CPB Discussion Paper

No 155

August 2010

Analyzing female labor supply -- Evidence from a Dutch tax reform

Nicole Bosch and Bas van der Klaauw

Abstract in Dutch

Vergeleken met andere OECD-landen is de participatiegraad van Nederlandse vrouwen gemiddeld, maar het aandeel deeltijdwerk veruit het hoogst. Dit paper onderzoekt in hoeverre vrouwen met partners reageren op financiële prikkels. We maken daarbij gebruik van de exogene variatie veroorzaakt door de belastingherziening van 2001. Onze belangrijkste conclusie is dat het positieve effect van de belastingherziening op arbeidsparticipatie het negatieve effect op uren domineert. Onze meest aannemelijke verklaring is dat vrouwen eerder reageren op veranderingen in belastingvrije sommen dan op veranderingen in marginale belastingtarieven.

Analyzing female labor supply – Evidence from a Dutch tax reform

Nicole Bosch[*]   Bas van der Klaauw[†]

July 2010

Abstract

Among OECD countries, the Netherlands has average female labor force participation, but by far the highest rate of part-time work. This paper investigates the extent to which married women respond to financial incentives. We exploit the exogenous variation caused by a substantial Dutch tax reform in 2001. Our main conclusion is that the positive significant effect of tax reform on labor force participation dominates the negative insignificant effect on working hours. Our preferred explanation is that women respond more to changes in tax allowances than to changes in marginal tax rates.

[*]CPB Netherlands Bureau of Economic Policy Analysis.
Address: CPB Netherlands Bureau of Economic Policy Analysis, P.O.Box 80510,
NL–2508 GM The Hague, The Netherlands. E-mail: n.m.bosch@cpb.nl
[†]VU University Amsterdam, Tinbergen Institute and CEPR.
Address: Department of Economics, VU University Amsterdam, De Boelelaan 1105,
NL–1081 HV Amsterdam, The Netherlands. E-mail: bklaauw@feweb.vu.nl

1   Introduction

Since the work by Heckman (1974) female labor supply is an important topic for economic research. A substantial contribution to this literature was made by Blundell, Duncan and Meghir (1998), who showed that changes in tax rules can be used to estimate wage elasticities. They investigated a series of modifications in marginal tax rates over a relatively long observation period. A more recent literature focusses on the importance of financial incentives, such as earned income tax credits, on labor force participation decisions (e.g. Eissa and Hoynes, 2004). We contribute to this literature by investigating the effects of a very substantial Dutch Tax reform in 2001.

In 2001, labor force participation rates of prime-age women in the Netherlands were close to those in the UK or US. However, part-time work among women is much more common in the Netherlands than in any other OECD country. Whereas on average in the OECD about 25% of the prime-age working women work less than 30 hours per week, this is over 55% in the Netherlands (OECD, 2004). The high degree of part-time work allows for substantial room of increasing labor supply. It is also generally believed that female labor supply is more responsive to financial incentives than male labor supply (e.g. Meghir and Phillips, 2008). Therefore, Dutch policymakers have a large interest in stimulating female labor supply. This should increase economic growth and deal with the costs of an ageing society. Indeed, one of the most important motivations for the Dutch tax reform in 2001 was to make work financially more attractive for women.

The key elements of the tax reform were reducing marginal tax rates, and the replacement of tax allowances by tax credits. Both elements might affect the force participation decision and hours of work decision. The second element, however, is more likely to affect the participation decision due to the elimination of perverse incentives to work (particularly for women with a high-income partner). Tax allowances and tax credits are both transferable between partners. Whereas the tax credit is a fixed amount, the benefit of the tax allowance depends on the marginal tax rate of the partner. For non-working women with a high-income partner, transferring the tax allowance to the partner is more profitable than start working at a low income. The reason was that when working (and earning more than 4000 euro

annually), women had to use their own allowance at a low marginal tax rate, while when not working it would be transferred to their partner with a higher marginal tax rate. The change from allowance to tax credit may thus be considered as a reduction in the fixed costs of working. Saez (2002) stresses the importance of financial incentives on the decision to participate in the labor force. There are two relevant margins to investigate, labor force participation (extensive margin) and hours of work (intensive margin).

In the empirical analyses, we focus on prime-age women who are either married or co-habiting. Our empirical model is similar to the model used in Blundell, Duncan and Meghir (1998). However, we study a much shorter time period with a substantial tax reform rather than a series of tax modifications. Tax reforms provide ideal natural experiments to study the effect of financial incentives on female labor supply (see Blundell and MaCurdy, 1999; Eissa, 1995; and Eissa and Hoynes, 2004). Because the tax reform generates exogenous variation in after-tax wages, it allows us to deal with the simultaneity of working hours and after-tax wages. This simultaneity can arise, for example, because unobserved preferences or ability affect both wages and working hours, or because working hours have a direct effect on after-tax wages due to the progressive tax system. Obviously, the tax reform does not depend on individual characteristics, past choices and working hours. In the empirical model, the after-tax marginal wage will be instrumented using the tax reform. Since the tax reform was introduced at one specific moment, we should control for aggregated calendar time effects on working hours. Furthermore, we exploit that the tax reform affected different groups differently to deal with self selection into employment.

In the empirical analyses, we use the Dutch Labor Force Survey collected by Statistics Netherlands, which is a repeated cross-section containing information on (weekly) working hours, and at detailed level information on the socioeconomic structure of the household. We link this to the Social Statistical Database on Jobs, which contains administrative information on jobs and gross income. Finally, we add taxable income registered by the tax offices. The overlapping period of the three databases is from 1999 to 2005. However, we restrict the observation period to 1999-2003.

We find that the estimated uncompensated wage elasticity is about -0.13, but not significantly different from 0. This suggests that the tax reform which increased after-tax

hourly wages did not increase female labor supply. However, the tax reform had a substantial positive effect on labor force participation, which we attribute to the shift from allowance to tax credit. Female labor force participation increased by 2.4 percentage-points. Simulations with our estimated model show that the positive effect on labor force participation dominates the negative effect of wages. Whereas working women, on average, reduced weekly working hours by 0.04 hours, average working hours increased by 0.37 in the full population. The effect of the tax reform is highest for the lowest-educated women and decreases in level of education.

Our empirical results contradict earlier studies finding that the uncompensated wage elasticity is between 0 and 0.3 (see Meghir and Phillips, 2008; for a survey). In particular, our results differ from those found in Blundell, Duncan and Meghir (1998). As a sensitivity analysis we apply their grouping estimator, which gives a positive and significant uncompensated wage elasticity in the same order as found in Blundell, Duncan and Meghir (1998). The grouping estimator might be less appropriate in our setting since our observation period contains a single tax reform rather than a series of tax modification. This might explain why the grouping estimator suffers from weak instrument problems, it contains many instrumental variables which are insignificant in the first-stage regressions.

The outline of this paper is the following. Section 2 provides details about the Dutch tax reform of 2001. Section 3 introduces the empirical model. In Section 4, we discuss our data. Section 5 presents the estimation results. And Section 6 concludes.

