives the behavioral responses and corresponding budgetary effects for the in-work benefit targeted at secondary earners.[46] In contrast to reform (1), the effect on hours worked by women with children aged 0–3 is smaller than the effect for women with older children. A larger part of the in-work benefits goes to mothers with the youngest child aged 4–11 who, due to their lower utilization of child care, were not among the main beneficiaries of the child care subsidy reform. It is interesting to note that the

[43]The increase in hours worked on the intensive margin may seem puzzling, since the subsidy is income independent. However, there is a substitution effect at work. Mothers that work more hours use more formal child care, and thus benefit more from the higher child care subsidy, ceteris paribus.

[44]The analysis by Bick (2016) also emphasizes the costs of parents switching from unsubsidized to subsidized child care in Germany.

[45]The 119 million euro is in addition to the initial expenditure of 100 million euro.

[46]Our structural model is based on a unitary utility function, in which case who is the first and who is the second earner is not important once they cross over. Results might be different if we would assume a collective household model, where individual incomes also matter for consumption patterns. Indeed, a secondary earner tax credit might lock women into second earnship.

intensive margin response is negative for women with children in both age groups, which is due to the income effect of the additional benefit for working women. The effect on the labor supply of men is again small, and since the in-work benefit does not affect the price of formal child care, the reform (2) has only a modest effect on the child care utilization.

Because reform (2) leads to lower overall stimulation of hours worked, the knock-on effect in terms of additional taxes and benefits is smaller than for reform (1). However, since we observe only a modest increase of child care subsidy expenditures, the reform (2) generates a positive overall knock-on effect of 20 million euro. For reform (2), ignoring knock-on effects, additional public spending per additional fte employed is 53 thousand euro. Taking into account the knock-on effects, this becomes 42 thousand euro. Hence, we find that the reforms (1) and (2) are about equally effective in raising additional labor supply per additional euro spent. Note, however, that reform (1) goes at the expense of greater inequality, whereas reform (2) actually reduces inequality.

Column (3) in Table 4 gives the results of the in-work benefit for both primary and secondary earners. In this scenario, a large part of the subsidies goes to men in couples with children, who were shown to hardly respond to fiscal incentives. As a result, the effects are much smaller than in reforms (1) and (2). We still see a positive effect on the extensive margin, and a negative effect on the intensive margin (due to the income effect). The increase in total hours worked is just 0.09%. The knock-on effects are therefore also small, only 3 million euro. This makes the third scenario the most expensive reform in terms of additional spending required per additional fte employed: 175 thousand euro.[47] We should note though, that this reform leads to a bigger drop in inequality than reform (2).

4.2 Income dependent stimuli for working parents

Next, we consider the trade-off between equity and efficiency by simulating reforms that are targeted more at middle and higher incomes. Specifically, we study the effects of the following three reforms:

(4) An increase in child care subsidies so that the parental fee falls by 41% for all in-

[47] Even less effective for labor supply by couples would be an EITC along the lines of the EITC in the US. The EITC in the US depends on household income and is targeted at low incomes (see e.g. Meyer, 2010, for an extensive discussion of the EITC in the US and the resulting behavioral responses). This is because the EITC increases the effective tax rate for the relatively elastic group of secondary earners, and it does so both on the extensive and the intensive margin. We simulate the introduction of the EITC in our sample, and we find that the policy actually reduces labor supply by couples with children by 2.9%(details of the parametrization of the award rates and further results are available upon request). These results are in line with the findings of Eissa and Hoynes (2004), who report a large negative effect of the EITC on labor participation by women in couples in the US.

comes.[48]

(5) An in-work benefit for secondary earners starting at zero and then rising after 4,000 euro.[49]

(6) An in-work benefit for primary and secondary earners in couples, starting at zero and then rising after 4,000 euro.[50]

Reforms (4) and (5) increase income inequality (as measured by the Gini-coefficient) more than reforms (1) and (2) (see Table 4.) And although reform (6) reduces inequality, the reduction is less pronounced than for the reform (3). Therefore, if there is a trade-off between equity and efficiency, we would expect these reforms to be more effective in terms of labor supply. The redistributional effects of these reforms are given in appendix M, see Figure M.2, M.4 and M.6, respectively.

Column (4) in Table 4 shows the effects of the child care stimulus targeted more at middle and high incomes. The effect on total labor supply is actually quite similar to reform (1), with the effect on women being more favorable and the effect for men being less favorable. There is no apparent trade-off for child care subsidies when looking only at hours worked. This suggests that the current system in the Netherlands, which targets subsidies mostly at low incomes, is not detrimental to hours worked.[51] However, the knock-on effects are more favorable for reform (4) than reform (1). The additional hours worked by middle and higher incomes generate more additional tax revenue per additional hour worked. Furthermore, substitution of other types of care for formal care is less costly for the government, as the baseline subsidy per hour of formal child care is lower for middle and higher incomes than for lower incomes. With an about equal effect on total hours worked and more favorable knock-on effects, the required public spending per additional fte employed is more favorable in reform (4) than in reform (1): 35 thousand euro per additional full time equivalent employed. However, the difference comes at the expense of additional income inequality.

Column (5) shows that the in-work benefit for secondary earners that rises with income has a bigger effect on hours worked than reform (2). The substitution effect at the intensive

[48] Given that middle and higher incomes pay a larger fee in the base, this reform targets mostly middle and high income families.

[49] Specifically, the in-work benefit starts at zero at an annual gross labor income of 4,000 euro, and then rising with 2.2% per euro of income up to a maximum of 581 euro at an income of 30,000 euro.

[50] Specifically, the in-work benefits starts at zero at an annual gross labor income of 4,000 euro, and then rising with 0.6% of income up to a maximum of 168 euro per year at an income of 30,000 euro.

[51] The case for targeting child care subsidies at low income households becomes even stronger when participation in child care benefits children from low income households more than children from middle and high income households, as suggested by the empirical evidence presented in e.g. Blau and Currie (2006) and Havnes and Mogstad (2015).

margin of this reform makes the intensive margin responses by women positive rather than negative. The effect on total hours worked is also considerably larger than reform (2), although still smaller than the child care reforms (1) and (4). However, because this reform does not generate a large response in the use of formal child care, the knock-on effects are rather favorable. When we calculate the expenditures required per additional fte employed, reform (5) proves to be the most cost-effective: the costs are 22 thousand euro per additional full time equivalent employed. However also in this case there is a trade-off with equity, as the additional hours worked come at the expense of additional income inequality. So, there is still a trade-off when it comes to targeting different income groups.

Finally, column (6) gives the results of the income dependent in-work benefit for primary and secondary earners. The overall effect on hours worked and government finances is slightly better than for the flat benefit for primary and secondary earners in column (1). Again, there is a trade-off between efficiency and equity. However, this reform still has only a marginal effect on overall hours worked and the costs per additional full time equivalent employed are still unfavorable: 126 thousand euro.

Reforms (4)–(6) show that there is indeed a trade-off between equity and efficiency, targeting fiscal stimuli more towards working parents with a middle or higher income leads to a larger increase in hours worked per additional public euro spent. However, the trade-off is less pronounced for child care subsidies than for in-work benefits.

Also of interest is the heterogeneity in responses not only by gender and age of the youngest child, but also by education and migration background. Table N.1 in the appendix shows the results of the reforms for various subgroups, based on education and ethnicity. Francesoni et al. (2009) find that there is hardly a differential effect of the WFTC reform in the UK (further discussed below) for mothers in couples with different levels of education (which may also in part be due to the ambiguous effects of the WFTC on couples labor supply in general). However, we do find a substantial difference in the effects of the reforms we consider in Table 5 by level of education for mothers (and to some extent also for fathers) with a youngest child 0–3 years of age (pre-primary school age). Differences by level of education are less pronounced for mothers and father with a youngest child in 4–11 years of age (primary school age). We also find substantial differences in the effects by migration background. We find larger labor supply effects for mothers and fathers with a non-Western migration background than for mothers and fathers with a Western migration background or with a Dutch (native) background. The differences are again the most pronounced among couples with a youngest child 0–3 years of age.[52]

The policy simulations discussed so far have two important limitations. First, we consider reforms starting out of the 2009 policy regime, in which the average child care subsidy rate is already 76%. This makes a further increase in child care subsidies rather costly for the government, because for each additional hour of formal child care the government has to pay 76% on top of the increase in the child care subsidy. To study the extent to which the marginal effectiveness of child care subsidies depends on the initial level of the subsidy rate we have simulated two more reforms: i) a decrease in the child care subsidy by 10.3 percentage points of the hourly price (resulting in an average child care subsidy rate of 67%), i.e. the mirror image of reform (1), and ii) a decrease in the child care subsidy rate by 20.6 percentage points of the hourly price (resulting in an average child care subsidy rate of 59%). When we decrease the child care subsidy by 10.3 percentage points of the hourly price, the budgetary savings per fte lost are 37,912 euro (compared to 42,514 euro required to increase the child care subsidy by 10.3 percentage points). When we decrease the child care subsidy by 20.6 percentage points of the hourly price, the savings per fte lost are 32,996 euro. This shows that when we start from a low initial child care subsidy rate, the marginal effectiveness of child care subsidies can be comparable or even higher than in-work benefits for secondary earners.[53]

The second limitation is that the income-dependent child care subsidies and in-work benefits for secondary earners may be more effective in the Netherlands than in other countries. As argued by e.g. Cattan (2016), it is important to realize that outcomes may still differ depending on e.g. the initial female participation and the extent to which parents can rely on informal care. Indeed, Bargain et al. (2014) show that intensive margin responses of women in couples in the Netherlands are relatively large when compared to other developed countries. However, it is a priori unclear to what extent this affects the relative effectiveness of child care subsidies versus in-work benefits for secondary earners.

[52] Francesoni et al. (2009) also find that the effects of the WFTC reform were quite similar by age of the youngest child. We find substantially bigger effects for the child care reform for couples with a youngest child 0–3 years of age, who also use much more hours of formal child care than couples with a youngest child 4–11 years of age. However, when it comes to the in-work benefits, differences by age of the youngest child are less pronounced in our simulations as well.

[53] This is in line with the findings of Blau (2003, pp. 506-507), who shows that starting from a base where there is no child care subsidy, a marginal increase in child care subsidies is likely to be more cost-effective than an in-work benefit for working mothers.

4.3 Comparison with related programs in the UK and US

Next, we compare the policies discussed so far to the policies that are in place in the UK and the US: the Working Tax Credit (WTC)[54] and the Earned Income Tax Credit (EITC). These policies have received much attention in the economics literature (an excellent starting point for the literature on these programs is Brewer and Hoynes, 2019), which makes them a particularly relevant comparison group for our policy simulations. In addition, we consider the related—yet distinct—programs of the Child Tax Credit (CTC) in the UK and the US.

For each of these policies, we highlight their main characteristics, discuss the key empirical findings in the literature and relate the Dutch policies and the associated findings to these programs and their effects. We also conduct further simulations on the potential effects of implementing these reforms in the Dutch context. Finally, we briefly discuss how our findings relate to the recent child care plan of the Biden administration in the US.

The two programs that are most closely related to the subsidies for working parents we considered before are the WTC in the UK and the EITC in the US.[55] Both programs are targeted largely at families with children, and a key characteristic is that they are both means-tested on family income rather than individual income. This is an important feature which distinguishes the two programs from the in-work credits in the Netherlands, and we will consider the implications of this feature for income inequality and labor supply below.

The means-testing of EITC involves a phase-in segment over which the credit grows with income, flat segment over which the credit is stable, and a phase out segment over which the credit declines (see e.g. Figure 2, Panel A, in Brewer and Hoynes, 2019). In the phase-in region the substitution effect promotes labor supply, while the income effect works in the other direction. In the phase-out region both the substitution and income effect discourage labor supply. In contrast, the WTC operates with no phase-in segment and the flat segment is very short. Starting from a relative low earnings level (of around 25% of median earnings) there is a steep phase-out (see e.g. Figure 1, Panel B, in Brewer and Hoynes, 2019). Another difference between the two policies is that there is a minimum hours requirement for the WTC, where eligible couples have to work at least 24 hours per week between them, with at least one person working 16 hours per week. Via a substitution effect this promotes labor supply for couples that do not satisfy these hours criteria.

[54] And its predecessor the Working Families Tax Credit (WFTC). We note that the WTC is being replaced by the Universal Credit, but here we focus on the W(F)TC because that is also the focus of the empirical research thus far.

[55] For the specific details of the WTC and the EITC we draw heavily on Brewer and Hoynes (2019).

