The Anatomy of the Extensive Margin Labor Supply Response*

Spencer Bastani† Ylva Moberg‡ Håkan Selin§

April 12, 2019

Abstract

We estimate how labor force participation among married women in Sweden responded to changing work incentives implied by a reform in the tax/transfer-system in 1997. Using rich, population-wide, administrative data we estimate an average participation elasticity of 0.13, thereby adding to the scarce literature estimating participation elasticities using quasi-experimental methods. We also highlight that estimated extensive margin responses necessarily are local to the observed equilibrium. Among low-income earners, elasticities are twice as large in the group with the lowest employment level as compared to the group with the highest employment level.

Keywords: labor supply; social assistance; housing allowance; in-work tax credits; take up of transfer programs

JEL Classification: H20; J22

*We are grateful to Andrea Weber and Björn Öckert as well as to the editor Søren Leth-Petersen, two anonymous referees, Lina Aldén, Mikael Elinder, Hilary Hoynes, Claus Kreiner, Che-Yuan Liang, Eva Mörk, Andreas Peichl, Jim Poterba, Mari Rege, Olof Åslund, seminar participants at MIT, Mannheim/ZEW, Stockholm Institute for Transition Economics (SITE), Uppsala University, University of Nuremberg, DIW Berlin, the Nordic Tax Workshop in Helsinki, the IIPF Conference in Taormina, the CESifo Public Sector conference in Munich, the Linnaeus Conference on Discrimination and Labour Market Research in Kalmar, IEB Barcelona, Nordic Summer Institute in Labour Economics (Helsinki), and VATT in Helsinki for helpful comments and suggestions. Financial Support from the Jan Wallander and Tom Hedelius Foundation is gratefully acknowledged. An earlier version of this paper circulated under the title 'Estimating participation responses using transfer program reform'.

†Email: spencer.bastani@lnu.se; Department of Economics and Statistics, Linnaeus University, SE-351 06 Växjö, Sweden; Linnaeus University Center for Discrimination and Integration Studies; Uppsala Center for Fiscal Studies; Uppsala Center for Labor Studies at the Department of Economics, Uppsala University, Sweden; CESifo, Germany.

‡Email: ylva.moberg@nek.uu.se; Uppsala Center for Fiscal Studies; Uppsala Center for Labor Studies at the Department of Economics, Uppsala University, Sweden.

§Email: hakan.selin@ifau.uu.se; Institute for Evaluation of Labour Market and Education Policy (IFAU) SE-751 20 Uppsala; Uppsala Center for Fiscal Studies; Uppsala Center for Labor Studies at the Department of Economics, Uppsala University, Sweden; CESifo, Germany.

1 Introduction

How the labor force participation of secondary earners responds to work incentives is a question of great academic and policy interest. Labor supply responses along the extensive margin are often quantified by participation elasticities, measuring the percentage change in labor force participation in response to a percentage change in the financial reward from working. Participation elasticities determine the efficiency gains from tax breaks for secondary earners and represent a key concept in the literature on optimal tax- and transfer systems, see, e.g., Immervoll et al. (2011). Nonetheless, there are very few quasi-experimental estimates of participation elasticities, as evident from the meta-analysis by Chetty (2012).[1]

In this paper we provide new quasi-experimental evidence on how labor force participation reacts when the secondary earners' work incentives change as well as provide a first systematic analysis of how participation elasticities differ across groups with different initial employment levels. The latter is a key contribution of our paper, as the relationship between the labor supply response and the employment level previously only has been highlighted in the structural labor supply literature, which typically has relied on small, survey-based data sets.[2] This is an important gap to fill, since knowledge about heterogenous elasticities is essential when designing tax reforms with the purpose of promoting labor force participation, since such reforms are often targeted to specific groups of the population.

Exploiting high-quality administrative data on the full population of Swedish taxpayers, we make two primary contributions. First, we present an estimate of 0.13 of the average participation elasticity in a population of women where the average labor force participation already is high. Second, exploiting our large sample size, we partition the sample and systematically investigate the participation responses for different subgroups of individuals with different base-

[1]The enormous literature on in-work tax credit policies focuses on singles. Eissa and Hoynes 2004, Francesconi et al. (2009), Bosch and van der Klaauw (2012) and Ellwood (2000) are notable exceptions. To our knowledge, the only previous quasi-experimental studies *explicitly* reporting the secondary earner's participation elasticity are Selin (2014) and Kosonen (2014). Related papers using quasi-experimental methods to estimate the effect of childcare prices on female labor supply are Lundin et al. (2008) for Sweden and Havnes and Mogstad (2011) for Norway. None of them found an effect of child-care prices.

[2]When surveying a large number of elasticity estimates from the structural labor supply literature, Bargain and Peichl (2016) noted that married women's elasticities tend to be larger in countries with low female labor force participation. Bargain et al. (2014) find a similar pattern when using a coherent structural estimation approach on micro data from 17 E.U. countries and the United States.

line employment rates. We divide the sample into four equal groups based on the wife's potential income (predicted income) and, interestingly, find elasticities that are monotonically falling in the potential income of the wife, ranging from 0.24 to 0.09. The results suggest that cross- and within country comparisons of participation elasticities always should be made with reference to the relevant employment level. Our work complements, and is also broadly in line with, earlier structural labor supply studies on Swedish data. Flood et al. (2004), for example, also found fairly low elasticities for Swedish married women.

For identification, we use a reform in the Swedish system for housing allowances for couples with children in 1997. Before 1997, the housing allowance was means-tested based on family income - a family received maximal housing allowance if the joint income of the household did not exceed SEK 117,000 (appr. USD 15,000). For every SEK of *household* income in excess of SEK 117,000, the housing allowance was reduced by 0.2 SEK. After the reform, the system was individualized so that every SEK of *individual* labor income earned in excess of SEK 58,500 reduced the housing allowance by 0.2 SEK. The overall effect of the reform was to substantially lower participation tax rates for secondary earners married to low- and middle income spouses, mainly by making it less attractive to not work.

Following earlier work on secondary earners' labor supply on survey data (e.g. Eissa and Hoynes 2004, Francesconi et al. 2009), we compare eligible households (with children) with ineligible households (without children) before and after the 1997 reform, and provide graphical evidence on the reactions to the reform. Since we have access to several pre-reform years of data, we can examine the parallel trends assumption. We focus on wives married to husbands with an income below the median and document that female employment increases in households with children relative to households without children in the post-reform period.[3] We carefully calculate participation tax rates in the treatment- and control groups before and after the reform, and use the reform-based variation in these tax rates to estimate participation elasticities.

The paper is organized as follows. In the next section we describe the 1997 reform in

[3]From a different angle the same reform has earlier been analyzed by Enström Öst (2012). Using data from the Swedish Social Insurance Agency she compares earnings growth in households with different income compositions in 1996. She estimates significant earnings responses for women. In an experimental study on U.S. data Jacob and Ludwig (2012) estimated a negative effect of housing assistance on labor supply.

the Swedish housing allowance system. In section 3 we describe our data sources. Section 4 develops a model for interpreting the evidence and section 5 presents the empirical strategy. We provide a graphical analysis, present regression results, and report elasticities in section 6. Finally, section 7 offers concluding remarks.

2 The reform

We begin by describing the reform in 1997 that we exploit to identify extensive margin labor supply responses.

2.1 General description of the transfer program

The housing allowance system can be characterized as a guaranteed income program, as there is no work-requirement for eligibility, and the associated cash transfer is reduced as a function of the income of the members of the household (means-testing). The program is administered by the Social Insurance Agency ("Försäkringskassan") and payments are given on a monthly basis. To receive the transfer (which is a cash transfer), the household has to apply for it by the end of each year. In 1996, 180,000 Swedish couples received the housing allowance and the transfer made up an important budget share of many low income households. The particular program that we analyze in this paper applies to low income families with children under the age of 20, and we will motivate our choice of control group in section 5.[4]

2.2 Incentive effects

To ease the description of the incentive effects of the housing allowance, we introduce some notation. The housing allowance can be written as a function $B(\tilde{z}^p, \tilde{z})$ where $\tilde{z}^p$ and $\tilde{z}$ are, respectively, the two spouses' *qualifying income* or "bidragsgrundande inkomst", which is the income concept used to assess eligibility for welfare programs in Sweden.[5] The function B is weakly decreasing in both its arguments which reflects that the housing allowance is a means-tested

[4]There is also a separate and different housing allowance system applying to young households without children that was not subject to reform and that we do not analyze in this paper.

[5]Qualifying income does not only include earnings, but also capital income and a fraction of wealth.

program. The maximal level of the housing allowance is obtained when neither spouse has any qualifying income and is equal to $B(0, 0)$, which we denote B^{00}. The value of B^{00} depends on a number of non-income characteristics, such as the number of children in the household, housing costs and the living space (sq.m.) of the household.[6]

Before the reform in 1997 the transfer was reduced as a function of *the sum* of the two spouses qualifying incomes, i.e. the housing allowance pre-reform could be written $B(\tilde{z}^p, \tilde{z}) = B^{pre}(\tilde{z}^p + \tilde{z})$ and took the following form:

$$B^{pre}(\tilde{z}^p + \tilde{z}) = \begin{cases} B^{00} & \text{if } \tilde{z}^p + \tilde{z} \leq 117,000 \\ \max\left\{B^{00} - h^{pre}(\tilde{z}^p + \tilde{z}), 0\right\} & \text{if } \tilde{z}^p + \tilde{z} > 117,000. \end{cases}$$

where $h^{pre}(x) = 0.2 \times (x - 117,000)$. Thus, a family received the maximum transfer if the joint income of the household did not exceed SEK 117,000 SEK. If the joint income exceeded this exemption level, the transfer was reduced at a phase-out rate of 20 percent. Hence, if say, family income was 118,000 SEK, the transfer was reduced by 200 SEK [$= 0.2 \times (118,000 - 117,000)$].

After the 1997 reform, the system was individualized so that the household received the maximum transfer only if the income of *neither* spouse exceeded SEK 58,500. The phase-out rate was kept at 20 %.[7] Thus the post-1997 housing allowance can be written as $B(\tilde{z}^p, \tilde{z}) = B^{post}(\tilde{z}^p, \tilde{z})$ defined as:

$$B^{post}(\tilde{z}^p, \tilde{z}) = \begin{cases} B^{00} & \text{if } \tilde{z}^p \leq 58,500 \quad \text{and} \quad \tilde{z} \leq 58,500 \\ \max\left\{B^{00} - h^{post}(\tilde{z}_p), 0\right\} & \text{if } \tilde{z}^p > 58,500 \quad \text{and} \quad \tilde{z} \leq 58,500. \\ \max\left\{B^{00} - h^{post}(\tilde{z}_p) - h^{post}(\tilde{z}), 0\right\} & \text{if } \tilde{z}^p > 58,500 \quad \text{and} \quad \tilde{z} > 58,500. \end{cases}$$

where $h^{post}(x) = 0.2 \times (x - 58,500)$.

How did the 1997 reform affect work incentives? To answer this question we need to make an assumption about how economic decisions within the family are organized. Even though

[6]In appendix B we describe in more detail how the value of B^{00} is determined.

[7]The reform implied no change to the income thresholds, the level of the housing allowance or the phase-out rates for single parents. Therefore, singles with children could *a priori* be considered to serve as a control group to married with children in the empirical analysis. However, owing to differential employment trends and levels we have not chosen this strategy.

there is individual taxation in Sweden, the transfer system depends on the income of both spouses. Hence, the total tax/transfer relevant for the labor force participation decision of one member of the family depends on the economic decision of his/her spouse. We analyze the incentive changes from the point of view of a sequential model, where the secondary earner decides whether to work or not conditional on the labor supply choice of the primary earner. For the moment we abstract from the take-up issue, and simply assume that the household always takes up the transfer when eligible. We discuss the model assumptions further in section 5.2.

In figure 1 we have illustrated the pre- and post-reform transfers $B^{pre}(\tilde{z}^p + \tilde{z})$ and $B^{post}(\tilde{z}^p, \tilde{z})$ for a family with two children as a function of the secondary earner's income $\tilde{z}$, while fixing $\tilde{z}^p$ to $170,000$ (a typical value of the primary earner's qualifying income in our estimation sample). We assume that if neither spouse would work, the household would be entitled to the maximum level of housing allowance for households with two children, $B^{00} = 38,100$. Given these assumptions, in the pre-reform scenario, the household is eligible for a transfer amounting to $38,100 - 0.2 \times (170,000 - 117,000) = 27,500$ when the secondary earner has zero earnings. According to the pre-reform rules, as soon as the secondary earners supplies any amount of positive earnings, the housing allowance is reduced. More specifically, it is reduced by 0.2 SEK for every SEK of secondary earnings up until the point where the total amount of 27,500 SEK is phased out (which happens at 137,500 SEK) . In the post-reform scenario, on the other hand, the transfer at zero earnings of the secondary earner is significantly smaller: $38,100 - 0.2 \times (170,000 - 58,500) = 15,800$ but the phase-out does not kick in until the secondary earner exceeds the income level of $58,500$. At this point the pre- and post-reform transfers are equal and the functions B^{pre} and B^{post} coincide for secondary earnings exceeding 58,500.

The important lesson from figure 1 is that if the potential earnings of the secondary earner is SEK 58,500 or more, the difference between the household's disposable income in the state of work and non-work, respectively, will entirely be driven by the difference in the transfer in the state of non-work. Since most married women earn annual incomes above SEK 58,500 when working, we therefore conclude that the variation used to recover participation elasticities in this paper is a variation in the housing allowance at zero earnings of the secondary earner. In summary, the reform makes not working much less attractive for the secondary earner. Accordingly, even though households may not be perfectly aware of the income splitting rules, one-earner households will certainly recognize that the size of the transfer will be reduced after the reform.