2   The Dutch tax reform

In this section, we provide details about the Dutch tax system. We mainly focus on elements relevant for this study, and particularly on changes which occurred during the tax reform in 2001.

The Dutch tax system is an almost fully individualized progressive tax system. Prior to 2001 all individuals had a general allowance and individuals only paid taxes on income above the allowance. There are also additional allowances for working and parenting, which we discuss below. Income above the allowances is taxed according to four income brackets with

Table 1: Main characteristics of the tax system.

	1999	2000	2001	2002	2003	2004	2005
General allowance* (in €)	4674	4646					
Tax credit** (in €)			1731	1752	1715	1881	1916
Marginal tax rates in % (including national Insurance premiums)							
Lowest income bracket	35.75	33.9	32.35	32.35	32.35	33.4	34.4
Income bracket 2	37.05	37.95	37.6	37.85	37.85	40.35	41.95
Income bracket 3	50	50	42	42	42	42	42
Highest income bracket	60	60	52	52	52	52	52

Explanatory notes:
* The general allowance reduces taxable income.
** The tax credit reduces the amount of tax paid.

increasing marginal tax rates (including national insurance premiums). The tax allowances thus yield a higher tax reduction for high-income individuals with a higher marginal tax rate.

An important feature of the general allowance is that if the allowance can not be fully used, the allowance is transferred to the partner.[1] Transferring the allowance is particularly beneficial if the partner's income falls in a brackets with a higher marginal tax rate. So working at a low income is financially relatively unattractive for women with a high-income partner.

The tax reform of 2001 replaced the general allowance by a tax credit. A tax credit is simply a reduction in the total amount of taxes that an individual should pay, so independent of the marginal tax rate. Like the general allowance also the tax credit is transferable between partners. However, if a woman increases labor supply, it can never be the case that the total amount of the tax reduction decreases. The tax reform thus removed some fixed costs of working, because the tax credit does not impose any disincentive effects of working at a low income.

The tax reform in 2001 also included a reduction of marginal tax rates. Table 1 shows for each year the marginal tax rates for the four different income brackets in the Dutch

---

[1]Married and cohabiting couples have the same status in the Dutch tax system. If a partner has an annual income above 4000 euro, the allowance can not be transferred.

Figure 1: Marginal tax rates before and after the tax reform of 2001.
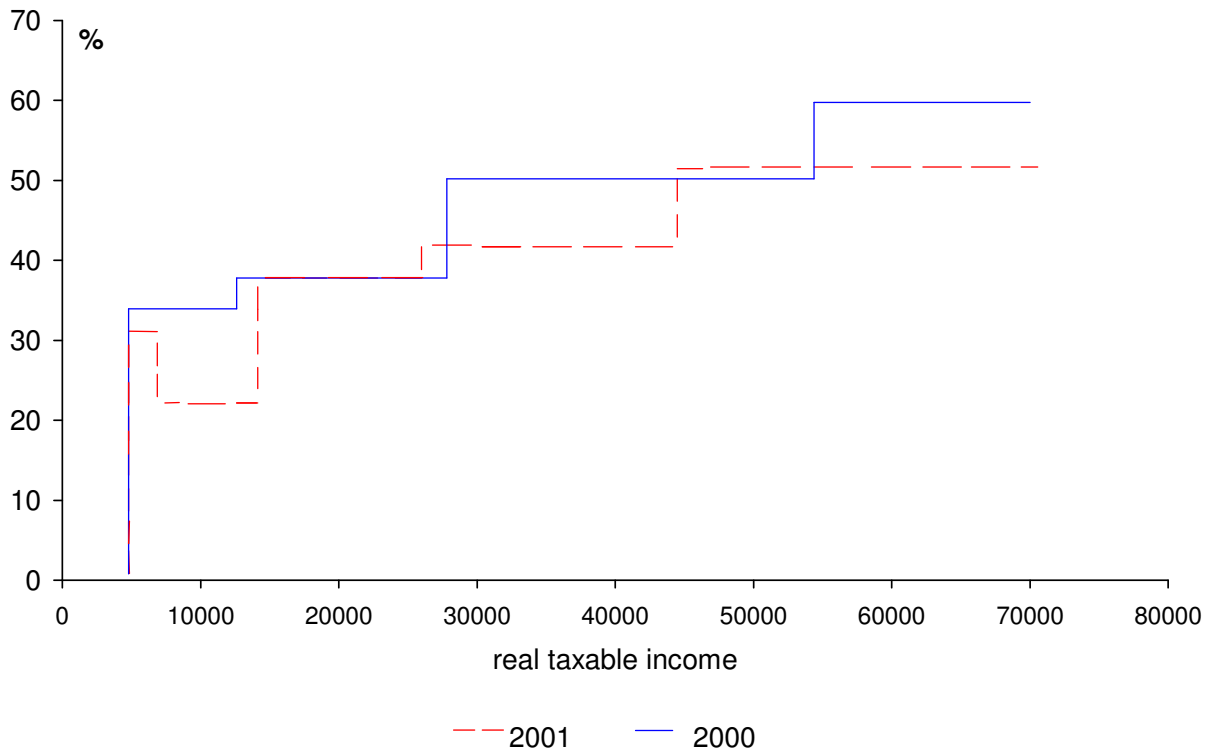
tax system. The most substantial reduction occurred in the highest two brackets, where marginal tax rates were reduced by 8 percentage-points. However, not only the marginal tax rates changed, also the cut-off points of the brackets shifted. Figure 1 shows the marginal tax rates for different taxable incomes in 2000 and 2001.

The tax reform of 2001 replaced the general allowance by a tax credit, and introduced new tax credits for parenting and the combination of working and parenting. The tax credit for parenting is transferable, but amounts to only 138 euro. The tax credit for working is more substantial, but not transferable. If an individual earns up till 50 percent of the annual (full-time) minimum wage, the tax credit increases to 150 euro. Above this, the tax credit further increases to 900 euro at an income level equal to the annual (full-time) minimum wage. The impact of the tax credit for working causes the drop in marginal tax rates between 8000 and 16,000 euro shown in Figure 1. The average taxable wage of working women is about 15,000 euro, which lies in the interval with the most substantial reduction in marginal tax rates.

Figure 2: Change in after-tax household income due to tax reform of 2001.

The tax reform of 2001 reduced labor tax for all individuals. As can be seen in Figure 2 at almost all taxable income levels the after-tax income was higher after the reform, and even more so for women with a high-income partner. However, it should be noted that some deductables were abolished. To compensate for the reduction in labor tax, the government increased consumer taxes from 17.5 to 19 percent leading to a higher inflation rate in 2001 (see Figure 3). The decision on the ta reform was made in 1999, when the economy was booming. After that, the economic growth slowed down to almost zero in the first quarter of 2002. Around the same time the unemployment rate started to increase for the first time since 1994.

3   Empirical model

This section presents our empirical model. The model describes weekly working hours and has strong similarities with the model used in Blundell, Duncan and Meghir (1998).