Empirical studies show that the WTC and EITC provide substantial income support for families with children. For low-income families, the WTC appears more generous than the EITC (Brewer and Hoynes, 2019).[56] Regarding the labor supply effects, studies that focus on lone parents typically find substantial positive labor supply effects of the WTC, see e.g. Francesoni and Van der Klaauw (2007); Blundell et al. (2008); Gregg et al. (2009), and the EITC, see e.g. Meyer and Rosenbaum (2001); Hoynes and Patel (2016).[57,] However, the results for couples, the focus of this paper, are more mixed, consistent with the aforementioned substitution effects (positive at the phase-in range, negative at the phase-out range) and income effects. The findings in Brewer and Browne (2006) suggest that the employment rate of British women in couples remained largely unaffected by the WFTC (the predecessor of the WTC), Francesoni et al. (2009) find a small increase in the employment rate of women in couples, with higher employment rates and hours worked for women whose partner did not work. Eissa and Hoynes (2004) find that the EITC in the US led to a decline in the employment rate of married women, with almost no effect on married men.

One of the key distinctive points of the Dutch in-work tax credits is that they are not targeted at low-income individuals. This has important implications for income inequality and labor supply responses. Indeed, when we consider the in-work benefit for secondary earners that increases with income (simulation (5) in Table 4, which is indicative of the reforms in the Netherlands over the past 15 years), we see that income inequality actually goes up after its implementation. This is because the primary beneficiaries of this policy are dual-earner couples with a relatively high family income. However, this policy also results in the highest increase in the labor supply of women with children, rising by 1.5% for women with a youngest child 0–3 years of age and 1.2% for women with a youngest child 4–11 years of age. The response for men is small.[58] We also note that targeting the in-work benefit solely at secondary earners (and using their individual income for means-testing) is more cost-effective than having a universal in-work benefit for both primary and secondary earners. This is because the former is targeted at the group that is more responsive to financial incentives, although this cost-efficiency is once again at the expense of inequality (compare simulations (5) and (6) in Table 4).

To ground our comparative analysis in empirics, we conduct a series of additional

[56]This is in part to compensate for the larger income support of non-working families in the UK (Brewer and Hoynes, 2019).

[57]Though see Kleven (2021) for a critical recent reappraisal of the impact of the EITC on the labor participation of lone parents in the US.

[58]Of interest here is the opposing effects on the extensive margin, as non-employed men in couples with children are stimulated to take up work, and the intensive margin, for men that are primary earners there is an income effect of higher household income.

Table 5: Reforms targeted at lower incomes

	'WTC'[a]	'CTC' 1[b]	'CTC' 2[c]
Inequality			
Gini-coefficient	-3.55	-0.57	-3.84
Labor supply			
Labor supply total	-1.44	-0.20	-3.26
Labor supply youngest child 0–3			
Women	-3.90	-0.36	-5.35
- Extensive margin	-3.48	-0.27	-4.58
- Intensive margin	-0.44	-0.09	-0.80
Men	-0.72	-0.20	-3.64
- Extensive margin	-0.29	-0.16	-3.45
- Intensive margin	-0.44	-0.04	-0.20
Labor supply youngest child 4–11			
Women	-1.75	-0.20	-2.61
- Extensive margin	-1.81	-0.14	-2.15
- Intensive margin	0.07	-0.06	-0.47
Men	-0.85	-0.13	-2.29
- Extensive margin	-0.58	-0.10	-2.16
- Intensive margin	-0.27	-0.02	-0.14
Formal child care			
Formal child care youngest child 0–3	-4.23	0.03	-5.74
Formal child care youngest child 4–11	-2.36	0.18	-3.04

[a] The WTC is an in-work benefit conditioned on the number of working hours. The maximum amount of the WTC is 2010 euro for the first child and an extra 1700 euro for the second child. Couples who work more than 24 hours per week, of which one partner works more than 16 hours per week, receive the WTC. The WTC is phased out for higher incomes at a rate of 37 percent.

[b] The CTC 1 is an income independent child benefit. Couples with 1 child receive a fixed amount of 146 euro for the first child and a fixed amount of 122 euro for the second child. The CTC 1 does not depend on income or labor force participation.

[c] The CTC 2 is an income dependent child benefit. The maximum amount of the CTC 2 is 1625 and 1350 euro for the first and second child, respectively. The WTC is phased out for higher incomes at a rate of 37 percent.

simulations that are presented in Table 5. In Column (1), we consider the labor supply effects of the introduction of a WTC type of in-work benefit in the Netherlands (starting from the tax-benefit system in 2009).[59,60] This results in a substantial decline in income inequality, with the Gini-coefficient dropping by more than 3.5%. This is to be expected, as working low-income families benefit the most from the reform, see Figure O.1 in appendix O. However, we also see a substantial drop in labor supply. The labor supply of women in couples with a youngest child 0–3 and 4–11 years of age drops by 3.9 and 1.8%, respectively. In the phase out region of the WTC, the substitution effect lowers labor supply. Labor supply of men in couples also declines, by 0.7 and 0.9%, for men in couples with a youngest child 0–3 and 4–11 years of age, respectively. The context of the Netherlands is different from the UK and US, in that the share of part-time work is much higher in the Netherlands. This may make intensive margin responses more important.[61] But in general, our results imply that we should be careful when we extrapolate results for the UK and the US to other countries, like continental European countries.

Table 5 also considers two other reform options in the Netherlands that are related to tax credits for families in the UK and the US. Column (2) considers the effects of a flat refundable credit/subsidy for both working and non-working couples with children. The subsidy only depends on the number of children.[62] This reform relates to the Child Tax Credit (CTC) in the US. Only working couples with a minimum earnings level qualify for the CTC, but this minimum earnings level is relatively low, so that almost all couples with children qualify, and the phase-out is at a very high income level (Brewer and Hoynes, 2019). We see that the simulation in column (2) leads to a slight reduction in income inequality and a slight reduction in labor supply, as 'only' an income effect is at work, see also Figure O.2 in appendix O.[63] In column (3) we consider a CTC that is related to the CTC in the UK. Both working and non-working couples with children qualify for the CTC in the UK, and there is a steep phase out starting from close to 80% of the median earnings (see Figure 1, Panel B, in Brewer and Hoynes, 2019). The results are

[59]Specifically, we consider an in-work credit based on family income, where the couples has to work at least 24 hours per week, with at least one of them working 16 hours per week, consistent with the WTC. The in-work credit is phased out at a steep rate of 37%, consistent with the WTC. The maximum amount is set so that the budgetary outlays are 100 million euro, consistent with the other simulations before.

[60]Simulating an EITC for low family incomes in the Netherlands would yield qualitatively similar results as the WTC simulation.

[61]Indeed, Bargain et al. (2014) find that intensive margin labor supply elasticities in the Netherlands are higher than in other countries.

[62]In this simulation couples with 1 child receive 146 euro, and couples with two or more children receive an additional 122 euro. The total budgetary outlays are again 100 million euro.

[63]For the average household, income effects are relatively low compared to substitution effects, see e.g. Bargain et al. (2014).

quite similar to column (1), a large drop in income inequality, but also a large drop in labor supply. This is to be expected, since in most couples with children at least one spouse is working. The decline in income inequality and labor supply is a bit bigger in column (3) than in column (1) because couples where both partners do not work also qualify, see also Figure O.3 in appendix O. Again, we should be careful with extrapolating results from one country to another, but these results suggest that there might be a substantial trade-off between income support for families with children and labor supply.

Our results also relate to the American Families Plan (AFP) of the Biden administration in the US.[64] The AFP plans to extend the EITC and CTC considered above. Furthermore, it also plans to increase government support for child care, targeted at low- and middle-income families, so that these families spend no more than seven percent of their income on child care. Our results on child care reforms (simulations (1) and (4) in Table 4) suggest that this may cause a substantial increase in labor supply of women in couples with children, as opposed to the expansion of the EITC. Furthermore, targeting the increase in child care subsidies more at low- and middle-income groups rather than across-the-board may not mitigate its effectiveness in terms of labor supply (compare the labor supply effect in simulations (1) and (4)).[65] The effects on income inequality will depend on the exact implementation of the proposal. Finally, also relevant is our point made earlier, that increasing child care subsidies is more effective in terms of the budgetary costs per additional labor supply unit (working persons or hours worked) when you start from a relatively low initial child care subsidy rate. Again, we should be careful translating results from one country to the other, but this suggests that increasing child care subsidies in the US may be relatively cost-effective, starting from a relative low child care subsidy rate.

5 Discussion and conclusion

In this study, we have estimated a structural model of labor supply and child care choices for couples with the youngest child aged 0–3 (pre-primary school age) and couples with the youngest child aged 4–11 (primary school age). Large exogenous variations in family tax policies benefit the identification of the structural parameters. The model accounts for unobserved heterogeneity, and produces labor supply responses to the family tax policy

[64] See e.g. https://www.whitehouse.gov/briefing-room/statements-releases/2021/04/28/fact-sheet-the-american-families-plan.

[65] Furthermore, studies that look at the impact of participation in formal child care on child development suggest that children in households with a relatively low income benefit more from participation in formal child care than children in households with a relatively high income Havnes and Mogstad (2015), which favors targeting child care subsidies at lower income groups.

reforms which are in line with outcomes of a difference-in-differences model.

We used the structural model to study the relative effectiveness of different types of fiscal policies, assessing their capacity to stimulate employment of parents with young children. The results show that the most effective fiscal stimulus for working parents is an in-work benefit targeted at secondary earners that rises with income. Child care subsidies prove to be less effective than in-work benefits for secondary earners, because substitution of other types of care for formal care drives up public expenditures. However, child care subsidies are still much more effective than in-work benefits that target both primary and secondary earners, as the labor supply of primary earners is rather unresponsive to additional financial incentives. The results also show that there is a trade-off between equity and efficiency of these fiscal stimuli: the effect on hours worked per additional public euro spent is bigger when the policy is targeted at middle- and high-income families. It is worth noting that this trade-off is less pronounced for child care subsidies than for in-work benefits.

Several limitations need to be acknowledged. First, our structural model is static in the sense that we do not consider potential lifecycle responses. Hence, our model could be considered a 'myopic' approximation of a true dynamic life cycle model as in Blundell et al. (2016).[66] The advantage of dynamic models is that they are able to incorporate the policy effects on career progression and habit formation (Bick, 2016; Bauernschuster et al., 2016), fertility patterns (Bick, 2016; Bauernschuster et al., 2016) and other dynamic processes influencing parental decision making, such as their marital stability. Naturally, some of these processes (career effects and habit formation) are likely to stimulate further growth of labor market participation and household incomes. Others are more complex. For example, higher labor force participation can empower women to leave toxic relationships, which can also stimulate labor market participation, but suppress household incomes. Subsidies that are linked to childbearing can also incentivize further fertility, although these policy effects are generally considered relatively small (e.g. Baker et al., 2008). Importantly, we believe that these dynamic effects are unlikely to distort the ranking of the reforms we considered in our counterfactual analyses, although this is yet to be seen and verified empirically.

Second, the model and its parametrization is necessarily specific to the Dutch social and institutional context. This can complicate the extrapolation of our findings to other national contexts, although we believe that the qualitative conclusions of our analyses should be meaningful for many developed countries. Third, it is also important to note

[66] And hence, we run the risk that the behavioral responses are misspecified when households are not myopic, which may result in biased estimates and invalid inferences.

that the promotion of parental labor participation is not the sole objective of child care subsidies and in-work benefits. The objectives of a policy maker can span multiple domains, including fostering child development and curbing income inequality, and this set of objectives can influence the relative desirability of the reforms analysed in this study. Indeed, if the key objective is the reduction of income inequality, then the policy maker will be better off using the policy designs akin to the EITC in the US or WTC in the UK. Alternatively, if the key objective is improving skill formation among disadvantaged children, then the preferred policy may very well be universal childcare.

References

Angrist, J. and Pischke, J.-S. (2009). Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press, Princeton.

Apps, P., Kabátek, J., Rees, R., and Van Soest, A. (2016). Labor supply heterogeneity and demand for child care of mothers with young children. Empirical Economics, 51(4):1641–1677.

Baker, M., Gruber, J., and Milligan, K. (2008). Universal child care, maternal labor supply, and family well-being. Journal of Political Economy, 116(4):709–745.

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

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

Bauernschuster, S., Hener, T., and Rainer, H. (2016). Children of a (policy) revolution: The introduction of universal child care and its effect on fertility. Journal of the European Economic Association, 14(4):975–1005.

Bernal, R. (2008). The effect of maternal employment and child care on children's cognitive development. International Economic Review, 9(4):1173–1209.

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

Bick, A. (2016). The quantitative role of child care and female labor force participation and fertility. Journal of the European Economic Association, 14(3):639–668.

Blau, D. (2003). Child care subsidy programs. In Moffitt, R., editor, Means-Tested Transfer Programs in the United States, pages 443–516. NBER.

Blau, D. and Currie, J. (2006). Preschool, day care, and after school care: Who's minding the kids? In Hanushek, E. and Welch, F., editors, Handbook of the Economics of Education, pages 1163–1278. Elsevier.

Blundell, R., Brewer, M., and Francesconi, M. (2008). Job changes and hours changes: understanding the path of labor supply adjustment. Journal of Labor Economics, 26:421–453.