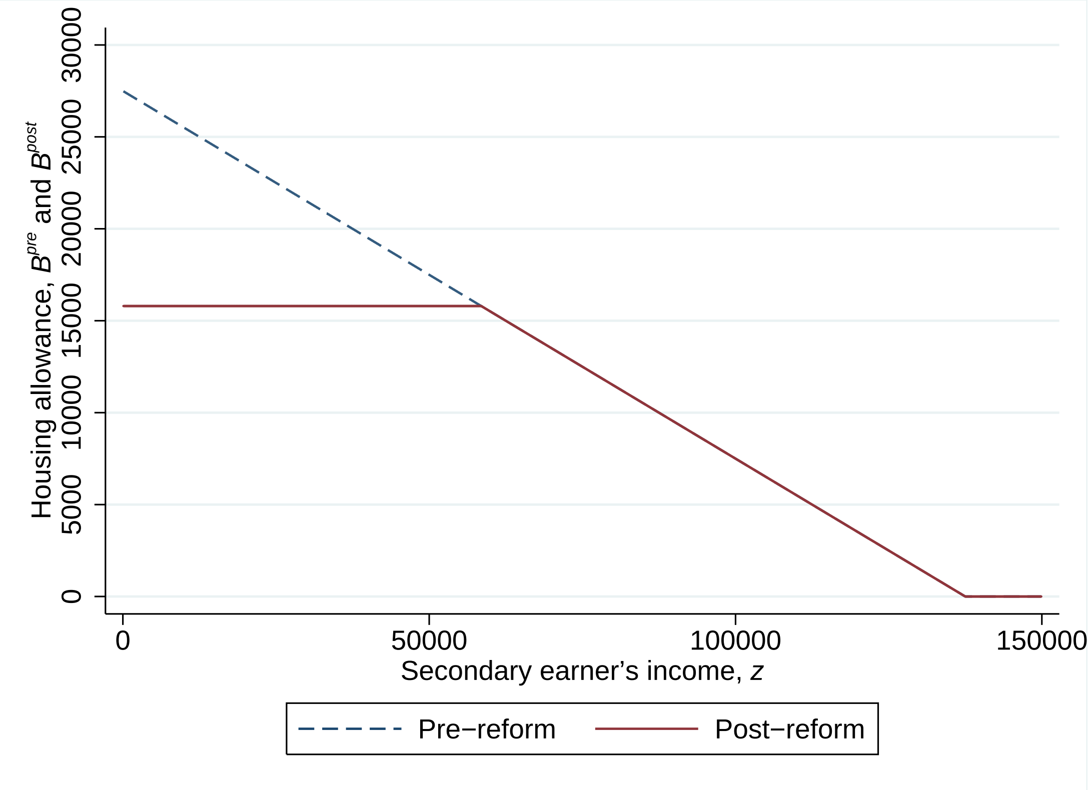

Figure 1: Housing allowance before and after the reform according to the functions $B^{pre}(\tilde{z}^p + \tilde{z})$ and $B^{post}(\tilde{z}^p, \tilde{z})$ as a function of secondary income $\tilde{z}$ for a family with two children. The primary earner's income is fixed at $\tilde{z}^p = 170,000$.

2.3 Time line and anticipation issues

The main objective of the 1997 reform was to cut government expenditures related to the housing allowance program. The size of the program more than doubled between 1990 and 1995 (Boverket 2006). In April 1995, when the annual expenditures were projected to amount to more than SEK 9 billion, the Social Democratic government appointed a government committee (Kommittédirektiv 1995:65). The mandate of the committee was straightforward: The committee was supposed to propose expenditure reductions, e.g. by changing the rules for means-testing. The committee issued their report in December, 1995. The committee's proposal was similar to the reform that was to be implented on January 1, 1997. The Social Democratic government presented a government bill in March 1996 and the bill was passed in parliament on May 8, 1996.[8]

Did households anticipate the 1997 reform? This is a key issue when interpreting the estimated elasticities (Blundell et al. 2011). In principle, well-informed households could have adjusted their behavior already in December 1995 when the committee's report became publicly known.[9]. However, we think that large-scale pre-reform anticipatory responses are unlikely. As far as we can tell, there was no public discussion about the income limits when the committee's report was presented.[10] According to Enström Öst (2012), the Social Insurance Agency ("Försäkringskassan") informed beneficiaries about the reform by sending out letters in June and October 1996. Accordingly, it is likely that the vast majority became aware of the new earnings limits close to the implementation of the reform on January 1, 1997.

3 Data

This study primarily exploits large population-wide administrative data sets provided by Statistics Sweden. We have access to all key variables from 1991 and onwards. These include earned income (which we define as the sum of wage income and self-employment income), education level, geographical indicators, the number of children in the household and region of origin. Our graphical analysis will cover the years 1991-2010, whereas, as we motivate in section 4 below, we focus on the years 1994-2001 in the regression analysis.

Since the variables that we use are collected from administrative registers, the overall quality is very good. There are however two caveats with this data. One is that the 1990's data on education level for many non-natives (who obtained their education degrees from other countries) is missing. We have been able to correct the missing values by using leads of the education vari-

[8]The Social Democratic party was in minority in the parliament, but was supported by the Centre (agrarian) party ("Centerpartiet").

[9]As discussed by Blundell et al., it is not *a priori* clear in which direction such anticipatory responses would go. If inter-temporal substitution is the dominating mechanism, we would observe people working less in anticipation of the reform. If, on the other hand, labor market frictions is the key mechanism we would expect people to start searching for new jobs already in the pre-reform period.

[10]A search on "bostadsbidrag" in the media archive "Newsline", suggests that the main media focus was on actions against fraud in the system for housing allowances, rather than work incentives when the committee presented their report. The media coverage was larger when the reform was legislated on May 8, 1996, but the focus was not on the earnings limits.

able. The Swedish authorities later on actively sent questionnaires to immigrants where they were asked to report their education level.[11] Another caveat is that non-married, cohabiting couples without common children are observed as singles in the administrative data. Therefore, even though the housing allowance system applies both to married and cohabiting couples, we limit the sample to formally married couples. We simply do not observe cohabiting couples *without* children.

In line with previous literature (e.g. Eissa and Hoynes 2004), we assume that the wife is the secondary earner and that the husband is the primary earner.[12] Accordingly, we restrict the sample to households where the husband has positive earnings. Our main identification strategy relies on comparing households with and without children. Therefore, we restrict the sample to households where the husband's actual qualifying income falls below the median level of *qualifying income*, which is an income concept used by the government to compute eligibility for transfers.[13] The cut-off at the median income was chosen because it corresponds to an income level of around 230,000 SEK in 1996, and households with levels of qualifying income exceeding this threshold were not eligible to any sizable housing allowances prior to the reform.[14] As described below in section (4), we will also run placebo regressions on a separate sample of high-income couples who were unaffected by the reform. This sample is identical to the main sample in all other respects.

We drop households where any of the two spouses are aged below 30 or above 55. As described in section 2, households with two spouses aged below 30 were subject to different housing allowance rules both before and after the reform. The upper age limit is imposed as we are interested in the labor supply behavior of prime-aged individuals and not in retirement behavior.

Our regression analysis will focus on the time period 1994 to 2001, while the graphical

[11] Unless the individual died or migrated between year t and year 2000 we use education information as of 2000 when constructing the variable for education level.

[12] In our data, the vast majority of secondary earners are women.

[13] In the register data, we compute qualifying income based on information on earnings and capital income and imputing financial assets from information on capital income.

[14] The upper limits of qualifying income (i.e. the income level where the entire housing allowance was phased out) differed depending on the number of children below 20 in the household. In 1997, the upper limit was SEK 267,000 for 1 child, SEK 307,500 for 2 children and SEK 351,000 for 3 or more children. Since we pool all households in the main analysis, we cannot use separate income cut-offs.

analysis covers the years 1991-2010. The reason for focusing on the time period 1994-2001 in the regression analysis is that reliable estimates from the micro-simulation FASIT, which we use to calculate participation tax rates, are available from 1994 and onwards. There was also a severe macro-economic crisis in the beginning of the 1990's in Sweden. The reason for not using years after 2001 is that a large childcare fee reform was implemented in 2002 (see Lundin et al. 2008).

In line with the theoretical framework presented in appendix A, we construct so-called *household types*, h, based on the two spouses' age (five groups) and education (four groups), giving rise to a total of $4^2 \times 5^2 = 400$ household types. In the empirical analysis, we use the household types as fully saturated controls for age and education. In section 5.4, we will also estimate heterogeneous elasticities based on dividing the potential (predicted) income of the secondary earner into four groups. As described in appendix A.3, these groups correspond to a partitioning of the household types according to the potential income of the secondary earner.

4 Difference-in-difference analysis

4.1 Model

Our empirical analysis consists of two parts. In this section we begin with the simplest and most transparent specification, a difference-in-difference analysis, by comparing the evolution of the average labor force participation for secondary earners in families with and without children before and after the reform. In section 5 below, we proceed to estimate participation elasticities by relating the change in labor force participation of secondary earners to the change in participation tax rates induced by the reform.

Here, we focus on the following specification:

$$e_{ihkt} = \mu_{kt} + \mu_t + \mu_k + \mu_h + \mu_{hk} + \mu_{ht} + \delta X_{ihkt} + \upsilon_{ihkt}, \tag{1}$$

where e_{ihkt} is a dummy equal to one if an individual i in household type h, with parental status k at time t is working, and zero otherwise. μ_k is a dummy variable for having at least one child aged below 20 in the household, μ_t is a time fixed effect, μ_h is a household type fixed effect, μ_{kt} is a shorthand for the interactions between the child dummy and the time dummies, and μ_{ht} represents the interaction between the household type dummies and the time dummies.

In equation (1), we are interested in the interactions μ_{kt} between the indicator variables for having children and the year dummies. For the post-reform years, these interactions capture the reform effect and its dynamics over time. The dynamic dimension is crucial: In the presence of adjustment costs we expect the long-run response to be larger than the short-run response.

The identifying assumption in the difference-in-difference specification is that the labor supply behavior of secondary earners with children (the treatment group) would have evolved in a similar way to the labor supply behavior of secondary earners without children (the control group), in the absence of the reform. This cannot, for obvious reasons, be verified directly (we do not know how treated individuals would have behaved if they had not been treated). Instead, as is customary in the literature, we proceed in the following way.

First, we use the fact that we have access to several years of pre-reform data to test if the labor force participation trends in the treatment and control group were similar in the years before the reform. The pre-reform trends are reflected in the coefficients of the μ_{kt} interactions in (1) for the years before the reform. If the trends are parallel in the years before the reform, this increases the likelihood that the post-reform trends also would have been parallel in the absence of the reform.

Second, we run 'placebo'-regressions on a sample of high income households (who were essentially all unaffected by the reform, independently of whether they had kids or not). If the labor force participation of secondary earners with and without children in *high income* households evolved similarly after the reform, this increases the likelihood that the post-reform trends for *low income* households with and without children would have evolved similar in the absence of the reform. Our placebo-test amounts to estimating equation (1) on the sample of women married to husbands with qualifying incomes above the 75th percentile. In 1996, this corresponded to an income level of around 310,000 SEK.[15]

The complete set of results for the reduced form effects analysis are presented in Table 1 and

[15]Here it should be noted that some households with 3 or more children could be eligible for housing allowance up to 351,100 SEK.

Figure 2 presents graphical evidence. The housing allowance reform occurred in 1997, which means that the estimation sample contains three pre-reform years and five post-reform years. We choose 1996 as the reference year.

4.2 Graphical evidence and regression results

Beginning with the graphical analysis in Figure 2, we can see how the employment of married women (defined as having positive earnings) evolved in couples with and without children between 1991 and 2010. A nice feature of Figure 2 is that it illustrates the evolution of employment outside the more narrow time period of our regression analysis.[16] We make the following observations. In the beginning of the 1990's, there was a sharp decline in employment due to a deep economic recession. Figure 2 suggests that female employment decreased slightly more among households with kids 1991-1993. However, between 1993 and 1996 the two lines moved in parallel. Note also that the employment *levels* are strikingly similar. After the 1997 reform, employment continued to evolve similarly until 1998. Then, there is a clear employment increase for women with children relative to women without children, an increase which continued until 2001. After 2001, the two graphs appear to again follow more or less similar trends. The fact that the response appears to be increasing in the post-reform years suggests that adjustment costs (e.g., costs associated with finding a new job) could be important. As discussed above in section 2.3, information about the reform became publicly available close before its implementation, and it probably took some time for households to adjust.

Columns 1-4 in Table 1 show the coefficients for the main 'low income' sample where most households with children were eligible for housing allowances (at zero earnings of the wife). First of all, it should be noted that the results appear to be quite robust across the different specifications in these four columns. The first column reports the results of a difference-in-difference specification without any control variables. In this column, the first thing to notice is that the coefficients for the pre-reform years, 1994 and 1995, are statistically insignificant, confirming the visual evidence of Figure 2 that the pre-reform trends were very similar for the treatment and

[16]In the graphical analysis and in the regressions we employ the same sample restrictions, i.e. we focus on households where the husband's qualifying income falls below the 50th percentile and where the husband reports positive earnings.

