The effect of children's disability on the labour supply of mothers in Hungary*

Katalin Bördős, Márton Csillag and Ágota Scharle†

Abstract

We evaluate the effect of child health impairment on maternal labour supply in Hungary using data from the Hungarian Labour Force Survey (2008-2012), where we can identify the presence of children with long-term illness based on claiming extended child care allowance. We find that the employment probability of mothers receiving child care allowance for long-term ill children is 40 percentage points lower than that of similar mothers raising healthy children. These results are similar across a large variety of estimation methods based on the ignorability assumption, and we do not find evidence of sensitivity due to the presence of selection on unobservables.

* We would like to thank Zsombor Cseres-Gergely and Péter Galasi for comments on an earlier version of this paper. We gratefully acknowledge the financial support of the Hungarian National Science Foundation (OTKA Grant No. 101925). All remaining errors are our own.
† Budapest Institute for Policy Analysis. Corresponding author: Márton Csillag; marton.csillag@budapestinstitute.eu

1 Introduction

Labour economists and sociologist have devoted considerable attention to disentangling the effect of the presence of children on maternal labour market status, but we know relatively little about how children's health influences their mother's employment. As the time commitment involved in caring for a child in poor health might inhibit the mother's ability to participate in the labour market, it is of primary importance to quantify the relationship between child disability and maternal labour supply to fully assess the economic consequences of raising health impaired children. The formulation of a sensible child-disability policy critically hinges on identifying the obstacles to maternal work activity.

In this paper, we investigate how poor health condition of a 4-9 year-old child affects maternal labour supply in Hungary, which is one of the countries where the employment rate of women with children is the lowest in the European Union. In order to estimate the effects, we used data from the Hungarian Labour Force Survey and applied propensity score matching and regression methods combined with inverse probability weighting. In the absence of direct data on child health, we identify children with a disability or permanent sickness by exploiting rules of child care allowance. Our results suggest that claiming extended child care allowance for long-term ill children reduces mothers' employment probability by around 40 percentage points.

The paper is structured as follows. First, we provide a short theoretical background on the topic and outline the possible empirical strategy which can be used to estimate the effects. Next we a review the existing empirical literature and results. Then, we briefly summarise the methodology we used and give a description of our data. Finally, we present our results and conclude with a discussion of our results.

Parental benefit system in Hungary

The most commonly claimed parental benefit in Hungary is called 'Child care home allowance' (Gyermekgondozási segély - GYES), which is regulated by Section 20 of Act LXXXIV of 1998. It is a universal and flat-rate benefit paid to one parent (or guardian) until the child reaches the age of 3 if the child is in good health, or until the age of 10 if the child is disabled or suffers from a long-term illness.[1] For twins, it is to be paid until the end of the year when the children reach schooling age. Its amount is the prevailing minimum old-age pension and is independent from the number of children (except in the case of twins). It is also worth noting that mothers are allowed to work full-time and also can claim child care allowance (between the first and the third birthday of their child)[2].

A government decree (ESzCsM 5/2003. (II. 19.)) by the Ministry of Health, Social and Family Affairs in 2003 listed all the health problems and disabilities which entitled the parents to the prolonged maternity benefit. The list included a variety of illnesses and disabilities ranging from asthma and food allergies to epilepsy, various chronic diseases, mental disorders, sensory and mobility limitations. In 2012, a new decree (EMMI 3/2013. (I. 7.)) was issued which tightened the conditions of claiming the

[1] Note that the clock is „reset" after the birth of each additional child, so if the spacing between the births of two children is less than three years, the mother can claim maternity benefits for at most 6 years.

[2] Before 2009 and between 2011 and 2013, the parent receiving the child home care allowance was only allowed to work a maximum of 30 hours per week between the child's first and third birthday. Parents who claimed the benefit after a long-term ill or disabled child were allowed to work full-time after the first birthday of the child during the whole period.

benefit: it reduced the list of illnesses and declared eligibility only in the most severe stages of some of the illnesses. This latter decision has been the much debated.

2 Theory & literature review

Basic theory on the labour supply of mothers with disabled or long-term ill children

Based on basic labour economic theory, one would expect some contradicting effects of poor child health on maternal labour supply. First, the disability or long-standing sickness of the child would reduce labour supply because of higher caring needs of the child. Second, the prolonged maternity leave benefit duration reduces the incentives to work via income effect as non-labour income is higher. Finally, there might be a long-term income effect because of the increased financial costs of caring duties and medication or special equipment: the family income expendable for consumption is reduced, which can lead to an increase in labour supply.

We expect potential heterogeneous effects and incentives among different subgroups in the population. Between single and married mothers, the elasticity might differ: for single mothers, the family income is usually lower as they do not have a spouse who could work, hence we would expect to see a larger (positive) income effect. However, if there is substitution between women and men in caring for children, then the reduction in labour supply will be smaller for women with a co-habiting partner.

