



CPB Discussion Paper | 229

# The dog that did not bark:

## The EITC for single mothers in the Netherlands

Leon Bettendorf  
Kees Folmer  
Egbert Jongen



# The dog that did not bark: The EITC for single mothers in the Netherlands

Leon J.H. Bettendorf\* Kees Folmer<sup>†</sup> Egbert L.W. Jongen<sup>‡</sup>

January 24, 2013

## Abstract

We study the extension of an EITC for single mothers in the Netherlands to mothers with a youngest child 12 to 15 years old. This reform increased net income for the treatment group by 5%. Using both DD and RD we show that this reform had a negligible effect on labour participation, with tight confidence intervals around zero. Our results are at odds with a number of related studies. This is likely to be due to their use of single women without children as the control group, which in our case is an invalid control group.

**JEL codes:** C21, H24, J22

**Keywords:** Policy evaluation, difference-in-differences, regression discontinuity, labour participation, single mothers

---

\*CPB Netherlands Bureau for Economic Policy Analysis. E-mail: L.J.H.Bettendorf@cpb.nl.

<sup>†</sup>CPB Netherlands Bureau for Economic Policy Analysis. E-mail: C.Folmer@cpb.nl.

<sup>‡</sup>CPB Netherlands Bureau for Economic Policy Analysis. Corresponding author: CPB, P.O. Box 80510, 2508GM, The Hague, The Netherlands. Phone: +31-70-3383468. E-mail: E.L.W.Jongen@cpb.nl. We have benefited from comments and suggestions by Joshua Angrist, Hans Bloemen, Richard Blundell, Chris van Klaveren, Pierre Koning, Erzo Luttmer, Hakan Selin, Arthur van Soest, participants of the IZA Labour Supply Workshop October 2011 in Bonn, participants of the IIPF Conference August 2012 in Dresden and seminar participants at CPB Netherlands Bureau for Economic Policy Analysis. Remaining errors are our own.

# 1 Introduction

The share of single parents is on the rise. In the Netherlands their number has increased from 360 thousand in 1995 to 500 thousand in 2011.<sup>1</sup> Single parents are of particular interest to policy makers as evidenced by the large number of subsidies and tax credits targeted at this group. In designing income support for single parents, the responsiveness of labour participation by single parents to financial incentives plays a crucial role.

Until 2001, working single parents received a tax credit, the *Aanvullende Alleenstaande Ouderkorting* (Additional Credit for Single Parents), if the youngest child was younger than 12 years of age. In 2002 this age limit was raised to 16 years. The goal was to stimulate the labour participation of single parents (Ministry of Finance, 2001).<sup>2</sup> We use this change in the age limit as a natural experiment to determine the labour supply responsiveness of single mothers to financial incentives.

We use difference-in-differences (DD) and regression discontinuity (RD) to estimate the effect of the policy reform on the participation rate of single mothers.<sup>3</sup> In the DD analysis we use single mothers with a youngest child that is younger (8-11 years of age) or older (16-19 years of age) than the treatment group (12-15 years of age) as the control group. In the RD analysis we focus on single mothers with a youngest child 14-17 years of age, with the cutoff at single mothers with a youngest child that turned 16 in December in the preceding year, using data on month of birth. Both the DD and RD analysis show that the policy reform had a small effect on the participation rate of single mothers. Indeed, we can not reject that the effect on the participation rate was zero. This is not due to a lack of statistical power, as the 95% confidence intervals of both the DD and RD estimates are quite tight.<sup>4</sup>

---

<sup>1</sup>Source: Statistics Netherlands.

<sup>2</sup>The former government (Rutte-I) had plans to reverse the policy change of 2002, to reduce the budget deficit (CPB, 2010), but for the moment these plans are on hold.

<sup>3</sup>Furthermore, we also consider a ‘difference-in-discontinuity’ analysis, where we allow for a potential pre-reform discontinuity (although we are unaware of a reason to expect a pre-reform discontinuity at the discontinuity we consider).

<sup>4</sup>The point estimate of the DD analysis is  $-0.2\%$ -points with a 95% confidence interval  $[-1.4, 0.9]$ . The point estimate of the RD analysis is  $-0.4$  (where we have reversed the sign of the coefficient since we are measuring the change in the participation rate of single mothers that do not qualify for the subsidy relative to single mothers that do qualify for the subsidy, see below), with a 95% confidence interval  $[-2.3, 1.6]$ .

Furthermore, an extensive robustness analysis shows that our results are robust.

Our results are at odds with the findings of a number of related studies on single mothers in other countries. Table 1 gives an overview of these studies. There is an extensive literature on the impact of the *Earned Income Tax Credit* (EITC) in the US, introduced in 1975. This literature is reviewed in Hotz and Scholz (2003). They conclude that the EITC appears to have had a substantial, positive effect on the participation rate of single mothers. Eissa and Liebman (1996) is one of the earlier studies that applies the DD methodology to labour supply responses. They estimate the impact of the EITC-expansion in 1987, combined with other elements of the Tax Reform Act of 1986, by comparing the change in labour supply of single mothers to the change in labour supply of single women without children. They find that the EITC-expansion increased the participation rate of single mothers by 2.8%-points (3.8%). According to Meyer and Rosenbaum (2001), the tax reform increased net-of-tax income of working single mothers by 2.7%. Hence, the implicit elasticity of participation with respect to net-of-tax income is  $3.8/2.7 = 1.4$ . Meyer and Rosenbaum (2001) examine the effects of changes in both welfare and tax policies in the US during the 1984-1996 period on the labour supply of single mothers. DD estimates suggest that the policy changes over this period have raised the participation rate of single mothers by 7.1%-points (9.7%),<sup>5</sup> using single women without children as the control group. The policy changes increased net-of-tax income by 5.4%. Hence, the implicit elasticity of participation with respect to net-of-tax income in this study is  $9.7/5.4 = 1.8$ .

Several studies examine the introduction of the *Working Families' Tax Credit* (WFTC) in 1999 in the UK. Brewer and Browne (2006) review the findings of studies on the WFTC. The increase in the participation rate varies from an insignificant 0.6%-points in Leigh (2005) to 7%-points in Francesconi and van der Klaauw (2004). Blundell et al. (2005) report an impact of 3.6%-points (7.7%) for single mothers, using single women without children as the control group. According to Brewer et al. (2006), the WFTC increased net-of-tax income by 6.5%. Hence, the implicit elasticity of participation with respect to net-of-tax income in for example Blundell et al. (2005) is  $7.7/6.5 = 1.2$ .

Stancanelli (2008) studies the impact of the *Prime Pour l'Emploi* (Work Pre-

---

<sup>5</sup>Estimates of a structural model suggest that changes in the EITC explain 62% of the total effect.

Table 1: Overview of related studies

Study	Country	Reform	Impulse (in %)	Treatment group	Control group	Initial part. rate	Treatment effect (SE)	Method	Placebo test	$[m, \rho]_o$
Eissa and Liebman (1996)	US	EITC 87	2.7 <sup>b</sup>	Single moth. child < 19y <sup>f</sup>	Single women	74	+2.8 (0.9)	DD	No	[0.81, 5.51]
Meyer and Rosenbaum (2001)	US	EITC 84-96 <sup>a</sup>	5.4 <sup>c</sup>	Single moth. child < 19y <sup>f</sup>	Single women	73	+7.1 <sup>k</sup>	DD	No <sup>m</sup>	[1.31, 2.86]
Gregg and Harkness (2003)	UK	WFTC 99	6.5 <sup>d</sup>	Single moth.	Single women <sup>h</sup>	47	+5.0	DD	No	[1.25, 2.36]
Francesconi et al. (2004)	UK	WFTC 99	6.5 <sup>d</sup>	Single moth. child < 17y <sup>g</sup>	Single women	41 <sup>i</sup>	+7.0 (2.1)	DD	No <sup>n</sup>	[2.01, 3.79]
Leigh (2005)	UK	WFTC 99	6.5 <sup>d</sup>	Single moth. child < 17y <sup>g</sup>	Single women	47	+0.6 (1.5)	DD	No	[0.15, 0.28]
Blundell et al. (2005)	UK	WFTC 99	6.5 <sup>d</sup>	Single moth.	Single women	47	+3.6 (0.5)	DD	Yes	[0.62, 1.18]
Stancanelli (2008)	FR	EITC 91	NA <sup>e</sup>	Single moth.	Single women	68	-0.1 (1.3)	DD	No	-
This study (2013)	NL	EITC 02	4.8	Single mth. youngest chd 12-15y	Single mth. young. chd 8-11/16-19y	71	-0.2 (0.6)	DD	Yes	[-0.04, -0.10]
				Single mth. young. chd 14-15y	Single mth. young. chd 16-17y	75 <sup>j</sup>	-0.4 <sup>l</sup> (1.0)	RD	Yes	[-0.08, -0.19]

<sup>a</sup>Component of broader tax reforms. <sup>b</sup>According to Meyer and Rosenbaum (2001), income taxes of working single mothers fell by \$491 over the period 1984-1988, while their net-of-tax income was \$18,426 in 1984. Net-of-tax income is calculated as earned income - income taxes + welfare benefit + medicaid (when working). <sup>c</sup>Tax reforms, mostly consisting of EITC changes, have increased over the period 1984-1996 net-of-tax earnings by around \$1,000 for single mothers, who earned \$18,426 in 1984. <sup>d</sup>According to Brewer et al. (2006), the WFTC raised mean net incomes of working lone parents by £16.24 a week from a base of £250.22. <sup>e</sup>For a single parent with 2 children in 2001, the tax credit reached a maximum of €500 for someone working full time at the minimum wage, resulting in gross earnings of about €11,000. The tax credit then equals 4.5% of gross earnings. <sup>f</sup>Children younger than 24 years if full-time student. <sup>g</sup>Working at least 16 hours per week. <sup>h</sup>Participation rate just above the discontinuity. <sup>i</sup>Control group is constructed using propensity matching techniques. Couples with children used as an alternative control group. <sup>j</sup>Children younger than 19 years if full-time student. <sup>k</sup>The EITC explains 62% of the total increase. <sup>l</sup>The discontinuity measures the difference in the participation rate of parents who do not qualify for the subsidy with parents who do qualify for the subsidy. <sup>m</sup>The paper includes a treatment dummy for every year both before and after the reform. <sup>n</sup>But allowing for different time trends in treatment and control group. <sup>o</sup>Lower and upper limit of the true elasticity given the point estimate and assuming 1% friction costs, using the methodology outlined in Chetty (2012, Section 3.3). The point estimate of the elasticity is the treatment effect over the initial participation rate divided by the impulse over 100. The lower limit of the true elasticity following the methodology of Chetty (2012) is then the point estimate of the elasticity over  $(1 + \rho)$  where  $\rho$  is defined as 2 times the friction costs (assumed to be 1%) over the impulse. The upper limit of the true elasticity following the methodology of Chetty (2012) is the point estimate of the elasticity over  $(1 - \rho)$ , with  $\rho$  defined as before.

mium) in France, introduced in 1991. She finds no significant effect of the reform on single mothers when compared to single women without children. She attributes the insignificant results partly to noticeable differences between the control and treatment group.