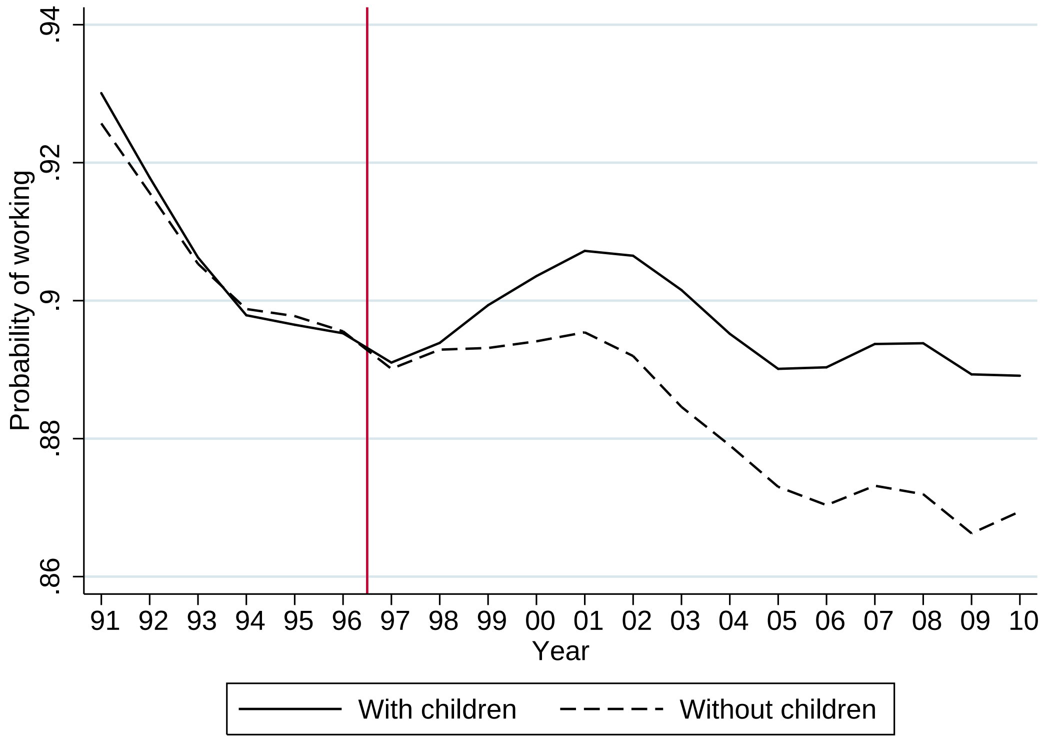

Figure 2: (*Graphical reduced form evidence and long term trends*). Female participation (share with positive earnings) in low income households where the husband participates in the labor force.

control group. In fact, the coefficients for the pre-reform years remain insignificant for all the specifications that we have considered, as evident by columns 1-4. Moreover, also consistent with Figure 2, we see that there is a statistically significant response to the reform in 1999 and that the response grows monotonically across the post-reform years. For 2001, the estimated effect amounts to 1.2 percentage points.

In column 2, we have added household type controls and the estimated effects become slightly larger. In column 3, we control for trends in a flexible way including the full set of interactions between the time dummies and the household type dummies as well as the interactions between the household type dummmies and the dummy for having children. Interestingly, in this specification, the reduced form effect estimates are also significant for the two post-treatment years 1996 and 1997 (at the 5 percent level). Finally, when the full set of controls are included in column 4, the overall pattern of coefficients is similar to column 3, but the reform effect estimate for 2001 is more in line with that obtained in the specification without controls in column 1. Our preferred estimate of the reform effect is the coefficient for 2001 in our most

Table 1: Reduced form effects (in percentage points)

	Low income	Low income	Low income	Low income	High income
	(1)	(2)	(3)	(4)	(5)
Year 1994 × children	-0.060	-0.152	0.000	-0.097	-0.264
	(0.130)	(0.129)	(0.163)	(0.159)	(0.171)
Year 1995 × children	-0.097	-0.121	-0.095	-0.140	0.016
	(0.111)	(0.110)	(0.140)	(0.137)	(0.149)
Year 1997 × children	0.120	0.154	0.348**	0.404***	-0.117
	(0.114)	(0.113)	(0.144)	(0.141)	(0.153)
Year 1998 × children	0.129	0.245*	0.331**	0.392**	0.000
	(0.134)	(0.132)	(0.169)	(0.164)	(0.178)
Year 1999 × children	0.652***	0.833***	0.681***	0.813***	0.189
	(0.145)	(0.144)	(0.181)	(0.177)	(0.192)
Year 2000 × children	0.976***	1.24***	0.790***	0.992***	0.245
	(0.154)	(0.152)	(0.189)	(0.185)	(0.202)
Year 2001 × children	1.214***	1.485***	0.863***	1.120***	0.385*
	(0.160)	(0.159)	(0.196)	(0.193)	(0.211)
Household type dummies	No	Yes	Yes	Yes	Yes
Household type × children	No	No	Yes	Yes	Yes
Household type × year dummies	No	No	Yes	Yes	Yes
Additional controls	No	No	No	Yes	Yes
Nr of observations	2,770,100	2,770,100	2,770,100	2,770,100	1,385,071

Note: Dependent variable: probability of having positive earnings. 'Low income' sample consists of wives married to husbands with a positive qualifying income, which falls below the 50th percentile. 'High income' sample consists of wives married to husbands with a positive qualifying income that falls above the 75th percentile. All specifications contain a dummy for having children and a full set of year dummies. 400 household types are defined based on 5 age dummies for each spouse and 4 education level dummies for each spouse. The additional control variables are specified in section 5. Standard errors reported below the estimates are robust to heteroscedasticity and clustered at the household level. * indicates significance at 10%. level, ** 5% level and *** at 1% level.

ambitious specification of column 4 and amounts to a 1.12 percentage point increase in the probability of married women to participate in the labor force.[17]

In column 5, we report the results from a 'placebo-regression' with the full set of controls, where we have estimated equation (1) on a sample consisting of women married to husbands with qualifying income above the 75th percentile (which were essentially all untreated by the reform). In all other respects, the selection criteria are identical to the main low-income sample. It is reassuring that all estimated coefficients are insignificant at the 5 percent level. One interaction, the interaction for 2001, is significant at the 10 percent level, but the coefficient estimate is considerably smaller than the corresponding point estimate in the low income sample. The results of this placebo regression, considered in conjunction with the results in column 1-4 (showing that the trends before the reform were parallel), and the visual evidence in 2, allow us to be reasonably confident that the identifying assumption in our difference-in-difference setup is satisfied.

4.3 Discussion

Before moving on to estimating participation elasticities, we briefly discuss a few aspects of the preceding analysis. First, we discuss some potential concerns with using child status as an indicator for treatment status and the possibility of using within-individual variation to estimate the effect of the reform. Second, we discuss the standard errors in the difference-in-difference setup. Finally, we discuss an analysis of male responses that we have done to test the validity of the primary-secondary-earner assumption.

A growing body of research has stressed the importance of children and the onset of parenthood for labor market outcomes at the individual level (see e.g. Kleven et al. (forthcoming) and Angelov et al. 2016). The strategy in this paper is essentially to follow groups rather than individuals, and we have demonstrated that the pre-reform trends are parallel at the group-level. When groups are defined in a coherent way over time, employment dynamics associated with child birth is not necessarily a problem. Moreover, in our analysis child status is associated with having a child under the age of 20, hence families with very small children represents only a

[17]These results are robust to excluding cells (defined based on year $\times$ children $\times$ household type) that contain less than 100 observations.

small subset of our treatment group.

An alternative empirical strategy would be to add individual fixed effects to equation (1). In this case, the identifying assumptions would be different as one would be comparing individuals with and without children before and after the reform. However, a major concern in this case is that individuals would be observed at different stages in their life cycle before and after the reform. We have estimated equation 1 adding individual level fixed effects and confirmed these concerns. In contrast to the baseline analysis, the results are in this case very sensitive to the inclusion of control variables (e.g. controls for age). When adding the full set of control variables, the post-reform results are reasonably similar to our baseline estimates of column 4 of Table 1, but we also estimate significant pre-reform trends. The details are contained in table A2 of online appendix D.[18]

Throughout the paper, we report standard errors that are clustered at the individual level. If employment shocks primarily operate on the individual level, our standard errors will be robust to serial correlation and arbitrary heteroskedasticity. Recall that we are using individual level data on the entire population, and individuals will move in and out from the treatment group over time (since new children are born and children grow up). If one instead worries about aggregate shocks to women with and without children, inference becomes more challenging as we essentially only have 16 clusters (two groups in eight time periods). In that case it is well-known that the cluster-robust covariance estimator is likely to be incorrectly estimated. Furthermore, if shocks are auto-correlated at the group level, the estimated standard errors would be wrong also in the presence of many clusters. Therefore, as a robustness check we performed a randomization inference exercise influenced by Bertrand et al. (2002), section 4.6, by comparing the observed treatment effect estimate with a distribution of placebo treatment effect estimates obtained in Monte Carlo-simulations. This exercise strongly indicates that inference is robust to both clustering and serial correlation at the group level. The details of the procedure are given in appendix F of the online appendix.

[18]As a robustness check, we have also estimated triple-difference models with individual fixed effects, incorporating also the high-income sample (which we previously exploited in the placebo regression reported in column 5 of Table 1). In these regressions, we did not obtain significant pre-reform trends, and with the full set of control variables we estimated treatment effects that are in the same ballpark as in the baseline difference-in-difference analysis. The details are contained in table A3 of online appendix D. Still, the estimates are more sensitive to control variables than in the baseline analysis.

In order to examine the validity of the 'primary-secondary earner' assumption we have estimated equation (1) on a sample of males. Our idea has been to construct the male sample as a *mirror image* of the female low income sample by conditioning the male sample on the wife's qualifying income falling below the 50th percentile. The results are presented in column 1 of Table A1 of Appendix D where it can be inferred that the estimated coefficients for this male sample are very different from the female sample. For 1994-2000, none of the interaction terms are statistically significantly different from zero. For 2001, we estimate a *negative* effect on male employment equal to -0.36 percentage points which is significant at the 5 percent level.

To dig deeper into the potential mechanisms at play we have also examined male earnings responses at the intensive margin. We transformed earnings into log earnings in the standard way, thereby *excluding* observations with zero earnings. We found no clear evidence of a response in log earnings after including the full set of controls, see columns 2 of Table A1. Finally, we have also estimated equation (1) on the main *female* sample with log earnings instead of employment on the left hand side. The estimation of this pure 'intensive margin' response resulted in significantly positive coefficients, especially for the years 2000 and 2001 (see column 3 of Table A1). It is worth noting that even a small intensive margin response may have important fiscal consequences when the baseline employment rate is high.[19]

5 Participation elasticities

In this section, we estimate participation elasticities by relating the change in labor force participation to the size of the change in the financial gain from working, as measured by the participation tax rate (PTR).

[19]We also experimented with log(earnings+1) as the dependent variable, thereby *including* females with zero earnings in the regression and found that the estimated coefficients were significant in *all* post-reform years and around three times as large in Column 3 of Table A1. Even though this log transformation is controversial, the results indicate that women primarily reacted to the reform along the extensive margin, i.e. they went from zero earnings to a positive amount of earnings.

5.1 Participation tax rates

To calculate participation tax rates, detailed information on individuals' budget sets is needed. As the housing allowance interacts with other parts of the transfer system, most notably social assistance, it is important to take into account the entire tax-and transfer system when constructing households' budget sets. To achieve this, we use the microsimulation model FASIT developed by the Swedish Ministry of Finance and Statistics Sweden. FASIT relies on a larger set of variables than what is available in the administrative registers, and therefore employs a smaller supplementary data set called HEK ('Hushållens ekonomi') that also contains survey data.[20] We therefore, as will be described in more detail below, compute participation tax rates for individuals in our population-wide data using an imputation procedure based on the variables that are available in both data sets.[21]

We let $T^{total}(z^p, z)$ refer to all taxes paid and benefits received by a household with primary earnings z^p and earnings of the secondary earner equal to z, assuming the household takes up all transfers.[22] The PTR for the secondary earner is defined as:

$$\tau(z^p, z) = \frac{T^{total}(z^p, z) - T^{total}(z^p, 0)}{z}. \tag{2}$$

This is the key independent variable that will appear in our estimation equations (4) and (5) below. Importantly, we compute PTR:s for all households assuming full take up of housing allowance and social assistance (for eligible households).

The PTR concept implies that the household chooses between two hypothetical disposable income states: The household disposable income when the secondary earner is working, and the household disposable income when the secondary earner is not working. Notice that to be

[20] After having imposed the same sample restrictions on HEK as on the administrative data, the size of the HEK sample varies between 1000 and 2000 observations each year. The sample is too small to be used in the labor supply analysis described in section 4, but is still very useful for the purpose of estimating PTR:s. Remember that the households' budget sets are given *deterministically* by the micro-simulation model and the variables in the HEK data. Of course, this does not mean that the sample size of HEK is unimportant, because the precision of the estimated group means become more precise the larger is the number of households represented in the HEK sample. A detailed comparison between the HEK and population-wide data is provided in the appendix for earnings (tables A5 and A6) and for labor force participation (tables A7 and A8).

[21] HEK includes the full set of variables that determine eligibility for the housing allowance program (such as housing costs and dwelling space) as well as the size of the benefit actually received (from registers), we also use HEK to compute the take-up of the housing allowance.

[22] The function T^{total} corresponds to $T + B$ below in section A in the appendix.

able to estimate the impact of PTR:s on labor force participation, we need to compute PTR:s for all individuals, both labor force participants (with positive earnings) and labor force non-participants (with zero earnings) in our population-wide register data. This implies that we need to compute the potential (hypothetical) income a secondary earner with zero income would have if she started working.