The second source of heterogeneity is the health condition and age of the child. Children with more severe health conditions would need more help in daily activities and caring can be more time-consuming, but the medication costs might also be higher. Younger children might also need more attention than older ones, and there might be more day-care institutions available for school-age children than for children below the age of 7 (e.g., special schools for children with disabilities).

Finally, we reckon that the effect of raising long-term ill children also differs by the education of mothers. Mothers who are less educated and are at the lower end of the income distribution face a lower opportunity cost of staying at home because they have lower wage expectations than women with higher education; hence we expect the short-term income effect of the maternity benefit to be stronger. Also, it is less likely that they could afford a formal carer for the child, whereas high-earner mothers have more opportunity to substitute homecare for formal care.

Results of previous research

There is an increasing number of empirical papers analysing the effects of child health on maternal labour supply in the United States. Salkever (1982a) and Salkever (1982b) used survey data from the 1970s to investigate whether the child's disability affects the labour supply and wage of the mother. He found strong and significant negative effects for the labour supply of married mothers (but no effect on single mothers), especially among the lower-income households and in the case of children with more severe disabilities. Porterfield (2002) used pooled cross-section data from two waves of the Survey of Income and Program Participation and assessed how child disability affects the mother's decision to work part-time, full-time or not at all. She also found heterogeneous effects on employment by the mother's education level and the age of the child (larger effects for mothers of 0-6 year-olds and lower educated mothers), and a stronger negative effect for single mothers; married mothers work in part-time jobs with a higher probability. Other studies (e.g. Lee et al. [2004], Wasi, den Berg, and Buchmueller [2012]) support the evidence of different impacts by subgroups (married mothers versus female household heads, mother's education level, child's age, type of limitation), although the results are somewhat mixed.

Powers (2001) pays attention to the potential two-way causation between children's health status and mothers' work effort, and she uses the type of impairment as instruments for the child's disability status. She also finds a significant negative relationship between child disability and maternal employment, but notes that accounting for the endogeneity of child health, as well as controlling for the health of the mother reduces the impact. Corman, Noonan, and Reichman (2005) also address the problem of endogeneity, but their results do not support the assumption of the endogeneity of child's health. [3] They also found a significant negative effect of bad child health on the mother's labour force participation and hours of work.[4]

Outside the USA, there are fewer studies which focus on this issue. Yamauchi (2012) using longitudinal data from Australia and a household fixed effects specification also finds a negative impact. On the contrary, Gupta, Das, and Singh (2013) in India estimated positive employment effects for married mothers in urban areas, although they also show a reduction in work hours for these same mothers and the impacts disappear in the cases of single mothers and mothers living in rural areas.[5]

To summarise the results and estimation strategies of the papers cited above, in all cases except for Gupta et al. (2013) the authors find significant and negative effects on maternal labour market participation, although it is difficult to compare the volume of the effects because of the different measures on disability or sickness, and the ages of the children examined also differ. They all analyse labour supply separately for single and married mothers, but the results are quite controversial regarding in which case is the impact larger. Most of the papers confirm the existence of heterogeneous effects, especially in the severity or type of health problem of the child but sometimes also the mother's education level or family income. They all emphasize the importance of having a sufficiently large sample for proper estimation, but many of them ignore the possible endogeneity problems.

To our best knowledge, the literature on the labour supply of the parents of disabled or long-term ill children in Hungary is scarce. Studies by Bass (2004a and 2004b) and Krémer and Rozsos (2009) describe the emotional, time and cost burden on families who raise children with severe and permanent disability. There also exist some research on the labour supply of mothers with young children and the impacts of the maternity benefit system in Hungary (see, for example, Köllő and Bálint [2007] who used the Hungarian Labour Force Survey, or Scharle [2007], Nagy [2009] , Szabó-Morvai

[3] The exclusion restrictions include the number of adoption agencies and the presence of a Level III neonatal intensive care unit in the hospital where the baby was delivered, and in the equations of the father's employment, the mother's health.

[4] As far as we know, the only authors who used propensity score matching in this field of study were Busch and Barry (2007), although their focus is slightly different from ours: they investigated the financial, labour market and time burden imposed on the parents of mentally disabled children relative to those raising children with other types of illnesses or disabilities.

[5] . In Europe, Dunkelberg and Spieß (2007) used panel data from Germany, but the sample they used had mothers with very young children (0-2 year-olds).

[2013]), but these papers do not address the special circumstances of mothers raising children with permanent health problems.