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

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

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

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

Brewer, M. and Browne, J. (2006). The effect of the Working Families' Tax Credit on labour market participation. IFS Briefing Note no. 69.

Brewer, M., Francesconi, M., Gregg, P., and Grogger, J. (2009). Feature: In-work benefit reform in a cross-national perspective - Introduction. Economic Journal, 119:F1–F14.

Brewer, M. and Hoynes, H. (2019). In-work credits in the UK and the US. Fiscal Studies, 40(4):519–560.

Cascio, E. (2009). Maternal labor supply and the introduction of kindergartens into American public schools. Journal of Human Resources, 44(1):140–170.

Cattan, S. (2016). Can universal preschool increase the labor supply of mothers? IZA World of Labor, 312.

De Boer, H. and Jongen, E. (2017). Optimal income support for lone parents in the Netherlands: are we there yet? CPB Discussion paper 361, The Hague.

De Boer, H., Jongen, E., and Koot, P. (2018). Optimal Taxation of Secondary Earners in the Netherlands: Has Equity Lost Ground. CPB Discussion paper 375, The Hague.

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

Deaton, A. and Paxson, C. (1994). Intertemporal choice and inequality. Journal of Political Economy, 102(3):437–467.

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

European Union (2014). Taxation Trends in the European Union, 2014 Edition. European Union.

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

Fitzpatrick, M. (2012). Revising our thinking about the relationship between maternal labor supply and preschool. Journal of Human Resources, 47(3):583–612.

Francesoni, M., Rainer, H., and Van der Klaauw, W. (2009). The effects of in-work benefit reform in Britain on couples: theory and evidence. Economic Journal, 119:F66–F100.

Francesoni, M. and Van der Klaauw, W. (2007). The socioeconomic consequences of in-work benefit reform for British lone mothers. Journal of Human Resources, 42:1–31.

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

Gong, X. and Breunig, R. (2017). Child care assistance: Are subsidies or tax credits better? Fiscal Studies, 38(1):7–48.

Gong, X. and Van Soest, A. (2002). Family structure and female labor supply in Mexico City. Journal of Human Resources, 37(1):163–191.

Gregg, P., Harkness, S., and Smith, S. (2009). Welfare reform and lone parents in the UK. Economic Journal, 119:F38–65.

Hansen, J. and Liu, X. (2015). Estimating labor supply responses and welfare participation: Using a natural experiment to validate a structural model. Canadian Journal of Economics, 48(5):1831–1854.

Havnes, T. and Mogstad, M. (2011). Money for nothing? Universal child care and maternal employment. Journal of Public Economics, 95:1455–1465.

Havnes, T. and Mogstad, M. (2015). Is universal child care leveling the playing field? Journal of Public Economics, 127:100–114.

Heckman, J. and Singer, B. (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data. Econometrica, 52:271–320.

Hoynes, H. and Patel, A. (2016). Effective policy for reducing poverty and inequality? The Earned Income Tax Credit and the distribution of income. Journal of Human Resources, 53:859–890.

Immervoll, H. and Pearson, M. (2009). A good time for making work pay? Taking stock of in-work benefits and related measures across the OECD. IZA Policy Paper No. 3, Bonn.

Kabátek, J. (2013). Iteration capping for discrete choice models using the EM algorithm. Discussion Paper 2013-019, Tilburg University, Center for Economic Research.

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

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

Kleven, H. (2014). How can Scandinavians tax so much? Journal of Economic Perspectives, 28(4):77–98.

Kleven, H. (2021). The EITC and the extensive margin: a reappraisel. NBER Working Paper 26405, Cambridge.

Kornstad, T. and Thoresen, T. (2006). Effects of family policy reforms in Norway: Results from a joint labour supply and childcare choice microsimulation analysis. Fiscal Studies, 27(3):339–371.

Kornstad, T. and Thoresen, T. (2007). A discrete choice model for labor supply and childcare. Journal of Population Economics, 20:781–803.

Kurowska, A., Myck, M., and Wrohlich, K. (2017). Making work pay: increasing labour supply of secondary earners in low income families with children. Contemporary Economics, 11(2):161–170.

Lefebvre, P. and Merrigan, P. (2008). Child-care policy and the labor supply of mothers with young children: A natural experiment from Canada. Journal of Labor Economics, 26(3):519–48.

Loeffler, M., Peichl, A., and Siegloch, S. (2014). Structural labor supply models and wage exogeneity. IZA Discussion Paper 8281, Bonn.

Lokshin, M. (2004). Household childcare choices and women's work behavior in Russia. Journal of Human Resources, 39(4):1094–1115.

Mastrogiacomo, M., Bosch, N., Gielen, M., and Jongen, E. (2017). Heterogeneity in Labour Supply Responses: Evidence from a Major Tax Reform. Oxford Bulletin of Economics and Statistics, 79(5):769–796.

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

Meyer, B. (2010). The effects of the Earned Income Tax Credit and recent reforms. In Brown, J., editor, Tax Policy and the Economy. University of Chicago Press.

Meyer, B. and Rosenbaum, D. (2001). Welfare, the Earned Income Tax Credit, and the labour supply of single mothers. Quarterly Journal of Economics, 116:1063–1114.

Ministry of Finance (2010). Het kind van de regeling, rapport Brede Heroverwegingen 5. Ministry of Finance, The Hague.

Mundlak, Y. (1978). On the pooling of time series and cross section data. Econometrica, 46(1):69–85.

OECD (2015a). Labour Force Statistics. OECD, Paris.

OECD (2015b). OECD Family Database. OECD, Paris.

Pacifico, D. (2013). On the role of unobserved preference heterogeneity in discrete choice models of labour supply. Empirical Economics, 45(2):929–963.

Romijn, G., Goes, J., Dekker, P., Gielen, M., and van Es, F. (2008). MIMOSI: Microsimulatiemodel voor belastingen, sociale zekerheid, loonkosten en koopkracht. CPB Document 161, The Hague.

Saez, E. (2002). Optimal income transfer programs: Intensive versus extensive labor supply responses. Quarterly Journal of Economics, 117(3):1039–1073.

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

Tekin, E. (2007). Child care subsidies, wages, and employment of single mothers. Journal of Human Resources, 42(2):453–487.

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

Thoresen, T. and Vatto, T. (2019). An up-to-date joint labor supply and child care choice model. European Economic Review, 112:51–73.

Todd, P. and 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. American Economic Review, 95(5):1384–1417.

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

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

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

Vermeulen, F. (2005). And the winner is... an empirical evaluation of unitary and collective labour supply models. Empirical Economics, 30(3):711–734.

Online appendix

A Wage equations

The wages used in the structural model are defined as follows. For employed parents, we use observed wages. For non-employed parents we use imputed wages. To this end, we estimate a series of Mincerian wage equations, splitting workers by their gender and education level. To account for unobserved individual-specific effects, we use panel data techniques. Specifically, we use the quasi-fixed effects model of (Mundlak, 1978).

To account for selection into employment, we first estimate the probability of participation using a pooled probit regression

$$p_{it} = x_{it}'\gamma + z_{it}'\theta + \nu_{it}, \tag{A.1}$$

where vector z_{it} contains variables that are expected to have an effect on the probability of participation but not on wages (exclusion restrictions). From this regression we determine the inverse Mills' ratio

$$invmills_{it} = \phi(p_{it})/\Phi(p_{it}). \tag{A.2}$$

The inverse Mills' ratio is then included in the quasi-fixed effects model

$$ln(w_{it}) = x_{it}'\beta + \omega_i + \bar{x_i}'\pi + \lambda_t invmills_{it} + \epsilon_{it}, \tag{A.3}$$

where the time-invariant individual specific effect consists of a random part, ω_i with $\sim IID(0, \sigma_\omega^2)$, and a part which is allowed to be correlated with regressors $\bar{x_i}'\pi$. Here, $\bar{x_i}$ is the average of time-varying variables. A significant coefficient for an element of π provides evidence that the individual specific effect is correlated with one of the regressors.

Table A.1 shows estimation results for all subgroups. We use age splines since we expect that the relationship between wage and age is nonlinear. Table A.1 shows that the wage increases with age but at a diminishing rate. We see that the age profile is steeper for higher educated individuals. We also include cohort and year dummies in the regression. Because of perfect collinearity between age, cohort and period we use transformed time dummies following Deaton and Paxson (1994), where all time effects are assumed to be transitory.[67] Year dummies are significant in most specifications while the cohort variables are jointly significant for most subgroups. Wages are lower on average for non-Western

[67] As a result of the Deaton and Paxon transformation we cannot include time dummies for the years 2006 and 2007, the time effects for 2006 and 2007 are calculated manually: t2006=–(t2007+t2008+t2009) and t2007=–2*t2008–3*t2009.

Table A.1: Results wage equations

	Women			Men		
	Lower educ.	Middle educ.	Higher educ.	Lower educ.	Middle educ.	Higher educ.
Age effect						
18-30	0.037***	0.039***	0.056***	0.039***	0.047***	0.054***
31-40	0.023***	0.024***	0.035***	0.020***	0.030***	0.045***
41-50	0.025***	0.019***	0.024***	0.019***	0.021***	0.028***
51-63	-0.006	-0.014	0.004	0.006	0.016***	0.016***
Cohort effect[a]						
1980-1989	0.114***	0.109***	0.232***	0.069*	0.106***	0.124***
1975-1980	0.043*	0.080***	0.174***	-0.005	0.056***	0.108***
1970-1975	0.024*	0.041***	0.118***	0.001	0.026***	0.079***
1960-1965	-0.035	-0.024**	-0.033**	-0.025*	-0.024**	-0.025**
1955-1960	-0.009	-0.069***	-0.070**	-0.038	-0.058***	-0.073***
<1955	0.268	0.009	-0.314***	-0.083**	-0.027	-0.062**
Year effect						
t2006	0.007	0.002	-0.003	0.011	0.006	0.003
t2007	-0.009	-0.003	-0.001	-0.012	-0.009	-0.004
t2008	-0.003	0.000	0.011***	-0.008***	0.000	-0.001
t2009	0.005**	0.001	-0.007**	0.010***	0.003***	0.002*
Ethnicitity						
Western immigrant	-0.011	-0.055***	-0.188***	-0.004	-0.008	-0.012
Non-western immigrant	-0.062**	-0.170***	-0.386***	-0.051**	-0.195***	-0.260***
Partner						
married	-0.001	-0.037***	-0.123***	-0.001	0.004	0.017***
Mundlak age averages						
18-30	-0.009	-0.009*	0.002	-0.004	-0.004	0.001
31-40	-0.012***	-0.008***	-0.005**	-0.010***	-0.006***	0.001
41-50	-0.011**	-0.008***	-0.014***	-0.008**	-0.008***	-0.012***
51-63	-0.023	-0.056**	-0.007	-0.001	-0.020***	-0.022***
Province[a]						
Groningen	-0.046***	-0.044***	-0.052***	-0.132***	-0.148***	-0.123***
Friesland	-0.036*	-0.079***	-0.102***	-0.116***	-0.145***	-0.141***
Drenthe	-0.056***	-0.046***	-0.115***	-0.107***	-0.137***	-0.089***
Overijssel	-0.060***	-0.054***	-0.080***	-0.110***	-0.112***	-0.113***
Flevoland	-0.058***	-0.047***	-0.109***	-0.029*	-0.050***	-0.079***
Gelderland	-0.051***	-0.040***	-0.074***	-0.076***	-0.077***	-0.081***
Noord-Brabant	-0.033	-0.074***	-0.116***	-0.047***	-0.064***	-0.079***
Zeeland	-0.047***	-0.041***	-0.051***	-0.048***	-0.069***	-0.056***
Limburg	-0.051***	-0.055***	-0.037***	-0.097***	-0.132***	-0.137***
Urban area						
>150.000 inhabitants	0.020***	0.025***	0.057***	-0.014**	-0.009*	0.009**
Inverse Mills' ratio	-0.027	0.225***	1.061***	-0.485***	-0.010	0.127
Constante	1.520***	1.529***	0.852***	1.642***	1.396***	1.181***
σ_ϵ wage equation	0.154	0.203	0.249	0.234	0.266	0.322
Observations	16,636	61,652	40,241	30,138	64,216	52,853
Number of individuals	6,277	21,667	13,878	10,574	21,975	17,916

*** p<0.01. ** p<0.05. * p<0.1. Reference group: born in 1965–1970. native and living in the Western region of the Netherlands.

Table A.2: Testing significance exclusion restrictions

	Women			Men		
	Lower educ.	Middle educ.	Higher educ.	Lower educ.	Middle educ.	Higher educ.
Youngest child 0–3 yrs	-0.099***	0.029	-0.038	0.030	0.005	-0.009
Partner lower educated	-0.124***	0.098***	0.071	-0.086	0.080*	-0.041
Partner middle educated	0.016	0.162***	0.188***	0.224***	0.108***	0.066
Wald	47.540***	77.210***	46.810***	65.400***	8.960**	4.980

immigrants. The coefficients for the Mundlak age averages are jointly significant in all specifications, but have no straightforward economic interpretation.