Figure 3: Macroeconomic outcomes.
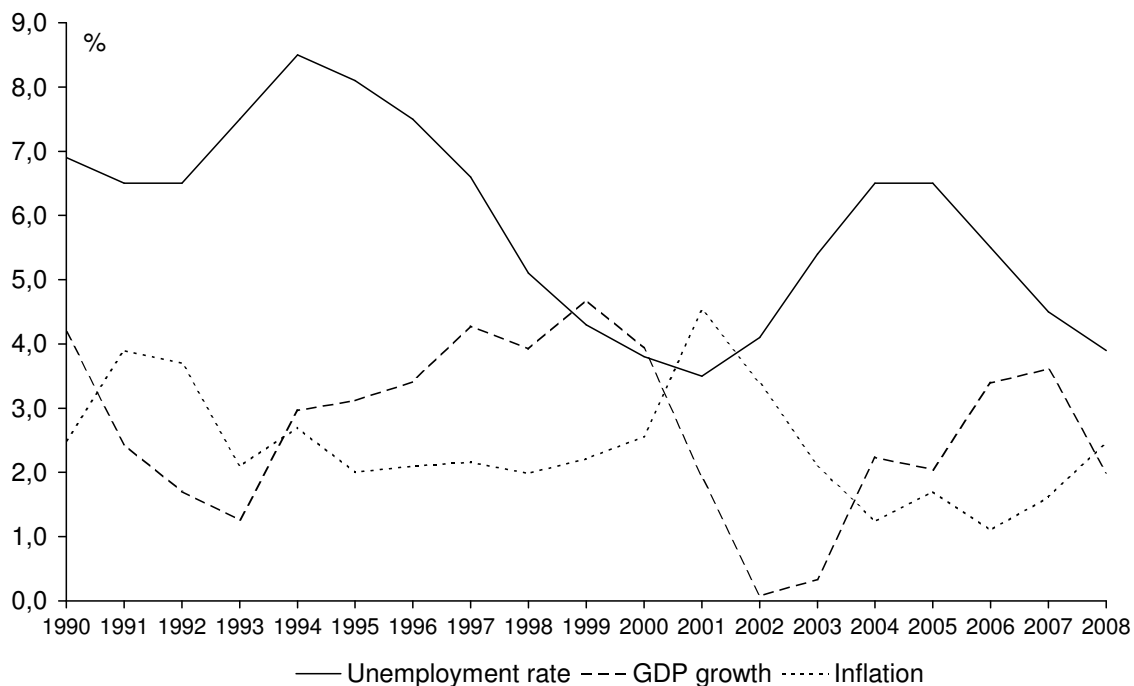
We focus on how the after-tax marginal hourly wage w_{it} of woman i in year t causally affects her weekly working hours h_{it}. Therefore, we investigate the traditional labor supply model (e.g. Heckman, 1974)

$$h_{i,t} = \beta_0 + \beta_1 \log(w_{i,t}) + x_{i,t}\beta_2 + \xi_t + \varepsilon_{i,t} \tag{1}$$

The vector $x_{i,t}$ contains observed individual characteristics such as demographic variables, cohort dummies, level of education, etc. And the function ξ_t describes common macroeconomic trends in female labor supply.

Obviously, the key parameter of interest in β_1. It is, however, well known that using OLS to estimate equation (1) yields inconsistent estimators for β_1. The logarithm of the after-tax marginal hourly wage may be correlated to the error term $\varepsilon_{i,t}$ for a number of reasons. First, there may be reversed causality. Working more hours increases annual taxable income and due to the progressity of the tax system individuals enter an income bracket with a higher marginal tax rate. Second, the vector $x_{i,t}$ probably does not capture all relevant

heterogeneity in individual preferences or ability. If there is unobserved ability and more able individuals earn higher wages and have a stronger preference for work, then there is a direct relation between $\log(w_{i,t})$ and $\varepsilon_{i,t}$.

We use the tax reform of 2001 to deal with endogeneity of $\log(w_{i,t})$. The tax reform provided some exogenous variation in after-tax marginal hourly wage. Therefore, we add the first-stage regression

$$\log(w_{i,t}) = \alpha_0 + \alpha_{i,1} \cdot I(t \geq 2001) + x_{i,t}\alpha_2 + \zeta_t + V_{i,t} \tag{2}$$

The indicator function $I(t \geq 2001)$ describing the period after the tax reform thus acts as instrumental variable. We allow the effect of the tax reform on wages $\alpha_{i,t}$ to be different for different types of individuals. In particular, we will allow for separate $\alpha_{i,1}$ for different educational groups. It should be noted that the indicator $I(t \geq 2001)$ is only a relevant instrumental variable if common macroeconomic trends ξ_t and ζ_t are smooth functions over calendar time. Our identifying assumption is thus that between 2000 and 2001 there are no general abrupt shocks in female labor supply other than due to the tax reform. This also implies that the tax reform should not be the response of policymakers to a sudden change in female labor force decisions or shifts in preferences. Although, indeed an important goal of the tax reform was to stimulate female labor force participation, there was no unexpected change in female labor supply in the period prior to the tax reform.

A second important issue is that both over time and due to the tax reform, the composition of working women might have changed. Recall that the tax reform included a shift from an allowance to a tax credit, which reduced the fixed costs of working. Furthermore, there is an increasing trend in employment rates among women. Finally, the decision to work might directly be related to unobserved preferences and ability. The self selection into work is thus most likely not random and cannot be ignored. To control for selective labor force participation $P_{i,t}$ we add a probit model

$$\Pr(P_{i,t} = 1) = \Phi\left(\gamma_0 + \gamma_{i,1} \cdot I(t \geq 2001) + x_{i,t}\gamma_2 + \psi_t\right) \tag{3}$$

Again we allow the tax reform to have a differential impact on the labor force participation

of women with different educational degrees. And we allow for a smooth trend ψ_t in female labor force participation.

We follow the estimation procedure of Blundell, Duncan and Meghir (1998), which is a control function approach. First, we estimate the reduced-form models (2) and (3) using OLS on the sample of employed women and maximum likelihood estimation on the full sample, respectively. This provides us the residuals $\hat{V}_{i,t}$ from the first-stage regression for wages and the inverse Mills ratios $\hat{\lambda}_{i,t} = \frac{\phi\left(\hat{\gamma}_0 + \hat{\gamma}_{i,1} \cdot I(t \geq 2001) + x_{i,t}\hat{\gamma}_2 + \hat{\psi}_t\right)}{\Phi\left(\hat{\gamma}_0 + \hat{\gamma}_{i,1} \cdot I(t \geq 2001) + x_{i,t}\hat{\gamma}_2 + \hat{\psi}_t\right)}$ from the participation probit. We add these as regressors in equation (1), which gives the second-stage equation

$$h_{i,t} = \beta_0 + \beta_1 \log(w_{i,t}) + x_{i,t}\beta_2 + \xi_t + \beta_3 \hat{V}_{i,t} + \beta_4 \hat{\lambda}_{i,t} + U_{i,t} \tag{4}$$

Estimating this model using OLS on the sample working women provides consistent parameter estimates for β_1. Because equation (4) includes two correction terms, at least two exclusion restrictions are required. The tax reform should thus not be restricted to having the same impact for all women on the after-tax marginal hourly wage and labor force participation (i.e. $\alpha_{i,1}$ and $\gamma_{i,1}$ should be allowed to be different for women with different levels of education). We return to this issue in Subsection 4.2 when we discuss the impact of the tax reform. Finally, we should include in the vector of regressors $x_{i,t}$ the after-tax income of the husband. The tax reform affects the income of the husband as well. If female labor supply decisions are related to the husband's after-tax income which is most likely the case, ignoring the husband's income would cause a direct correlation between the error terms $\varepsilon_{i,t}$ and the tax reform $I(t \geq 2001)$. This would violate the validity condition for using the tax reform as instrumental variable.