We proceed in the following way. We start by calculating PTR:s for all secondary earners with *positive* earnings in the HEK data. This is done by first computing the disposable income of each household at zero earnings of the secondary earner. We then subtract this number from the household's actual disposable income to obtain the household's financial gain from secondary earner employment. Finally, we divide this financial gain by the secondary earner's earnings to obtain the PTR according to equation (2).[23]

Next, pooling the HEK data for the years 1994-2001, we regress PTR:s on four dummies based on the actual qualifying income of the husband (year-specific quartiles), four dummies based on the number of children in the household, eight year dummies as well as the full set of interactions between the income, children and year dummies. Additionally, we include three dummies for the educational level of the wife, which we of course also interact with the year dummies. These variables explain a very large share of the variation in PTR:s, the R2 value exceeds 90 percent. The estimated coefficients from these regressions are then used to impute PTR:s for all secondary earners (with either zero or positive earnings) in the population-wide register data.[24]

[23]We acknowledge that earnings in the state of work might differ between employed and unemployed women, even conditional on observable characteristics, which may induce a selection bias. We have however not been able to find any valid instruments that enable us to use a selection correction term. In this respect, our approach bares some similarities with Gelber and Mitchell (2011) and Meyer and Rosenbaum (2001). When researchers can use variation from several tax reforms it is, in principle, feasible to estimate the extensive and intensive margins simultaneously (see Alpert and Powell 2014).

[24]Since the main purpose of FASIT has been to assess revenue effects of changes in the tax- and transfer system, we had to rewrite some parts of the code to serve our purposes. Most importantly, there were no modules for computing social assistance benefits for the years 1994-1995. Hence, for these years, we wrote the code ourselves based on national guidelines for social assistance. Rules for social assistance differ across municipalities. For some, but not all, years we can compute social assistance both as a function of municipality-specific parameters and national guidelines. For coherency, we have chosen to use national guidelines for all years. We have verified that the two methods produce similar results for the years that both methods are available to us.

5.2 Econometric framework

Our aim is to estimate the following relationship on secondary earners in (formally) married couples where both spouses are aged 30-55, using data for the years 1994 to 2001:

$$e_{ihkt} = \alpha + \beta\tau_{ihkt} + \eta_{ihkt}. \tag{3}$$

In the above equation, the dependent variable e_{ihkt} is a dummy which takes on the value of 1 if individual i with k children in household type h in year t is employed and is zero otherwise. In our baseline specification we define employment as having positive earnings. Moreover, k will be binary in the analysis and equal to 1 if there is at least one child aged below 20 in the household and 0 otherwise. The independent variable τ_{ihkt} is individual i's PTR which is calculated assuming that eligible households take up the housing allowance. Finally, η_{ihkt} is an error term. The parameter of interest is β, the participation elasticity.

Notice that the participation elasticity of secondary earners is a structural parameter, and is defined in relation to a particular model of household labor supply behavior. We outline such a model in appendix A, where we formally define β (see equations 13-15). A key assumption in our framework is that households behave sequentially, namely, the secondary earner decides about his/her labor supply taking the behavior of the primary earner as given. In line with earlier literature (see e.g. Eissa 1995; Eissa and Hoynes 2004) we treat primary earner behavior as exogenous.[25]

The model also features an endogenous take-up decision and we specify sufficient conditions under which reforms in transfers (that are subject to take-up decisions) can directly be used to assess the sensitivity of employment to taxes (see Proposition 1 in appendix A). In

[25]The model of household behavior builds on Immervoll et al. (2011) and assumes Pareto efficiency, no income effects, and a sharing rule (dictating how resources are divided in the family) that is unaffected by taxes. To simplify the interpretation of our empirical results, we assume that the extensive margin of the primary earner is inelastic. This seems ex-ante reasonable given the high participation rate of primary earners in Sweden. The non-responsiveness of primary earners along the extensive margin is also supported by our empirical results in Table A1 in appendix D. The omission of income effects is not without loss of generality, but simplifies the analysis considerably and has also become a standard practice in the literature (see Brewer et al. 2010). The secondary earner assumption is also common in the literature, but has been criticized by Gelber (2014). He questions the implication of the unitary model that a married individual's pre-tax earnings should react equally to an increase in that individual's unearned income as it reacts to an increase in the unearned income of his or her spouse. In our model, both of these effects are assumed to be zero since utility is linear in consumption.

this paper, we transform transfer-elasticities to participation elasticities by scaling the transfer-elasticities with the inverse of the take-up rate. Our model clarifies the conditions under which this approach is valid.[26]

As already described in section 5.1, we estimate τ_{ihkt} on a smaller survey data set that contains all variables necessary to compute the household's taxes and transfers accurately. Let W denote a vector of variables that are contained both in the main (population wide) data set and in the smaller survey data set (W is a subset of the variables needed to compute the PTR). We refer to the coefficient vector in the regression of τ_{ihkt} on W_{it} on the smaller data set as ρ and focus on the following regression model for the population wide data set:

$$e_{ihkt} = \alpha + \beta\hat{\tau}_{ihkt} + \eta_{ihkt}, \tag{4}$$

where $\hat{\tau}_{ihkt} = \hat{\rho}W_{it}$. To account for the fact that $\hat{\rho}$ is estimated with uncertainty we have checked that the standard errors are robust to the corrections suggested by Murphy and Topel (1985), see section 5.3 below. The fundamental condition for the validity of this imputation procedure is that ρ, and its covariance matrix, is consistently estimated on the auxiliary data source (HEK).

If we were to estimate (4) in a cross section without any control variables, one would fear that $\widehat{\beta}$ would be biased. The reason is of course that $\widehat{\beta}$ also would capture direct effects of W on e. If, on the other hand, one would include controls for W in a very flexible way, identification would be lost. The leading idea of our paper is to exploit the 1997 housing allowance (HA) reform to address the potential endogeneity of $\hat{\tau}_{ihkt}$ in equation (4). The HA reform substantially reduced PTRs for households with children in certain income intervals, but left households without children unaffected. Hence, if there are no direct effects on the outcome variable of the interactions between the children dummy, λ_k, and the time dummies, λ_t, (conditional on λ_k and λ_t) the HA reform can be used as an instrument for τ.

The richness of the data enables us to control for covariates and time trends in a very flexible way. We let λ_{kt}, the full set of interactions between the child and time dummies, be the vector

[26] A similar model of labor force participation and take-up has recently been developed by Gelber et al. (2018).

of excluded instruments. Ultimately, we wish to estimate the equation

$$e_{ihkt} = \alpha + \beta\hat{\tau}_{ihkt} + \lambda_t + \lambda_k + \lambda_h + \lambda_{hk} + \lambda_{ht} + \gamma X_{ihkt} + \eta_{ihkt}, \tag{5}$$

where X_{ihkt} is a rich set of pre-determined control variables not used to construct the household types. In the X vector we include seven dummies for region of origin as it is well-known that foreign-born on average exhibit lower employment rates than natives.[27] In addition, we include 21 dummies for county of residence to account for regional employment differences. Moreover, we interact the dummies for region of origin and the county dummies with the children and the time dummies. Finally, we also include detailed age dummies (one dummy per age), which we interact with the children dummmy.

Notice that, at the *individual* level, the imputed participation tax rate $\hat{\tau}$ in equation (4) will be measured with error. The reason is that the imputations are made at the group level (see section 5.1). However, since we instrument $\hat{\tau}$ with λ_{tk}, the requirement for consistent estimation of β in equation (5) is that the year-specific group averages are correct.

Before moving on to the results, we mention that an alternative identification strategy would have been to focus only on households with children and define treatment status according to the income of the husband. That is, wifes with low income husbands would be assigned to the treatment group and wives married to high income husbands would be assigned to the control group. A main reason for why we have not taken this route is that in order for the structural interpretation of β to hold, we need to impose the assumption that the marginal effect of τ on e is the same in the treatment and control groups. In practice, this means that we will not only have to consider common trends for households with and without children, but we also need to check that the employment *levels* are reasonably similar between the groups.[28] As was apparent from figure 2 above, this is indeed the case for couples with and without children. In contrast, female employment is systematically higher in high income households than in low income households.[29]

[27]These regions are (i) Sweden, (ii) Western Europe, North America and Oceania, (iii) Eastern Europe and former Soviet Union, (iv) South America, (v) Sub-Saharan Africa, (vi) Northern Africa and Middle East and (vii) Asia.

[28]As emphasized in Section A.3, we expect the employment response to depend on the employment level.

[29]For this reason, as was explained in Section 4, we instead exploit untreated high-income households for

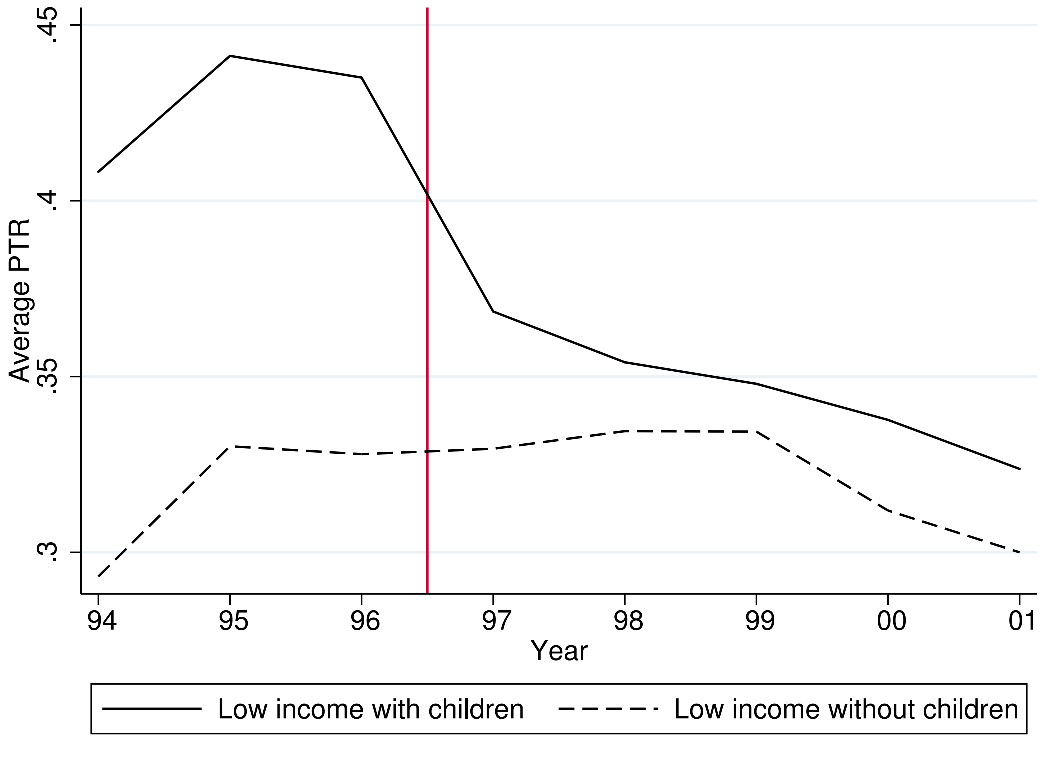

Figure 3: (*Graphical first-stage*) Average participation tax rates (PTR) by child status on HEK data. PTR:s are calculated in FASIT. The sample is restricted to households where the husband's qualifying income falls below the 50th percentile and where the husband reports positive earnings.

5.3 Results

In Figure 3 we plot the evolution of the average PTR for the treatment and control groups (households with and without children) over the time period 1994-2001.[30] As can be seen, the reform in 1997 implied a sharp drop in the average PTR for the treatment group. This drop was caused by the housing allowance reform and demonstrates the strength in the first stage of our IV strategy. Before the housing allowance reform of 1997, the gap in the average PTR:s for couples with and without children, respectively, exceeded 10 percentage points. After the reform, the gap was substantially smaller.

We now turn to our IV estimates of participation elasticities. In appendix A.4, we describe how we construct the participation elasticities based on the regression coefficients (marginal effects). The results are presented in Table 2. Columns 1-4 show estimates using different sets of control variables. The instruments are strongly correlated with the PTR, since a lot of the varia-

making placebo tests. Reduced form results are, however, quite similar if we keep 'low income households with children' as the treatment group, but instead use 'high income households with children' as the control group.

[30]We maintain the same sample restrictions as section 4.

tion in PTR:s comes from the interaction between time and child status. In the 2SLS regressions presented in Table 2, the first-stage F-statistic of the excluded instruments is always very large. In each case, we obtain precise estimates of the participation elasticity. Our preferred estimate is obtained for our most ambitious set of controls (column 4), in which case the elasticity estimate is 0.13. The exact magnitude of the elasticity estimate varies somewhat depending on the set of control variables used in the regressions. This is perhaps not too surprising in light of the results for the reduced form effects in Table 1.

Before closing this section we would like to remark on the generated regressor problem. The PTR:s have been estimated in a separate step with some uncertainty, and our standard errors in the main analysis might be biased due to the presence of the generated regressor in equation (5). Therefore, we made a correction of the standard errors á la Murphy and Topel (1985) for the specification without control variables reported in column 1.[31] The correction did not have any substantial impact on the standard errors. The standard error increased only slightly from 0.013 to 0.014.[32] This does not come as a surprise given the large R2-value from the imputation regression (the included variables explain more than 90 percent of the variation, hence the uncertainty in the PTR predictions is small). Thus, we conclude that the imputation procedure is not problematic from the perspective of statistical inference.

5.4 Heterogeneous responses

In the past, extensive margin responses to taxes have been estimated on relatively small data sets. Since we have access to population wide registers we are able to examine how the elasticity differs across subpopulations with different baseline employment rates in a systematic way. More specifically, we divide the low income sample into four quartile groups based on the imputed log earnings of the secondary earner. As already mentioned in section 3, and formally explained in appendix A.3, this corresponds to a particular grouping of the household types. Of course, predicted earnings may reflect other differences than differences in skills, and our

[31]Performing a proper correction of the covariance matrix for the full specification, which contains a huge amount of dummy variables, would be computationally very burdensome.

[32]We make this correction while estimating (5) with the control function method. By construction, 2SLS and the control function method give identical coefficients under linearity, but in general the standard errors differ. In our case, the standard errors are very similar with 2SLS and control function method.

Table 2: Participation elasticity estimates.