Most studies using DD abroad find sizeable labour supply responses by single mothers to changes in financial incentives. Indeed, there appears to be a consensus in the literature that the participation elasticity for single mothers is among the highest of all demographic groups (Meghir and Phillips, 2010). However, we find only a small response by single mothers to the reform we consider. There are a number of possible explanations for this discrepancy (a more extensive discussion is given in Section 7).

First, we consider single mothers with a youngest child that is relatively old, 12 to 15 years of age. As a result our treatment group already had relatively high participation rates before the reform was implemented. However, the treatment group in the US studies has an initial participation rate quite similar to the treatment group we consider, but still they find much higher elasticities.

Second, the tax incentive might have been too small to induce a significant response or may not have been salient enough since it was not a new policy but only a change in the eligibility conditions. However, the impulse we consider is comparable in size to those in the other countries, and the utility loss of inattention is relatively large for the participation margin (as opposed to the hours per week margin conditional on participation), see Chetty (2012).

Third, demand side restrictions may have prevented single mothers from realizing their preferred labour supply choice, due to unfavourable business cycle conditions during the time of the reform or due to minimum wage legislation. However, the treatment effect we find is small not only in the first years after the reform but also in later years. Furthermore, the initial relatively high participation rate of the treatment group suggests that minimum wages are not an important hurdle for our treatment group. Furthermore, even if minimum wages reduce the level of participation, it is not clear if this reduces the elasticity of participation as well.

Finally, all studies discussed above use DD with single women without children as the control group.<sup>6</sup> Because our reform targets only a subsample of single mothers

---

<sup>6</sup>Gregg and Harkness (2003) also consider women in couples with children as an alternative control group.

we can use other single mothers as the control group in our DD analysis, and the targeted reform allows us to do a RD analysis as well. The use of single women without children as a control group for single mothers is criticized by Blundell and MaCurdy (1999) and Meghir and Phillips (2010) because single mothers and single women may have different participation trends and differ substantially in observed characteristics and hence presumably also in unobserved characteristics. We show that in our case single women are indeed not a valid control group. A placebo test indicates that single mothers and single women do not share the same pre-reform trend.<sup>7</sup> Furthermore, our descriptive statistics show that these groups differ substantially in observable characteristics. The use of a different control group seems a plausible explanation for why we find much smaller effects of financial incentives on the participation rate of single mothers than the other studies. Indeed, when we use single women as a control group, to make our case, we also find sizeable treatment effects, in line with the other studies.

The rest of the paper is structured as follows. Section 2 discusses the policy reform that we use in our empirical strategy. Section 3 outlines the estimation methods. Section 4 describes the data used in the analysis. Section 5 presents the estimation results of the DD analysis, and Section 6 presents the estimation results of the RD analysis. Section 7 discusses our findings and concludes. Supplementary material is given in the appendix.

## 2 The natural experiment

All working individuals in the Netherlands receive a general tax credit, the *Arbeidskorting* (Working Credit). Working single parents in the Netherlands receive an additional tax credit, the *Aanvullende Alleenstaande Ouderkorting* (Additional Credit for Single Parents). Until 2001, only single parents with a (dependent) child younger than 12 years old<sup>8</sup> received this additional tax credit. In 2002 this age limit

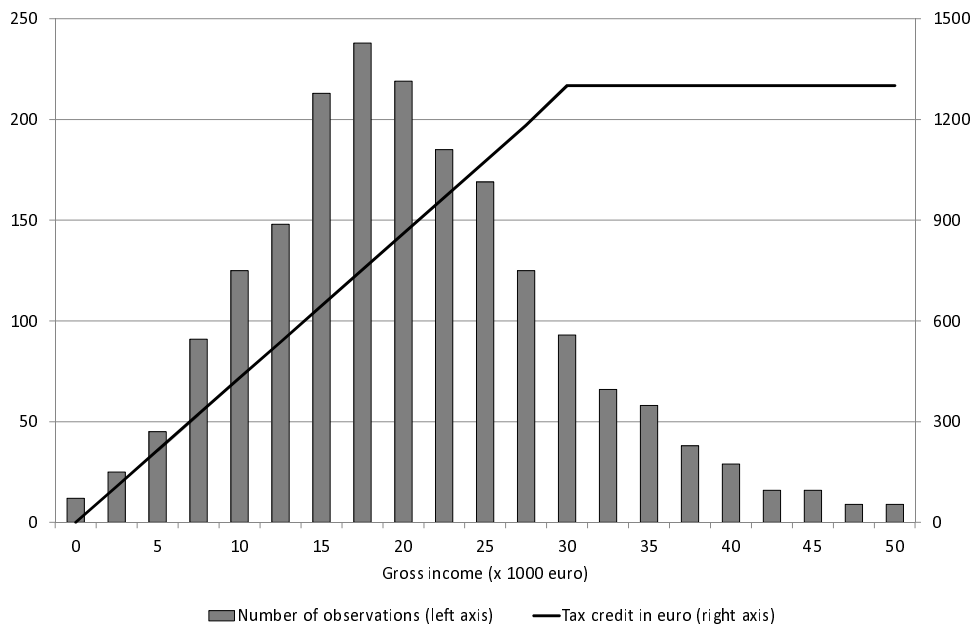
---

<sup>7</sup>The last column of Table 1 shows that most studies do not report placebo tests. Blundell et al. (2005) is the only study that explicitly tests for a treatment effect before the reform. They find an insignificant placebo effect, which supports the choice of single women without children as a valid control group in their case. Meyer and Rosenbaum (2001) estimate a treatment effect for every year before and after the reform, while Francesconi and van der Klaauw (2004) correct for different time trends.

<sup>8</sup>On the 1st of January.



Figure 1: Tax credit and income distribution of working mothers in the treatment group in 2002



Source: Labour Market Panel (Statistics Netherlands).

was raised to 16 years old, to promote the labour participation of single parents (Ministry of Finance, 2001).<sup>9</sup>

The tax credit for working single parents is income dependent and amounts to 4.3% of gross income up to a maximum credit of 1,301 euro in 2002, see Figure 1. As can be seen, the phase-in is up to a gross income of about 30,000 euro, which is about twice the minimum wage. The credit is not phased out. Figure 1 also shows the distribution of earnings for working single mothers with a youngest child aged 12 to 15 years of age. A single mother at the mode of the income distribution of the treatment group has gross income of 17,500 euro, her corresponding net income is 16,282 euro. The tax credit for this mother is 753 euro, an increase in net income of 4.8%.

<sup>9</sup>Apart from the increase of the age limit in 2002, the credit is adjusted to the growth of average gross wages annually. There are no jumps in the credit in real terms in our data period.

### 3 Empirical methodology

We consider a difference-in-differences (DD) and a regression discontinuity (RD) approach<sup>10</sup> to estimate the impact of the expansion of the tax credit to working single mothers with a youngest child 12 to 15 years of age.

The DD approach identifies the labour supply response of a policy reform by combining two types of comparisons. First, we calculate the difference in labour supply before and after the reform for the treatment group (the first difference). Second, we subtract the change in labour supply before and after the reform for a control group (the second difference). By taking a double difference, the approach removes common time effects and time-invariant group effects. The treatment group in our study consists of single mothers with a youngest child aged 12 to 15 years old. Our preferred control group consists of single mothers with a youngest child aged 8-11 or 16-19 years old. As an alternative, we also consider single women without children as a control group, the control group used in related studies on single mothers.

To explain the participation rate we estimate a linear probability model (Angrist and Pischke, 2009). We regress participation status  $part_{igt}$  on a year fixed effect ( $\beta_t$ ), a group fixed effect ( $\beta_g$ ) that applies if the youngest child is 12 to 15 of age, a treatment effect ( $\beta_{DD}$ ) that applies if the youngest child is 12 to 15 years of age in the post-reform period (year>2001), and individual and household characteristics ( $X_{it}$ )

$$part_{igt} = \beta_t + \beta_g + \beta_{DD} + X_{it}\gamma + \epsilon_{igt}. \quad (1)$$

The participation dummy equals 1 if the mother works and 0 otherwise. The common time effects are captured by the year fixed effects, while the constant difference in participation between the treatment and control group is captured by the group effect. We further add controls for the age of the parent (or single) (5-year classes), the level of education (lower, medium or higher educated), ethnicity (native, non-Western allochtonous or Western allochtonous) and the number of children in the household (two children, three or more children). We are primarily interested in the treatment coefficient  $\beta_{DD}$ . Since we have panel data, we also estimate a fixed effects model to control for unobserved fixed characteristics.

---

<sup>10</sup>See *e.g.* Imbens and Lemieux (2008), Imbens and Wooldridge (2009), Blundell and Dias (2009) and Lee and Lemieux (2010).

The RD approach identifies the labour supply response by comparing single mothers with a youngest child just older than the cutoff that determines entitlement to the credit with single mothers with a child just younger than this cutoff. The idea is that in the absence of the tax credit, participation is a smooth function in the age of the youngest child, and the tax credit introduces a discontinuity in this function.

Again we use a linear probability model to explain the participation rate, now using only observations after the reform. We regress participation status  $part_{it}$  on a year fixed effect ( $\beta_t$ ), the age of the youngest child in months<sup>11</sup> ( $\beta_{agechild}$ ), a treatment effect that applies if the age of the youngest child is above 15 ( $\beta_{RD}$ ) capturing the discontinuity, individual and household characteristics ( $X_{it}$ ), and in some specifications also an interaction term that captures the additional effect of the age of the youngest child when the child is older than 15 ( $\beta_{agechild>15}$ ) to allow for a different slope to the right of the discontinuity<sup>12</sup>

$$part_{it} = \beta_t + \beta_{agechild}agechild_{it} + \beta_{agechild>15}1(agechild > 15)agechild_{it} + \beta_{RD} + X_{it}\gamma + \epsilon_{it}. \quad (2)$$

In an extension we consider what might be called a ‘difference-in-discontinuity’ setup, using both the pre and post-reform data. Using observations both before and after the policy reform, we can control for a potential discontinuity before the reform, due to *e.g.* discontinuities in other policies, although we are unaware of a reason for a pre-reform discontinuity around the cutoff we consider. In this specification we include a treatment effect that captures the pre-reform discontinuity ( $\beta_{PRD}$ ), and an additional treatment effect for the post-reform discontinuity relative to the pre-reform discontinuity  $\beta_{DRD}$  (i.e. the discontinuity before the reform equals  $\beta_{PRD}$

---

<sup>11</sup>On the 1st of January.