The parameter of interest β_1 should be interpreted as the uncompensated wage coefficient. This can be translated into the uncompensated wage elasticity of labor supply by dividing by working hours. In particular, we divide by mean working hours of employed women $\bar{h}$, so $\frac{\beta_1}{\bar{h}}$. The uncompensated wage elasticity includes both the substitution and the income effect. A positive value of β_1 implies that the substitution effect dominates the income effect.

4 Data

4.1 Sample

In the empirical analyses, we use a data set constructed from three separate databases. The Dutch Labor Force Survey collected by Statistic Netherlands is a repeated cross-section containing detailed information on socioeconomic and demographic characteristics of households. This database contains information on employment status and weekly working hours, but lacks wages and income. Therefore, we use a unique identification code to merge the Dutch Labor Force Survey with the Social Statistical Database on Jobs. This database contains information reported by employers on gross wages and annual working hours for about a random one-third of the working population. We link a third database containing taxable income registered by the tax office. The three databases share the observation period from 1999 until 2005.

We restrict the data to married or cohabiting women between age 20 and 50 whose education level is observed. Women should have a working husband with a taxable income above 9000 euro. The latter restriction is made to avoid complications with husbands transferring tax allowances and tax credits to their working wives. Furthermore, if the husband works the wive is not entitled to means-tested welfare benefits. In total 116,553 women in the Labor Force Survey satisfy the criteria above.

The employment status of women reflects the moment of the survey. If a woman is employed, she is asked to report her weekly working hours. Individuals should report their contractual hours, or if they are working without a contract the average number of hours they work during a week. In total 75 percent of the women are employed, and they work, on average, 25 hours per week.

Next, we use the Social Statistical Database on jobs to compute the gross hourly wage, which is the gross annual wage divided by annual working hours. Because the Social Statistical Database on Jobs contains one-third of the working population, we get three subsamples: (*i*) working women with observed hourly gross wages, (*ii*) working women without observed wages and (*iii*) non-working women. To reduce measurement errors in wages, we transfer women from subsample (*i*) to subsample (*ii*) if the gross hourly wage is below 90 percent

of the mandatory minimum wage or above 200 euro per hour. [2] Furthermore, we dropped women with more than one job and whose employment status in Social Statistical Database and Labor Force Survey diverged.

We add taxable income to our data set, and apply the existing tax rules to translate taxable income in after-tax income. By dividing the (annual) after-tax income by annual working hours we obtain the average after-tax hourly wage. The taxable income data is collected by the tax office for all employees. Information on annual working hours is retrieved from the Social Statistical Database on Jobs containing information for a random one-third of the employees. Therefore, we can compute the average after-tax hourly wage for one-third of the working population. In our empirical model the relevant wage rate is not the average after-tax hourly wage, but the marginal after-tax hourly wage. To obtain the marginal tax rate we increase taxable income by 10 euro. Next, we translate taxable income again in after-tax income and divide the increase in after-tax income by 10. Again, this is only for the one-third of the sample for which we observe annual working hours in the Social Statistical Database on jobs. To remove outliers from the data, we transfer women from subsample (i) to subsample (ii) if their after-tax hourly wage is below 54 percent of the gross minimum wage or above 200 euro per hour. Finally, women reporting to work over 48 hours, which is the legal maximum number of working hours per week, and women having an annual taxable income above 150,000 euro are transferred from (i) to (ii). In total for 38,423 women we observe wage information. However, recall that in the estimation we also take into account the remaining women, who are either not working or employed but without an observed wage. We have corrected wages for inflation. This takes account of the increased consumer tax in 2001.

Table 2 reports some labor market outcomes. Female labor force participation rates steadily increased between 1999 and 2003, and remained relatively constant afterwards. During most of the period in which participation rates increased, average working hours decreased. This suggests that those women who newly enter the workforce, work relatively

---

[2]In the Netherlands, the minimum wage is the same for all workers above age 23 and is formulated in terms of gross hourly wages. Below age 23 the minimum wage is age dependent. The minimum wage is corrected for inflation twice a year. In 2000 the minimum wage for those above age 23 was 6.41 euro per hour.

Table 2: Female labor market outcomes.

	1999	2000	2001	2002	2003	2004	2005
Participation (in %)	70	72	74	76	78	78	79
Weekly working hours	25.6	25.1	24.9	25.5	25.1	24.7	24.9
Average after-tax hourly wage (in €)	9.46	9.64	10.27	10.46	10.44	10.68	10.43
Marginal after-tax hourly wage (in €)	7.79	7.89	8.49	8.63	8.60	8.74	8.49

Explanatory note: Weekly working hours and hourly wage are the average for employed women. Wages are corrected for inflation.

few hours. Finally, both average and marginal after-tax hourly wages express a substantial increase in 2001. In 2004, wages again show some increase, which might have been the result of the minor changes in the tax system implemented in 2004. In the empirical analyses we only focus on the years 1999 until 2003, which reduces the sample size by 35,551 women (of which for 13,488 we observe wage information).

Individual characteristics are obtained from the Labor Force Survey. Table 3 provides summary statistics both for working and non-working women. Working women are, on average, slightly younger, have less (young) children and are higher educated. Also the partners of working women are slightly younger and, on average, higher educated. However, the after-tax annual income of partners with working women is, on average, more than 2000 euro lower.

Table 4 shows for the different educational groups some characteristics of the distribution of working hours by year. First, it should be noted that more educated women are more likely to work and also more than 32 hours. For all groups there is a negative trend in the percentage not working, however, the fraction of women working more than 24 hours per week remains constant.

4.2 Impact tax reform on wages and participation

This subsection investigates the impact of the tax reform on after-tax wages and on labor force participation. First, we provide some descriptive evidence. Next, we focus on estimating the reduced-form models for the logarithm of the after-tax marginal hourly wags and

Table 3: Weighted summary statistics of our data set.