3 Methods

Empirical specification issues and potential sources of bias

The problem of selection usually arises in observational studies where the assignment to treatment (that is, the variable of interest) is not random. In evaluation studies, where the researcher is interested in the effect of the treatment on the outcome variable, the goal is to estimate two objects of interest. First, the average treatment effect (ATE) can be defined as $= E[Y(1) - Y(0)]$, where E[.] denotes expected value, Y(1) is the outcome (e.g., labour force participation or wage) if the treatment happened (e.g., if a mother has a disabled or steadily ill child) and Y(0) is the outcome for the same individual in the absence of treatment. Second, the average treatment effect on the treated (TOT) is given by $= E(Y \mid T = 1) = E[Y(1) \mid T = 1) - E[Y(0) \mid T = 1]$, which is the average treatment effect using the distribution of the X variables for the treated units only. While the first quantity represents the average effect of the treatment in the population, the second stands for the effect of the treatment on those who actually received it. The problem of the counterfactual means that between Y(1) and Y(0), only one outcome can be observed as an individual has either received treatment or has not. Because of that, a comparable group called control group is needed to estimate the impact. If assignment to treatment is random (every family has the same chance of having a child with health problems), then ATE=TOT and both can be estimated by a simple comparison of the means, $[Y(1) \mid T = 1] - E[Y(0) \mid T = 0]$.

With respect to the estimated impacts, bias can emerge if having a sick or disabled child is not random in the population. If the assignment to treatment is not random, $E[Y(0) \mid T = 0]$ might not be equal of the counterfactual outcome of the treated, $[Y(0) \mid T = 1]$. Some factors (unobserved to the researcher) might simultaneously affect the probability of having a long-term ill or disabled child and labour market chances and labour supply. If these unobserved factors correlate positively with the probability of having a sick child and negatively with prospects in the labour market, then we would overestimate the impact of child health on maternal labour supply. For example, mothers who are less health-conscious (e.g. those who smoke, have alcohol or drug problems, or lower sense of responsibility etc.) might bear sick or disabled children more likely but would have worse chances at the labour market even without a sick child because of these unobserved factors. Those who have strong preferences for family life and thus have weaker ties to the labour market might be less prone to abortion or giving up the child to adoption if it is discovered that the (expected) child is disabled. On the other hand, it is also possible that some unobserved factors correlate negatively with both the prevalence of child disability and labour market prospects. In families with less advantageous backgrounds, a child born with a serious illness might have less chances to survive in her first few years, and as a result, those who raise disabled or sick children between the age of 4 and 9 are families who are less deprived and thus have better chances in the labour market. Corman, Noonan, and Reichman (2005) point out that the correlation of poor child health and the mother's labour market participation can be positive if mothers with strong preferences for work also work more during pregnancy and may experience higher level of prenatal stress and as a consequence, the child is born with worse health condition. [6]

Empirical methods and estimation strategy

In our analysis, we rely on methods based on ignorability of treatment: propensity score matching, regression adjustment and regression adjustment combined with inverse probability weighting. We briefly present the main ideas and issues of these methods in turn.

The application of all three estimation techniques requires two assumptions:

- the *unconfoundedness criterion*: it states that the potential outcomes of the treated and controls do not depend on treatment after conditioning on the covariates: $Y(0), Y(1) \perp T \mid X$. This implies that selection into treatment must only depend on observable characteristics X. For example, if unobserved factors such as lifestyle or income influence the probability of giving birth to a child with ill health (or the observed take-up decision), this unconfoundedness assumption is violated and a hidden bias emerges. Although this assumption cannot be directly tested, there are some possibilities to check the sensitivity of the estimated results with respect to deviations of these assumptions.

- the *common support criterion*: for each value of X, the probability of receiving treatment must be positive and less than one. A region of common support is defined where distributions of the propensity score for treated and comparison units overlap. Observations which do not fall into the range of common support must be discarded. Having a weak common support (dropping a substantial share of the observations) may cause biased estimates in the average treatment effect if the observations discarded are significantly different from the ones in the remaining sample.

In our study, we focus on estimating the average treatment effect for the treated (TOT). This is due to two reasons. First, from a policy perspective, estimating the burden imposed by raising long-term ill children on those with such children rather than on all families with children is more relevant.[8] Second, - as shown in Wooldridge (2010) for example – the treatment effect on the treated can be consistently

[6] In his study on the circumstances of families with severely disabled children in Hungary, Bass (2004a) states that, before the birth of the disabled child, the socio-economic status of these families does not differ from that of other families and their circumstances only begin to worsen after the birth of the disabled child. The impoverishment of these families usually begins with the mother losing or giving up her job because of caring duties and escalates further with the increasing medical and caring costs. This suggests that the occurrence of a severely disabled child is random in the population. However, Bass's statements only refer to families with severely disabled children, and it is not known whether the birth of a steadily sick child can be associated with socio-economic status.