<sup>12</sup>We also estimated relations with a quadratic term in age of the youngest child in months, and a quadratic term in age of the youngest child in months interacted with a dummy which equals 1 if the youngest child is older than 15, but these terms were insignificant and made the linear terms insignificant as well.

and the discontinuity after the reform equals  $\beta_{PRD} + \beta_{DRD}$ )<sup>13</sup>

$$\begin{aligned} part_{it} = & \beta_t + \beta_{agechild}agechild_{it} + \beta_{agechild>15}1(agechild > 15)agechild_{it} \\ & + \beta_{PRD} + \beta_{DRD} + X_{it}\gamma + \epsilon_{it}. \end{aligned} \quad (3)$$

## 4 Data

We use data from the Labour Market Panel (*Arbeidsmarktpanel*) of Statistics Netherlands (Statistics Netherlands, 2012). The Labour Market Panel is an administrative household panel dataset, starting in 1999. We use data for the period 1999-2008. The dataset combines information from municipalities (*Gemeentelijke Basisadministratie*) on demographic individual and household characteristics, from the Social Statistical Panel (*Sociaal Statistisch Bestand*) on income from employment and benefits, and from the Labour Force Survey (*Enquete Beroepsbevolking*) on the level of education.

From the Labour Market Panel we select single mothers and single women. We drop all individuals under 20 years old and over 57 years old. The maximum age is set at 57 years old because we do not want outcomes to be influenced by changes in early retirement benefits in the data period. We also drop self-employed, disabled and students. To determine whether or not an individual participates we use the social economic classification (*Sociaal Economische Categorie*) of Statistics Netherlands (Statistics Netherlands, 2011) in our base results. This variable classifies individuals according to their main source of income. After the selections above, the remaining sample contains single mothers and singles that are wage earners, on welfare benefits, on unemployment benefits or without wage or benefit income. Single mothers and singles that are wage earners are defined as participating, the rest is defined as not participating. As a robustness check we consider participation as having taxable labour income in excess of a certain percentage (we consider 50, 70 or 100%) of the minimum wage. The results for this alternative definition of participation are similar to the base results. Descriptive statistics are given in the analysis below.

---

<sup>13</sup>We also estimated relations where we allowed for a different relation between the participation rate and the age of the youngest child after the reform, and/or a different relation between the participation rate and the age of the youngest child to the right of the discontinuity after the reform, using interaction terms, but these were never significant.

## 5 DD analysis

We first consider the results for the DD analysis. The treatment group consists of single mothers with a youngest child 12 to 15 years of age. Our preferred control group consists of single mothers with a youngest child 8 to 11 or 16 to 19 years of age. By combining single mothers with a youngest child that is somewhat younger and single mothers with a youngest child that is somewhat older than the treatment group we obtain a control group that is comparable in observable characteristics to the treatment group. As an alternative, we also consider single women without children, the control group used in related studies on single mothers.

Table 2 gives descriptive statistics of the treatment and control groups. In the DD analysis we focus on the period up to 2005. After 2005 there were some changes in childcare subsidies and tax credits for parents with a youngest child up to 12 years old that might influence our control group, although the results for the period up to 2008 are similar to the results for the period up to 2005.<sup>14</sup> Table 2 gives the mean and standard deviation of the left and right hand side variables for the DD regression for the treatment group, and the differences in the means between the treatment and control groups before and after the reform (up to 2005).

The change in the mean difference in the participation rate before and after the reform already gives an indication of the reform effect, not controlling for changes in the covariates in the treatment and control group. Using our preferred control group, single mothers with a somewhat younger or older youngest child than the treatment group, we find a negative treatment effect of  $-1.5\%$ -points. Using the alternative control group, single women without children, we find a positive treatment effect of  $+2.7$  ppt.

Turning to the controls, single mothers in the treatment group are on average slightly older than single mothers in our preferred control group.<sup>15</sup> Single women without children in our sample are much younger than single mothers.<sup>16</sup> Differences in the level of education and ethnicity with the treatment group are small for our preferred control group, but much larger for single women without children. The

---

<sup>14</sup>Compare the results in Table A.4 in the Appendix to the results in Table 3 below.

<sup>15</sup>There are more single mothers with a youngest child 8-11 than single mothers with a youngest child 16-19, because at some point these children leave the household.

<sup>16</sup>It takes time to find a partner to have children with and before parents divorce in the case the single mother was part of a couple before.

Table 2: Descriptive statistics treatment and control groups DD analysis

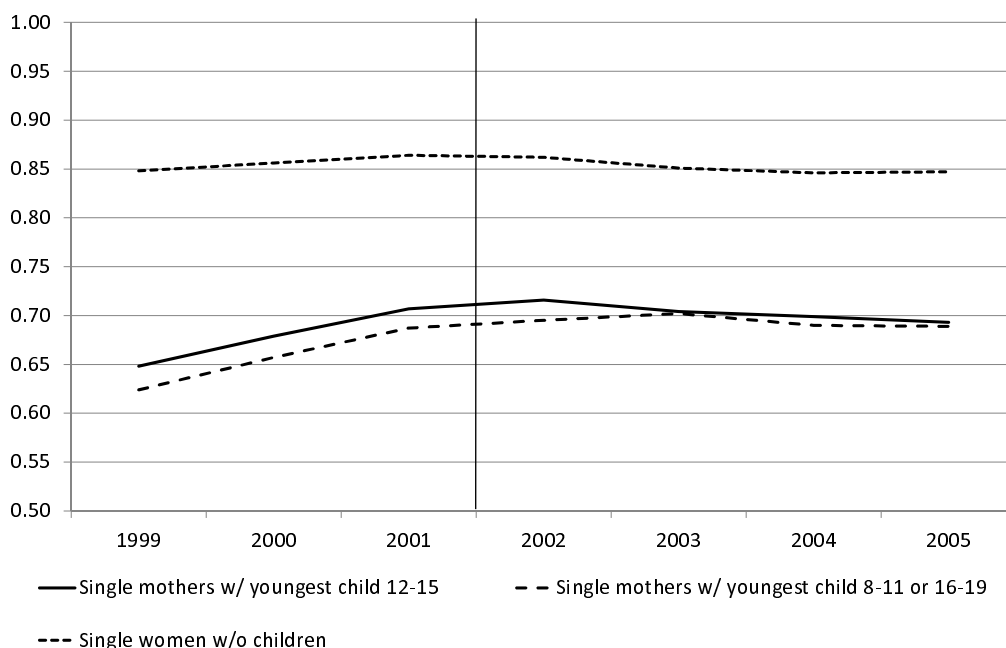
	Single mothers w/ young. child 12–15		Single mothers w/ young. child 8–11 or 16–19		Single women w/o children	
	1999-2001		Difference in means (Treat–Control)		Difference in means (Treat–Control)	
	Mean	SD	1999–2001	2002–2005	1999–2001	2002–2005
Participation	0.679	0.467	0.023	0.008	–0.176	–0.149
Age	44.12	4.956	0.944	0.974	7.083	6.561
Higher educated	0.230	0.421	0.026	0.011	–0.177	–0.174
Medium educated	0.390	0.488	–0.006	–0.009	–0.004	–0.005
Lower educated	0.380	0.485	–0.020	–0.002	0.181	0.180
Native	0.712	0.453	0.006	0.002	–0.128	–0.128
Non–Western Immigrant	0.167	0.373	–0.008	0.001	0.102	0.111
Western Immigrant	0.120	0.325	0.003	–0.003	0.026	0.017
One child	0.434	0.496	–0.039	–0.030	–	–
Two children	0.441	0.497	0.037	0.031	–	–
Three or more children	0.125	0.331	0.002	–0.001	–	–
Observations 1999–2005	19,358		37,206		161,686	

Source: Labour Market Panel (Statistics Netherlands).

shares of single mothers with one, two or three or more children in the treatment group are comparable to these shares for single mothers in our preferred control group. Changes in the differences in means are small for all controls in our preferred control group, which indicates that the composition of the groups is rather stable so that there is a fixed group effect. Indeed, we will see that the fixed effects estimates, which capture changes in the composition of unobserved fixed characteristics, are very close to the DD estimates. Changes in the differences in means are small for most controls in the control group of single women without children as well, though there is a drop in the average age of the control group relative to the treatment group.

In the appendix we present so-called normalized differences for both control groups, which are mean differences divided by the square root of the sum of variances (see Table A.1 and Table A.2). Imbens and Wooldridge (2009) argue that this is an informative way to see if the treatment and control group have sufficient overlap in

Figure 2: Participation pre and post-reform treatment and control groups



Source: Labour Market Panel (Statistics Netherlands).

the covariates. As a rule of thumb they suggest that when the normalized difference exceeds a value of .25, linear regression becomes sensitive to the specification. The normalized differences stay well below .25 for all our covariates for our preferred control group. However, this condition is not satisfied for the control group of single women without children for the age of the parent/single and the shares of lower and higher educated (and some covariates have a normalized difference just below .25). Next to differential trends (see below), this points at another potential problem for using this group as a control group for single mothers.

Figure 2 gives the participation rates for the treatment and control groups over the period 1999–2005. We see that the change in the participation rate of the treatment group and our preferred control group is very similar up to the reform, they move in tandem. Hence, an eyeball test suggests that this control group is a valid control group, but this is not controlling for differences in covariates and changes in the composition of unobserved fixed effects. We also see that the profile

of the participation rate of single women without children before the reform is much more flat than for single mothers. Hence, when we use this group as a control group the ‘treatment effect’ is likely to capture also differential trends. Below we will formally test this using a pre-reform placebo treatment dummy.