The lower part of Table A.1 shows that the inverse Mills' ratio is significant for most groups, emphasizing the importance of selection into employment. We also include an attrition indicator in order to test for the presence of attrition bias.[68] The indicator proves insignificant in all subgroups.

As exclusion restrictions, we use the presence of young children aged between 0 and 3 years in the participation equation. Here, we follow the approach by Van Soest (1995) and Bargain et al. (2014). In addition, we use the education level of the partners, similar to Blundell et al. (2007). The exclusion restrictions are jointly significant, as indicated by the Wald test in Table A.2.

[68]The attrition indicator is a dummy which equals 1 if an individual leaves the sample in our data period 2006–2009.

B Price equations formal child care

For families who do not use formal child care we simulate a price for child care. Here, a distinction is made between the gross hourly price of daycare (children 0–3 years of age) and the gross hourly price of after school care (children 4–11 years of age). For the gross hourly child care prices we also estimated a Heckman selection model. However, the parameter of the inverse Mill's ratio in the price equation was insignificant and close to zero. Therefore, our preferred specification does not include the inverse Mill's ratio.

For gross hourly child care prices we estimate a random effects model for the prices of daycare and after school care. We estimate the following price equation:

$$p_{it} = x'_{it}\beta + \omega_i + \epsilon_{it} \tag{B.1}$$

where the the random effects are captured by ω_i with $\sim IID(0, \sigma^2_\omega)$. Our dependent variable is the hourly real price (deflated by the consumer price index).

We focus on households since child care is consumed at the household level. As it turns out, characteristics of mothers are more important in predicting the use and gross price of child care than characteristics of fathers. Hence, we only include characteristics of mothers in the regressions.

Table B.1 shows estimation results for daycare and after school care.[69] Estimation results show that year dummies are increasing for daycare and decreasing for after school care. Households with higher educated women or older women pay a higher price on average. We also find some heterogeneity across provinces, and prices are higher in large cities.

[69] Including ethnicity, a dummy for age of the youngest child or a dummy for multiple children one at a time, leads to insignificant coefficients for each of these variables.

Table B.1: Results price equations formal child care

	Daycare	After school care
Year effect		
2007	0.051***	-0.004
2008	0.109***	-0.017**
2009	0.126***	-0.034***
Higher educated	0.005	0.027***
Age	0.039***	0.018*
Age squared	-0.001***	0.000**
Province		
Groningen	0.044**	0.053**
Friesland	0.157***	0.148***
Drenthe	0.162***	0.127***
Overijssel	0.124***	0.188***
Flevoland	-0.234***	-0.121***
Gelderland	0.124***	0.172***
Noord-Brabant	-0.025	-0.060
Zeeland	-0.030***	0.014
Limburg	-0.019	0.119***
Urbanisation		
>150.000 inhabitants	0.090***	0.047***
Constant	4.719***	5.277***
Observations	32.978	26.447
Individuals	13.705	11.023

[a] Reference group: lower educated women in couples.
*** p<0.01. ** p<0.05. * p<0.1.

C Likelihood base model

The corresponding (log-)likelihood function has no closed-form solution and therefore we approximate it using the maximum simulated likelihood method. The functional form is as follows:

$$\mathcal{L} = \sum_{i=1}^{I} \log \left(\frac{1}{R} \sum_{r=1}^{R} \sum_{j=1}^{J} D_{ij} \left(\frac{\exp\left(U^s_{ij}(\nu_r, \gamma_r)\right)}{\sum_{j'=1}^{J} \exp\left(U^s_{ij'}(\nu_r, \gamma_r)\right)} \right) \right), \tag{C.1}$$

where D_{ij} is 1 for the observed choice, and zero otherwise. To approximate the aforementioned heterogeneities (of wages, child care prices, and preferences), the likelihood function is integrated over $R = 50$ draws from the corresponding distributions.[70] This means that the utilities U^s_{ij} presented in Equation 6 are specific to the given draw, which we reflect by the fact that the vectors ν and γ are subscripted by r.

D Likelihood latent class model

With the latent class model we estimate a finite mixture model with K parameterizations of the utility function. All the preference parameters therefore become class-specific, which is equivalent to the assumption of Heckman and Singer (1984) that the parameters are drawn from a mass-point distribution. The full set of parameters to be estimated is then:

$$\theta = (\theta_1, ..., \theta_K) = (\mathbf{A}_1, \mathbf{b}_1, \mathbf{d}_1, ..., \mathbf{A}_K, \mathbf{b}_K, \mathbf{d}_K). \tag{D.1}$$

Since the classes are by definition unobservable, we cannot determine whether a given household belongs to a specific class or not. Instead, we have to construct household-level probabilities of class membership $P_i(class = k)$, which reflect the likelihood that the household i exhibits preferences corresponding to class k. These probabilities are conditional on the household's choices and other observable characteristics, and they are used as individual weights for a set of k multinomial logit models with separate parameter vectors θ_k. Again, the estimation is a repeated cross-section although we use the panel dimension to ensure that the probabilities of the class membership do not vary over time.

[70] The number of draws proves sufficient to ensure consistency of our estimates. Increasing the number of draws from 10 to 50 did not change the predictions of our model, see Table J.7 in the online appendix.

The resulting (log-)likelihood function of the finite mixture model has the following form:

$$\mathcal{L} = \sum_{i=1}^{I} \log \left(\frac{1}{R} \sum_{r=1}^{R} \sum_{k=1}^{K} P_i(class = k) \sum_{j=1}^{J} D_{ij} \left(\frac{\exp\left(U_{ij}^{s}(\nu_r, \theta_k)\right)}{\sum_{j'=1}^{J} \exp\left(U_{ij'}^{s}(\nu_r, \theta_k)\right)} \right) \right), \tag{D.2}$$

where R again denotes the number of draws from the estimated wage and price equation for non-workers and non-users of formal child care, as before. The number of draws in our latent classes specifications is 10, and it is kept relatively low to further limit the computational complexity of the model. The model is estimated using Expectation-Maximization (EM) algorithm.[71] In order to assess how many latent classes should be used, we have estimated a set of models allowing for 1 (no unobserved heterogeneity), 2 or 3 latent classes.

[71] Train (2008) has shown that the EM algorithm offers a tractable way of estimating latent class discrete choice models. For a discussion of the benefits of latent class models within the domain of structural labor supply modelling, see Apps et al. (2016). For an overview of their implementation and potential computational improvements, see Kabátek (2013).

E Descriptive statistics data for the structural model

Table E.1 shows that fathers in our sample are on average a few years older than mothers. Both mothers and fathers are predominantly born in the Netherlands, and most of them have a level of education classified as middle. Fathers with the youngest child 0–3 years of age are slightly more likely to be higher educated than fathers with the youngest child 4–11 years of age. However, mothers with the youngest child 0–3 years of age are considerably more likely to be higher educated than mothers with the youngest child 4–11 years of age. The majority of couples lives in smaller cities and towns (<150,000 inhabitants). There is a considerable gap in the gross hourly wage between fathers and mothers, with fathers earning on average 4 and 6 euro per hour more than mothers in couples with the youngest child 0–3 and 4–11 years of age, respectively. The participation rate is higher for fathers than for mothers. Furthermore, the participation rate of mothers with the youngest child 0–3 years of age is higher than the participation rate of mothers with the youngest child 4–11 years of age, the higher education level of mothers with the youngest child 0–3 years of age contributes to this difference. Finally, couples with the youngest child 0–3 years of age are more likely to use formal child care than couples with older children. 50% of the couples with the youngest child 0–3 years of age sends their children to formal child care, compared to 13% for couples with the youngest child 4–11 years of age.[72] A typical school day is from 8:30 to 15:00, and many families are able to cover the remaining hours with parental time or informal care. This is also reflected in the average hours of formal child care used per week by couples that do use formal child care, which is much lower for couples with the youngest child 4–11 years of age.

[72]The share of households using formal child care is higher than the share of children in formal child care in Figure 1. Figure 1 also includes children from households who are not eligible for child care subsidies such as households with disability benefits. Furthermore, households that have many children typically use less formal child care on average.

Table E.1: Descriptive statistics by gender and age of young. child (2006–2009)

	Women				Men			
	0-3 yrs		4-11 yrs		0-3 yrs		4-11 yrs	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age	34.1	4.37	40.9	4.57	36.8	4.86	43.3	5.11
Native	0.84	0.37	0.84	0.37	0.84	0.36	0.86	0.35
Western immigrant	0.09	0.28	0.09	0.28	0.08	0.27	0.07	0.26
Non-western immigrant	0.08	0.27	0.07	0.26	0.08	0.26	0.07	0.25
Lower educated	0.14	0.34	0.22	0.42	0.19	0.39	0.21	0.41
Middle educated	0.46	0.50	0.51	0.50	0.44	0.50	0.44	0.50
Higher educated	0.40	0.49	0.26	0.44	0.38	0.48	0.35	0.48
Large city	0.16	0.37	0.16	0.36	0.16	0.37	0.16	0.36
Small city	0.84	0.37	0.84	0.36	0.84	0.37	0.84	0.36
Hourly gross wage	16.3	6.30	16.1	7.60	20.2	10.0	22.2	11.2
Participation rate	0.82	0.39	0.75	0.43	0.96	0.19	0.95	0.21
Hours worked per week	23.0	8.20	21.2	8.53	38.7	5.20	38.7	5.48
Using formal child care	0.50	0.50	0.14	0.34	0.50	0.50	0.14	0.34
Hours formal child care per week	27.1	16.2	14.4	11.1	27.1	16.2	14.4	11.1
Observations	4,170		5,013		4,170		5,013	

[a]Education is classified as follows (using the Dutch abbreviations): i) lower educated = BO and VMBO, ii) middle educated = MBO, HAVO and VWO, iii) higher education = HBO and WO. [b]A city is defined as large (small) when it has 150,000 inhabitants or more (less than 150,000 inhabitants). [c]Hours worked per week per employed. [d]Hours of formal child care per week per couple using formal child care.

F Estimates model with random preference heterogeneity

Table F.1: Estimated preferences by age of the youngest child

	Youngest child 0–3	Youngest child 4–11
Income	14.990***	3.538***
Income^2	-2.225***	0.865***
Leisure woman	-12.320***	-24.000***
X (age woman-38)/10[a]	1.981***	0.965***
X (age woman-38)^2/100	2.173***	1.080***
Leisure woman^2	-149.300***	-113.600***
St. dev. random pref. heterogeneity leisure woman	0.665***	0.876***
Leisure man	-71.910***	-96.350***
X (age man-38)/10	1.071***	1.550***
X (age man-38)^2/100	-0.603***	2.040***
Leisure man^2	-121.900***	-149.500***
St. dev. random pref. heterogeneity leisure man	2.772***	7.828***
Fixed costs of work woman	-2.515***	-1.617***
X 1(low educated woman)	-0.237***	-0.544***
X 1(medium educated woman)	0.356***	-0.223***
X 1(non-Western immigrant woman)	-1.308***	-0.586***
X 1(Western immigrant woman)	-0.597***	-0.156***
Fixed costs of work man	-10.500***	-10.730***
X 1(low educated man)	1.177***	0.847***
X 1(medium educated man)	1.516***	1.262***
X 1(non-Western immigrant man)	-0.745***	-1.804***
X 1(Western immigrant man)	-1.756***	-1.193***
Hours of formal child care	-2.042***	-2.026***
X 1(non-Western immigrant woman)	0.990***	0.225***
X 1(Western immigrant woman)	0.262***	-0.706***
X 1(non-Western immigrant man)	-0.100***	0.720***
X 1(Western immigrant man)	0.328***	0.899***
X 1(>=150.000 inhabitants)	0.806***	-0.330***
Hours of formal child care^2	-0.015**	-0.795***
St. dev. random pref. heterogeneity hours child care[b]	–	0.327***
Fixed costs of child care	0.525***	-2.596***
X 1(low educated woman)	-0.933***	-1.116***
X 1(medium educated woman)	-0.531***	-0.359***
X 1(non-Western immigrant woman)	-1.398***	0.311***
X 1(Western immigrant woman)	-0.123***	0.749***
X 1(low educated man)	-0.384***	-0.126***
X 1(medium educated man)	-0.364***	-0.233***
X 1(non-Western immigrant man)	-0.300***	-1.320***
X 1(Western immigrant man)	-0.256***	-0.968***
X 1(>=150.000 inhabitants)	-1.204***	0.799***
Income X leisure woman	-7.561***	0.932***
Income X leisure man	2.772***	7.496***
Leisure man X leisure woman	-21.870***	-16.710***
Income X hours of formal child care	0.530***	0.697***
Leisure woman X hours of formal child care	-6.978***	-7.934***
Leisure man X hours of formal child care	0.823***	-3.546***

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. [a] We standardize age by subtracting 38 and dividing by 10. [b] The parameter was not significant for couples with the youngest child 0–3 years of age (excluded from this model).