[8] Consider a simple example to highlight the difference between ATE and TOT. Suppose that the prevalence of a certain type of illness is higher in micro-regions where incidentally labour market conditions are also poor in the sense that the female employment rate is low. Then, if in these regions the effect of raising long-term ill children is smaller than in regions with relatively good labour market conditions, the TOT will be smaller than the ATE. If we were interested in the potential gains from providing families with ill children additional nursing facilities, we would overestimate the effect of this policy change if we were to base our conclusions on the ATE.

estimated under relatively weaker conditions than the average treatment effect (ATE). Instead of having to assume full conditional independence of outcomes from treatment status (conditional on the observable variables), identification of TOT only requires mean independence in the base (untreated) state. In our case this means that if the children of those mothers currently raising long-term ill children were healthy, then their employment outcomes would be similar to that of mothers with the same observable characteristics but currently raising healthy children. By contrast, we can allow for the losses in terms of employment probability to be correlated with the presence of long-term ill children.[9]

Propensity score matching (PSM) - developed by Rosenbaum and Rubin (1983) - is a method for causal inference in observational (non-experimental) studies which ensures the similarity of the control group based on observable characteristics. In the first step, one has to predict the probability of receiving treatment (in the case of binary treatment, it is usually done via probit or logit regression) and then, each observation in the treatment group is matched to one or more observations in the control group based on the estimated probabilities (the propensity score). The propensity score is thus defined as the conditional probability of treatment, given the characteristics X: $P(X) = \Pr(T = 1 \mid X)$. Under these assumptions, $E(Y(0)|P(X), T = 1) = E(Y(0)|(P(X), T = 0)$, so PSM solves the problem of estimating the counterfactuals by imputing the employment probability of mothers of healthy children to mothers raising long-term ill children with the same propensity score. In order to do that, first we need to select one (or more) suitable "matches" from among the control group to members of the treated group based on the similarity of their propensity scores. In order to ensure the similarity of treated and control units in terms of observable characteristics X, balancing tests need to be done to check whether the means of the covariates (and the propensity score) are the same for the treated and the control observations with a similar propensity score. Finally, the average treatment effect on the treated (TOT) can be estimated by the difference in outcome means over the common support, weighting control observations by the propensity score distribution of the treated units.

The second method we used is (parametric) regression adjustment. With this method, the TOT is estimated as follows: a regression model on the outcome variable (the mother's employment status) is run separately for the treated and the control units, and the predicted outcomes are computed for each observation by each treatment level. Then the difference in the averages of the two different potential outcomes is calculated. The difference in the predicted values over the subsample of treated individuals represents the average treatment effect on the treated, the TOT:

$$TOT_{RA} = \left(\sum_{i=1}^{N} T_i\right)^{-1} \left\{\sum_{i=1}^{N} T_i * [\, E(Y \,\widehat{|X, T} = 1) - E(Y \,\widehat{|X, T} = 0)]\right\},$$

where $E(Y \,\widehat{|X, T} = 1)$ and $E(Y \,\widehat{|X, T} = 0)$ denote the predicted values (in our case, probabilities of employment) in case of treatment and in the absence of treatment, respectively.

This regression adjustment can be combined with propensity score methods. This methodology, the inverse-probability weighted regression adjustment (IPWRA) uses the inverse of the predicted

[9] We need to point out that this is an important advantage in our case. Please keep in mind that the presence of long-term ill children is based on the receipt of extended child care allowance, and that there is a large heterogeneity in health conditions that potentially entitle families to this benefit. It might be the case that it is those families where the child suffers from a more severe health condition – which limits the employment prospects of mothers more – are the ones who apply for the extended allowance.

probabilities (obtained from the propensity score regression) as weights when performing regression adjustment. Inverse probability weighting, as proposed by Hirano and Imbens (2001), can be applied as follows:

- when estimating TOT, treated observations are given the weight of 1 and control observations the weight of $\frac{p(X)}{1-p(X)}$ (where $p(X)$ is the estimated propensity score);

IPWRA is called double-robust estimator as it allows for potential misspecifications in either the propensity model or the regression model predicting the two potential outcomes. If one of the two models is specified incorrectly, the IPWRA estimator still produces a consistent estimate of the treatment effect.[10]

We performed the regression adjustment and the inverse-probability weighted regression adjustment only for the subsample of married and co-habiting women, since the sample size for single mothers was too small for obtaining reliable regression coefficients.

For the outcome equations, we estimated probit, logit and linear probability regression models with the following covariates as independent variables:

- age and square age of the mother
- education level (four categories)
- county of residence
- year of the observation
- the natural logarithm of the quarterly unemployment rate in the micro-region
- the natural logarithm of the taxable income divided by the number of 15-61 years old in settlement (in HUF)
- a dummy whether the father of the child was employed
- and dummies indicating the father's education level.

There is no consensus in the literature on matching methods about the proper usage of sampling weights in case the data are based on a complex survey sampling design. DuGoff, Schuler, and Stuart (2014) suggest using a combination of sampling weights and inverse probability weighting when running regressions on survey data, in order to ensure that the conclusions can be generalised to the population studied. We experimented with this approach, but as the results we very similar to the unweighted case, we will not present them to save space.