We first estimate equation (1) with our preferred control group. Table 3 reports the estimated treatment effects (full estimation results can be found in Table A.3 in the appendix). In column (1) we first present DD estimates without demographic controls, this is the simple DD treatment effect that we calculated before using the descriptive statistics in Table 2. The treatment of  $-1.5\%$ -points is insignificant. When we add demographic controls, column (2), the treatment effect becomes  $-0.7\%$ -points. In column (3) we fully exploit the panel nature of our dataset by including individual fixed effects. The treatment effect is now only  $-0.2\%$ -points and there is a substantial drop in the standard error. This is our preferred specification. Note that the insignificance of the coefficient is not due to a lack of statistical power. Indeed, the 95% confidence interval of  $[-0.014, 0.009]$  is quite tight. To check the common trend assumption, column (4) adds a pre-reform placebo treatment dummy for 2001. We find an insignificant placebo treatment effect of  $-0.1\%$ -points, while the treatment effect for the post-reform period is hardly affected, supporting our common trend assumption for our preferred control group.

Table 4 gives the estimated treatment effects using the alternative control group of single women without children. In the regression without demographic controls, column (1), we again find the simple DD treatment effect we calculated before,  $+2.7\%$ -points, and more importantly, this treatment effect is highly significant. When we add demographic controls, column (2), the treatment effect remains similar and highly significant. When we add fixed effects in column (3), the treatment effect rises to  $+3.8\%$ -points. However, when we add a pre-reform placebo treatment dummy for 2001 to the regression we find that with  $4.3\%$ -points it is large and also highly significant, while the treatment effect in the post-reform is inflated to  $6.4\%$ -points. This indicates that our single women without children do not share the same trend as our treatment group, and are therefore not a valid control group for our treatment group.

An extensive robustness analysis of the DD analysis is given in the appendix. First, Table A.4 and A.5 show that running the regressions for the longer period 1999-2008 yields similar results, although the treatment effects for the regressions



Table 3: Treatment effect DD using single mothers w/ youngest child 8–11 or 16–19 as the control group

	(1)	(2)	(3)	(4)
	DD w/o controls <sup>a</sup>	DD w/ controls <sup>b</sup>	DD w/ FE <sup>c</sup>	DD w/ FE and placebo <sup>d</sup>
TreatmentDD	-0.0149 (0.0119)	-0.0066 (0.0110)	-0.0024 (0.0057)	-0.0030 (0.0074)
Placebo				-0.0012 (0.0082)
Observations	56,564	56,564	56,564	56,564

Sample period 1999-2005. Robust standard errors clustered at the individual level in parentheses, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. <sup>a</sup>Including year dummies and a group dummy for parents with a youngest child 12–15 years old. <sup>b</sup>The additional control variables are listed in Table 2. Results for control variables are in Table A.3 in the Appendix. <sup>c</sup>With individual fixed effects. <sup>d</sup>Including a placebo treatment dummy for the pre-reform year 2001.

Table 4: Treatment effect DD using single women w/o children as the control group

	(1)	(2)	(3)	(4)
	DD w/o controls <sup>a</sup>	DD w/ controls <sup>b</sup>	DD w/ FE <sup>c</sup>	DD w/ FE and placebo <sup>d</sup>
TreatmentDD	0.0270*** (0.0090)	0.0260*** (0.0084)	0.0380*** (0.0063)	0.0642*** (0.0084)
Placebo				0.0432*** (0.0074)
Observations	181,044	181,044	181,044	181,044

Sample period 1999-2005. Robust standard errors clustered at the individual level in parentheses, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. <sup>a</sup>Including year dummies and a group dummy for parents with a youngest child 12–15 years old. <sup>b</sup>The additional control variables are listed in Table 2. <sup>c</sup>With individual fixed effects. <sup>d</sup>Including a placebo treatment dummy for the pre-reform year 2001.

with single women without children are somewhat larger (because the differential trend widens the gap in the participation rate with the treatment group further in later years). In the baseline regressions participation is determined by using the social economic category indicator of Statistics Netherlands. In Table A.6 we consider alternative definitions which are based on whether annual taxable wage income exceeds 50%, 70% or 100% of the annual minimum wage, respectively. These indicators may also capture short employment spells or incomplete work histories which are not captured by the social economic category indicator. These alternative definitions yield similar results. Finally, Table A.7 shows that running the DD equation for subgroups yields small insignificant treatment effects for all subgroups as well.

## 6 RD analysis

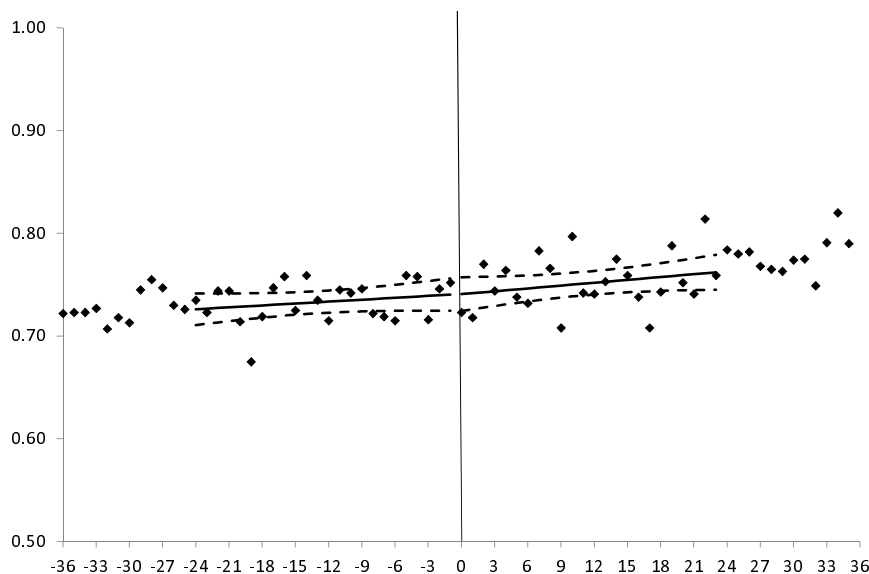
Next, we turn to the RD analysis. In the RD analysis we focus on parents with a youngest child 14 to 17 years of age, using data from the post-reform period. We know the age of the youngest child in months on the 1st of January in each year. Single mothers with a youngest child 192 months of age on the 1st of January still qualify for the subsidy, whereas single mothers with a youngest child 193 months of age on the 1st of January do not.<sup>17</sup> Since we are looking at children 14 to 17 years of age, the childcare reform and the reform of the tax credit for parents with a child less than 12 years old after 2005 are unlikely to affect our results. Therefore, to maximize the statistical power of our tests, we focus on the RD results for the full post-reform period 2002-2008, although using the shorter period 2002-2005 yields similar results (compare Table A.10 and Table 5).

First, we again perform an eyeball test. Figure 3 shows the relation between the participation rate of the single mother and the age of the youngest child in months. We plot the age of the youngest child relative to the discontinuity (we subtract 193 months). Single mothers to the left of the discontinuity, marked by the solid vertical

---

<sup>17</sup>A child that is 16 years of age on the 1st of January and born in December will be  $16 \cdot 12 + 1 = 193$  months old on the 1st of January. This child was born 1 month ‘too early’ to still qualify for the subsidy 16 years later. A child that is 15 years of age on the 1st of January and born in January will be  $15 \cdot 12 + 12 = 192$  months on the 1st of January. This child was born 1 month ‘late enough’ to still qualify for the subsidy almost 16 years later.

Figure 3: Participation rate by month of birth youngest child relative to discontinuity



The solid lines give the predicted values of a reduced-form RD without year dummies and demographic control variables, estimated separately on the left and right hand side of the discontinuity. The dotted lines represent the 95% confidence intervals.

line, still qualify for the subsidy, single mothers to the right do not. We also plot a regression line (solid lines) along with the 95% confidence interval (dotted lines) of a simple linear regression of the participation rate of the single mother on the age of the youngest child in months, estimated separately on data to the left and to the right of the discontinuity. We do not observe a discontinuity in the relation between participation and age of the child.

In Table 5 we present a more formal analysis, where we estimate variants of equation (2). In column (1) we present results for the treatment effect without (year and demographic) controls, and assuming the same linear relation between the age of the youngest child and the participation rate of the single mother to the left and to the right of the cutoff. We find an insignificant negative treatment effect of  $-0.01\%$ -points. Note that to arrive at the effect of the subsidy, we need to reverse the sign of this coefficient, since we are measuring the drop in the participation

rate to the right of the discontinuity relative to the left of the discontinuity. In column (2) we add a quadratic term in the age of the youngest child, but the coefficient is insignificant and also renders the linear term in the age of the youngest child insignificant. In column (3) we add an interaction term between the age of the youngest child and a dummy which is 1 when the age of the youngest child exceeds the cutoff, to allow for a different slope in the relation between the age of the youngest child and the participation rate of the parent to the right of the discontinuity. The coefficient is insignificant. The treatment effect becomes more negative, but is (highly) insignificant. When we add controls to the specification in column (1), column (4), the slope coefficient for the age of the youngest child becomes less significant. The treatment effect switches sign but is still small and insignificant.<sup>18</sup> Adding a quadratic term in column (5) still renders all estimates insignificant, as does adding a separate slope coefficient for the age of the youngest child to the right of the discontinuity in column (6). Our preferred specification is column (4), with an insignificant treatment effect of 0.4%-points, which again measures the change in the participation rate for individuals that do not qualify for the subsidy relative to individuals that do qualify for the subsidy. Note that the insignificance of the coefficient is again not due to a lack of statistical power, as the 95% confidence interval of  $[-0.016, 0.023]$  is still quite tight.

Table 6 gives the results of a difference-in-discontinuity analysis, see equation (3). We now take the full sample period 1999-2008, and allow for a pre-reform discontinuity at the same age of the youngest child as the post-reform discontinuity. Although we are unaware of *e.g.* a policy that would cause this pre-reform discontinuity, we can still test for it using this setup. In column (1) we indeed do not find a significant pre-reform discontinuity, looking at the coefficient of TreatmentRD. Furthermore, we do not find a significant difference in this discontinuity in the post-reform period, looking at the coefficient of TreatmentRD  $\times$  1(year>2001). In column (2) we add a separate slope coefficient for the relation with the age of the youngest child to the right of the discontinuity, but again this coefficient is insignificant.

An extensive robustness analysis of the RD analysis is given in the appendix. First, Table A.10 shows that running the regressions for the shorter period 1999-2005

---

<sup>18</sup>Adding fixed effects leads to similar small and insignificant treatment effects, see Table A.9 in the appendix. However then also the age of the youngest child variable becomes insignificant, therefore we decided to present results for the RD using the specifications without FE.