	(1)	(2)	(3)	(4)
Participation elasticity	0.088***	0.116***	0.098***	0.127***
	(0.013)	(0.013)	(0.020)	(0.019)
Household type dummies	No	Yes	Yes	Yes
Household type × children	No	No	Yes	Yes
Household type × year dummies	No	No	Yes	Yes
Additional controls	No	No	No	Yes
Nr of observations	2,770,100	2,770,100	2,770,100	2,770,100

Note: Elasticities are evaluated at the mean values of employment (0.897) and (1-PTR) (0.659) over the years 1994-2001 in the total 'low income sample'. 2SLS regressions are run on 'low income sample', which consists of wives married to husbands with a qualifying income below the 50th percentile. The average take-up rate is set to 0.6. The interactions between the year dummies and the dummy for having children are the excluded instruments. All specifications contain a dummy for having children and a full set of year dummies. 400 household types are defined based on 5 age dummies for each spouse and 4 education level dummies for each spouse. The additional control variables are specified in section 5. Standard errors reported below the estimates are robust to heteroscedasticity and clustered at the household level. * indicates significance at 10%. level, ** 5% level and *** at 1% level. Standard errors for elasticities are obtained by the delta method.

analysis should not be interpreted as an attempt to assess the causal effect of the skill level on the elasticity.

After partitioning the sample into four quartile groups, we rerun equation (5) on each group and evaluate the elasticity at the *subsample-specific* mean values of employment and $(1-\tau)$.[33] The results are shown in Table 3 where we use the full set of control variables. As we move across the four quartile groups, we see that the elasticities fall monotonically in the wife's potential income, mirrored by a corresponding monotonic increase in the employment level. In line with our expectations, the elasticity is the largest in the first quartile group, where the employment level is substantially smaller than in the other three quartile groups. The elasticity estimate for the first quartile group (0.226) and the fourth quartile group (0.088) are statistically different at a level of 95 percent.[34]

[33]For details, see equation (16) in appendix A.3.

[34]Following e.g. Clogg et al. (1995), p.1276, we test this using the fact that differences between the coefficients from a regression run on two independent large samples x and y can be assessed by the statistic $Z = \left(\hat{\beta}_x - \hat{\beta}_y\right) \Big/ \sqrt{se_x^2 + se_y^2}$, which follows a standard unit normal distribution. $\hat{\beta}_j$ and se_j are the coefficient and the standard error of sample $j = x, y$. Since we are interested in testing for differences in elasticities, we have made the proper adjustments by multiplying the coefficients and standard errors by different constants. Using the values for the elasticities and standard errors in column 1 and 4 of Table 2 we obtain a Z-ratio of 2.235, which is larger than the critical value 1.96.

Table 3: Heterogenous response

	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Participation elasticity	0.226***	0.119*	0.110**	0.088***
	(0.056)	(0.046)	(0.037)	(0.026)
Mean employment level	0.808	0.903	0.923	0.955
PTR coefficient	-0.178	-0.102	-0.094	-0.078
Household type dummies	Yes	Yes	Yes	Yes
Household type × children	Yes	Yes	Yes	Yes
Household type × year dummies	Yes	Yes	Yes	Yes
Additional controls	Yes	Yes	Yes	Yes
Nr of observations	692,559	692,542	692,476	692,523

Note: Elasticities are evaluated at the mean values of each subsample. 2SLS regressions are run on 'low income sample', which consists of wives married to husbands with a qualifying income below the 50th percentile. Quartiles are created based on tbe wife's predicted income. The average take-up rate is set to 0.6. The interactions between the year dummies and the dummy for having children are the excluded instruments. All specifications contain a dummy for having children and a full set of year dummies. 400 household types are defined based on 5 age dummies for each spouse and 4 education level dummies for each spouse. The additional control variables are specified in section 5. Standard errors reported below the estimates are robust to heteroscedasticity and clustered at the household level. * indicates significance at 10%. level, ** 5% level and *** at 1% level. Standard errors for elasticities are obtained by the delta method.

Before closing this section, it is useful to compare our results to Selin (2014) who exploited the switch from joint to individual taxation in Sweden in 1971 to estimate participation elasticities for married Swedish women. Selin found estimates in the range 0.5-1.0, which are well above our estimates. However, the results are actually completely consistent when adopting the perspective of our paper. Selin (2014) reports that the pre-reform employment level for married women was 67% (Table 8) whereas the corresponding share in our study is 90%. This further highlights the important relationship between the participation elasticity and employment level that we have emphasized in this paper.

6 Concluding remarks

In this paper, we have investigated how the labor force participation of secondary earners responded to a large reform in the tax-transfer system in Sweden. Using detailed information about individuals' budget sets, and a specific economic framework, we have estimated participation elasticities, exploiting the reform for identification. Our central estimate of the participation elasticity is 0.13, arguably a lower value than many earlier estimates obtained in the literature. Crucially, we have also presented quasi-experimental estimates of participation elasticities for subgroups of the population with different employment levels. This exercise was made possible by virtue of our large sample size. Dividing up the population into four quartile groups based on the wife's potential income, we find participation elasticities ranging from 0.23 at the bottom to 0.09 at the top. The point estimates of the elasticities fall monotonically across these groups, and the elasticity differences between the bottom and the top are statistically significant. These results are quite intuitive: The higher the employment level, the smaller is the pool of unemployed that can be incentivized to enter the labor force. In line with the public finance literature, we have assumed that employment is voluntary and focused on the decision to enter the labor force. If involuntary unemployment is more common among those with low potential incomes, we potentially underestimate the participation elasticity in this group.

The key insight from this paper is that the participation elasticity is fundamentally different in nature from the intensive margin labor elasticity. When designing tax reforms targeted to specific groups, it is important to consider the employment level in the subpopulation of interest.[35] This point has been made before, see e.g. Chetty et al. (2012); our contribution is to examine this feature of the participation response using administrative data and a quasi-experimental identification strategy.

[35]Our quasi-experimental estimates provide a useful contrast against estimates obtained using microsimulation models. Immervoll et al. (2007) analyze welfare reforms in 15 European countries including Sweden, and calibrate the average participation elasticity for the whole economy to 0.2, but decreasing across deciles. In a related exercise, which is more focused on participation responses, Immervoll et al. (2011) assume participation elasticities for secondary earners in the range 0.3-0.7. In light of this paper these elasticities appear to be too large, at least for a country like Sweden.

References

Alpert, A. and D. Powell (2014). Estimating Intensive and Extensive Tax Responsiveness Do Older Workers Respond to Income Taxes? Working paper, Working Paper, Rand Corporation.

Angelov, N., P. Johansson, and E. Lindahl (2016). Parenthood and the gender gap in pay. *Journal of Labor Economics 34*(3), 545–579.

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

Bargain, O. and A. Peichl (2016). Own-wage labor supply elasticities: variation across time and estimation methods. *IZA Journal of Labor Economics 5*(1), 10.

Bertrand, M., E. Duflo, and S. Mullainathan (2002). How much should we trust differences-in-differences estimates? Working Paper 5023, National Bureau of Economic Research.

Bhargava, S. and D. Manoli (2015). Psychological frictions and the incomplete take-up of social benefits: Evidence from an IRS field experiment. *American Economic Review 105*(11), 3489–3529.

Blundell, R., M. Francesconi, and W. van der Klaauw (2011). Anatomy of welfare reform evaluation: Announcement and implementation effects. IZA Discussion Papers 6050, Institute for the Study of Labor (IZA).

Bosch, N. and B. van der Klaauw (2012). Analyzing female labor supply. Evidence from a Dutch tax reform. *Labour Economics 19*(3), 271–280.

Boverket (2006). Bostadsbidrag, ett rättvist bostadsstöd för barnen? - långsiktiga effekter av 1990-talets besparingar. *Diarienummer: 212-3795/2006*.

Brewer, M., E. Saez, and A. Shephard (2010). *Dimensions of Tax Design: the Mirrlees Review*. Oxford University Press.

Chetty, R. (2012). Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply. *Econometrica 80*(3), 969–1018.

Chetty, R., A. Guren, D. Manoli, and A. Weber (2012). Does indivisible labor explain the difference between micro and macro elasticities? A meta-analysis of extensive margin elasticities. In *NBER Macroeconomics Annual 2012, Volume 27*, NBER Chapters, pp. 1–56. National Bureau of Economic Research, Inc.

Clogg, C. C., E. Petkova, and A. Haritou (1995). Statistical methods for comparing regression coefficients between models. *American Journal of Sociology 100*(5), pp. 1261–1293.

Eissa, N. (1995). Taxation and labor supply of married women: The tax reform act of 1986 as a natural experiment. Working Paper 5023, National Bureau of Economic Research.

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*(9-10), 1931–1958.

Ellwood, D. T. (2000). The impact of the earned income tax credit and social policy reforms on work, marriage, and living arrangements. *National Tax Journal 53*(4), 1063 – 1105.

Enström Öst, C. (2012). Ekonomiska drivkrafter i bostadsbidragssystemet. En utvärdering av individuella inkomstgränser för makar med barn. Technical report. Rapport 2012:6. Inspektionen för socialförsäkringen.

Flood, L., J. Hansen, and R. Wahlberg (2004). Household labor supply and welfare participation in Sweden. *Journal of Human Resources 39*(4).

Francesconi, M., H. Rainer, and W. van der Klaauw (2009). The effects of in-work benefit reform in britain on couples: Theory and evidence. *Economic Journal 119*(535), 66–100.

Gelber, A. M. (2014). Taxation and the earnings of husbands and wives: Evidence from sweden. *The Review of Economics and Statistics 96*(2), 287–305.

Gelber, A. M., D. Jones, D. W. Sacks, and J. Song (2018). Using non-linear budget sets to estimate extensive margin responses: Method and evidence from the social security earnings test. Technical report, NBER Working Paper No. 23362.

Gelber, A. M. and J. W. Mitchell (2011). Taxes and time allocation: Evidence from single women and men. *The Review of Economic Studies*.

Havnes, T. and M. Mogstad (2011). Money for nothing? Universal child care and maternal employment. *Journal of Public Economics 95*(11), 1455–1465.

Immervoll, H., H. J. Kleven, C. T. Kreiner, and E. Saez (2007). Welfare reform in European countries: a microsimulation analysis. *The Economic Journal 117*(516), 1–44.

Immervoll, H., H. J. Kleven, C. T. Kreiner, and N. Verdelin (2011). Optimal tax and transfer programs for couples with extensive labor supply responses. *Journal of Public Economics 95*, 1485 – 1500.

Jacob, B. A. and J. Ludwig (2012). The effects of housing assistance on labor supply: Evidence from a voucher lottery. *American Economic Review 102*(1), 272–304.

Kleven, H., C. Landais, and J. E. Søgaard (forthcoming). Children and gender inequality: Evidence from Denmark. *American Economic Journal: Applied Economics (forthcoming)*.

Kosonen, T. (2014). To work or not to work? The effect of childcare subsidies on the labour supply of parents. *The B.E. Journal of Economic Analysis & Policy 14*(3), 32.

Lundin, D., E. Mörk, and B. Öckert (2008). How far can reduced childcare prices push female labour supply? *Labour Economics 15*(4), 647 – 659.

Meyer, B. D. and D. T. Rosenbaum (2001). Welfare, the earned income tax credit, and the labor supply of single mothers. *The Quarterly Journal of Economics 116*(3), 1063–1114.

Murphy, K. M. and R. H. Topel (1985). Estimation and inference in two-step econometric models. *Journal of Business & Economic Statistics 3*(4), 370–79.

Selin, H. (2014). The rise in female employment and the role of tax incentives. An empirical analysis of the Swedish individual tax reform of 1971. *International Tax and Public Finance 21*(5), 894–922.

Online Appendix (not for publication)

A A model to interpret the evidence

A.1 The model

To support the interpretation of our empirical evidence we sketch a simple model that will allow us to (i) clarify conditions under which there is a very simple relationship between elasticities describing the responsiveness to transfers with imperfect take-up and elasticities with respect to changes in taxes (which by assumption have perfect take-up) (see section A.2 below) and (ii) highlight how participation elasticities depend on the employment level of the subgroup under consideration (see section A.3 below).

We consider a model with a discrete set of household types $\mathcal{H}$ indexed by $h \in \mathcal{H}$. There are π_h number of households of each household type. Each household consists of two agents with earnings capacaties z_h^p and z_h, where $z_h^p > z_h$, making one household member the "primary earner" and the other household member the "secondary earner". In a given household type all households are identical with respect to their potential earnings z_h^p and z_h. We focus on the optimal decision-making of the secondary earner from the perspective of the household, treating the primary earner as a passive agent with fixed income z_h^p.