	Working	Non-working
Age (in years)	36.3	37.9
No children (in %)	43	26
Two minor children (in %)	27	34
Three or more minor children (in %)	9	17
Age youngest child 0-3 (in %)	22	30
Age youngest child 4-11 (in %)	22	30
Age youngest child 12-17 (in %)	13	14
Highest completed level of education		
Primary + lower secondary (in %)	23	47
Higher secondary (in %)	49	41
College or university (in %)	28	12
Age partner (in years)	38.7	40.7
Highest completed level of education partner		
Primary + lower secondary (in %)	24	34
Higher secondary (in %)	45	41
College or university (in %)	32	25
After-tax annual income partner (in €)	18,446	20,535
Observations	87,858	28,695

Table 4: Fraction of women with particular working hours by year and educational level.

	No work	≤ 16 hours	16-24 hours	24-32 hours	> 32 hours
			Low education		
1999	44%	13%	20%	10%	13%
2000	42%	14%	20%	9%	14%
2001	39%	13%	22%	10%	15%
2002	36%	15%	22%	12%	15%
2003	37%	13%	21%	12%	18%
			Medium education		
1999	26%	11%	24%	14%	25%
2000	25%	11%	25%	16%	24%
2001	22%	13%	26%	15%	24%
2002	22%	13%	28%	16%	21%
2003	21%	12%	27%	18%	22%
			High education		
1999	16%	5%	21%	19%	39%
2000	15%	6%	21%	19%	38%
2001	14%	6%	22%	21%	38%
2002	13%	6%	23%	20%	38%
2003	11%	6%	21%	21%	41%

labor force participation (equations (2) and (3) respectively).

Table 5 summarizes for each year the marginal tax rate, the after-tax marginal hourly wage and the participation rate. Recall from Section 2 that the tax reform has different impacts on marginal tax rates at different parts of the income distribution. Therefore, we distinguish between low-educated, medium-educated and high-educated women. For all three groups marginal tax rates are substantially reduced due to the tax reform. The impact of the tax reform is, however, smallest for the highest educated women.

The drop in marginal tax rates due to the tax reform causes a substantial increase in after-tax marginal hourly wage for all groups. Participation rates steadily increase during the observation period, but show for the low and medium educated a jump in 2001. The impact of the tax reform thus seems to be different for different educational groups. As we argued earlier the tax reform is not correlated to individual preferences or unobserved ability. However, for drawing inference it is not only necessary that the instrument variables are not correlated to the error terms in the labor supply equation (1), but they should also be relevant. The latter implies that the instruments should have a sufficiently large impact on wages and labor force participation.

Table 6 reports the estimation results for the reduced-form equations (2) and (3). The F-test for joint significance of the instrumental variables are 33 and 16, which is above the critical value for weak instruments mentioned by Stock and Staiger (1997).[3] In the participation equation there are some differential impacts of the tax reform by educational group, but the differences are not significantly (the p-value for similarity is 0.81). In the wage equation the coefficients for the different educational groups are very similar and the p-value for similarity equals 0.74.

5   Estimation results

In this section, we first present the estimation results for the labor supply model. Next, we perform simulations with the model. Finally, we discuss some sensitivity analyses.

---

[3]Note that the weak instrument literature focuses on linear models, while our first-stage model for labor force participation is non-linear.

Table 5: Impact of the tax reform.

	marginal tax rate (in %)	marginal wage (in €)	participation rate (in %)
Low education			
1999	34.0	6.71	56
2000	33.0	6.90	58
2001	29.4	7.47	61
2002	29.4	7.33	64
2003	29.7	7.45	64
1999-2000	33.5	6.81	57
2001-2003	29.5	7.41	63
change	−4.0	0.61	6
Medium education			
1999	36.1	7.48	74
2000	36.0	7.61	75
2001	32.1	8.27	78
2002	32.4	8.48	78
2003	32.5	8.34	79
1999-2000	36.0	7.55	75
2001-2003	32.3	8.37	79
change	−3.7	0.82	4
High education			
1999	40.2	8.85	84
2000	40.5	8.86	85
2001	37.8	9.42	86
2002	38.0	9.56	87
2003	38.0	9.50	89
1999-2000	40.4	8.85	85
2001-2003	38.0	9.50	87
change	−2.4	0.64	3

Table 6: Estimation results for the reduced-form equations.

	participation		log hourly wage	
instrumental variables				
low education post reform	0.103	(0.037)	0.073	(0.018)
medium education post reform	0.091	(0.033)	0.071	(0.010)
high education post reform	0.058	(0.058)	0.061	(0.011)
F-test for instruments	16.17		33,31	
Observations	72,599		24,935	

Explanatory note: Standard errors in parentheses. Controls added for education, cohort, calendar time, household situation, education and income of partner (see Table 10 for all parameter estimates).

5.1   Parameter estimates

Table 7 presents the estimation results for the labor supply model. In the full specification of the model (column (1)), the uncompensated wage elasticity equals $\frac{-3.197}{25.00} = -0.13$. A negative wage elasticity implies that the income effect dominates the substitution effect. Our finding does not coincide with the literature surveyed by Meghir and Phillips (2008). They mention that most studies find an elasticity between 0 and 0.3. However, our wage elasticity is not significant at any commonly used significance level. Earlier studies for the Netherlands found large and positive wage elasticities, for example, Van Soest, Woittiez and Kapteyn (1990) estimated the wage elasticity to be 0.45. And Evers, De Mooij and Van Vuuren (2008) conclude, based on surveying the literature, that the wage elasticity for women in the Netherlands should be around 0.5.

It is important to stress that the instrumental variables approach measures a local average treatment effect (Imbens and Angrist, 1994). So we estimate the uncompensated wage elasticity at the margin where the tax reform affects the after-tax marginal hourly wage. Recall that the tax reform had two important elements, marginal tax rates were reduced, and allowances were transformed into tax credits. The second element caused that work became financially more attractive for women with high income partners. Our data show

Table 7: Estimation results for the labor supply model.

	(1)		(2)		(3)		(4)	
Log hourly wage	−3.197	(2.821)	−0.994	(2.790)	−0.100	(0.217)	−0.085	(0.217)
Wage residual	3.109	(2.832)	0.912	(2.802)				
Inverse Mills	−4.761	(1.011)			−4.592	(0.999)		
Observations	24,935		24,935		24,935		24,935	

Explanatory note: Robust standard errors in parentheses. Column (1) adds controls for endogenous wages and selective labor force participation. Column (2) and (3) add controls for one of the two, whereas column (4) adds no additional controls and reports the conditional correlation. Furthermore controls are added for education, sector, cohort, calendar time, household situation, education and income of partner (see Table 11 for all parameter estimates).

a large degree of assortative matching on education (the correlation in level of education between partners is 0.45). However, if we look at income, the correlation between partners is much smaller, only 0.06 in the period before the tax reform. After the tax reform this correlation increased to 0.08. After the tax reform women with high-income partners started to earn more relative to women with low-income partners. However, we saw from Table 4 that the fraction of women with many working hours remained constant during the observation period. The tax reform thus caused women with high-income partners to enter the labor force, but to work only few hours. We attribute this mainly to the change from allowance to tax credit, which removed some perverse incentives for working. Since the women who newly enter the labor market have relatively good skills (due to assortative matching), they can also earn relatively good wages, but devote fewer hours to work than would be expected. This might explain why we estimate the wage elasticity to be negative.