Table 5: Treatment effect using RD

	Without controls			With controls		
	(1)	(2)	(3)	(4)	(5)	(6)
TreatmentRD	-0.0001 (0.0105)	-0.0004 (0.0105)	-0.0596 (0.1186)	0.0038 (0.0099)	0.0031 (0.0099)	-0.1265 (0.1141)
Age y. child in months	0.0008* (0.0004)	-0.0020 (0.0049)	0.0006 (0.0005)	0.0010** (0.0004)	-0.0047 (0.0047)	0.0007 (0.0005)
(Age y. child in months) <sup>2</sup> /100		0.0007 (0.0013)			0.0015 (0.0012)	
(Age y. child in months) x 1(Y. child>15 yrs old)			0.0003 (0.0006)			0.0007 (0.0006)
Observations	19,778	19,778	19,778	19,778	19,778	19,778

Sample period 2002–2008. Robust standard errors clustered at the individual level in parentheses, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table 6: Treatment effect using difference-in-discontinuity

	(1)	(2)
TreatmentRD	0.0161 (0.0159)	-0.0935 (0.1021)
TreatmentRD x 1(year>2001)	-0.0143 (0.0155)	-0.0145 (0.0155)
Age young. child in months	0.0011*** (0.0004)	0.0009** (0.0004)
(Age young. child in months) x 1(young. child>15 yrs old)		0.0006 (0.0005)
Observations	23,952	23,952

Sample period 1999–2008. Robust standard errors clustered at the individual level in parentheses, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

yields similar results. Table A.11 shows the results of a placebo RD regression for the period 1999-2001. In line with the difference-in-discontinuity analysis above, we do not find a significant discontinuity in the pre-reform period either. Table A.12 shows that the alternative definitions of participation based on whether taxable wage income exceeds 50%, 70% or 100% of the annual minimum wage, respectively, yields similar insignificant treatment effects. Table A.13 shows that running the RD equation for subgroups yields small insignificant treatment effects for all subgroups as well. When we consider a more narrow age window for the age of the youngest child of 15 to 16 years old, we get similar results for the treatment coefficients, but the age of the youngest child becomes insignificant, see Table A.14. When we consider a wider age window for the age of the youngest child of 13 to 18 years old, we get similar results for the treatment coefficients, see Table A.15. Here, the proverbial exception is the specification where we include a separate slope for the age of the youngest child to the right of the discontinuity. This results in large, significant negative treatment effects, but this is driven by the observations far away from the discontinuity and renders the coefficient for the age of the youngest child insignificant. We also consider the results when we use a wider bandwidth for the age of the youngest child, the age of the youngest child in quarters rather than months. Table A.16 shows that the results are similar to the specification in months. Also, running the difference-in-discontinuity specification using quarter of birth yields results similar to the base specification, see Table A.17. Although we do not find a significant treatment effect using RD, it is common in RD analysis to check for discontinuities in density of the forcing variable (age of the youngest child) to consider *e.g.* behavioural responses to the introduction of the discontinuity and to check for discontinuities in the covariates around the discontinuity. Figure A.1 shows that there are no abrupt changes in the number of observations close to the discontinuity.<sup>19</sup> Figure A.2 and Figure A.3 show that there are no sudden changes in the relation between the age of the youngest child and the covariates.

Finally, before the reform in 2002 the discontinuity in the subsidy was at the age

---

<sup>19</sup>There is a drop in the observations from 1 month before the discontinuity to 1 month after the discontinuity. However, this is because more children are born in January (just after the cutoff) than in December (just before the cutoff) every year, as can be seen when we look at the number of observations for single mothers with a youngest child 12 months younger or older than the discontinuity, marked by the dotted lines.

of 11 rather than at the age of 15. Hence, we can also study the effect of this subsidy for parents with a youngest child around the cutoff of 11 years of age. Table A.18 shows the results of a RD analysis using data for the period 1999-2001 and an age window of 10 to 13 years of age for the youngest child. Also for this discontinuity we do not find a significant coefficient for TreatmentRD. Table A.19 presents the results of the difference-in-discontinuity analysis, indicating that the removal of the discontinuity also did not have a significant effect; both the TreatmentRD and the TreatmentRD interacted with year>2001 are insignificant.

## 7 Discussion and concluding remarks

In 2002 the eligibility conditions of the tax credit for working single parents were relaxed to stimulate the labour force participation of single parents. We use this natural experiment to study the responsiveness of labour supply by single mothers to financial incentives using difference-in-differences and regression discontinuity. Both the DD and RD analysis point to a small if not zero effect. This is not due to a lack of statistical power, as the confidence intervals are quite tight around zero. A number of robustness checks show that our results are robust. These results are at odds with the findings of related studies on single parents abroad. Below we consider a number of potential explanations for this discrepancy and whether or not they seem plausible.

First, in contrast to other studies, we examine a policy change targeted at a group of single mothers with a relatively old youngest child, 12 to 15 years old. The other studies consider all single mothers, including those with small children. Single mothers with an older child have a higher participation rate. Indeed, in our case, 71% of the treatment group was already working, potentially leaving less room for a sizeable increase. A negative relation between the participation rate of women and their labour supply elasticity can be considered a stylized fact in the empirical labour supply literature, over time and over countries, see e.g. Heim (2007) and Bargain et al. (2012). However, Table 1 shows that the EITC reforms in the US had a sizeable effect on the participation rate, notwithstanding an initial participation rate of 74%. Hence, this explanation for the discrepancy with the related studies seems questionable.

Second, the tax change might not have been big or salient enough to generate

a significant effect. Chetty (2012) shows that small tax changes can have little or no effect on labour supply due to friction costs or inattention,<sup>20</sup> while larger tax changes induce much larger responses. Regarding the size of the reform, the reform increased net income of a modal single mother by around 5% in our case. This is quite similar to the impulse in the reforms considered in other studies, see Table 1. Hence, this explanation seems questionable. Furthermore, Chetty (2012) shows that the utility loss of ignoring changes in financial incentives on the extensive margin are of first-order, whereas the utility loss of ignoring changes on the intensive margin are only of second-order. As a result, frictions and inattention are less likely to generate a small elasticity when the actual elasticity is large for the extensive margin than the intensive margin. In the last column of Table 1 we calculate the lower and upper bound of the true elasticity, given the point estimate of the participation rate elasticity and assuming friction costs of 1%, using the methodology outlined in Chetty (2012, Section 3.3). The lower bound of the true elasticity in related studies is always above the upper bound of the true elasticity in our study, because the elasticities differ substantially between our study and the related studies, and the impulses of the reforms are sufficiently big. Hence, frictions and inattention are an unlikely explanation for the discrepancy with the related studies.

Third, demand side restrictions may have prevented single mothers from realizing their preferred labour supply choice, due to unfavourable business cycle conditions during the time of the reform or due to minimum wage legislation. Indeed, around 2002, the Netherlands, like most other countries, experienced a business cycle downturn. However, the treatment effect we find is small and not significantly different from zero not only in the first years after the reform, but also in later years. As for minimum wages, these are relatively high in the Netherlands. Immervoll (2007, Table 2) shows that gross earnings of a full-time minimum-wage worker in the Netherlands as a % of gross average wages was 49% in 2002, compared to only 33% in the UK and 37% in the US.<sup>21</sup> However, although minimum wages are relatively high in the Netherlands, the initial relatively high participation rate of the treatment group suggests that this is not an important hurdle for our treatment

---

<sup>20</sup>We have been looking for information on potential non take-up of the tax credit we consider, but so far have come up empty.

<sup>21</sup>Furthermore, minimum wage levels relative to average wage levels are stable for the Netherlands, the UK and the US over the data period we consider.



group. Also, even if minimum wages reduce the level of participation, it is then not clear whether they reduce the elasticity of participation as well. Hence, demand side restrictions do not provide a likely explanation for the discrepancy with the related studies.

Finally, we may find smaller effects because we use other single mothers as a control group rather than single women without children. The other studies use single women without children as the control group. The use of this control group for single mothers has been criticized by Blundell and MaCurdy (1999) and Meghir and Phillips (2010). Single women are quite different from single mothers in their labour supply behaviour. Indeed, our data suggest that they have a different pre-reform trend than single mothers and they also differ significantly in characteristics like age and level of education. When we use this invalid control group we also find sizeable effects, but this is not due to the reform. Using our preferred control group of single mothers with a somewhat younger or older child than the treatment group we find a small effect, not significantly different from zero. Furthermore, our DD results are supported by a RD and a difference-in-discontinuity analysis, that also suggest a small effect, not significantly different from zero but with tight confidence intervals. Hence, the use of a different control group seems a plausible explanation for the discrepancy with the related studies.

An interesting avenue for future research would be to scrutinize these differences more closely by re-estimating the treatment effects of reforms in different countries using different control groups. Furthermore, it would be interesting to look more closely at the role played by the size of the shock and the salience of the reform on the effect on the participation rate across different reforms.

## References

- Angrist, J. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics*. Princeton University Press.
- Bargain, O., Orsini, K., and Peichl, A. (2012). Comparing labor supply elasticities in Europe and the US: new results. mimeo, IZA, Bonn.
- Blundell, R., Brewer, M., and Shephard, A. (2005). Evaluating the labour market

- impact of working families' tax credit using difference-in-differences. HM Revenue and Customs, Working Papers 4.
- Blundell, R. and Dias, M. C. (2009). Alternative approaches to evaluation in empirical microeconomics. *Journal of Human Resources*, 44(3):565 – 640.
- Blundell, R. and MaCurdy, T. (1999). Labor supply: a review of alternative approaches. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 3A. North-Holland.
- 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., Duncan, A., Shephard, A., and Suarez, M. (2006). Did the working families' tax credit work? The impact of in-work support on labour supply in Great Britain. *Labour Economics*, 13:699–720.
- Chetty, R. (2012). Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply. *Econometrica*, 80(3):969–1018.
- CPB (2010). Analyse economische effecten financieel kader (economic impact coalition agreement). CPB Notitie 2010/33.
- Eissa, N. and Liebman, J. (1996). Labor supply response to the earned income tax credit. *Quarterly Journal of Economics*, 111:605–637.
- Francesconi, M. and van der Klaauw, W. (2004). The consequences of in-work benefit reform in Britain: new evidence from panel data. Institute for Social and Economic Research, Working Paper 2004-13.
- Gregg, P. and Harkness, S. (2003). Welfare reform and lone parents employment in the UK. Centre for Market and Public Organisation, Working Paper 03/072.
- Heim, B. T. (2007). The incredible shrinking elasticities: married female labor supply, 1978–2002. *Journal of Human Resources*, 42(2):881–918.
- Hotz, V. and Scholz, J. (2003). The earned income tax credit. In Moffitt, R., editor, *Means-Tested Transfer Programs in the U.S.* . University of Chicago Press.