The household decides whether the secondary earner should enter the labor force or not and whether the household should take up the transfer or not. There is no intensive margin hours choice in the theoretical model. As the reform changed marginal work incentives at very low earnings levels of the secondary earner we *a priori* consider the extensive margin to be the important one.[36] Within a given household type households differ along two dimensions, 'fixed costs of working', q_h, and 'take-up costs', χ_h. Each household i of household type h makes a draw from the joint distribution of q_h and χ_h with the associated bi-variate probability density function $f_h(q_h, \chi_h)$. In the tradition of Cogan (1981) and Hausman (1980) the fixed cost of working, q_h, can be interpreted broadly to accommodate the utility costs (stemming from foregone leisure or the psychological costs associated with leaving a child under the supervision

[36]We have also conducted a reduced form analysis which points in this direction, see Table A1 and the discussion in section 4.

of a non-parent) or monetary costs (such as commuting or child care costs) associated with secondary earner labor market entry. The take-up cost, χ_h, can be interpreted as a cost from gathering information about the transfer program, a time-cost associated with filling out the paperwork, a complexity cost (understanding, and gathering the correct information about how to fill out the paperwork) or simply the social stigma associated with accepting transfers from the government.[37]

The two binary decisions at the household level implies that each household selects between four different states: (i) working without transfers, (ii) working with transfers, (iii) not-working and not taking up transfers, and, finally, (iv) not working and taking up transfers. We denote the decision of the household by $(M, L) \in \{0, 1\} \times \{0, 1\}$ where M is the take-up decision and L is the labor force participation decision of the secondary earner. Let c_{ih} denote household consumption of household i in household type h. The utility function for each household is:

$$u_{ih}(c_{ih}, M_{ih}, L_{ih}) = c_{ih} - q_{ih}L_{ih} - \chi_{ih}M_{ih}, \tag{6}$$

and the budget constraint of the household is given by:

$$c_{ih} \leq z_h^p + z_h L_{ih} - T(z_h^p, z_h L_{ih}) + B(z_h^p, z_h L_{ih})M_{ih} \tag{7}$$

where $T(z_h^p, z_h L_{ih})$ is the total tax liability (possibly negative) and $B(z_h^p, z_h L_{ih})$ is a non-negative transfer received from the government. It is a standard practice in the public finance literature to treat the nonlinear income tax T as representing the complete tax system (including transfers). In this paper we follow this approach with the exception that we leave out the *particular components* of the transfer system that are associated with costly take-up and designate these to the B-function.

Each household of type h chooses, based on its realized characteristics $(q_{ih}, \chi_{ih}) \in \mathbb{R}_+^2$, one out of the four different alternative states to maximize their utility (6) subject to the budget constraint (7). The mass of individuals choosing each state (M, L) correspond to different regions

[37] Using a large-scale policy experiment, conducted in collaboration with the Internal Revenue Service (IRS) in the US, Bhargava and Manoli (2015) find that incomplete take-up among low-income earners can at least partially be attributed to lack of program awareness and understanding combined with an aversion to program complexity.

in the (q,χ)-space. We denote the share of households of household type h in each state with e_h^{ML}, $M=0,1; L=0,1$. Employment in household h is defined as $e_h = e_h^{11} + e_h^{01}$.

A.2 Participation elasticites with imperfect take-up

We now introduce the following simplified notation based on the T and B functions introduced in the budget constraint (7): $T_h^1 = T_h(z_h^p, z_h)$, $T_h^0 = T(z_h^p, 0)$, $T_h = T_h^1 - T_h^0$ and $B_h^1 = B_h(z_h^p, z_h)$, $B_h^0 = B_h(z_h^p, 0)$. We assume $B_h^0 > B_h^1$ and $T_h^1 > T_h^0$, which is the relevant case that applies when transfers are means-tested and participation taxes are less than 100%. In terms of the variables above, the participation tax introduced in (2) can be decomposed as:

$$\tau_h = \frac{T_h(z_h^p, z_h) - T(z_h^p, 0) + \left[B_h(z_h^p, 0) - B_h(z_h^p, z_h)\right]}{z_h} = \frac{T_h + B_h^0 - B_h^1}{z_h}. \tag{8}$$

This is the relevant participation tax rate for an individual who takes up both the work-related transfer and the non-work transfer and allows us to distinguish, for theoretical purposes, between three possible sources of variation in the incentives to participate in the labor force. These are, (i) a variation in T_h (the difference in taxes between the work and non-work state), (ii) a variation in transfer in the state of non-employment B^0, and, (iii) a variation in the transfer in the state of employment B^1.[38]

We define $\epsilon_h = -\frac{de_h}{dT_h}\frac{z_h - T_h - B_h^0 + B_h^1}{e_h}$ as the *participation elasticity* which yields the percentage increase in employment following a one percent increase in the financial reward from working $z_h - T_h - B_h^0 + B_h^1$ due to a change in T_h. Moreover, we define $\epsilon_h^{B^0} = -\frac{de_h}{dB_h^0}\frac{z_h - T_h - B_h^0 + B_h^1}{e_h}$ and $\epsilon_h^{B^1} = \frac{de_h}{dB_h^1}\frac{z_h - T_h - B_h^0 + B_h^1}{e_h}$ as the *transfer elasticities*, i.e. the elasticities obtained when using variation in the transfer system (which are subject to take-up costs).[39] We can then derive the following proposition which is very useful:

Proposition 1. *Suppose that at the household-type level, namely, for each $h \in \mathcal{H}$, (i) the random variables q_h and χ_h are independent, and, (ii) q_h is locally uniform on the open interval $(z_h -$*

[38]The difference between T^{total} entering equation (2) and T entering (8) is that T excludes those components of the transfer system that are associated with costly take-up which we instead capture with the B-function. In our empirical analysis the variation in T^{total} stems mainly from variation in B^0.

[39]Notice that we have chosen to evaluate all elasticities at the point $z_h - T_h - B_h^0 + B_h^1$ which is the financial reward from work for a person who takes up transfers both in the state of work and non-work.

$T_h - B_h^0, z - T_h) \subset \mathbb{R}^+$ and unrestricted elsewhere. Then, letting G_h denote the CDF of χ_h,

$$\epsilon_h = \frac{\epsilon_h^{B^0}}{G_h(B_h^0)} = -\frac{\epsilon_h^{B^1}}{G_h(B_h^1)},$$

where $G_h(B_h)$ is the take-up rate in household type h when the level of transfers is B_h, or, equivalently, the fraction of type-h workers with take-up costs less than B_h.

Proof

First we characterize the fractions of the population in each of the four household states emphasized on page 33 (i.e. $e^{ij}, i, j = 0, 1$) without making any distributional assumptions. Thereafter we impose the assumptions in Proposition 1 to derive the relevant derivatives of e with respect to T, B^0, and, B^1, that can be used to establish the relationship between the two key elasticities given in the proposition. To simplify the exposition in this appendix we omit the h index. All calculations apply at the household-type level.

Part 1

We describe the decision-making of the household by considering the labor-market entry conditions for the secondary earner depending for different values of the take-up cost χ.

If $0 \leq \chi \leq B^1$ the household always takes up the transfer (both when working and not working) and therefore participates in the labor force when the following condition is met:

$$z - (T^1 - T^0) - (B^0 - B^1) \geq q \quad \text{(low)} \tag{9}$$

If $\chi > B^0$ the household does not take up the transfer in the state of work nor in the state of non-work, and the participation equation becomes:

$$z - (T^1 - T^0) \geq q \quad \text{(high)} \tag{10}$$

If $B^1 < \chi \leq B^0$ the household takes up the transfer when unemployed, but not when working, which implies that the participation equation becomes:

$$z - (T^1 - T^0) - B^0 \geq q - \chi \quad \text{(intermediate)} \tag{11}$$

Note that this last condition depends on χ. That is, the incentive to enter the labor force depends on the size of the take-up cost. It appears in (11) because households with $B^1 < \chi \leq B^0$ only experience the take-up cost when they are outside the labor force.

As the above conditions only depend on the difference between T^1 and T^0 we set $T = T^1 - T^0 \geq 0$ without loss of generality. We denote the threshold values of q which cause inequalities (9), (10), and (11) to bind by q^L, q^H, and q^I, respectively. We have that $q^L \leq q^I \leq q^H$ by virtue of the fact that $B^0 > B^1$ (and the fact that q^I only applies for values of χ satisfying $B^1 < \chi \leq B^0$). Notice that q^L and q^H are fixed and can be expressed in terms of observable quantities as $q^L \equiv q^L(z, T, B^1, B^0)$ and $q^H \equiv q^H(z, T)$ [specifically, $q^L = z - (T^1 - T^0) - (B^0 - B^1)$ and $q^H = z - (T^1 - T^0)$] whereas q^I depends on the take-up cost χ and takes on the value $q^I = q^L$ when $\chi = B^1$ and $q^I = q^H$ when $\chi = B^0$.[40]

In the following we assume q and χ are jointly distributed according to the probability density function $f(q, \chi)$.

Based on conditions (9)-(11) we can write down the number of workers in each state e^{ML}, $M = 0, 1; L = 0, 1$. Note that the division of agents into the four categories above based on their innate characteristics (q, χ) completely characterizes the optimal behavior of agents.

The number of households who work and take-up transfers are:

$$e^{11} = \int_0^{q^L} \int_0^{B^1} f(q, \chi) d\chi dq$$

The number of households who work and do not take up transfers are:

$$e^{10} = \int_0^{q^L} \int_{B^1}^{\infty} f(q, \chi) d\chi dq + \int_{q^L}^{q^H} \int_{q-q^I}^{\infty} f(q, \chi) d\chi dq$$

The number of households who do not work and take up transfers are:

$$e^{01} = \int_{q^H}^{\infty} \int_0^{B^0} f(q, \chi) d\chi dq + \int_{q^L}^{q^H} \int_0^{q-q^I} f(q, \chi) d\chi dq$$

[40]Notice that q^I will be a line in the (χ, q)-space.

Finally, the number of households who neither work nor take up transfers are:

$$e^{00} = \int_{q^H}^{\infty} \int_{B^0}^{\infty} f(q, \chi) d\chi dq.$$

It follows by construction that the total number of workers is $e^1 = e^{11} + e^{10}$ and that the total number of unemployed agents is $e^0 = e^{00} + e^{01}$ with $e^0 + e^1 = 1$.

<u>Part 2</u>

Assuming q and χ are independent we can write the number of individuals in each group as follows:

$$\begin{aligned}
e^{11} &= F(q^L)G(B^1) \\
e^{10} &= F(q^L)[1 - G(B^1)] + \int_{q^L}^{q^H} f(q) \left[\int_{q-q^I}^{\infty} g(\chi) d\chi \right] dq \\
&= F(q^L)[1 - G(B^1)] + \int_{q^L}^{q^H} f(q)[1 - G(q - q^I)] dq \\
e^{01} &= [1 - F(q^H)]G(B^0) + \int_{q^L}^{q^H} f(q) \left[\int_{0}^{q-q^I} g(\chi) d\chi \right] dq \\
&= [1 - F(q^H)]G(B^0) + \int_{q^L}^{q^H} f(q)G(q - q^I) dq \\
e^{00} &= [1 - F(q^H)][1 - G(B^0)]
\end{aligned}$$

To establish Proposition 1 we need to compute the derivatives of $e = e^{11} + e^{10}$ with respect to the tax/transfer instruments T, B^0 and B^1. That is, we are interested in computing:

$$\begin{aligned}
\frac{de}{dB^0} &= \frac{de^{11}}{dB^0} + \frac{de^{10}}{dB^0} \\
\frac{de}{dB^1} &= \frac{de^{11}}{dB^1} + \frac{de^{10}}{dB^1} \\
\frac{de}{dT^1} &= \frac{de^{11}}{dT^1} + \frac{de^{10}}{dT^1}.
\end{aligned}$$

To make progress we impose the additional assumption that $F(q)$ is *locally* uniform on the open interval $(z - T - B^0, z - T)$ in the sense that it has constant pdf with density γ on this interval and is unrestricted elsewhere. In the derivations below, recall that $q^L = z - (T^1 - T^0) - (B^0 - B^1)$ and $q^H = z - (T^1 - T^0)$.

Then, we first notice that:

$$
\begin{aligned}
\frac{de^{11}}{dB^0} &= -\gamma G(B^1) \\
\frac{de^{11}}{dB^1} &= \gamma G(B^1) + G'(B^1)F(q^L) \\
\frac{de^{11}}{dT^1} &= -\gamma G(B^1).
\end{aligned}
$$

For example, the first condition above states that as B^0 is marginally increased, there will be an outflow from the group of workers who take-up transfers according to their number $G(B^1)$ times the marginal density of the fixed-cost distribution γ (which simply reflects the number of individuals who are indifferent between working and not working).[41] In the second condition, the first term states that as B^1 is increased, the fraction of workers who take up the transfer when working will be incentivized to join the labor force, according to the marginal density γ. In addition, there will be an increase in take-up represented by the second term.

Applying slightly more effort we can apply Leibniz integral rule and derive:

$$
\begin{aligned}
&\frac{de^{10}}{dB^0} = \frac{d}{dB^0}\Big(F(q^L)[1 - G(B^1)]\Big) + \\
&\gamma \int_{q^L}^{q^H} \frac{d}{dB^0}[1 - G(q - q^I)]dq + \gamma\frac{dq^H}{dB^0}[1 - G(q^H - q^I)] - \gamma\frac{dq^L}{dB^0}[1 - G(q^L - q^I)] = \\
&= -\gamma[1 - G(B^1)] + \gamma[-G(q - q^I)]_{q^L}^{q^H} + \gamma[1 - G(B^1)] = \\
&= -\gamma[G(B^0) - G(B^1)]
\end{aligned}
$$

This condition gives the change in the group who works and does not take up transfers in response to an increase in the guaranteed income (out-of-work transfer) B^0. An increase in B^0 increases non-participation proportionally to $[G(B^0) - G(B^1)]$ which is the fraction of workers with intermediate take-up costs in the sense that they only take-up transfers when unemployed.

[41] Notice that any worker who belongs to the group e^{11} will by assumption also take up the transfer when not-working since $B^1 \leq B^0$.

Similarly, we can derive:

$$\frac{de^{10}}{dB^1} = \frac{d}{dB^1}\Big(F(q^L)[1 - G(B^1)]\Big) +$$
$$\gamma \int_{q^L}^{q^H} \frac{d}{dB^1}[1 - G(q - q^I)]dq + \gamma\frac{dq^H}{dB^1}[1 - G(q^H - q^I)] - \gamma\frac{dq^L}{dB^1}[1 - G(q^L - q^I)] =$$
$$= -F(q^L)G'(B^1).$$

This expression states that as B^1 increases, there will be a dynamic take-up response. Some who previously worked without transfers will now work and take up transfers.