G Fit structural model

Fit labor supply women

Figure G.1: Age youngest child 0–3 yrs

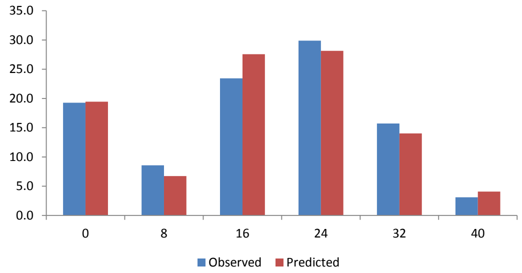
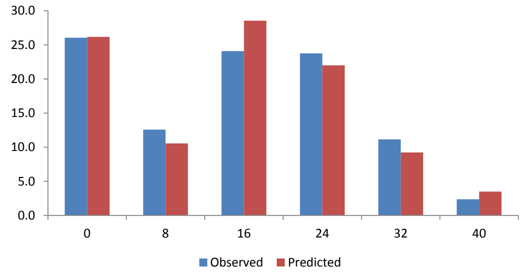


Figure G.2: Age youngest child 4–11 yrs



Fit labor supply men

Figure G.3: Age youngest child 0–3 yrs

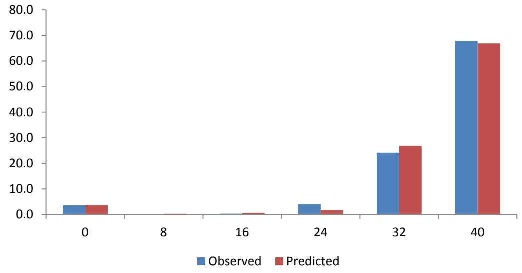
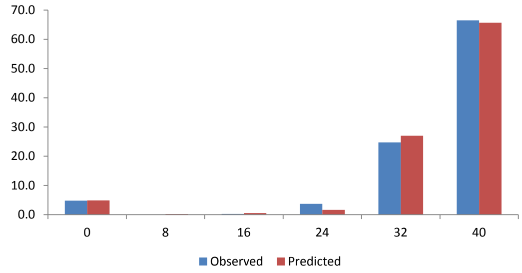


Figure G.4: Age youngest child 4–11 yrs



Fit formal child care use

Figure G.5: Age youngest child 0–3 yrs

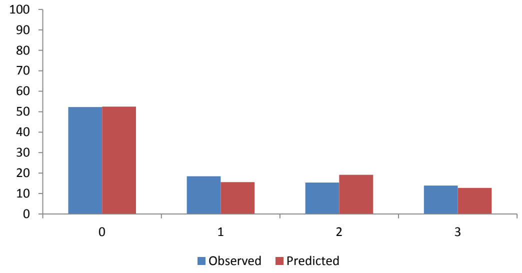


Figure G.6: Age youngest child 4–11 yrs

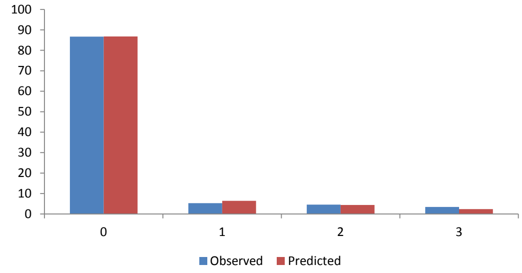


Table G.1: Model fit joint outcomes labor supply

Couples youngest child 0-3

Predicted probabilities (shares in %)								
		Labor supply women						
		0	8	16	24	32	40	
Labor supply men	0	1.8	0.1	0.5	0.4	0.2	0.2	3.3
	8	0.1	0.0	0.0	0.1	0.0	0.0	0.2
	16	0.1	0.0	0.0	0.1	0.1	0.1	0.6
	24	0.3	0.1	0.3	0.4	0.5	0.3	1.8
	32	3.3	0.9	6.1	9.7	5.8	0.6	26.4
	40	12.9	5.4	21.2	18.4	7.0	2.8	67.7
		18.5	6.6	28.2	29.1	13.5	4.0	100.0
Observed probabilities (shares in %)								
		Labor supply women						
		0	8	16	24	32	40	
Labor supply men	0	1.8	0.2	0.3	0.4	0.4	0.1	3.3
	8	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	16	0.0	0.0	0.1	0.1	0.1	0.0	0.3
	24	0.3	0.2	0.9	1.5	1.0	0.2	4.2
	32	3.0	1.8	5.8	8.1	4.6	1.0	24.3
	40	13.5	6.3	16.5	20.0	9.8	1.8	67.9
		18.7	8.6	23.5	30.1	15.9	3.2	100.0

Couples youngest child 4-11

Predicted probabilities (shares in %)								
		Labor supply women						
		0	8	16	24	32	40	
Labor supply men	0	2.0	0.1	0.6	0.5	0.5	0.5	4.2
	8	0.1	0.0	0.0	0.0	0.0	0.1	0.2
	16	0.1	0.1	0.1	0.2	0.0	0.1	0.5
	24	0.3	0.2	0.5	0.4	0.3	0.0	1.7
	32	5.3	2.4	7.9	7.1	3.6	0.6	26.9
	40	17.9	7.9	19.9	14.1	4.6	2.1	66.5
		25.6	10.7	29.0	22.3	9.0	3.4	100.0
Observed probabilities (shares in %)								
		Labor supply women						
		0	8	16	24	32	40	
Labor supply men	0	1.6	0.5	0.7	0.9	0.7	0.3	4.6
	8	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	16	0.0	0.0	0.1	0.1	0.0	0.0	0.3
	24	0.6	0.4	1.0	1.1	0.6	0.1	3.8
	32	5.3	3.0	6.3	6.6	3.1	0.7	24.9
	40	18.2	8.8	16.2	15.2	6.7	1.3	66.4
		25.7	12.7	24.2	23.9	11.1	2.4	100.0

H Estimates child care price elasticities

Table H.1: Gross and net price of formal child care elasticities

	Price of formal child care +1%			
	0–3 yrs		4–11 yrs	
	Mean	SE	Mean	SE
Gross price elasticities				
Formal child care	–0.748	0.004	–0.415	0.003
Labor supply women	–0.151	0.001	–0.022	0.000
Labor supply men	–0.002	0.000	–0.001	0.000
Net price elasticities				
Formal child care	–0.434	0.004	–0.252	0.003
Labor supply women	–0.087	0.001	–0.014	0.000
Labor supply men	–0.001	0.000	–0.001	0.000

Notes: Bootstrapped standard errors based on 200 draws of the estimated preference parameters. The gross price of formal child care elasticities relate the percentage change in the use of formal child care and labor supply by women and men to the percentage change in the full price of formal child care. The net price of formal child care elasticities relate the percentage change in the use of formal child care and labor supply by women and men to the percentage change in the parental fee for formal child care.

Table H.1 presents the price elasticities of formal child care. We consider both gross and net prices of child care. We see a substantial negative gross price elasticity of formal child care: –0.75 for couples with the youngest child 0–3 years of age, and –0.42 for couples with the youngest child 4–11 years of age. There is hardly any effect on the labor supply of men, but a significant negative effect on the labor supply of women. This is particularly true for women with the youngest child 0–3 years of age, who use formal child care much more than women with the youngest child 4–11 years of age. The net price elasticities of child care are lower than the gross price elasticities, which is caused by the cap on the maximum price of child care.[73] The net elasticities are more comparable to other studies, which typically focus on the elasticity with respect to the parental fee.[74]

[73] For families who pay for childcare at prices that are above the cap, a 10% increase of the gross childcare price is more costly than a 10% increase of the net childcare price. This is because the latter is moderated by the capped subsidy, whereas the former is not. Without the cap on the maximum price of child care, the gross and net elasticities would be equivalent.

[74] Our results for the net price elasticity of labor supply by women are in line with the review presented in Blau (2003, p. 492). For the studies that explicitly allow for a formal child care choice next to a labor supply choice, and hence do not impose a 1-to-1 link between the two, the elasticity of labor supply of women with respect to the net price of formal child care is relatively low, ranging from –0.09 to –0.20. For mothers with the youngest child 0–3 years of age, we find a similar low elasticity of –0.09. For mothers with the youngest child 4–11 the elasticity is even lower (–0.01), which is partly a result of the lower share of women using formal care in this group.

I Robustness check: out-of-sample prediction 2009

Fit labor supply women 2009

Figure I.1: Age youngest child 0–3 years

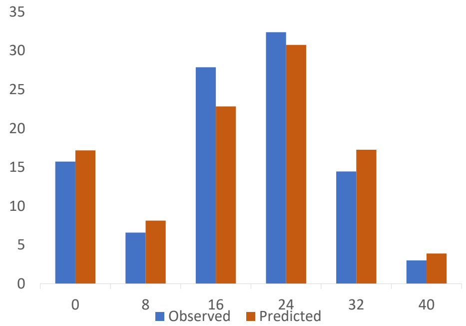
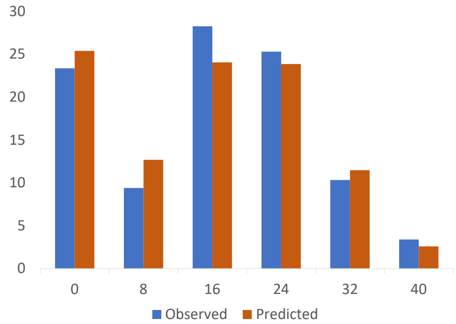


Figure I.2: Age youngest child 4–11 years



Fit labor supply men 2009

Figure I.3: Age youngest child 0–3 years

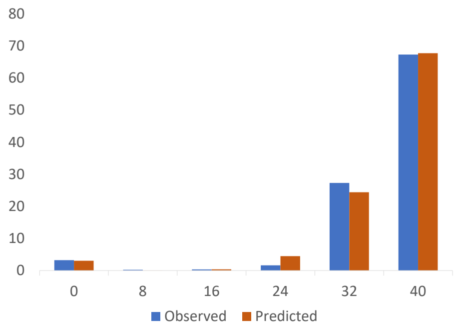
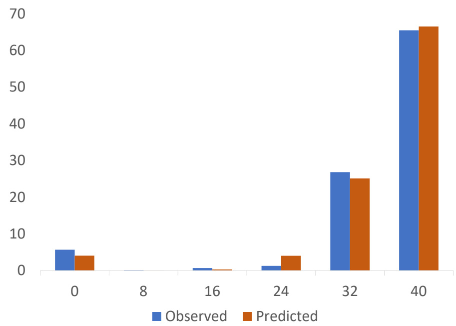


Figure I.4: Age youngest child 4–11 years



Fit formal child care use 2009

Figure I.5: Age youngest child 0–3 years

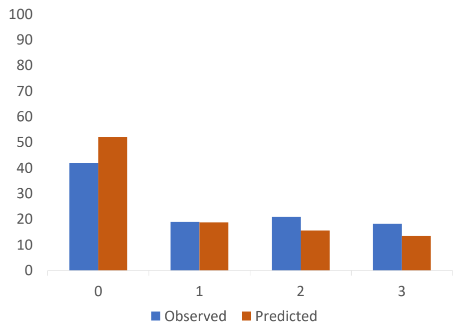
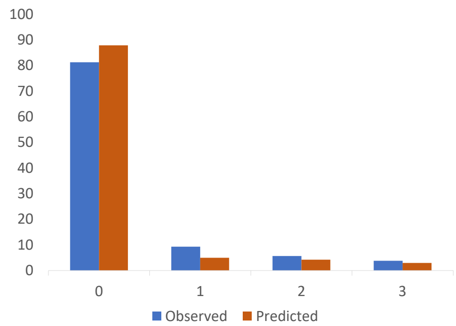


Figure I.6: Age youngest child 4–11 years



Note: only using data for 2006–2008.