- Imbens, G. and Lemieux, T. (2008). Regression discontinuity designs: a guide to practice. *Journal of Econometrics*, 142:615–635.
- Imbens, G. and Wooldridge, J. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47:5–85.
- Immervoll, H. (2007). Minimum wages, minimum labour costs and the tax treatment of low-wage employment. OECD Social, Employment and Migration Working Papers No. 46, Paris.
- Lee, D. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, 48:281–355.
- Leigh, A. (2005). Optimal design of earned income tax credits: evidence from a British natural experiment. Australian National University, Centre for Economic Policy Research, Discussion Paper 488.
- Meghir, C. and Phillips, D. (2010). Labour supply and taxes. In Mirrlees, J., editor, *The Mirrlees Review*, chapter 3. Oxford University Press.
- Meyer, B. and Rosenbaum, D. (2001). Welfare, the earned income tax credit, and the labor supply of single mothers. *Quarterly Journal of Economics*, 116:1063–1114.
- Ministry of Finance (2001). Wetsvoorstel Belastingplan 2002: Arbeidsmarkt en inkomensbeleid. Press release 01/247.
- Stancanelli, E. G. (2008). Evaluating the impact of the French tax credit on the employment rate of women. *Journal of Public Economics*, 92(10-11):2036–2047.
- Statistics Netherlands (2011). Documentatierapport Sociaal Economische Categorie (SEC) 1999-2008V2.
- Statistics Netherlands (2012). Documentatierapport Arbeidsmarktpanel 1999-2009V1.

Table A.1: Descriptive statistics single mothers with youngest child 12–15 vs. single mothers with youngest child 8–11 or 16–19

	Treatment Group		Differences (Treat–Control)		Normalized differences (Treat–Control)	
	Mean	SD	1999–2001	2002–2005	1999–2001	2002–2005
Participation	0.679	0.467	0.023	0.008	0.035	0.012
Age	44.12	4.956	0.944	0.974	0.122	0.126
Higher educated	0.230	0.421	0.026	0.011	0.044	0.020
Medium educated	0.390	0.488	–0.006	–0.009	–0.008	–0.013
Lower educated	0.380	0.485	–0.020	–0.002	–0.029	–0.003
Native	0.712	0.453	0.006	0.002	0.009	0.003
Non-Western Immigrant	0.167	0.373	–0.008	0.001	–0.016	0.001
Western Immigrant	0.120	0.325	0.003	–0.003	0.006	–0.006
One child	0.434	0.496	–0.039	–0.030	–0.055	–0.043
Two children	0.441	0.497	0.037	0.031	0.053	0.044
Three or more children	0.125	0.331	0.002	–0.001	0.003	–0.001
Observations 1999–2005	Treatment group	19,358	Control group	37,206		

Treatment group: single mothers with youngest child 12–15 years old; control group: single mothers with youngest child 8–11 or 16–19 years old. Normalized differences are mean differences divided by the square root of the sum of variances (see Imbens and Wooldridge, 2009). Source: Labour Market Panel (Statistics Netherlands).

Table A.2: Descriptive statistics single mothers with youngest child 12–15 vs. single women without children

	Treatment Group		Differences (Treat–Control)		Normalized differences (Treat–Control)	
	Mean	SD	1999–2001	2002–2005	1999–2001	2002–2005
Participation	0.679	0.467	–0.176	–0.149	–0.300	–0.258
Age	44.12	4.956	7.083	6.561	0.612	0.555
Higher educated	0.230	0.421	–0.177	–0.174	–0.274	–0.273
Medium educated	0.390	0.488	–0.004	–0.005	–0.006	–0.008
Lower educated	0.380	0.485	0.181	0.180	0.288	0.285
Native	0.712	0.453	–0.128	–0.128	–0.219	–0.217
Non-Western Immigrant	0.167	0.373	0.102	0.111	0.228	0.240
Western Immigrant	0.120	0.325	0.026	0.017	0.058	0.038
One child	0.434	0.496	–	–	–	–
Two children	0.441	0.497	–	–	–	–
Three or more children	0.125	0.331	–	–	–	–
Observations 1999–2005	Treatment group	19,358	Control group	161,686		

Treatment group: single mothers with youngest child 12–15 years old; control group: single women without children. Normalized differences are mean differences divided by the square root of the sum of variances (see Imbens and Wooldridge, 2009). Source: Labour Market Panel (Statistics Netherlands).

Table A.3: Treatment effect DD using single mothers w/ youngest child 8–11 or 16–19 as a control group

	(1)	(2)	(3)	(4)
	DD w/o controls <sup>a</sup>	DD w/ controls <sup>b</sup>	DD w/ FE <sup>c</sup>	DD w/ FE and placebo <sup>d</sup>
TreatmentDD	-0.0149 (0.0119)	-0.0066 (0.0110)	-0.0024 (0.0057)	-0.0030 (0.0074)
Placebo				-0.0012 (0.0082)
Group dummy	0.0231*** (0.0087)	0.0057 (0.0081)	0.0045 (0.0051)	0.0050 (0.0068)
Age parent 25–29		-0.0995 (0.4643)	-0.8324*** (0.0049)	-0.8326*** (0.0051)
Age parent 30–34		-0.0910 (0.4629)	-0.8602*** (0.0338)	-0.8605*** (0.0339)
Age parent 35–39		-0.0205 (0.4628)	-0.8505*** (0.0367)	-0.8507*** (0.0367)
Age parent 40–44		0.0218 (0.4628)	-0.8366*** (0.0378)	-0.8368*** (0.0378)
Age parent 45–49		0.0140 (0.4628)	-0.8416*** (0.0387)	-0.8419*** (0.0388)
Age parent 50–54		-0.0352 (0.4628)	-0.8605*** (0.0399)	-0.8607*** (0.0400)
Age parent 50–57		-0.1188 (0.4631)	-0.8722*** (0.0423)	-0.8724*** (0.0424)
Two children		0.0229*** (0.0069)	0.0151** (0.0064)	0.0151** (0.0064)
Three or more children		-0.0314*** (0.0106)	0.0177 (0.0113)	0.0177 (0.0113)
Medium educated		0.2438*** (0.0084)		
Higher educated		0.3301*** (0.0092)		
Non-Western immigrant		-0.1471*** (0.0103)		
Western immigrant		-0.0640*** (0.0115)		
Observations	56,564	56,564	56,564	56,564

Sample period 1999–2005. Robust standard errors clustered at the individual level in parentheses, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Year dummies included but not reported.

Table A.4: Treatment effect DD using single mothers w/ youngest child 8–11 or 16–19 as the control group: longer data period 1999–2008

	(1)	(2)	(3)	(4)
	DD w/o controls <sup>a</sup>	DD w/ controls <sup>b</sup>	DD w/ FE <sup>c</sup>	DD w/ FE and placebo <sup>d</sup>
TreatmentDD	-0.0111 (0.0104)	-0.0031 (0.0096)	0.0091 (0.0058)	0.0098 (0.0074)
Placebo				0.0017 (0.0084)
Observations	86,404	86,404	86,404	86,404

Sample period 1999–2008. Robust standard errors clustered at the individual level in parentheses, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. <sup>a</sup>Including year dummies and a group dummy for parents with a youngest child 12–15 years old. <sup>b</sup>The additional control variables are listed in Table 2. <sup>c</sup>With individual fixed effects. <sup>d</sup>Including a placebo treatment dummy for the pre-reform year 2001.

Table A.5: Treatment effect DD using single women w/o as the control group: longer data period 1999–2008

	(1)	(2)	(3)	(4)
	DD w/o controls <sup>a</sup>	DD w/ controls <sup>b</sup>	DD w/ FE <sup>c</sup>	DD w/ FE and placebo <sup>d</sup>
TreatmentDD	0.0415*** (0.0087)	0.0365*** (0.0081)	0.0447*** (0.0064)	0.0679*** (0.0083)
Placebo				0.0425*** (0.0074)
Observations	259,760	259,760	259,760	259,760

Sample period 1999–2008. Robust standard errors clustered at the individual level in parentheses, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. <sup>a</sup>Including year dummies and a group dummy for parents with a youngest child 12–15 years old. <sup>b</sup>The additional control variables are listed in Table 2. <sup>c</sup>With individual fixed effects. <sup>d</sup>Including a placebo treatment dummy for the pre-reform year 2001.

Table A.6: Treatment effect DD using different definitions of participation

	(1)	(2)	(3)	(4)
	Base <sup>a</sup>	Wages>50% min. wage <sup>b</sup>	Wages>70% min. wage <sup>c</sup>	Wages> min. wage <sup>d</sup>
TreatmentDD	-0.0024 (0.0057)	0.0013 (0.0056)	0.0056 (0.0059)	0.0071 (0.0061)
Observations	56,564	56,564	56,564	56,564

Sample period 1999–2005. Robust standard errors clustered at the individual level in parentheses, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. <sup>a</sup>Using social economic classification variable (SEC) of Statistics Netherlands to determine participation. <sup>b</sup>Participation is defined as annual taxable wage income > 50% annual minimum wage. <sup>c</sup>Participation is defined as annual taxable wage income > 70% annual minimum wage. <sup>d</sup>Participation is defined as annual taxable wage income > annual minimum wage.