Our interpretation of the results is that the shift from general allowance to tax credit was the part of the tax reform which yields the most substantial incentive. In-work tax credits are often found to have stimulating effects on the labor force participation decision of women (e.g. Aaberge and Flood, 2008; Eissa and Hoynes, 2004; Eissa and Liebman, 1996; and Meyer and Rosenbaum, 2001). If indeed this shift was the key element of the tax reform, the impact of the reform might be different for women with partner's with high and low income. Therefore, we added as instrumental variable an interaction between the reform and the partner's income. However, while having the expected sign the coefficients of this instrument in the first-stage regressions were not significant, and there were no changes in the results of the second-stage regression.

The coefficient for the wage residuals is positive but not significant. Therefore, we cannot reject that the after-tax marginal wage is exogenous. However, ignoring possible endogeneity in wages reduces the wage elasticity to almost 0, i.e. $\frac{-0.100}{25.00} = -0.0$ (see column (3)). The significant coefficient of the inverse Mills ratio (in column (1)) means that there is self selection into work. The negative coefficient of the inverse Mills ratio implies that women who participate in the labor market are also women who are more likely to work relatively many hours. Ignoring selective labor market participation also reduces the wage elasticity (see column (2)). Ignoring both selective labor market participation and endogenous wages

has about the same effect as ignoring only endogeneity in wages (column (4)).

In the baseline specification, other exogenous regressors (see Table 11) are almost always significant and have the expected signs. Higher-educated, younger and cohabiting women work more hours. Women with children work less. However, working hours increase in the age of the youngest child, but reduce in the number of children present in the family. Women with a high-income partner work fewer hours. We have tried to include the partner's income squared to capture that the association between female working hours and partner's income might be nonlinear. But this term was not significant and adding this term did not change the main parameters of interest.

5.2   Model simulations

Next, we perform some simulations to get insight in the effect of the tax reform on female working hours. Table 8 shows the results of these simulations for all women and also separately for low, medium and high-educated women. In the simulations we use all women who are observed in the data in the year 2001.

The first row of the table shows the situation in 2001 (so after the tax reform). The labor force participation model predicts that 77.8 percent of the women has a job. The average after-tax marginal hourly wage is 7.88 euro. Women who are working have on average 24.46 working hours per week. However, in the full population the average number of working hours is 19.72 hours per week. The second row of the table describes the case where the tax reform only affects participation, but marginal wages remain unaffected. To some extent this mimics the situation where marginal tax rates remained as in 2000 and the tax reform only included a shift from allowance to tax credit. Of course, participation rates remain at their post reform level, but after-tax marginal hourly wages are lower. Since the wage elasticity is negative, the reduced after-tax marginal hourly wages cause a slight increase in working hours and also of working hours of labor force participants. The third row describes the case in which the tax reform would not have been implemented in 2001. It shows that labor force participation rates are substantially lower and working hours of working women are slightly higher. The last column of the table shows that the labor force participation effect of the tax reform dominates the effect of wages. So the tax reform caused average weekly working

Table 8: Model simulations.

	Participation rate	After-tax hourly wage	Hours of participants	Hours of all women
All women				
After tax reform	77.8%	7.88	24.46	19.72
Only tax reform on participation	77.8%	7.35	24.68	19.89
No tax reform	75.4%	7.35	24.50	19.35
Low educated				
After tax reform	63.7%	7.00	22.00	14.53
Only tax reform on participation	63.7%	6.51	22.24	14.68
No tax reform	60.0%	6.51	21.97	13.89
Medium educated				
After tax reform	81.2%	7.86	24.26	20.23
Only tax reform on participation	81.2%	7.32	24.48	20.41
No tax reform	78.9%	7.32	24.32	19.89
High educated				
After tax reform	89.2%	9.06	28.17	25.53
Only tax reform on participation	89.2%	8.53	28.36	25.70
No tax reform	88.2%	8.53	28.29	25.46

hours (in the full population) to increase from 19.35 to 19.72.[4]

The remainder of the table shows the same simulation separately for different education groups. For all three groups the tax reform increases labor force participation, although the effect becomes smaller as the educational level increases (and also labor force participation is already higher). In nominal terms the tax reform has about the same impact on after-tax marginal hourly wages of all three groups. However, since low-educated women have, on average, the lowest wages, the relative impact on wages of the tax reform decreases as education increases. The positive effect of the increase labor force participation dominates the negative effect of higher wages on average working hours in the full population.

5.3 Sensitivity analyses

In this subsection, we present a number of sensitivity analyses. First, we split the sample according to the presence of children. We distinguish three groups of married women. The first group consists of women with at least one child younger than age 12. The second group describes women with the youngest child above age 12 and the third group contains women without dependent children in the household. In the upper plane of Table 9 we show the estimation results for these three groups.

It should be noted that only for the group with young children, there do not seem to be problems with weak instruments. For both other groups, the instruments seem particularly weak in the probit model for labor force participation. Furthermore, it should be noted that for all groups we find an insignificant wage elasticity. We take this as evidence that splitting the sample into different groups generates too small subsamples and causes estimators to become imprecise.

Next, we split the sample according to the age of the women. We distinguish between women under age 40 and above age 40. The second plane of Table 9 shows the estimation results by age. Whereas for the younger age group the effect of wages on hour of work is negative and significant, it is positive and significant for the older age group. For younger women both

---

[4]Van Soest and Das (2001) estimate a structural model and simulate the tax reform before it was implemented. They predict positive effects on the participation rate (1.48 percent) and working hours (4 percent).

Table 9: Sensitivity analyses.

	(1)		(2)		(3)	
Log hourly wage	−1.510	(4.500)	−2.103	(7.421)	−5.622	(4.046)
Wage residual	1.073	(4.511)	2.207	(7.448)	5.396	(4.067)
Inverse Mills	0.977	(1.893)	−5.965	(4.278)	−17.047	(2.491)
	F-test statistic for instruments					
Wage equation	12.8		6.0		20.5	
Participation	21.9		0.9		5.8	
# of instruments	3		3		3	
Observations	12,240		3170		9525	

Explanatory note: Robust standard errors in parentheses.

(1) Married women with young children (youngest child under age 12)

(2) Married women with older children (youngest child between age 12 and 17)

(3) Married women without dependent children.

	(4)		(5)	
Log hourly wage	−11.205	(3.306)	4.880	(0.338)
Wage residual	10.677	(3.318)	−4.812	(5.114)
Inverse Mills	−9.673	(1.163)	−2.162	(2.150)
	F-test statistic for instruments			
Wage equation	21.3		13.1	
Participation	6.9		8.5	
# of instruments	3		3	
Observations	15,109		9826	

Explanatory note: Robust standard errors in parentheses.

(4) Women below age 40.

(5) Women above age 40.

Table 9: (Continued).