J Robustness check: labor supply elasticities alternative models

Table J.1: Elasticities couples 0–3: repeated cross-section vs. cross-section

	Baseline		2006		2007		2008		2009	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Labor supply women	0.425	0.002	0.469	0.005	0.641	0.005	0.275	0.005	0.341	0.004
-Extensive margin	0.276	0.001	0.322	0.003	0.407	0.003	0.186	0.003	0.223	0.003
-Intensive margin	0.144	0.001	0.143	0.002	0.225	0.002	0.088	0.002	0.115	0.002
Labor supply men	0.082	0.001	0.078	0.002	0.074	0.002	0.069	0.002	0.069	0.002
-Extensive margin	0.071	0.001	0.067	0.001	0.067	0.001	0.065	0.001	0.060	0.001
-Intensive margin	0.010	0.000	0.011	0.000	0.007	0.001	0.004	0.001	0.009	0.001
Gross price elasticity child care	-0.748	0.004	-0.914	0.009	-1.067	0.007	-0.518	0.006	-0.664	0.008

Table J.2: Elasticities couples 4–11: repeated cross-section vs. cross-section

	Baseline		2006		2007		2008		2009	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Labor supply women	0.411	0.003	0.311	0.006	0.414	0.003	0.415	0.003	0.423	0.003
-Extensive margin	0.269	0.002	0.222	0.004	0.271	0.002	0.274	0.002	0.276	0.002
-Intensive margin	0.138	0.001	0.088	0.002	0.139	0.001	0.138	0.001	0.142	0.001
Labor supply men	0.080	0.001	0.044	0.001	0.082	0.001	0.083	0.001	0.083	0.001
-Extensive margin	0.069	0.001	0.039	0.001	0.072	0.001	0.073	0.001	0.073	0.001
-Intensive margin	0.011	0.000	0.005	0.001	0.010	0.000	0.011	0.000	0.011	0.000
Gross price elasticity child care	-0.415	0.003	-1.154	0.006	-0.371	0.002	-0.396	0.005	-0.395	0.008

Table J.3: Elasticities: test set vs. training set

	Couples with children 0–3				Couples with children 4–11			
	Training set		Test set		Training set		Test set	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Labor supply women	0.425	0.002	0.436	0.002	0.411	0.003	0.412	0.003
-Extensive margin	0.276	0.001	0.289	0.001	0.269	0.002	0.268	0.002
-Intensive margin	0.144	0.001	0.142	0.001	0.138	0.001	0.140	0.001
Labor supply men	0.082	0.001	0.085	0.001	0.080	0.001	0.084	0.001
-Extensive margin	0.071	0.001	0.075	0.001	0.069	0.001	0.073	0.001
-Intensive margin	0.010	0.000	0.010	0.000	0.011	0.000	0.011	0.000
Gross price elasticity child care	-0.748	0.004	-0.759	0.004	-0.415	0.003	-0.406	0.003

Table J.4: Elasticities: without and with time dummies

	Couples with children 0–3				Couples with children 4–11			
	No time dummies		Time dummies		No time dummies		Time dummies	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Labor supply women	0.425	0.002	0.436	0.002	0.408	0.003	0.408	0.003
-Extensive margin	0.276	0.001	0.289	0.001	0.273	0.002	0.269	0.002
-Intensive margin	0.144	0.001	0.142	0.001	0.131	0.001	0.135	0.001
Labor supply men	0.082	0.001	0.085	0.001	0.069	0.001	0.079	0.001
-Extensive margin	0.071	0.001	0.075	0.001	0.061	0.001	0.069	0.001
-Intensive margin	0.010	0.000	0.010	0.000	0.008	0.000	0.010	0.000
Gross price elasticity child care	-0.748	0.004	-0.753	0.004	-0.415	0.003	-0.400	0.003

Table J.5: Elasticities: model with vs. without child care

	Couples with children 0–3				Couples with children 4–11			
	With child care		Without child care		With child care		Without child care	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Labor supply women	0.425	0.002	0.451	0.002	0.411	0.003	0.361	0.003
-Extensive margin	0.276	0.001	0.285	0.001	0.269	0.002	0.235	0.002
-Intensive margin	0.144	0.001	0.162	0.001	0.138	0.001	0.123	0.001
Labor supply men	0.082	0.001	0.082	0.001	0.080	0.001	0.080	0.001
-Extensive margin	0.071	0.001	0.071	0.001	0.069	0.001	0.070	0.001
-Intensive margin	0.010	0.000	0.010	0.000	0.011	0.000	0.001	0.000
Gross price elasticity child care	-0.748	0.004	–	–	-0.415	0.003	–	–

Table J.6: Elasticities: model with vs. without cross effects leisure/child care

	Couples with children 0–3				Couples with children 4–11			
	With cross effects		Without cross effects		With cross effects		Without cross effects	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Labor supply women	0.425	0.002	0.502	0.003	0.411	0.003	0.389	0.003
-Extensive margin	0.276	0.001	0.329	0.001	0.269	0.002	0.250	0.002
-Intensive margin	0.144	0.001	0.168	0.001	0.138	0.001	0.136	0.001
Labor supply men	0.082	0.001	0.078	0.001	0.080	0.001	0.081	0.001
-Extensive margin	0.071	0.001	0.066	0.001	0.069	0.001	0.071	0.001
-Intensive margin	0.010	0.000	0.013	0.000	0.011	0.000	0.010	0.000
Gross price elasticity child care	-0.748	0.004	-0.980	0.005	-0.415	0.003	0.0378	0.003

Table J.7: Elasticities w/o unobserved heterogeneity and for different number of draws

	No unobserved heterogeneity 10 draws	Random preference heterogeneity 10 draws	Random preference heterogeneity 50 draws
Couples youngest child 0–3 years of age			
Gross hourly wage women +1%			
Labor supply women	0.40	0.40	0.43
– Extensive margin	0.26	0.26	0.28
– Intensive margin	0.13	0.13	0.14
Gross hourly wage men +1%			
Labor supply men	0.09	0.09	0.08
– Extensive margin	0.08	0.08	0.07
– Intensive margin	0.01	0.01	0.01
Gross price formal child care +1%			
Child care	–0.69	–0.70	–0.75
Couples youngest child 4–11 years of age			
Gross hourly wage women +1%			
Labor supply women	0.36	0.36	0.41
– Extensive margin	0.24	0.24	0.27
– Intensive margin	0.12	0.12	0.14
Gross hourly wage men +1%			
Labor supply men	0.09	0.09	0.08
– Extensive margin	0.08	0.08	0.07
– Intensive margin	0.01	0.01	0.01
Gross price formal child care +1%			
Child care	–0.37	–0.37	–0.42

Table J.8: Elasticities by number of latent classes: youngest child 0–3 yrs

	1 LC	2 LC	3 LC
Gross hourly wage women +1%			
Labor supply women	0.40	0.32	0.39
– Extensive margin	0.26	0.17	0.26
– Intensive margin	0.13	0.15	0.13
Labor supply men	–0.05	–0.03	–0.02
Gross hourly wage men +1%			
Labor supply men	0.09	0.08	0.08
– Extensive margin	0.08	0.06	0.07
– Intensive margin	0.01	0.02	0.01
Labor supply women	–0.15	–0.15	–0.15
Gross price formal child care +1%			
Formal child care	–0.69	–0.63	–0.70
Labor supply men	0.00	0.00	0.00
Labor supply women	–0.14	–0.09	–0.15

Table J.9: Elasticities by number of latent classes: youngest child 4–11 yrs

	1 LC	2 LC	3 LC
Gross hourly wage women +1%			
Labor supply women	0.36	0.39	0.41
– Extensive margin	0.24	0.26	0.26
– Intensive margin	0.12	0.13	0.13
Labor supply men	–0.04	–0.03	–0.02
Gross hourly wage men +1%			
Labor supply men	0.09	0.09	0.07
– Extensive margin	0.08	0.08	0.05
– Intensive margin	0.01	0.01	0.02
Labor supply women	–0.11	–0.08	–0.08
Gross hourly price formal child care +1%			
Formal child care	–0.36	–0.36	–0.45
Labor supply men	0.00	0.00	0.00
Labor supply women	–0.02	–0.04	–0.03

Table J.10: Elasticities for model with and without proxy informal child care

	Couples 0-3 yrs		Couples 4-11 yrs	
	Base	With proxy informal care	Base	With proxy informal care
Labor supply elasticity women	0.43	0.45	0.41	0.42
Labor supply elasticity men	0.08	0.07	0.08	0.08
Price elasticity formal child care	–0.75	–0.83	-0.42	-0.41

K Robustness check: allowing for involuntary unemployment

In our structural model we implicitly assume individuals can freely choose their preferred alternative from the discrete choice set. In reality, demand side restrictions may limit these discrete choice sets. We extend our model by taking into account the possibility that individuals are restricted in their choices. In our data set for 2006–2009, we observe individual job search behaviour, and we can use this information to identify whether individuals are involuntarily unemployed or not. We follow the approach by Bargain et al. (2010) who proceeds in two steps.[75] The first step is to estimate the determinants of involuntary employment by a standard probit regression. Next, we use this information to simulate a probability of involuntary unemployment for all individuals in our data set. The second step is to include these probabilities in the simulated maximum likelihood function. For individuals, there are three possible states on the labor market: voluntary unemployed, involuntary unemployed and employed. So we end up with nine possible states for the household at the labor market:

1) Man and woman voluntarily unemployed, where P_i^{VOL} is the probability of being voluntary unemployed:

$$P_i^{VOL_m,VOL_f} = \frac{\exp\left(U_{i1}\right)}{\sum_{j=1}^{J} \exp\left(U_{ij}\right)}, \tag{K.1}$$

2) Man involuntarily unemployed and woman voluntarily unemployed, where P_i^{INVOL} is the probability of being involuntary unemployed

$$P_i^{INVOL_m,VOL_f} = \Phi_m(\beta \boldsymbol{X}) \sum_{k \in (h_m > 0, h_f = 0)}^{J} \frac{\exp\left(U_{ik}\right)}{\sum_{j=1}^{J} \exp\left(U_{ij}\right)}, \tag{K.2}$$

3) Man voluntarily unemployed and woman involuntarily unemployed

$$P_i^{VOL_m,INVOL_f} = \Phi_f(\beta \boldsymbol{X}) \sum_{k \in (h_m = 0, h_f > 0)}^{J} \frac{\exp\left(U_{ik}\right)}{\sum_{j=1}^{J} \exp\left(U_{ij}\right)}, \tag{K.3}$$

[75] De Boer (2018) conducts a similar analysis for the Netherlands and entails a more detailed description of the framework.

4) Man and woman involuntarily unemployed:

$$P_i^{INVOL_m,INVOL_f} = \Phi_m(\beta \boldsymbol{X}) \sum_{k\in(h_m>0,h_f>0)}^{J} \frac{\exp(U_{ik})}{\sum_{j=1}^{J}\exp(U_{ij})}, \tag{K.4}$$

5) Man employed and woman voluntarily unemployed, where P_i^{EMP} is the probability of being employed:

$$P_{i,k\in(h_m>0,h_f=0)}^{EMP_m,VOL_f} = (1-\Phi_m(\beta \boldsymbol{X}))\frac{\exp(U_{ik})}{\sum_{j=1}^{J}\exp(U_{ij})}. \tag{K.5}$$

6) Man voluntarily unemployed and woman employed:

$$P_{i,k\in(h_m=0,h_f>0)}^{VOL_m,EMP_f} = (1-\Phi_f(\beta \boldsymbol{X}))\frac{\exp(U_{ik})}{\sum_{j=1}^{J}\exp(U_{ij})}. \tag{K.6}$$

7) Man employed and woman involuntarily unemployed:

$$P_{i,k\in h_m>0}^{EMP_m,INVOL_f} = (1-\Phi_m(\beta \boldsymbol{X}))\Phi_f(\beta \boldsymbol{X}) \sum_{k\in h_f>0}^{J} \frac{\exp(U_{ik})}{\sum_{j=1}^{J}\exp(U_{ij})}. \tag{K.7}$$

8) Man involuntarily unemployed and woman employed:

$$P_{i,k\in h_f>0}^{INVOL_m,EMP_f} = \Phi_m(\beta \boldsymbol{X})(1-\Phi_f(\beta \boldsymbol{X})) \sum_{k\in h_m>0}^{J} \frac{\exp(U_{ik})}{\sum_{j=1}^{J}\exp(U_{ij})}. \tag{K.8}$$

9) Man and woman employed:

$$P_{i,k\in(h_m>0,h_f>0)}^{EMP_m,EMP_f} = (1-\Phi_m(\beta \boldsymbol{X}))(1-\Phi_f(\beta \boldsymbol{X}))\frac{\exp(U_{ik})}{\sum_{j=1}^{J}\exp(U_{ij})}. \tag{K.9}$$

Table K.1 gives the elasticities for the model without and with involuntary unemployment. We do not observe job search behaviour for all individuals in our original sample. We drop the individuals for who we do not observe this information, so we end up with less observations than in the original sample. The first column presents the elasticities for this new baseline model without involuntary unemployment. Allowing for the possibility of involuntary unemployed lowers the labor supply response which is in line with Bargain et al. (2010) and De Boer (2018). The main reason for the small impact of incorporating involuntary unemployment on behavioural responses is the small share of involuntary unemployment in our data period. Only 3% of the individuals is involuntary unemployed in the period 2006–2009.