Table A.7: Treatment effect DD for subgroups

	(1)	(2)	(3)	(4)
	Lower educated	Medium educated	Higher educated	Native
TreatmentDD	-0.0146 (0.0101)	0.0110 (0.0089)	-0.0034 (0.0105)	-0.0035 (0.0065)
Observations	21,861	22,745	11,958	39,942
	(5)	(6)	(7)	(8)
	Non-Western immigrant	Western immigrant	One child	Two children
TreatmentDD	-0.0062 (0.0149)	0.0085 (0.0183)	-0.0057 (0.0088)	-0.0054 (0.0093)
Observations	10,077	6,545	25,915	23,603

Sample period 1999–2005. Robust standard errors clustered at the individual level in parentheses, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table A.8: Treatment effect using regression discontinuity

	Without controls			With controls		
	(1)	(2)	(3)	(4)	(5)	(6)
TreatmentRD	-0.0001 (0.0105)	-0.0004 (0.0105)	-0.0596 (0.1186)	0.0038 (0.0099)	0.0031 (0.0099)	-0.1265 (0.1141)
Age y. child in months	0.0008* (0.0004)	-0.0020 (0.0049)	0.0006 (0.0005)	0.0010** (0.0004)	-0.0047 (0.0047)	0.0007 (0.0005)
(Age y. child in months) <sup>2</sup> /100		0.0007 (0.0013)			0.0015 (0.0012)	
(Age y. child in months) x 1(Y. child>15 yrs old)			0.0003 (0.0006)			0.0007 (0.0006)
Age parent 30–34				0.0587 (0.0641)	0.0585 (0.0641)	0.0585 (0.0641)
Age parent 35–39				0.1145*** (0.0271)	0.1144*** (0.0271)	0.1144*** (0.0271)
Age parent 40–44				0.1354*** (0.0220)	0.1355*** (0.0220)	0.1354*** (0.0220)
Age parent 45–49				0.1306*** (0.0209)	0.1307*** (0.0209)	0.1306*** (0.0209)
Age parent 50–54				0.0790*** (0.0200)	0.0790*** (0.0200)	0.0790*** (0.0200)
Two children				0.0493*** (0.0092)	0.0494*** (0.0092)	0.0494*** (0.0092)
Three or more children				0.0245 (0.0156)	0.0245 (0.0156)	0.0245 (0.0156)
Medium educated				0.2224*** (0.0114)	0.2224*** (0.0114)	0.2224*** (0.0114)
Higher educated				0.2858*** (0.0126)	0.2858*** (0.0126)	0.2858*** (0.0126)
Non-Western immigrant				-0.1855*** (0.0146)	-0.1855*** (0.0146)	-0.1855*** (0.0146)
Western immigrant				-0.0544*** (0.0163)	-0.0544*** (0.0163)	-0.0544*** (0.0163)
Observations	19,778	19,778	19,778	19,778	19,778	19,778

Sample period 2002–2008. Robust standard errors clustered at the individual level in parentheses, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Year dummies included but not reported in columns (2), (3), (5) and (6).



Table A.9: Treatment effect RD: including fixed effects

	Without controls			With controls		
	(1)	(2)	(3)	(4)	(5)	(6)
TreatmentRD	0.0052 (0.0057)	0.0060 (0.0057)	0.1797*** (0.0875)	0.0052 (0.0058)	0.0058 (0.0058)	0.1398 (0.0872)
Age y. child in months	0.0008*** (0.0002)	0.0089*** (0.0036)	0.0012*** (0.0003)	0.0011 (0.0014)	0.0075* (0.0039)	0.0014 (0.0014)
(Age y. child in months) <sup>2</sup> /100		-0.0021** (0.0009)			-0.0017* (0.0009)	
(Age y. child in months) x 1(Y. child>15 yrs old)			0.0009** (0.0005)			0.0007 (0.0005)
Observations	19,778	19,778	19,778	19,778	19,778	19,778

Sample period 2002–2008. Robust standard errors clustered at the individual level in parentheses, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table A.10: Treatment effect RD: shorter data period 2002-2005

	Without controls			With controls		
	(1)	(2)	(3)	(4)	(5)	(6)
TreatmentRD	-0.0201 (0.0153)	-0.0211 (0.0153)	-0.1714 (0.1947)	-0.0164 (0.0144)	-0.0185 (0.0144)	-0.3857** (0.1141)
Age y. child in months	0.0016*** (0.0006)	-0.0062 (0.0080)	0.0013* (0.0008)	0.0019*** (0.0006)	-0.0148* (0.0076)	0.0010 (0.0007)
(Age y. child in months) <sup>2</sup> /100		0.0020 (0.0021)			0.0044** (0.0020)	
(Age y. child in months) x 1(Y. child>15 yrs old)			0.0008 (0.0010)			0.0019** (0.0010)
Observations	9,446	9,446	9,446	9,446	9,446	9,446

Sample period 2002–2005. Robust standard errors clustered at the individual level in parentheses, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table A.11: Placebo treatment effect using RD: pre-reform period 1999-2001

	Without controls			With controls		
	(1)	(2)	(3)	(4)	(5)	(6)
TreatmentRD	0.0048 (0.0242)	0.0040 (0.0243)	-0.1363 (0.3278)	0.0096 (0.0225)	0.0093 (0.0225)	-0.0181 (0.3077)
Age y. child in months	0.0011 (0.0009)	-0.0048 (0.0134)	0.0008 (0.0012)	0.0017* (0.0009)	-0.0001 (0.0126)	0.0016 (0.0011)
(Age y. child in months) <sup>2</sup> /100		0.0015 (0.0035)			0.0005 (0.0033)	
(Age y. child in months) x 1(Y. child>15 yrs old)			0.0007 (0.0017)			0.0001 (0.0016)
Observations	4,174	4,174	4,174	4,174	4,174	4,174

Sample period 1999–2001. Robust standard errors clustered at the individual level in parentheses, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table A.12: Treatment effect RD using different definitions of participation

	(1)	(2)	(3)	(4)
	Base <sup>a</sup>	Wages>50% min. wage <sup>b</sup>	Wages>70% min. wage <sup>c</sup>	Wages> min. wage <sup>d</sup>
TreatmentRD	0.0038 (0.0099)	-0.0061 (0.0101)	0.0005 (0.0104)	-0.0157 (0.0108)
Observations	19,778	19,778	19,778	19,778

Sample period 1999–2008. Robust standard errors clustered at the individual level in parentheses, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. <sup>a</sup>Using social economic classification variable (SEC) of Statistics Netherlands to determine participation. <sup>b</sup>Participation is defined as annual taxable wage income > 50% annual minimum wage. <sup>c</sup>Participation is defined as annual taxable wage income > 70% annual minimum wage. <sup>d</sup>Participation is defined as annual taxable wage income > annual minimum wage.

Table A.13: Treatment effect RD for subgroups

	(1)	(2)	(3)	(4)
	Lower educated	Medium educated	Higher educated	Native
TreatmentRD	0.0075 (0.0191)	0.0086 (0.0137)	-0.0108 (0.0175)	-0.0033 (0.0111)
Observations	7,375	8,176	4,227	14,209
	(5)	(6)	(7)	(8)
	Non-Western immigrant	Western immigrant	One child	Two children
TreatmentRD	0.0237 (0.0266)	0.0241 (0.0310)	0.0034 (0.0146)	-0.0029 (0.0146)
Observations	3,385	2,184	9,880	8,175

Sample period 2002–2008. Robust standard errors clustered at the individual level in parentheses, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table A.14: Treatment effect RD: narrower age range 15–16

	Without controls			With controls		
	(1)	(2)	(3)	(4)	(5)	(6)
TreatmentRD	-0.0066 (0.0205)	-0.0067 (0.0206)	-0.1168 (0.3028)	0.0006 (0.0191)	0.0003 (0.0191)	-0.1631 (0.2902)
Age y. child in months	0.0016 (0.0016)	-0.0036 (0.0267)	0.0013 (0.0018)	0.0014 (0.0015)	-0.0146 (0.0255)	0.0010 (0.0017)
(Age y. child in months) <sup>2</sup> /100		0.0013 (0.0069)			0.0042 (0.0066)	
(Age y. child in months) x 1(Y. child>15 yrs old)			0.0006 (0.0016)			0.0008 (0.0015)
Observations	9,969	9,969	9,969	9,969	9,969	9,969

Sample period 2002–2008. Robust standard errors clustered at the individual level in parentheses, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table A.15: Treatment effect RD: wider age range 13–18

	Without controls			With controls		
	(1)	(2)	(3)	(4)	(5)	(6)
TreatmentRD	-0.0011 (0.0079)	-0.0029 (0.0079)	-0.1712** (0.0794)	0.0029 (0.0075)	0.0001 (0.0075)	-0.2620*** (0.0768)
Age y. child in months	0.0008*** (0.0002)	-0.0038*** (0.0021)	0.0004 (0.0003)	0.0011*** (0.0002)	-0.0061*** (0.0021)	0.0005* (0.0003)
(Age y. child in months) <sup>2</sup> /100		0.0012*** (0.0006)			0.0019*** (0.0005)	
(Age y. child in months) x 1(Y. child>15 yrs old)			0.0009** (0.0004)			0.0014*** (0.0004)
Observations	29,004	29,004	29,004	29,004	29,004	29,004

Sample period 2002–2008. Robust standard errors clustered at the individual level in parentheses, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table A.16: Treatment effect RD: wider bandwidth, quarter of birth

	Without controls			With controls		
	(1)	(2)	(3)	(4)	(5)	(6)
TreatmentRD	0.0016 (0.0105)	0.0014 (0.0105)	-0.0475 (0.1195)	0.0055 (0.0098)	0.0049 (0.0099)	-0.1170 (0.1152)
Age y. chd in quarters	0.0021 (0.0013)	-0.0043 (0.0148)	0.0017 (0.0015)	0.0027** (0.0012)	-0.0122 (0.0143)	0.0019 (0.0014)
(Age y. chd in quarters) <sup>2</sup> /100		0.0050 (0.0115)			0.0117 (0.0111)	
(Age y. chd in quarters) x 1(Y. child>15 yrs old)			0.0008 (0.0018)			0.0019 (0.0018)
Observations	19,778	19,778	19,778	19,778	19,778	19,778