[10] The treatment effects were estimated using Stata 13's 'teffects' commands.

4 Data

The Hungarian Labour Force Survey

We relied on the Hungarian version of the Labour Force Survey (LFS) which is administered quarterly by the Hungarian Central Statistical Office (KSH). The LFS contains data on about 22-34 thousand households with about 60-80 thousand individuals every quarter, and the questionnaire is filled by every member in the household (although some questions only refer to individuals between the age of 15-74). The LFS has a rotating panel design where individuals are tracked for six quarters and in each wave, one-sixth of the participating households are dropped from the sample. We used data from the first quarter of 2008 to the first quarter of 2013 and appended the 21 waves into one dataset. As the treatment variable (whether someone in the household received maternal benefit) and most of the independent variables we used did not vary much over the different waves within households, we chose a random sample where every household in the appended dataset was only represented by one random wave[11]. Our sample thus had a pooled cross section design where each household and individual had one observation in time.[12]

The LFS does not have any data on health but it includes questions about the types of welfare benefits received, hence we know which individuals receive maternity leave benefit. As we also know the ages of children living in the same household and, for of a healthy child, maternity leave is only awarded to the parent until the child's third birthday, we assume that those who receive maternity leave and have children between the age of 3 and 9 (and no children younger than 3) have a child with a severe sickness or disability.

To select our final sample, we had to drop households which

- did not include children between the age of 4 and 9[13] (since this is the only cohort for which we could tell whether the child had a long-term sickness or disability based on maternity benefit receipt)
- included children between the age of 0 and 3 (since their parents are entitled to maternity leave regardless of their child's health condition)
- had more than two children under the age of 18 (since their parents are entitled to a different kind of maternity benefit, 'Child raising support' [GYET[14]])

[11] We repeated our main estimations by choosing a different wave for each household randomly to see if our results differ. Choosing a different random sample did not confound with any of our estimates and the results remained the same.

[12] We also included characteristics of the micro-region of residence in our analysis. Regional data was merged into our dataset using KSH's Regional Statistics Database System (T-STAR) and in the case of quarterly unemployment rates per micro-region, data from the National Labour Office (NMH).

[13] We had to drop households with three-year old children as well since a substantially higher proportion of their parents claimed that they were receiving maternity leave as opposed to parents with children of 4-9 years old. One possible reason for that is that sometimes the benefit is paid with a few months delay, after the child's third birthday.

[14] GYET imposes restrictions on the number of working hours when received so mothers with more than two children were dropped.

- included twins (since their parents are also entitled to maternity leave for a longer period).

We also had to identify the mother and father of the child(ren) living in each of the households. In most of the cases[15] we found the parents with the help of two variables, 'relationship to household head' and 'family status'. Finally we trimmed our sample by excluding families where the mother was less than 23 or more than 47 years old and households with more than 9 people.

We need to have some remarks about the quality of the data we used. Since we do not have direct information on child health and we can only observe whether a family received maternal benefit or not, there might be measurement errors caused by the take-up rate of the benefit not being complete. If a parent does not take up the benefit even though she is eligible (i.e., she has a child with a long-term illness or disability) - possibly because the administrative costs (filling the application form and sending the medical certifications on the illness) and the "social stigma" associated with welfare benefits might outweigh the benefits, or simply due to lack of information - the comparison group might include observations who should be instead in the treatment group. This can cause our estimations of the effect to be biased either upwards or downwards, depending on the characteristics of these individuals who do not take up the benefit. On the one hand, if higher-income parents who have better chances in the labour market do not claim the benefit because they do not stand in need of it, we overestimate the effect. On the other hand, if parents living in deprived areas do not know of the opportunity to receive the benefit after their sick child for a longer period, we might underestimate the effect, as in this case parents with particularly bleak labour market opportunities "contaminate" the comparison group. Unfortunately, there is no other representative survey in Hungary available to us where child health can be directly observed and has a sample large enough for our estimates to be efficient.

Finally, the LFS does not contain information about household income and earnings, nor the health condition of adults. The first lacuna means that we are not able to separate out income effects and the childcare supply need effects. Concerning the latter, we had data on the illnesses and health limitations of the respondents between the ages of 15-64 in one wave, as the 2012q2 ad-hoc questionnaire contained questions about health. Although the sample is quite small, we checked whether the parents of ill-health and healthy children differ in terms of health condition and distribution of health problems, but we did not find a significant dissimilarity across the two groups.

[15] In some cases, mostly in large, intergenerational households, the identity of the mother or father was not clean-cut as the family status of many of the family members was labelled 'other relative'. In these cases, we identified the mother as a woman who is at least 14 years older but at most 45 years older than the eldest child. Households where no such woman was found were discarded from the sample. We checked the accuracy of our identification method using the 2010q2 ad-hoc questionnaire which contained explicit information on the identity of the parents in each family. In more than 99,9% of the cases we identified the mothers correctly and measurement error was the same across treatment and control families.