Table K.1: Elasticities: model without vs. with involuntary unemployment

	Youngest child 0–3				Youngest child 4–11			
	Without		With		Without		With	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Labor supply women	0.481	0.005	0.433	0.004	0.400	0.004	0.360	0.004
Extensive margin	0.294	0.003	0.246	0.002	0.260	0.002	0.220	0.002
Intensive margin	0.182	0.002	0.187	0.002	0.130	0.001	0.132	0.001
Labor supply men	0.090	0.001	0.077	0.001	0.070	0.001	0.060	0.001
Extensive margin	0.090	0.001	0.076	0.001	0.060	0.001	0.051	0.001
Intensive margin	0.010	0.000	0.000	0.000	0.010	0.000	0.009	0.000

L Differences-in-differences analysis

L.1 Empirical methodology

For the differences-in-differences (DD) analysis we build on Bettendorf et al. (2015). Specifically, we use their data set and their specification, but we impose the same sample restrictions as in the structural model. The treatment group consists of parents with the youngest child up to 11 years of age. The control group consists of parents with the youngest child 12 to 18 years of age (living at home). Bettendorf et al. (2015) show that another potential control group, persons without children, is more problematic because it does not share the same pre-reform time effects as the treatment group.

Following Bettendorf et al. (2015), we estimate the effect of the reform on the participation rate and hours worked per week. Specifically, we regress participation status and hours worked per week on year fixed effects (α_t), group fixed effects (γ_g), individual characteristics (X_i), a treatment dummy for 2005–2007 (short run) ($D_{g,2005-2007}$) and a treatment dummy for 2008–2009 (medium to long run) ($D_{g,2008-2009}$):

$$y_{igt} = \alpha_t + \gamma_g + X_i\beta + \delta_{2005-2007}D_{g,2005-2007} + \delta_{g,2008-2009}D_{g,2008-2009} + \epsilon_{igt}. \tag{L.1}$$

The common time effects of the treatment and control group are captured by the year fixed effects, while the constant difference between the treatment and control group is captured by the group fixed effects. Following Bettendorf et al. (2015) we include different group fixed effects for parents with the youngest child aged 0–3, 4–7 and 8–11. Individual characteristics are included to control for observable changes in the composition of households over time. Participation is a discrete variable, so equation (L.1) is a linear probability model for participation (Angrist and Pischke, 2009). For hours worked per week, we estimate a linear model using the same sample of individuals and the same regression specification. Non-employed parents are coded to work zero hours. All presented models use robust standard errors.[76]

L.2 Data

We use data from the Labor Force Survey (LFS) of Statistics Netherlands. The LFS is an annual survey of approximately 80 thousand individuals per year. We use the cross-sections

[76] Bettendorf et al. (2015) also consider standard errors clustered at the group and year-group level. However, in the majority of cases, the clustered standard errors are smaller than the robust standard errors. We prefer to be conservative and report the larger (robust) standard errors (as suggested by Angrist and Pischke, 2009). Furthermore, the conclusions are robust across the alternative specifications for the standard errors.

for the period 1995–2009. This gives us 10 years of pre-reform data (1995–2004) and 5 years of post-reform data (2005–2009). The survey contains information on labor supply (participation and hours worked), along with individual characteristics (age, education level, native/immigrant) and household characteristics (number of children, ages of the children). Child care information is not collected by the LFS, and so we focus only on the labor supply outcomes. We restrict the sample to individuals in couples that are between 20 and 50 years of age. We note that the average participation rate in the DD sample in the years 2006–2009 is very similar to the participation rate in the sample for the structural model. Descriptive statistics of the sample for the DD analyses are provided in Table L.1.

L.3 Results

The estimates of the reform effects are presented in Table L.2. First, we consider mothers with the youngest child aged 0–3 (pre-primary school age). In terms of participation rate, we find a treatment effect of 1.1 percentage points in the short run (2005–2007), rising to a statistically significant 2.0 percentage points in the medium to long run (2008–2009). The effects are increasing over time, and so we focus our discussion on the medium to long run effects. For hours worked, we find a statistically significant increase of 1.1 hours per week. For mothers with the youngest child aged 4–11 (primary school age), we find an increase in the participation rate of 2.2 percentage points, and an increase in hours worked of 0.7 hours per week, both statistically significant.

The effects for fathers are much smaller. We do not find a statistically significant effect on the paternal participation rate or hours worked, neither in the families with the youngest child aged 0–3 nor with the youngest child aged 4–11. If anything, the results hint at a very small increase in the participation rate and a small decline in hours worked per week.

Table L.3 presents a robustness check for the differences-in-differences analyses, where we include placebo reform dummies for the years 2003–2004, following Bettendorf et al. (2015). These placebo reform dummies are informative about the assumption of common time effects and potential anticipation effects. Reassuringly, we find that all placebo effects are insignificant.[77]

[77]The exception is the model of hours worked per week for men with the youngest child aged 0–3, for which the placebo treatment dummy is ‘borderline’ significant at the 10% level.

Table L.1: Descriptive statistics treatment and control groups (1995–2009)

	Women			Men		
	Treatment group 0–3	Treatment group 4–11	Control group 12–18	Treatment group 0–3	Treatment group 4–11	Control group 12–18
Age	32.51	38.79	44.09	35.02	40.79	45.34
Lower educated	0.240	0.315	0.399	0.261	0.281	0.304
Middle educated	0.472	0.538	0.426	0.428	0.430	0.435
Higher educated	0.288	0.223	0.175	0.311	0.289	0.261
Immigrant	0.200	0.179	0.143	0.192	0.163	0.124
Household size	3.868	4.194	4.006	3.863	4.187	4.008
Age youngest child	1.304	7.224	14.348	1.298	7.167	14.27
Participation rate	0.671	0.685	0.717	0.941	0.942	0.941
Hours worked per week	14.84	14.19	16.22	38.32	38.63	38.71
Observations	89,114	94,883	52,779	88,055	89,824	43,548

Source: own calculations using the Labor Force Survey (Statistics Netherlands). Values are means weighted with sample weights.

Table L.2: Full estimation results differences-in-differences analyses

	Women				Men			
	0–3 yrs		4–11 yrs		0–3 yrs		4–11 yrs	
	PR	Hrs/W	PR	Hrs/W	PR	Hrs/W	PR	Hrs/W
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treat 05–07	0.011	0.607***	0.015**	0.431**	0.006	–0.144	0.009*	0.212
	(0.007)	(0.218)	(0.007)	(0.216)	(0.005)	(0.254)	(0.005)	(0.252)
Treat 08-09	0.020***	1.099***	0.022***	0.741***	0.006	–0.364	0.003	–0.141
	(0.007)	(0.215)	(0.007)	(0.212)	(0.004)	(0.229)	(0.004)	(0.227)
Immigrant	–0.170***	–1.359***	–0.129***	–0.048	–0.143***	–7.187***	–0.133***	–6.688***
	(0.004)	(0.130)	(0.004)	(0.138)	(0.004)	(0.162)	(0.004)	(0.171)
Lower	–0.173***	–4.803***	–0.154***	–4.493***	–0.050***	–2.495***	–0.050***	–2.542***
education	(0.004)	(0.098)	(0.003)	(0.093)	(0.002)	(0.115)	(0.002)	(0.112)
Higher	0.105***	5.642***	0.106***	5.477***	0.016***	–0.851***	0.017***	–0.453***
education	(0.003)	(0.100)	(0.003)	(0.107)	(0.002)	(0.092)	(0.002)	(0.095)
20–24 yrs	–0.095***	–1.241***	0.037	3.060***	–0.047***	–2.949***	–0.136***	–8.226***
of age	(0.012)	(0.330)	(0.034)	(1.008)	(0.014)	(0.605)	(0.048)	(1.835)
25–29 yrs	0.023***	1.202***	0.026**	2.357***	0.019***	0.499**	–0.019	–1.217**
of age	(0.008)	(0.222)	(0.012)	(0.362)	(0.005)	(0.249)	(0.013)	(0.574)
30–34 yrs	0.088***	2.966***	0.048***	1.953***	0.036***	1.387***	0.010**	0.790***
of age	(0.007)	(0.204)	(0.006)	(0.184)	(0.004)	(0.199)	(0.005)	(0.235)
35–39 yrs	0.089***	3.463***	0.073***	2.225***	0.030***	1.063***	0.017***	0.992***
of age	(0.006)	(0.194)	(0.004)	(0.134)	(0.004)	(0.192)	(0.003)	(0.141)
40–44 yrs	0.060***	2.404***	0.053***	1.670***	0.016***	0.356**	0.016***	0.667***
of age	(0.005)	(0.146)	(0.004)	(0.113)	(0.003)	(0.153)	(0.002)	(0.110)
One child	0.192***	7.062***	0.071***	3.347***	0.015***	–1.222***	0.003	–1.670***
	(0.004)	(0.121)	(0.004)	(0.129)	(0.003)	(0.144)	(0.003)	(0.144)
Two children	0.133***	3.761***	0.062***	1.740***	0.022***	–0.616***	0.015***	–1.076***
	(0.004)	(0.114)	(0.004)	(0.105)	(0.003)	(0.137)	(0.002)	(0.119)
Youngest	–0.081***	–2.993***	–	–	–0.014***	–0.823***	–	–
chd 0–3 yrs	(0.006)	(0.181)			(0.003)	(0.174)		
Youngest	–	–	–0.080***	–3.549***	–	–	–0.005*	–0.756***
chd 4–7 yrs			(0.005)	(0.145)			(0.003)	(0.147)
Youngest	–	–	–0.042***	–1.996***	–	–	–0.006**	–0.564***
chd 8–11 yrs			(0.004)	(0.133)			(0.003)	(0.137)
Observations	141,893	141,893	147,662	147,662	131,603	131,603	133,372	133,372

Heterogeneity robust standard errors in parentheses, * denotes significant at 10% level, ** at 5% level and *** at 1% level. PR denotes participation rate and Hrs/W denotes hours worked per week. Year fixed effects are included but not reported. Base group are native Dutch, middle educated, 45–50 years of age and with three or more children.

Table L.3: Estimation results differences-in-differences analyses with placebo

	Women				Men			
	0–3 yrs		4–11 yrs		0–3 yrs		4–11 yrs	
	PR	Hrs/W	PR	Hrs/W	PR	Hrs/W	PR	Hrs/W
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Placebo 03–04	–0.006	–0.114	–0.001	0.113	0.006	0.546*	0.002	0.378
	(0.008)	(0.257)	(0.008)	(0.259)	(0.005)	(0.283)	(0.006)	(0.288)
Treat 05–07	0.010	0.583***	0.014*	0.453**	0.008	–0.037	0.009*	0.285
	(0.007)	(0.226)	(0.007)	(0.216)	(0.005)	(0.261)	(0.005)	(0.258)
Treat 08-09	0.018***	1.075***	0.022***	0.764***	0.007*	–0.257	0.003	–0.067
	(0.007)	(0.223)	(0.007)	(0.221)	(0.004)	(0.236)	(0.004)	(0.234)
Observations	141,893	141,893	147,662	147,662	131,603	131,603	133,372	133,372

Heterogeneity robust standard errors in parentheses, * denotes significant at 10% level, ** at 5% level and *** at 1% level. PR denotes the participation rate and Hrs/W denotes hours worked per week. Placebo dummy equals 1 for the treatment group in the pre-reform years 2003 and 2004.

Table L.4: Comparison of structural model with DD analysis: policy reforms 2005–2009

	Structural model				DD analysis[a]	
	(1)	(2)	(3)	**(4)**	**(5)**	
	Child care subsidies	Combination Credit	Income-depend. Combi. Credit	**Total effect**	**Coefficient**	SE
Model with random preference heterogeneity			Changes in levels			
Youngest child 0-3 yrs						
Participation rate women	0.019	–0.006	0.020	**0.033**	**0.020**	0.007
Hours worked per week women	0.752	–0.106	0.630	**1.275**	**1.099**	0.215
Participation rate men	0.003	–0.002	0.004	**0.005**	**0.006**	0.004
Hours worked per week men	0.067	–0.033	0.046	**0.080**	**–0.364**	0.229
Youngest child 4-11 yrs						
Participation rate women	0.003	–0.005	0.016	**0.014**	**0.022**	0.007
Hours worked per week women	0.105	–0.086	0.453	**0.471**	**0.741**	0.212
Participation rate men	0.000	–0.001	0.003	**0.002**	**0.003**	0.004
Hours worked per week men	0.018	–0.024	0.051	**0.045**	**–0.141**	0.227
Model with three latent classes						
Youngest child 0-3 yrs						
Participation rate women	0.016	–0.006	0.020	**0.030**	**0.020**	0.007
Hours worked per week women	0.719	–0.103	0.602	**1.218**	**1.099**	0.215
Participation rate men	0.003	–0.002	0.003	**0.003**	**0.006**	0.004
Hours worked per week men	0.049	–0.039	0.020	**0.030**	**–0.364**	0.229
Youngest child 4-11 yrs						
Participation rate women	0.003	–0.005	0.016	**0.014**	**0.022**	0.007
Hours worked per week women	0.115	–0.087	0.440	**0.468**	**0.741**	0.212
Participation rate men	0.000	–0.001	0.002	**0.001**	**0.003**	0.004
Hours worked per week men	0.026	–0.013	0.029	**0.042**	**–0.141**	0.227
Model without unobserved heterogeneity						
Youngest child 0-3 yrs						
Participation rate women	0.017	–0.005	0.019	**0.031**	**0.020**	0.007
Hours worked per week women	0.703	–0.097	0.583	**1.189**	**1.099**	0.215
Participation rate men	0.003	–0.002	0.004	**0.005**	**0.006**	0.004
Hours worked per week men	0.079	–0.040	0.056	**0.095**	**–0.364**	0.229
Youngest child 4-11 yrs						
Participation rate women	0.002	–0.004	0.014	**0.012**	**0.022**	0.007
Hours worked per week women	0.097	–0.074	0.396	**0.419**	**0.741**	0.212
Participation rate men	0.000	–0.001	0.003	**0.002**	**0.003**	0.004
Hours worked per week men	0.020	–0.027	0.061	**0.055**	**–0.141**	0.227

[a] DD treatment effect for 2008–2009, see Table L.2.