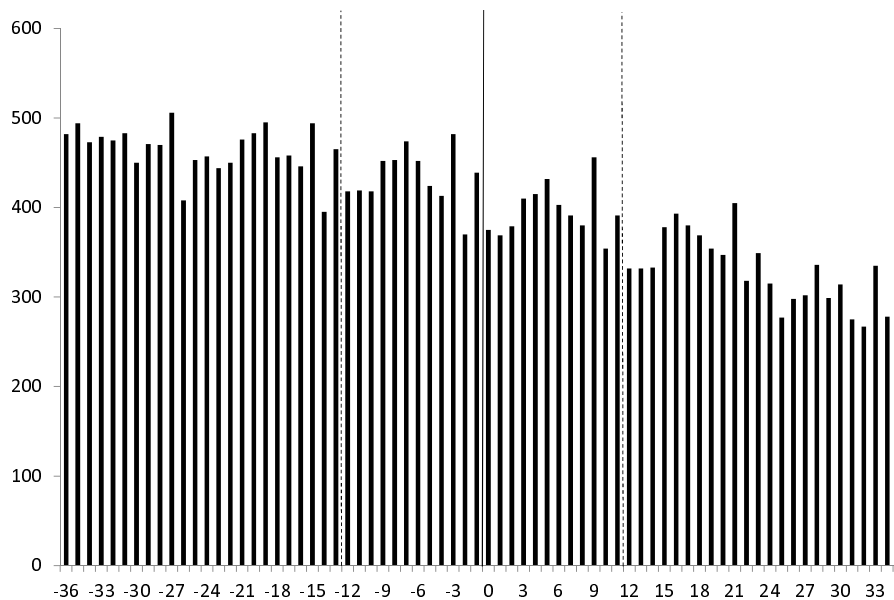
Sample period 2002–2008. Robust standard errors clustered at the individual level in parentheses, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table A.17: Treatment effect using difference-in-discontinuity: wider bandwith, quarter of birth

	(1)	(2)
TreatmentRD	0.0173 (0.0159)	-0.0955 (0.1031)
TreatmentRD x 1(year>2001)	-0.0144 (0.0155)	-0.0145 (0.0155)
Age y. child in quarters	0.0033*** (0.0011)	0.0025* (0.0013)
(Age y. child in quarters) x 1(y. child>15 yrs old)		0.0017 (0.0016)
Observations	23,952	23,952

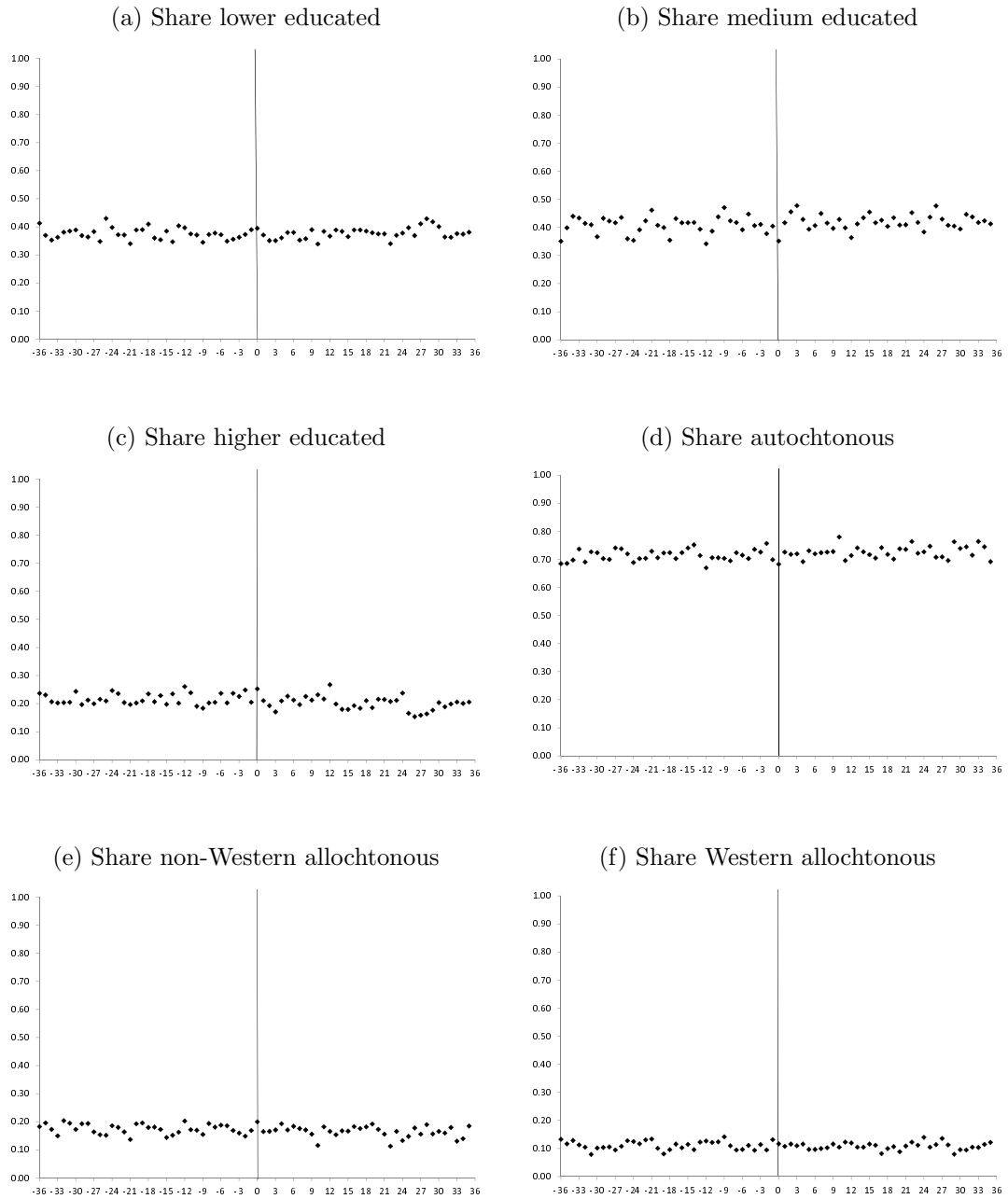
Sample period 1999–2008. Robust standard errors clustered at the individual level in parentheses, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Figure A.1: Observations by month of birth youngest child



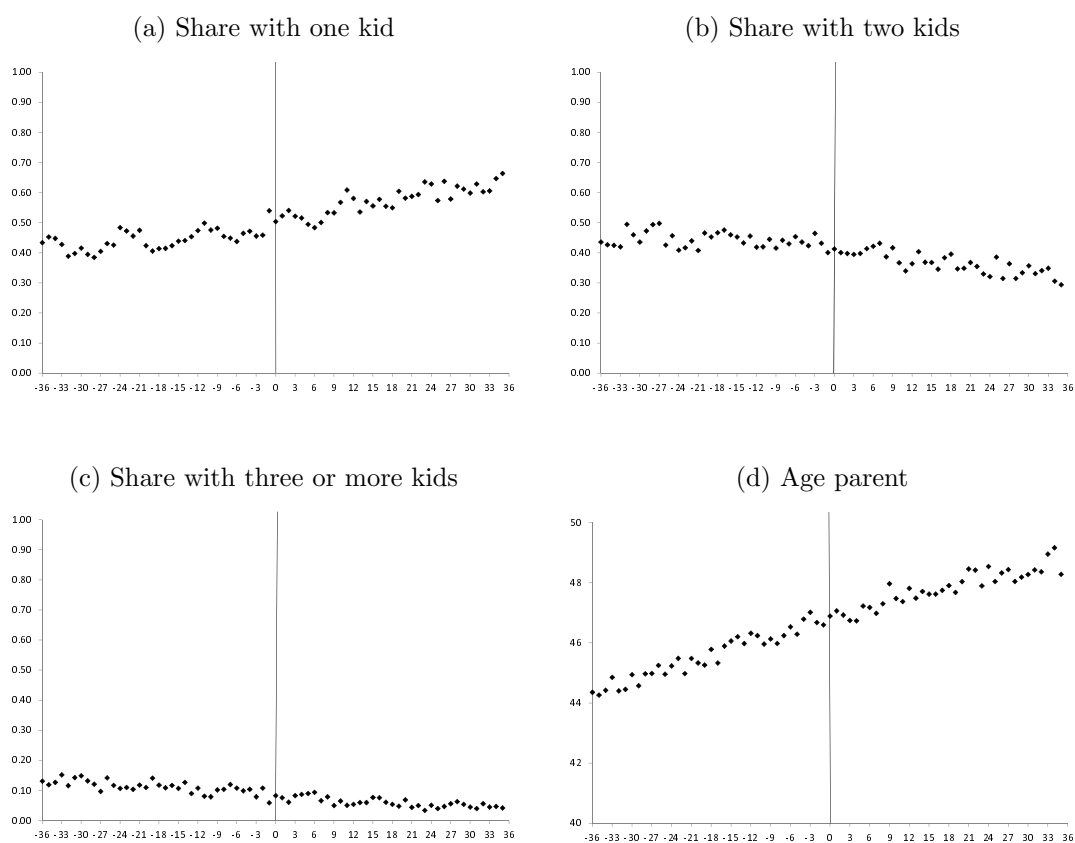
Sample period: 2002–2008. Source: Labour Market Panel (Statistics Netherlands).

Figure A.2: Control variables by month of birth youngest child



Sample period: 2002-2008. Source: Labour Market Panel (Statistics Netherlands).

Figure A.3: Control variables by month of birth youngest child



Sample period: 2002-2008. Source: Labour Market Panel (Statistics Netherlands).

Table A.18: Treatment effect RD: discontinuity at 11

	Without controls			With controls		
	(1)	(2)	(3)	(4)	(5)	(6)
TreatmentRD	-0.0032 (0.0194)	-0.0026 (0.0194)	0.1244 (0.2027)	-0.0046 (0.0183)	-0.0039 (0.0183)	0.1213 (0.1922)
Age y. chd in months	0.0016** (0.0008)	0.0060 (0.0084)	0.0020* (0.0010)	0.0016** (0.0007)	0.0067 (0.0079)	0.0020** (0.0010)
(Age y. chd in months) <sup>2</sup> /100		-0.0016 (0.0029)			-0.0018 (0.0027)	
(Age y. chd in months) x 1(Young. child>11 yrs old)			-0.0009 (0.0014)			-0.0009 (0.0013)
Observations	6,674	6,674	6,674	6,674	6,674	6,674

Sample period 1999–2001. Robust standard errors clustered at the individual level in parentheses, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table A.19: Treatment effect using difference-in-discontinuity: discontinuity at 11

	(1)	(2)
TreatmentRD	0.0011 (0.0133)	0.0454 (0.0714)
TreatmentRD x 1(year>2001)	0.0043 (0.0130)	0.0044 (0.0130)
Age young. child in months	0.0011*** (0.0003)	0.0012*** (0.0004)
(Age young. child in months) x 1(young. child>11 yrs old)		-0.0003 (0.0005)
Observations	30,221	30,221

Sample period 1999–2008. Robust standard errors clustered at the individual level in parentheses, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.







Publisher:

CPB Netherlands Bureau for Economic Policy Analysis  
P.O. Box 80510 | 2508 GM The Hague  
T (070) 3383 380

January 2013 | ISBN 978-90-5833-580-7