Descriptive statistics

The final sample contained information on 9 116 families (7 457 where the mother was either married or was living with a partner, and 1 659 with a single mother). Both in the cases of two-parent and single-mother households the proportion of families who received maternal leave benefit was around 5% (0.0495 and 0.0494 respectively). Summary statistics of the relevant variables are provided in Table 1 for two-parent households and Table 2 for single-mother households. Based on the last columns of the tables, one can see that according to t-tests on the equality of the averages, the treated and control families differ in terms of many characteristics, especially in the cases of non-single mothers. Married mothers with a disabled or long-term ill child tend to be slightly younger (but their children are also younger), less educated and live in regions with higher unemployment rate compared to married mothers with healthy children. Also, their partners are less often employed. These summary statistics are quite consistent with other papers' descriptives which notice that the prevalence of disability is somewhat higher in lower-educated households (e.g. Powers [2001]). It is interesting that households which contain a child with a disability or sickness are not significantly larger than those which do not, as one would expect that the higher caring needs of the child would induce the moving in of other relatives who could help taking care of the child.

If we compare the average outcomes for the treated and the untreated in the full sample, we get a difference of -0.5036 for married mothers: 23.31 percent of non-single mothers who receive maternity benefit are employed (n = 369) and the rest of them are either unemployed or inactive; and among those who do not receive the benefit (n = 7 088), 73.67 percent has a job. For single mothers, this difference in the means is even larger, -0.552.

	Treat		Control		p-value of t-test
	Mean	Standard deviation	Mean	Standard deviation	
Mother is employed	0.233	0.423	0.737	0.44	0.000
Mother's age (years)	33.92	5.50	35.32	5.18	0.000
Father is employed	0.77	0.42	0.85	0.36	0.000
Mother's education level					
- max. 8 years	0.27	0.44	0.18	0.38	0.000
- lower secondary school	0.32	0.47	0.27	0.44	0.038
- higher secondary school	0.33	0.47	0.35	0.48	0.402
- university	0.09	0.29	0.21	0.41	0.000
Other adult(s) living in household	0.23	0.42	0.24	0.43	0.567
Number of children (1 or 2)	1.57	0.50	1.56	0.50	0.647
Age of youngest child	5.99	1.82	6.41	1.81	0.000
Age of oldest child (in households with 2 children)	7.33	1.11	7.73	1.15	0.000
Unemployment rate in micro-region (quarterly, %)	11.37	5.48	10.38	5.51	0.001
Taxable income/no. of 15-61 years old in settlement (HUF)	419498.55	520575.45	463607.37	568082.79	0.144
Number of observations	*369*		*7088*		

1. Table: Descriptive statistics on the full sample. Households with two parents.

	Treat		Control		p-value of t-test
	Mean	Standard deviation	Mean	Standard deviation	
Mother is employed	0.134	0.343	0.686	0.464	0.000
Mother's age (years)	34.45	5.91	34.53	5.73	0.905
Mother's education level					
- max. 8 years	0.39	0.49	0.25	0.43	0.004
- lower secondary school	0.29	0.46	0.26	0.44	0.546
- higher secondary school	0.26	0.44	0.33	0.47	0.162
- university	0.06	0.24	0.16	0.37	0.015
Other adult(s) living in household	0.41	0.50	0.43	0.50	0.802
Number of children (1 or 2)	1.35	0.48	1.39	0.49	0.525
Age of youngest child	6.38	1.74	6.58	1.78	0.318
Age of oldest child (in households with 2 children)	8.33	0.52	7.68	1.26	0.211
Unemployment rate in micro-region (quarterly, %)	11.35	5.61	10.52	5.58	0.192
Taxable income/no. of 15-61 years old in settlement (HUF)	387329.31	538440.35	457650.21	562563.95	0.269
Number of observations	*82*		*1577*		

2. *Table: Descriptive statistics on the full sample. Households with single mothers.*