	(6)		(7)		(8)	
Log hourly wage	−3.268	(2.889)	−5.921	(1.269)	10.658	(3.813)
Wage residual			5.852	(1.279)	−9.723	(3.818)
Inverse Mills	−4.899	(1.045)	−4.814	(0.852)	1.151	(1.171)
F-test statistic for instruments						
Wage equation	31.5		202.8		2.7	
Participation	16.2		164.8		27.35	
# of instruments	3		3		24	
Observations	24,935		38,367		24,935	

Explanatory note: Robust standard errors in parentheses.
(6) Inverse Mills ratio included in wage equation.
(7) Observation extended to 1999-2005 (including quadratic time trends).
(8) Grouping estimator.

wages and labor force participation is endogenous, while this is not the case for older women. A possible explanation is that older women are less flexible on the labor market, they change jobs less often and have, therefore, also less possibilities to change working hours (although it should be noted that in the Netherlands there is some regulation that provides workers flexibility to change working hours within their job).

Now, we slightly modify the estimation method. First, we estimate the participation equation. Next, we perform 2SLS on the labor supply equation (and the wage equation) in which we include the inverse Mills ratio as exogenous regressor. The only difference with our baseline approach is that now also the inverse Mills ratio is included in the wage equation. If the inverse Mills ratio would be linear in its index, this estimation procedure would be equivalent to the control function approach used so far. So if estimated coefficients are very different, this is evidence that important non-linearities in the model are ignored. As is shown in column (6) of Table 9 there is only a very minor difference in the estimated wage elasticity.

Next, we extend the observation period by also including data from 2004 and 2005, and include quadratic terms in the time trends. This improves the predictive power of the

24

instruments in the first-stage regressions. It results in a more negative wage coefficient. and because standard errors are reduced, the wage coefficient now becomes significant. Selectivity in labor force participation remains important, but also wages become endogenous.

Finally, we follow Blundell, Duncan and Meghir (1998) and apply the grouping estimator. As instruments we use education interacted with cohort and year dummies, which yields in total 24 instruments. When applying the grouping estimator wages are endogenous, while there is no longer selective labor force participation. This is similar to what is found by Blundell, Duncan and Meghir (1998), and also our estimates wage elasticity (0.43) resembles their estimated wage elasticity (0.30). The wage residual in the labor supply equation becomes negative (and significant). Women with unobserved characteristics that provide them higher wages, have thus a lower preference for working many hours. This is not what one would expect if unobservables would capture, for example, motivation, work attitude or unobserved skills. Also Blundell, Duncan and Meghir (1998) do not provide an explanation for this type of endogeneity. A possible explanation for the difference between the results from the grouping estimator and our estimates presented in Table 7 is that the grouping estimator contains many instruments, and that the resulting wage elasticity describes a mixture of many local average treatment effects.

We should also consider the first-stage regressions. Most instruments are not significant in the first-stage regressions. However, the F-test for joint significance of the instruments has both in the wage equation and the labor force participation equation a p-value of less than 0.001 (which is what Blundell, Duncan and Meghir (1998) report). The F-test statistic is, however, small and in case of having many instruments indicates problems with weak instruments (e.g. Stock and Yogo, 2005). Extending the observation period by including also 2004-2005 (and thus taking account of the tax modifications in these years) does not solve the weak instrument problem and also does not change any of the estimation results.

6   Conclusions

In this paper, we investigate how financial incentives affect female labor supply. We focus on the Netherlands, which has average female labor force participation, but by far the highest

rate of part-time work among women within the OECD. This suggests that women carefully choose their working hours (and within the Netherlands there is also some regulation that allows for flexibility in adapting hours of work). We exploited exogenous variation from a substantial tax reform in 2001. An important reason for the tax reform was to induce women to increase labor supply. The tax reform included a reduction in marginal tax rates and a change from allowance to tax credit (which are both transferable). The latter removed some perverse incentives for working. In the empirical analyses we combine three different databases to construct a large data set covering the period around the tax reform. The empirical results show a negative (but insignificant) uncompensated wage elasticity, implying that the income effect dominates the substitution effect.

The tax reform does have a significant effect on labor force participation and we also find significant selection effects. We attribute this to the change from allowance to tax credit. This finding is in agreement with a large literature on earned income tax credits, which finds that the extensive margin of female labor supply is more sensitive to financial incentives than the intensive margin. We performed some simulations with the estimated model and found that the tax reform increased average weekly hours of work by 0.37, which is about 2 percent of average working hours in the population.

The uncompensated wage elasticity and the labor force participation effect are used in many economic models measuring welfare effects of (tax) policy. Nowadays, these models include a positive wage elasticity, meaning that individuals supply more hours to the market if their after-tax wages increase. We find a negative (but insignificant) wage elasticity suggesting that lower marginal taxes do not lead to higher hours (intensive margin). We do however find that the tax cut induced more women to work (extensive margin). These conclusions are based on a natural experiment and they only hold at that part of the labor supply curve that is affected by the tax reform. One direction for further research is to estimate labor supply responses using a structural model taking the complete tax system into account. Another direction is to look at other margins of adjustment. A recent survey study by Saez et al. (2009) highlights the usefulness of the elasticity of taxable income in calculating welfare effects of taxes.

References

Aaberge, R. and L. Flood (2008), Evaluation of an in-work tax credit reform in Sweden: effects on labor supply and welfare participation of single mothers, Working Paper 3736, IZA Bonn.

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

Blundell, R. and T. MaCurdy (1999), Labor supply: a review of alternative approaches, in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics, Volume 3A*, North-Holland, Amsterdam.

Eissa, N. (1995), Taxation and labor suply of married women: the tax reform of 1986 as a natural experiment, Working paper 5023, NBER.

Eissa, N. and H.W. Hoynes (2004), Taxes and the labor market participation of married couples: the earned income tax credit, *Journal of Public Economics* 88, 1931–1958.

Eissa, N. and J.B. Liebman (1996), Labor supply responses to the earned income tax credit, *Quarterly Journal of Economics* 111, 605–637.

Evers, M., R. de Mooij and D. van Vuuren (2008), The wage elasticity of labour supply: a synthesis of empirical estimates, *De Economist* 156, 25–43.

Heckman J.J. (1974), Shadow prices, market wages and labor supply, *Econometrica* 42, 679–694.

Imbens, G.W. and J.D. Angrist (1994), Identification and estimation of local average treatment effects, *Econometrica* 62, 467–475.

Meghir, C. and D. Phillips (2008), Labour Supply and Taxes, Discussion Paper 3405, IZA Bonn.

Meyer, B.D. and D.T. Rosenbaum (2001), Welfare, The earned income tax credit, and the labor supply of single mothers, *Quarterly Journal of Economics* 116, 1063–1114.

OECD (2004), Female labour force participation: past trends and main determinants in OECD countries, Economics Department Working Paper.

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

Saez, E., J. Slemrod and S. Giertz (2009), The elasticity of taxable income with respect to marginal tax rates: a critical review, Working paper 15012, NBER.

Staiger, D. and J. Stock (1997), Instrumental variables regression with weak instruments, *Econometrica* 65, 557–586.