Finally, we derive:

$$\frac{de^{10}}{dT^1} = \frac{d}{dT^1}(\gamma q^L + \rho)[1 - G(B^1)] +$$
$$\gamma \int_{q^L}^{q^H} \frac{d}{dT^1}[1 - G(q - q^I)]dq + \gamma\frac{dq^H}{dT^1}[1 - G(q^H - q^I)] - \gamma\frac{dq^L}{dT^1}[1 - G(q^L - q^I)] =$$
$$= -\gamma[1 - G(B^1)]$$

To understand this effect note that e^{10} is the fraction of workers who do not take up the transfer while working represented by the fraction $[1 - G(B^1)]$ of the population. A number of these individuals will drop out of the labor force in response to the tax increase according to the marginal density γ.

Putting things together we get:

$$\frac{de}{dB^0} = \frac{de^{11}}{dB^0} + \frac{de^{10}}{dB^0} = -\gamma G(B^1) - \gamma[G(B^0) - G(B^1)] = -\gamma G(B^0)$$
$$\frac{de}{dB^1} = \frac{de^{11}}{dB^1} + \frac{de^{10}}{dB^1} = \gamma G(B^1) + G'(B^1)F(q^L) - F(q^L)G'(B^1) = \gamma G(B^1)$$
$$\frac{de}{dT^1} = \frac{de^{11}}{dT^1} + \frac{de^{10}}{dT^1} = -\gamma G(B^1) - \gamma[1 - G(B^1)] = -\gamma.$$

This completes the proof. □

The above proposition specifies sufficient conditions under which reforms in transfers (that are subject to take-up decisions) can readily be used to assess the sensitivity of employment to taxes. The only necessary adjustment in this case is to scale the transfer-elasticities with the inverse of the take-up rate. Notice that the distributional assumptions in Proposition 1 are not very restrictive since they apply at the *household-type* level. Even though we in this paper study an out-of-work (guaranteed income) program (a variation in B^0), Proposition 1 can also be fruitfully applied when studying in-work tax credits (variations in B^1).

A.3 Heterogeneous responses and aggregate elasticities

It is well-known that the responsiveness along the extensive margin is not captured by a single structural parameter but instead by the number of workers who are, at the margin, indifferent between working and not working. To illustrate this in the simplest possible way, consider our model while assuming *identical* fixed cost distributions for all $h \in \mathcal{H}$, with pdf $f(q)$ and cdf $F(q)$. In this simple example we abstract from the take-up decision. Hence, employment in household type h can be written $e_h = \int_0^{z_h - T_h} f(q)dq = F(z_h - T_h)$. Notice that when the fixed cost functions are identical across h, the employment level will solely depend on disposable income in the state of work, $z_h - T_h$, and employment will be larger in household types with larger potential earnings. We have that $z_h - T_h = F^{-1}(e_h)$ where $F^{-1}(e_h)$ is the generalized inverse distribution function defined as $F^{-1}(e_h) = \inf\{x \in \mathbb{R} \mid F(x) \geq e_h\}$. Moreover,

$$\frac{de_h}{dT_h} = -F'(z_h - T_h) = -F'(F^{-1}(e_h)). \tag{12}$$

This shows that the employment effect depends on the mass (density) of the fixed cost distribution at the *quantile* $F^{-1}(e_h)$. Specifically, $\frac{de_h}{dT_h}$ will depend on e_h, unless F is uniform. A related observation is made by Chetty et al. 2012 who notes that the size of the extensive margin responses depend on the density of the distribution of reservation wages around the economy's equilibrium and that these elasticities vary with the wage rate unless the density of the reservation wage distribution happens to be uniform.[42]

In the empirical analysis we will recover participation elasticities for different subgroups

[42]The model analyzed by Chetty et al. (2012) is isomorphic to ours. The reservation wage corresponds to the fixed-cost threshold for labor force participation that appear in the proof of Proposition 1. Moreover, in a perfectly competitive labor market equilibrium, there is a one-to-one relationship between the wage rate and the employment level.

by using variation in the secondary earner's PTR. Recall that the PTR conditional on taking up the transfer is $\tau_h = \frac{T_h + B_h^0 - B_h^1}{z_h}$. As explained in section 2, the variation in τ_h mainly originates from changes in transfers received in the state of non-work, B^0. We now assume that there are Θ subsets of $\mathcal{H}$ and denote each subset by $\mathcal{H}_\theta$. In the empirical analysis in the main text, we group household types into four groups $\{\mathcal{H}_\theta\}_{\theta=1}^4$ based on the secondary earners' predicted income. The average employment in each set $\mathcal{H}_\theta$ is $\overline{e}_\theta = \sum_{h\in\mathcal{H}_\theta} \frac{\pi_h e_h}{\sum_{h\in\mathcal{H}_\theta} \pi_h}$. Consider now how this quantity responds to a marginal increase in the PTRs $\{\tau_h\}_{h\in\mathcal{H}_\theta}$ induced by marginal increases in B_h^0, $h \in \mathcal{H}_\theta$. The marginal effect on $\overline{e}_\theta$ of such a change can, invoking the assumptions in Proposition 1, be written as:

$$\nabla_{\mathbf{v}} \overline{e}_\theta = -\sum_{h\in\mathcal{H}_\theta} \frac{\pi_h}{\sum_{h\in\mathcal{H}_\theta} \pi_h} \frac{de_h}{dB_h^0} z_h \tag{13}$$

$$= -\sum_{h\in\mathcal{H}_\theta} \frac{\pi_h}{\sum_{h\in\mathcal{H}_\theta} \pi_h} \gamma_h z_h G_h(B_h^0) \tag{14}$$

$$= \beta_\theta, \tag{15}$$

where $\nabla_{\mathbf{v}} \overline{e}_\theta$ is the directional derivative of the average employment in group $\mathcal{H}_\theta$ along the direction $\mathbf{v}$ specified by the change in the PTRs $\{\tau_h\}_{h\in\mathcal{H}_\theta}$ (which operate through changes in $\{B_h^0\}_{h\in\mathcal{H}_\theta}$). G_h is the CDF of the take-up cost distribution and γ_h is the density of the fixed cost of work distribution (see part 2 in the proof of Proposition 1 for details). The parameter of interest that we will estimate is β_θ. It is, however, more in line with previous literature to transform marginal effects into elasticities. We define the average participation elasticity in subpopulation $\mathcal{H}_\theta$ as:

$$\overline{\epsilon}_\theta^T = -\sum_{h\in\mathcal{H}_\theta} \frac{\pi_h}{\sum_{h\in\mathcal{H}_\theta} \pi_h} \frac{de_h}{dT_h} \frac{z_h - T_h - B_h^0 + B_h^1}{e_h} = -\sum_{h\in\mathcal{H}_\theta} \frac{\pi_h}{\sum_{h\in\mathcal{H}_\theta} \pi_h} \frac{de_h}{dT_h} z_h \frac{(1-\tau_h)}{e_h}.$$

Using equations (13)-(15), we can approximate the average participation elasticity in subgroup $\mathcal{H}_\theta$ as

$$\overline{\epsilon}_\theta^T \approx \beta_\theta \frac{(1-\bar{\tau}_\theta)}{\overline{e}_\theta \overline{G}_\theta(B^0)}, \tag{16}$$

where for a variable x, $\bar{x}_\theta$ denotes an average over the subset $\mathcal{H}_\theta$. Finally, note that we could use the same reasoning as that behind (16) to aggregate over the entire treated population.

A.4 Constructing participation elasticities from marginal effects

The elasticities are calculated according to equation (16) in appendix A.3, where we multiply the estimate $\hat{\beta}$ of β of equation (5) with the ratio $\frac{\overline{1-\tau}}{\bar{e}\bar{G}(B^0)}$. In this expression, $\overline{1-\tau}$ and $\bar{e}$ are the averages of $1-\tau$ and e (the employment rate) over the years 1994-2001 in the low income sample and $\bar{G}(B^0)$ is the average take-up rate of one-earner households in the pre-reform period, which is observed to be around 0.6 in the HEK sample.[43] Moreover, if the conditions specified in Proposition 1 in section A.2 are satisfied, the participation elasticities that we construct based on the marginal effect in regression (5), can be given a structural interpretation.

B Formula for calculating the HA

Both for 1996 and 1997 the maximum monthly housing allowance (MMHA) can be written

$$MMHA = 600 + \max\{0, (\min(QHE, 3000) - 2000) \times 0.75\} + \max(0, (QHE - 3000) \times 0.50) \quad (17)$$

where HA = household housing allowance [SEK/month], QHE = qualifying household housing expenses [SEK/month], and I = household income before tax [SEK/month]. However, the qualifying housing expenses changed between 1996 and 1997.

In 1996 QHE was simply the rent paid by the tenant. There was also a minimum guaranteed housing expense level (which was a function of the number of children).

For 1997 the QHE can be written

$$QHE = \max\left(MHE, HE \times \frac{\min(SC, AS)}{AS}\right), \quad (18)$$

where MHE =minimum guaranteed housing expense level , HE = actual housing expense (rent),

[43]We obtained this figure by pooling the pre-reform years, 1994-1996. Due to the fact that the sample is restricted to only include households where the wife does not work, the sample size is too small to provide a more disaggregated estimate of the take-up rate.

SC = space constraint and AS = actual space constraint. The space constraint depends on the number of kids in the houshold.[44] [45]

C Other components of the reform

In the discussion of section (2.2) we only considered the individualization of the exemption level, which is the main focus of our paper. However, two other potentially important components of the reform deserve to be mentioned as well; the new space restriction and the *ex post* adjustment of the allowance.

Although the upper cap on the transfer before phase-out, B^{00}, did not change, many households nevertheless experienced a decrease in B^{00}. In the 1997 reform package the government introduced an upper limit to the qualifying housing space, i.e. the number of square-meters of dwelling space the household could be compensated for. We take this space restriction into account when calculating the participation tax rates. It lowered the transfers, especially for couples who tend to live in larger apartments than singles.

Both before and after 1997 the beneficiary had to repay the benefit if the household's qualifying income substantially increased and the household did not report this increase in income.[46] However, before 1997 the household never had to repay an allowance it was eligible for at the month of the monthly benefit payment. From 1997 and onwards, the monthly allowance receipt was labeled as 'preliminary'. In the new system, the beneficiary applies in December year t for housing allowance in year $t + 1$. In year $t + 1$ the beneficiary each month receives the housing allowance based on the qualifying income reported in the application in December year t. In year $t + 2$ the two spouses file their tax returns. By the end of $t + 2$ the Social Insurance Agency receive information from the Tax Agency on the household's *ex post* qualifying incomes in $t + 1$. Finally, in the spring of year $t + 3$ the Social Insurance Agency charge/reimburse households where the incomes reported in year t deviate from the realized income in year $t + 1$.

From the point of view of fiscal sustainability, the reform was a great success, to say the

[44] 1 child: 80 sqm, 2 children: 100 sqm, 3 children: 120 sqm, 4 children: 140 , 5 or more: 160 sqm.

[45] The yearly rent per square meter was approximately SEK 700, 1996-97. Rent statistics: http://www.boverket.se/Global/Webbokhandel/Dokument/2011/Hyror-i-Sverige-1975-2009.pdf, figure 2.1.

[46] See Boverket (2006) (in Swedish) for a description of these pretty complex rules.

least. As can be seen from Figure A1, the government's expenditures on the program fell dramatically in the years following 1997 (marked with a vertical line). Moreover, we see that there was a huge decrease in the number of couples receiving the transfer between 1996 and 1997. The decrease among singles was arguably more modest. Single households were affected both by the space restriction (but to a smaller degree than couples as their dwellings typically were smaller) and by the new rules for *ex post* repayments/reimbursements. However, the income limits of singles were unchanged. Why did the size of the program decrease in the years following 1997? In the post-reform period the benefit levels and the income limits were kept at their nominal levels of 1996; they were not indexed. Accordingly, with inflation and real wage growth, a growing fraction of couples and singles became eligible only for small amounts, or became ineligible.

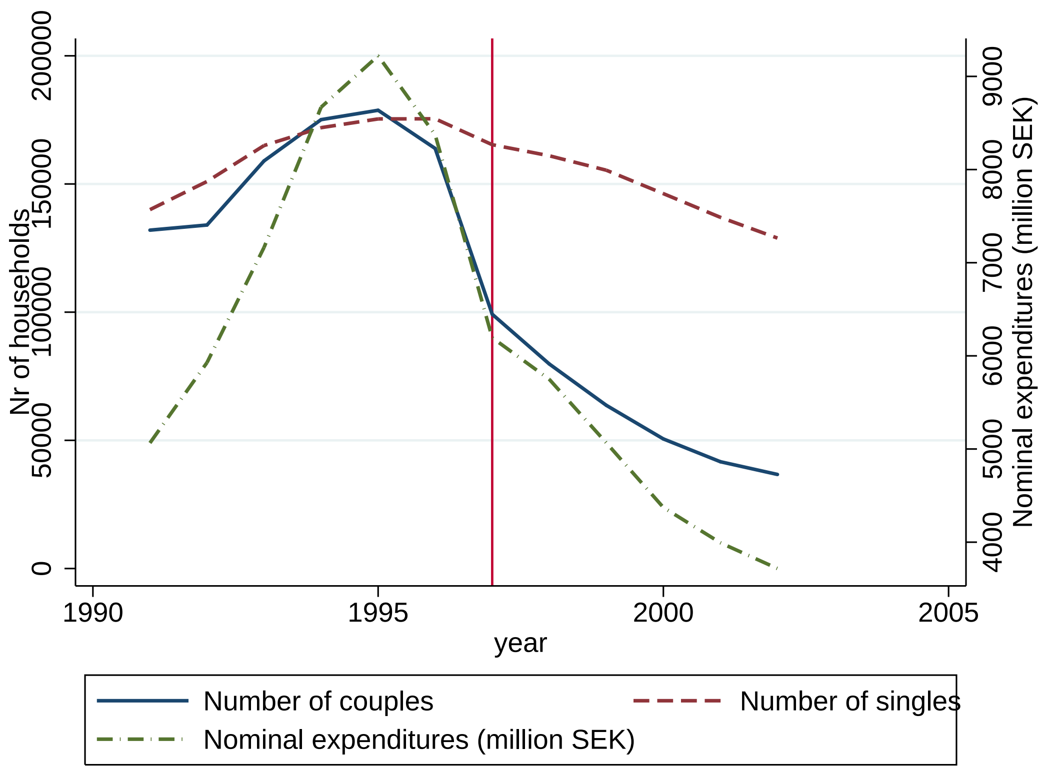

Figure A1: Number of couples and singles receiving housing allowances, as well as nominal expenditures on housing allowances in million SEK. Source: Boverket (2006), Table A.