5 Results

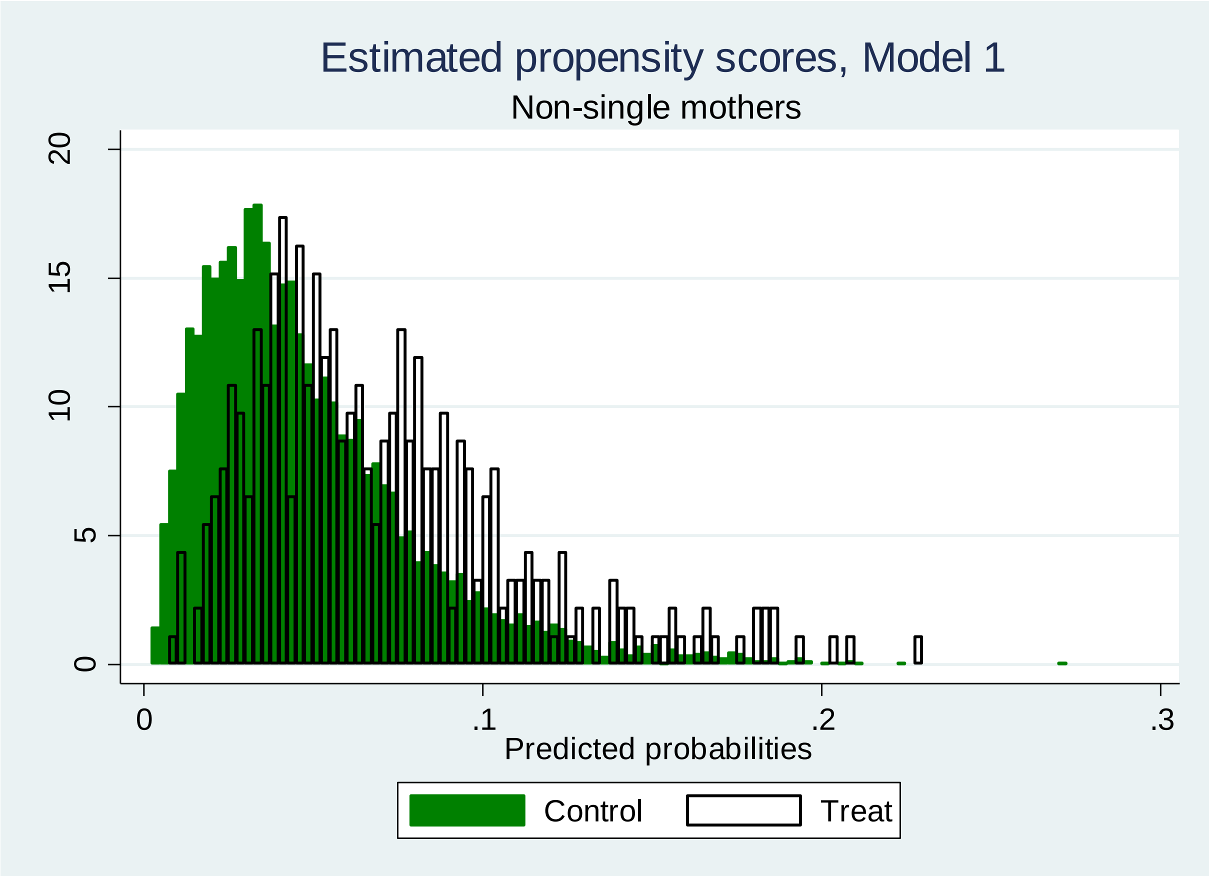
Estimation of the propensity score and matching quality

We estimated three probit models to predict the probability of receiving maternal leave benefit in each household. We ran the regressions separately for households with two parents (where the mother is either married or living with a partner) and for households with a single mother. The models show that higher educated mothers and those who have a working partner tend to receive maternal benefit with a somewhat lower probability. This suggests that in families with less income, either there is an increased chance of having a child with long-term illness or disability, or they take up the benefit more likely because of their less advantageous background. The full list of regressors and the estimated coefficients obtained from the regressions can be seen in Appendix 1.[16]

We defined common support by the minima and maxima approach, meaning we discarded observations whose estimated propensity score was smaller than the minimum and larger than the maximum in the opposite treatment group (Caliendo and Kopeinig [2008]). Depending on the model,