M Policy simulations: effects on initial income

Effect on initial incomes: child care subsidies

Figure M.1: Not targeted

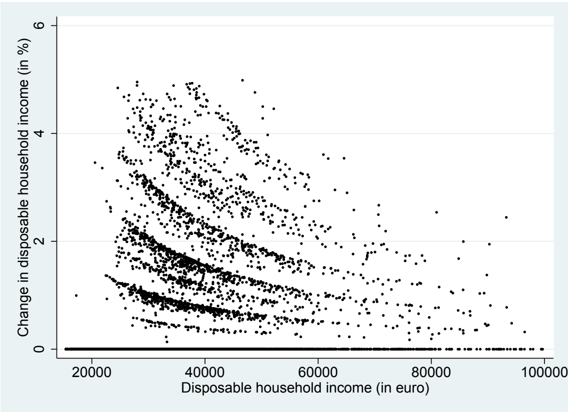
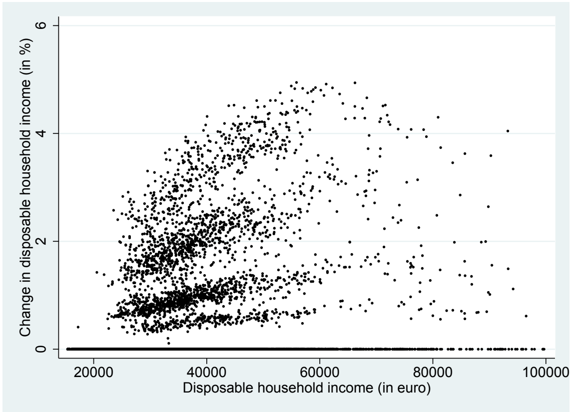


Figure M.2: Targeted more at higher incomes



Notes: On the horizontal axis is initial disposable household income, and on the vertical axis is the percentage change in disposable household income before behavioral responses. We see a number of 'lines' or 'groups' in terms of the change in disposable income, corresponding to families which differ in terms of the number of children, and the intensity of formal child care use.

Effect on initial incomes: in-work benefit for secondary earners with children

Figure M.3: Not targeted

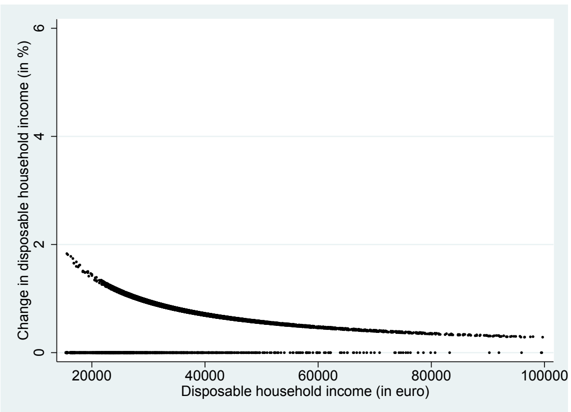
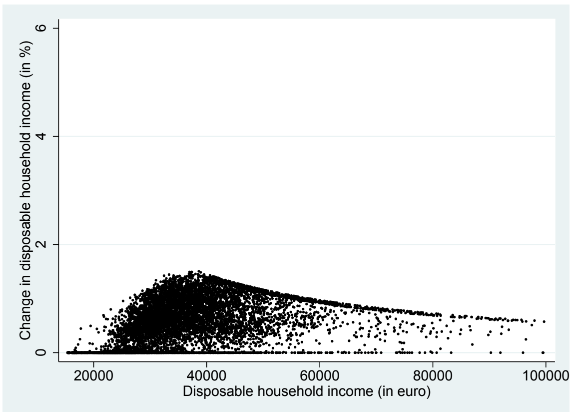


Figure M.4: Targeted more at higher incomes



Notes: On the horizontal axis is initial disposable household income, and on the vertical axis is the percentage change in disposable household income before behavioral responses. On the left we see two lines for the in-work benefit for secondary earners. The line close to the horizontal axis corresponds to households in which the secondary earner does not work. The higher line corresponds to two-earner couples. On the right we see more scatter, as the subsidy depends on the income of the secondary earners, which are coupled to primary earners with different income levels.

Effect on initial incomes: in-work benefit for primary and secondary earners with children

Figure M.5: Not targeted

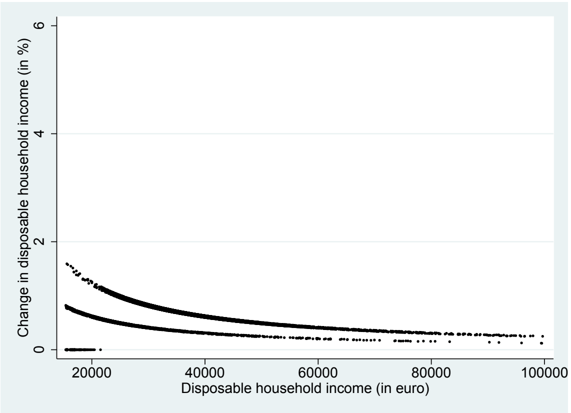
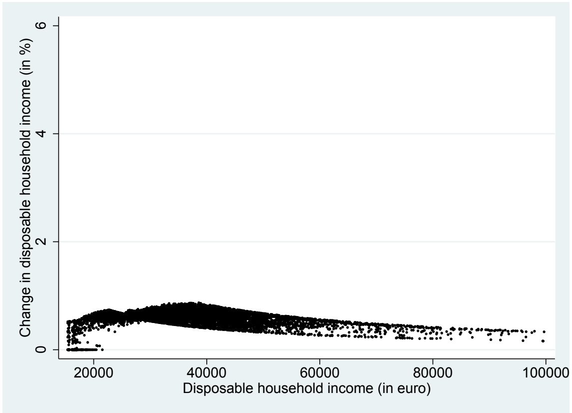


Figure M.6: Targeted more at higher incomes



Notes: On the horizontal axis is initial disposable household income, and on the vertical axis is the percentage change in disposable household income before behavioral responses. On the left we see three lines, one for two-earner households, one for one-earner households and one for households in which neither of the two parents is working (their disposable income remains the same). On the right we see again more scatter, as the subsidy depends on the income of the primary and secondary earners.

N Policy simulations: effects for subgroups

Table N.1: Labor supply effects fiscal stimuli: subgroups[a,b]

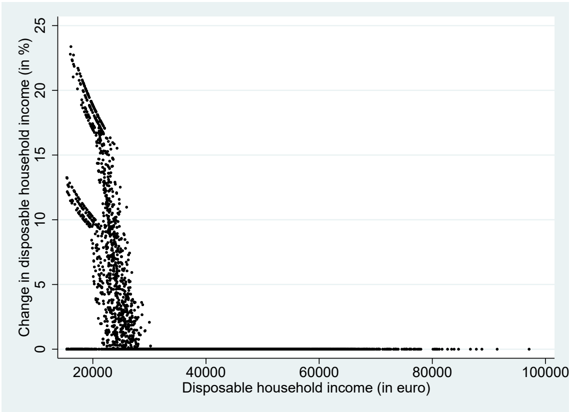
	Not targeted			Targeted more at higher incomes		
	(1) Child care subsidy	(2) In-work benefit second. earners	(3) In-work ben. all parents	(4) Child care subsidy	(5) In-work benefit second. earners	(6) In-work ben. all parents
Youngest child 0–3	Percentage changes					
Labor supply women	2.03	0.93	0.29	2.17	1.70	0.35
Labor supply men	0.13	0.03	0.00	0.18	0.03	0.03
Women						
Lower educated	4.15	2.55	0.95	2.61	2.63	0.54
Middle educated	2.13	1.01	0.32	1.99	1.79	0.37
Higher educated	1.46	0.50	0.12	2.26	1.40	0.29
Native	1.73	0.81	0.24	1.97	1.61	0.33
Non-western immigrant	6.58	2.88	1.14	4.67	3.09	0.75
Western immigrant	2.43	1.06	0.35	2.74	1.82	0.39
Men						
Lower educated	0.23	0.09	0.06	0.17	0.07	0.10
Middle educated	0.11	0.01	-0.01	0.15	0.02	0.03
Higher educated	0.10	0.02	-0.01	0.21	0.01	-0.01
Native	0.07	0.00	-0.02	0.14	0.00	0.01
Non-western immigrant	0.31	0.13	0.16	0.21	0.10	0.17
Western immigrant	0.63	0.24	0.11	0.61	0.25	0.12
Youngest child 4–11						
Labor supply women	0.20	0.70	0.24	0.28	1.09	0.27
Labor supply men	0.02	0.06	-0.01	0.02	0.05	0.02
Women						
Lower educated	0.17	1.18	0.45	0.21	1.25	0.23
Middle educated	0.28	0.74	0.25	0.30	1.16	0.30
Higher educated	0.10	0.35	0.10	0.28	0.88	0.25
Native	0.20	0.68	0.21	0.28	1.08	0.26
Non-western immigrant	0.33	1.13	0.61	0.33	1.32	0.41
Western immigrant	0.18	0.60	0.30	0.27	1.07	0.33
Men						
Lower educated	0.03	0.10	0.03	0.01	0.09	0.10
Middle educated	0.01	0.04	-0.01	0.02	0.03	0.02
Higher educated	0.02	0.05	-0.04	0.03	0.05	-0.02
Native	0.01	0.03	-0.02	0.02	0.03	0.01
Non-western immigrant	0.05	0.30	0.14	0.05	0.23	0.19
Western immigrant	0.08	0.15	0.00	0.07	0.13	0.09

[a]For the details of the simulations see Table 4 in the main text.

O Additional policy simulations: effects on initial income

Effect on initial incomes: 'Working Tax Credit'

Figure O.1: Targeted at lower incomes



Notes: On the horizontal axis is initial disposable household income, and on the vertical axis is the percentage change in disposable household income before behavioral responses. We see two 'groups' that differ in whether they have one or more children.

Effect on initial incomes: 'Child Tax Credit'

Figure O.2: Not targeted

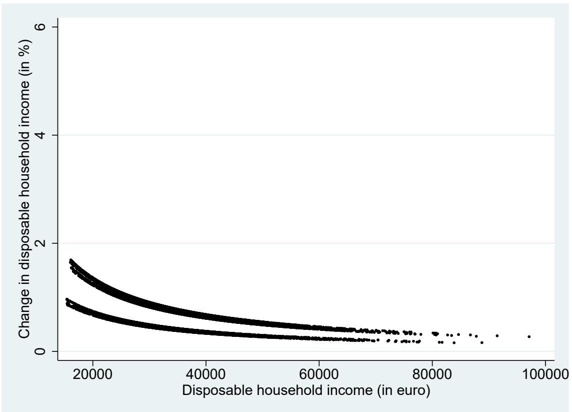
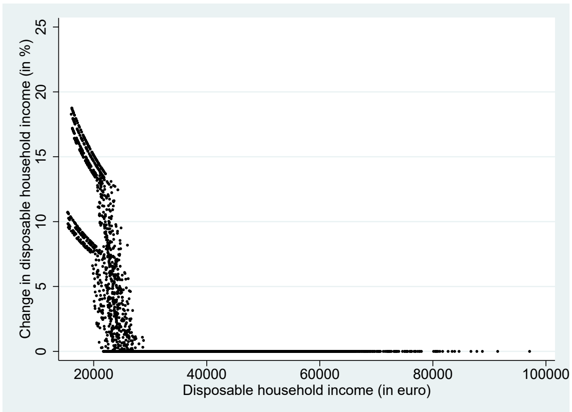


Figure O.3: Targeted at lower incomes



Notes: On the horizontal axis is initial disposable household income, and on the vertical axis is the percentage change in disposable household income before behavioral responses. On the left we see two lines, depending on whether they have one or more children. The figure on the right is very similar to the figure above, with the two 'groups' that differ in whether they have one or more children.