D Alternative empirical specifications

Table A1: Reduced form effects in alternative specifications

	Male employment (1)	Male log(earnings) (2)	Female log(earnings) (3)
Year 1994 × children	0.045	-0.000	-0.008
	(0.159)	(0.006)	(0.005)
Year 1995 × children	-0.061	0.005	-0.004
	(0.139)	(0.005)	(0.005)
Year 1997 × children	0.049	-0.000	0.002
	(0.141)	(0.005)	(0.005)
Year 1998 × children	-0.036	0.0039	0.000
	(0.161)	(0.006)	(0.006)
Year 1999 × children	-0.081	0.007	0.012*
	(0.170)	(0.006)	(0.006)
Year 2000 × children	-0.096	0.011*	0.029***
	(0.177)	(0.006)	(0.006)
Year 2001 × children	-0.364**	0.009	0.038***
	(0.182)	(0.007)	(0.007)
Household type dummies	Yes	Yes	Yes
Household type × children	Yes	Yes	Yes
Household type × year dummies	Yes	Yes	
Additional controls	Yes	Yes	Yes
Observations	2,658,815	2,521,767	2,485,259

Note: Dependent variable: probability of having positive earnings. 'Male sample' consists of husbands married to wives with a positive qualifying income, which falls below the 50th percentile.'Female sample' sample consists of wives married to husbands with a positive qualifying income, which falls below the 50th percentile. All specifications contain a dummy for having children and a full set of year dummies. 400 household types are defined based on 5 age dummies for each spouse and 4 education level dummies for each spouse. The additional control variables are specified in section 5. Standard errors reported below the estimates are robust to heteroscedasticity and clustered at the household level. * indicates significance at 10%. level, ** 5% level and *** at 1% level.

Table A2: Regression with individual fixed effects

	(1)	(2)
Year 1994 × Children	-1.96***	-0.532***
	(0.107)	(0.135)
Year 1995 × Children	-0.980***	-0.341***
	(0.0876)	(0.111)
Year 1997 × Children	0.880***	0.298***
	(0.0887)	(0.112)
Year 1998 × Children	2.08***	0.776***
	(0.104)	(0.132)
Year 1999 × Children	3.49***	1.22***
	(0.116)	(0.147)
Year 2000 × Children	4.77***	1.52***
	(0.125)	(0.161)
Year 2001 × Children	5.78***	1.64***
	(0.135)	(0.177)
Indivdiual fixed effects	Yes	Yes
Full set of controls	No	Yes
Observations	2,770,100	2,770,100

Note: The specifications correspond to those reported in columns 1 and 4 of Table 1, but here we also add individual fixed effects.

Table A3: Triple difference estimates

	(1)	(2)
Low income × Year 1994 × Children	-0.131	0.0554
	(0.164)	(0.171)
Low income × Year 1995 × Children	-0.127	-0.100
	(0.135)	(0.141)
Low income × Year 1997 × Children	0.213	0.135
	(0.136)	(0.141)
Low income × Year 1998 × Children	0.678***	0.370**
	(0.158)	(0.163)
Low income × Year 1999 × Children	1.31***	0.717***
	(0.173)	(0.178)
Low income × Year 2000 × Children	1.68***	0.639***
	(0.187)	(0.192)
Low income × Year 2001 × Children	2.23***	0.967***
	(0.202)	(0.207)
Individual fixed effects	Yes	Yes
Full set of controls	No	Yes
Observations	4,155,170	4,155,170

Note: In addition to the "low income sample" from the main analysis we also include a "high income sample", see Column 5 of Table 1. "Low income" is a dummy for being in the low income sample. We control for the full set of double interactions between the dummies for low income, children and year.

E Summary Statistics

Summary statistics are reported in Table A4.

Table A4: Summary Statistics

	With children		Without children	
Labor force participation	0.898	(0.303)	0.895	(0.307)
Net of tax rate $1 - \tau$	0.619	(0.064)	0.680	(0.021)
Age of secondary earner	39.720	(5.962)	47.649	(5.765)
Age of primary earner	42.201	(6.246)	49.286	(5.560)
Earnings	1245.602	(841.517)	1395.388	(868.631)
Qualifying income of primary earner	1891.367	(1121.539)	1895.504	(758.296)
Education				
At most 9 years of education	0.156	(0.363)	0.282	(0.450)
At most high school education	0.573	(0.495)	0.536	(0.499)
College education	0.265	(0.441)	0.172	(0.377)
Country of origin				
Sweden	0.920	(0.271)	0.949	(0.220)
Western Europe, North America and Oceania	0.058	(0.233)	0.039	(0.194)
Eastern Europe and former Soviet Union	0.015	(0.121)	0.010	(0.100)
South America	0.001	(0.030)	0.000	(0.016)
Sub-Saharian Africa	0.000	(0.016)	0.000	(0.010)
Northern Africa and Middle East	0.005	(0.069)	0.001	(0.025)
Asia	0.001	(0.033)	0.000	(0.019)
Number of observations	2,069,793		700,307	

Note: Standard deviations reported in parenthesis. Incomes are expressed in 100 SEK. Summary statistics refer to the period 1994-2001.

Table A5: Earnings, HEK-sample

	Pre-reform, with children	Pre-reform, without children	Post-reform, with children	Post-reform, without children
Quartile 1	1019.821	1198.739	1233.113	1419.183
	(19.73331)	(34.10906)	(18.21582)	(28.35034)
	1373	*538*	*2520*	*1026*
Quartile 2	1191.636	1367.007	1451.973	1666.108
	(18.67924)	(31.35867)	(17.48671)	(27.67966)
	1384	*526*	*2366*	*1179*
Quartile 3	1314.855	1535.593	1584.352	1744.547
	(22.4985)	(30.99836)	(19.24418)	(26.25316)
	1309	*601*	*2299*	*1247*
Quartile 4	1462.458	1688.859	1881.471	2084.528
	(26.65222)	(40.10846)	(26.63977)	(34.78893)
	1290	*620*	*2340*	*1203*
Total	1242.934	1459.059	1531.536	1740.829
	(11.18809)	(17.77397)	(10.59878)	(15.16933)
	5356	*2285*	*9525*	*4655*

Note: Standard error of the mean in parenthesis. Number of observations in *italics*.
Incomes are expressed in hundreds of SEK. The sample is selected according to the criteria described in Section 3, but we include all wives (not only the low income sample). "Quartile" refers to quartile group with respect to the husband's income.

Table A6: Earnings, Population data

	Pre-reform, with children	Pre-reform, without children	Post-reform, with children	Post-reform, without children
Quartile 1	1013.048	1183.985	1216.25	1395.403
	(1.171407)	(2.090849)	(1.142551)	(2.077657)
	405419	*136602*	*635267*	*207778*
Quartile 2	1205.643	1331.754	1450.299	1573.232
	(1.110416)	(1.944903)	(1.085439)	(1.93327)
	399465	*142557*	*629674*	*213378*
Quartile 3	1302.709	1450.671	1580.583	1731.482
	(1.206462)	(2.077368)	(1.203234)	(2.146596)
	399083	*142938*	*629224*	*213838*
Quartile 4	1495.891	1694.832	1834.816	2028.685
	(1.595235)	(2.634307)	(1.67026)	(2.977806)
	399077	*142937*	*632142*	*210904*
Total	1253.356	1417.96	1520.11	1683.113
	(0.6566866)	(1.130526)	(0.6633737)	(1.186372)
	1603044	*565034*	*2526307*	*845898*

Note: Standard error of the mean in parenthesis. Number of observations in *italics*.
Incomes are expressed in hundreds of SEK. The sample is selected according to the criteria described in Section 3, but we include all wives (not only the low income sample). "Quartile" refers to quartile group with respect to the husband's income.

Table A7: Labor force participation, HEK-sample

	Pre-reform, with children	Pre-reform, without children	Post-reform, with children	Post-reform, without children
Quartile 1	0.8871085	0.8643123	0.8710317	0.8762183
	(0.0085436)	(0.0147781)	(0.006678)	(0.0102866)
	1373	*538*	*2520*	*1026*
Quartile 2	0.9212428	0.9239544	0.9349112	0.927905
	(0.007243)	(0.0115687)	(0.0050725)	(0.0075358)
	1384	*526*	*2366*	*1179*
Quartile 3	0.9381207	0.953411	0.9430187	0.9198075
	(0.0066619)	(0.0086041)	(0.0048356)	(0.0076941)
	1309	*601*	*2299*	*1247*
Quartile 4	0.9286822	0.9612903	0.9397436	0.9418121
	(0.0071681)	(0.0077534)	(0.0049203)	(0.0067522)
	1290	*620*	*2340*	*1203*
Total	0.9184093	0.9277899	0.9211549	0.9179377
	(0.0037408)	(0.005416)	(0.0027615)	(0.0040231)
	5356	*2285*	*9525*	*4655*

Note: Standard error of the mean in parenthesis. Number of observations in *italics*. The sample is selected according to the criteria described in Section 3, but we include all wives (not only the low income sample). "Quartile" refers to quartile group with respect to the husband's income.

Table A8: Labor force participation, Population data

	Pre-reform, with children	Pre-reform, without children	Post-reform, with children	Post-reform, without children
Quartile 1	0.8667428	0.8814659	0.8657919	0.8753477
	(0.0005338)	(0.0008746)	(0.0004277)	(0.0007247)
	405419	*136602*	*635267*	*207778*
Quartile 2	0.9268422	0.9126595	0.9322649	0.9102813
	(0.000412)	(0.0007478)	(0.0003167)	(0.0006187)
	399465	*142557*	*629674*	*213378*
Quartile 3	0.9374742	0.9268354	0.9416599	0.9255651
	(0.0003832)	(0.0006888)	(0.0002955)	(0.0005676)
	399083	*142938*	*629224*	*213838*
Quartile 4	0.9411241	0.9412468	0.9406763	0.9394274
	(0.0003726)	(0.000622)	(0.0002971)	(0.0005194)
	399077	*142937*	*632142*	*210904*
Total	0.917845	0.915936	0.9199943	0.9128311
	(0.0002169)	(0.0003691)	(0.0001707)	(0.0003067)
	1603044	*565034*	*2526307*	*845898*

Note: Standard error of the mean in parenthesis. Number of observations in *italics*.The sample is selected according to the criteria described in Section 3, but we include all wives (not only the low income sample). "Quartile" refers to quartile group with respect to the husband's income.

F Randomization inference test

To begin with, we estimate the following equation on our data:

$$e_{ihkt} = \mathrm{Tr}_{kt} + \mu_t + \mu_k + \mu_h + \mu_{hk} + \mu_{ht} + \delta X_{ihkt} + \upsilon_{ihkt}, \tag{19}$$

where Tr_{kt} is a treatment indicator equal to one if $\mu_k = 1$ and $\mu_t = 1$, $t \in [1997, 2001]$. This specification corresponds to column 4 of Table 1, but solely contains one treatment indicator rather than the full set of interactions. Estimating this equation, we obtain a treatment effect estimate of 0.809 percentage points (with an individual-level clustered standard error of 0.121). If the aggregates of women with and without children are considered to be the relevant clusters in inference, the specification in (19) leaves us with 16 groups: 5 with $\mathrm{Tr} = 1$ and 11 with $\mathrm{Tr} = 0$. We now generate 500 placebo samples by randomly allocating 5 groups into $\mathrm{Tr} = 1$, and we re-estimate model (19) on each placebo sample. All randomization is conducted at the group level. This procedure generates a distribution of placebo estimates, and we can e.g. perform a two-tailed test at the 5 percent level by checking if our observed estimate is larger (lower) than the 97.5 percent (2.5 percent) values. It turned out that the placebo estimates were ranging between −0.863 and 0.786 percentage points. Accordingly, the observed estimate of 0.809 percentage points was larger than 100 percent of the placebo estimates. We also did the same exercise without controls, corresponding to column 1 of Table 1. Then the observed treatment effect estimate amounted to 0.650 percentage points, and we obtained 7 out of 500 placebo estimates (1.4 percent) which were larger.

Appendix References

Bhargava, S. and D. Manoli (2015). Psychological frictions and the incomplete take-up of social benefits: Evidence from an IRS field experiment. *American Economic Review* 105(11), 3489–3529.

Boverket (2006). Bostadsbidrag, ett rättvist bostadsstöd för barnen? - långsiktiga effekter av 1990-talets besparingar. Diarienummer: 212-3795/2006.

Chetty, R., A. Guren, D. Manoli, and A.Weber (2012). Does indivisible labor explain the difference between micro and macro elasticities? A meta-analysis of extensive margin elasticities. In NBER Macroeconomics Annual 2012, Volume 27, NBER Chapters, pp. 1–56. National Bureau of Economic Research, Inc.

Cogan, J. F. (1981). Fixed costs and labor supply. *Econometrica* 49(4), pp. 945–963.

Hausman, J. (1980). The effects of wages, taxes and fixed costs on women's labor force participation. *Journal of Public Economics* 14, 61–194.