Stock, J.H. and M. Yogo (2005), Testing for Weak Instruments in Linear IV Regression, in D.W.K. Andrews and J.H. Stock (eds), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, Cambridge University Press.

Van Soest, A., I. Woittiez and A. Kapteyn (1990), Labor supply, income taxes, and hours restrictions in the Netherlands, *Journal of Human Resources* 25, 517–558.

Van Soest, A. and M. Das (2001), Family Labor Supply and Proposed Tax Reforms in the Netherlands, *De Economist* 149, 191–218.

Table 10: Estimation results for the reduced-form equations.

	participation		log hourly wage	
instrumental variables				
low education post reform	0.103	(0.037)	0.073	(0.018)
medium education post reform	0.091	(0.033)	0.071	(0.010)
high education post reform	0.058	(0.058)	0.061	(0.011)
exogenous regressors				
low education and linear trend	0.021	(0.013)	0.006	(0.006)
medium education and linear trend	0.025	(0.012)	0.009	(0.003)
high education and trend	0.053	(0.020)	0.015	(0.004)
coh1949-54. lower education	−0.392	(0.044)	0.021	(0.016)
coh1949-54. medium education	0.199	(0.049)	0.106	(0.013)
coh1949-54. higher education	0.713	(0.060)	0.281	(0.013)
coh1955-64. lower education	0.040	(0.040)	−0.015	(0.014)
coh1955-64. medium education	0.540	(0.044)	0.089	(0.012)
coh1955-64. higher education	0.983	(0.050)	0.220	(0.010)
coh1965-74. lower education	0.106	(0.040)	−0.067	(0.015)
coh1965-74. medium education	0.667	(0.042)	0.039	(0.011)
coh1965-74. higher education	1.031	(0.050)	0.129	(0.009)
coh1975-85. lower education	0.610	(0.050)	−0.194	(0.023)
coh1975-85. medium education	0.891	(0.078)	−0.106	(0.013)
cohabiting. been married	0.196	(0.040)	−0.019	(0.011)
cohabiting. never married	0.300	(0.022)	−0.016	(0.005)
youngest child. 0-3 years	−0.560	(0.020)	0.092	(0.006)
youngest child. 4-11 years	−0.318	(0.021)	0.034	(0.006)
youngest child. 12-17 years	−0.073	(0.021)	−0.018	(0.006)
two minor children	−0.147	(0.016)	0.002	(0.005)
three or more minor children	−0.413	(0.020)	−0.011	(0.008)
presence child 18+	−0.187	(0.018)	−0.036	(0.006)
medium education partner	−0.333	(0.019)	0.018	(0.005)
higher education partner	−0.104	(0.016)	0.039	(0.006)
fiscal net yearly wage partner	−0.890	(0.022)	0.072	(0.007)
public sector			0.032	(0.004)
intercept	9.468	(0.222)	1.176	(0.067)
Observations	72,599		24,935	

Explanatory note: Robust standard errors in parentheses.

Table 11: Estimation results for the labor supply model.

	(1)		(2)		(3)		(4)	
log hourly wage	−3.197	(2.821)	−0.994	(2.790)	−0.100	(0.217)	−0.085	(0.217)
wage residual	3.109	(2.832)	0.912	(2.802)				
Inverse Mills	−4.761	(1.011)			−4.592	(1.000)		
			exogenous regressors					
lower ed. and trend	0.134	(0.119)	0.188	(0.119)	0.052	(0.094)	0.163	(0.091)
med. ed. and trend	−0.113	(0.095)	−0.086	(0.095)	−0.201	(0.051)	−0.113	(0.047)
high ed. and trend	−0.160	(0.106)	−0.123	(0.106)	−0.256	(0.061)	−0.152	(0.057)
coh1949-54. lower ed.	−6.539	(0.659)	−8.812	(0.432)	−6.700	(0.642)	−8.836	(0.426)
coh1949-54. med. ed.	−3.841	(0.536)	−4.905	(0.482)	−4.211	(0.417)	−5.005	(0.374)
coh1949-54. higher ed.	−0.229	(0.872)	−0.771	(0.865)	−1.094	(0.388)	−1.026	(0.389)
coh1955-64. lower ed.	−4.869	(0.460)	−6.122	(0.361)	−4.880	(0.460)	−6.112	(0.359)
coh1955-64. med. ed.	−2.845	(0.402)	−3.269	(0.390)	−3.137	(0.306)	−3.352	(0.301)
coh1955-64. higher ed.	0.783	(0.674)	0.786	(0.675)	0.116	(0.297)	0.585	(0.286)
coh1965-74. lower ed.	−3.940	(0.472)	−4.944	(0.412)	−3.787	(0.450)	−4.887	(0.371)
coh1965-74. med. ed.	−2.556	(0.309)	−2.681	(0.308)	−2.690	(0.285)	−2.720	(0.286)
coh1965-74. higher ed.	1.225	(0.437)	1.364	(0.438)	0.840	(0.264)	1.246	(0.258)
coh1975-85. lower. ed.	−3.688	(0.783)	−4.357	(0.771)	−3.144	(0.599)	−4.185	(0.557)
coh1975-85. medium ed.	−2.381	(0.442)	−2.321	(0.443)	−2.070	(0.335)	−2.228	(0.334)
cohabiting. been married	3.118	(0.338)	3.505	(0.329)	3.187	(0.333)	3.521	(0.326)
cohabiting. never married	1.908	(0.153)	2.253	(0.136)	1.968	(0.143)	2.267	(0.129)
youngest child. 0-3	−6.800	(0.357)	−7.804	(0.296)	−7.112	(0.220)	−7.888	(0.151)
youngest child. 4-11	−6.606	(0.226)	−7.114	(0.202)	−6.727	(0.199)	−7.145	(0.179)
youngest child. 12-17	−3.908	(0.201)	−3.908	(0.201)	−3.853	(0.195)	−3.891	(0.195)
two minor children	−2.159	(0.156)	−2.460	(0.143)	−2.177	(0.155)	−2.463	(0.143)
three or more kids	−2.771	(0.272)	−3.629	(0.199)	−2.768	(0.272)	−3.619	(0.197)
presence child 18+	−1.875	(0.226)	−2.132	(0.219)	−1.775	(0.205)	−2.099	(0.193)
med. ed. partner	−0.189	(0.157)	0.153	(0.141)	−0.228	(0.152)	0.137	(0.132)
high ed. partner	0.751	(0.211)	1.199	(0.189)	0.655	(0.192)	1.165	(0.158)
net yearly wage part	−1.619	(0.435)	−3.219	(0.285)	−1.904	(0.353)	−3.287	(0.194)
public sector	−2.347	(0.137)	−2.406	(0.136)	−2.447	(0.104)	−2.436	(0.104)
intercept	56.834	(4.324)	67.009	(3.751)	53.696	(3.221)	65.956	(1.879)
Observations	24,935		24,935		24,935		24,935	

Explanatory note: Robust standard errors in parentheses. Ed = education, med = medium.