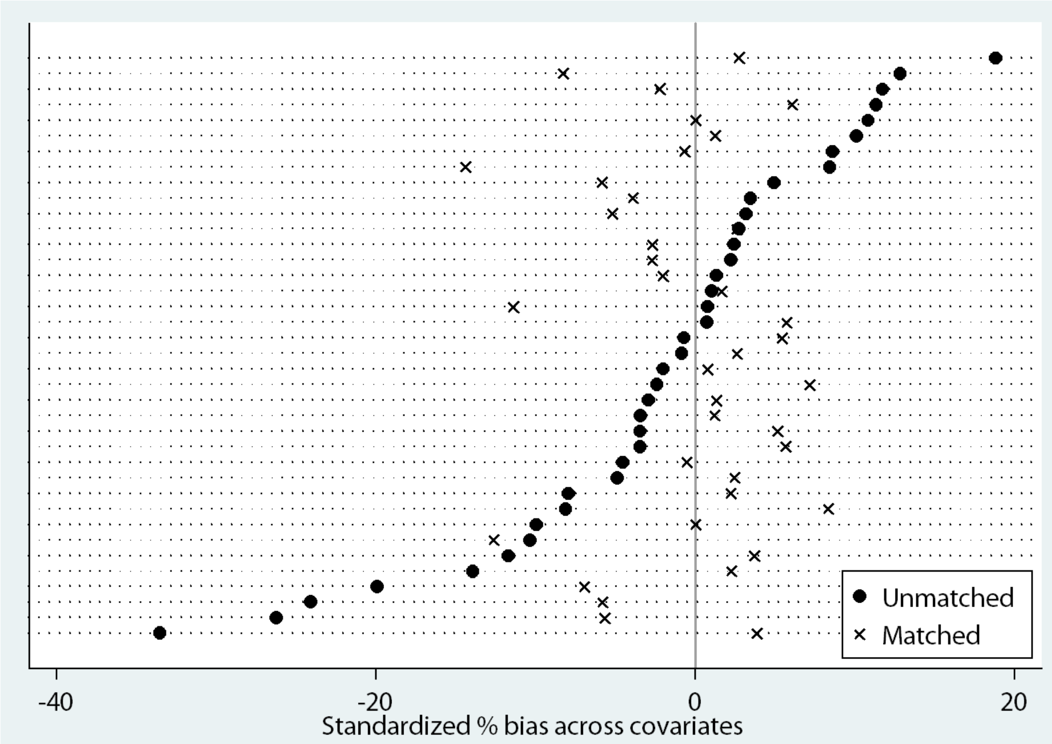
[16] For the estimation of the propensity scores, we used a user-written Stata package called 'psmatch2' developed by Leuven and Sianesi (2014) and 'pscore' by Becker and Ichino (2002) for the balancing tests to ensure that the averages of the independent variables are the same for the treated and control observations in each stratum of the propensity score. We also used 'nnmatch' by Imbens et al. (2004) to do nearest-neighbour matching on the propensity score combined with exact matching on county, the mother's education level, and whether the father was employed (in case of non-single mothers).

the estimated probabilities ranged from 0.0055-0.009 to 0.23-0.25 in the case of treated and 0.002-0.0046 to 0.27-0.29 in the case of control observations. In the case of propensity score model 1 and non-single mothers, the distribution of the estimated propensity scores by treatment level can be seen on Figure 1; the figures on the predicted probabilities of the other models can be found in Appendix 2. As one can see on the figures, all of the treated observations fall into the overlap and only control observations had to be dropped, and even in the cases of control observations, less than 3% of the sample was discarded. This means that the common support criterion was satisfied. It is also visible that the distribution of the propensity scores for the non-treated are more skewed to the right compared to the distribution for the treated.



1. Figure: Distribution of the propensity scores

After having estimated the propensity scores and discarded those (control) observations that fell outside the common support, we performed matching. We used a variety of matching procedures to ensure ourselves that our results were robust to the matching algorithm. These were: one-to-one nearest neighbour matching with and without replacement, five-to-one nearest neighbour matching with replacement, radius matching, kernel matching and stratification matching. We also did some tests to see how much the matching resulted in the reduction of the standardised bias, which is the difference of the sample means in the treated and non-treated (full or matched) sub-samples as a percentage of the square root of the average of the sample variances in the treated and non-treated groups, as defined by Rosenbaum and Rubin (1985)[17]. Overall for the covariates, the mean of the absolute value of the bias before matching was 8.2% and it was reduced to 4.3% after the matching. Figure 2 shows the change in standardised bias for each covariate in Model 1 before matching (the whole sample) and after matching (one-to-one nearest neighbour algorithm). As can be seen, in most covariates, the bias was reduced and is closer to 0 than before matching.[18]



2. Figure: Standardised bias for the covariates before and after the matching, model 1, non-single mothers.

Results of the matching

Using a difference in means type of approach and estimating the treatment effect on the treated (TOT), we got significant results ranging from -0.46 to -0.39 in case of mothers living with a spouse and from -0.55 to -0.42 in case of single mothers, depending on the matching method used. This means that non-single mothers who raise a child with a permanent sickness or disability work about 40 percentage points less likely than mothers with similar characteristics who raise one or two children without serious health problems, and single mothers with disabled children are employed about 50 percentage points less likely than mothers of healthy children. Some of the results (nearest-neighbour

[17] The formula for standardised bias is $= \frac{|\bar{X}_c - \bar{X}_t|}{\sqrt{0.5*[Var(X_c)+Var(X_t)]}}$, where $\bar{X}_c$ is the mean of the variable among the controls, $\bar{X}_t$ the mean for the treated units and Var(.) denotes variance. The standardised bias is calculated before the matching (on the full sample) and after matching (on the matched sample).

[18] While in terms of characteristics of mothers (education, age) and the household (partner's employment status, number and age of children), as well as the income level and unemployment rate in the micro-region of residence, the matching resulted in a substantial reduction in bias, there were a few variables where matching lead to an increase in bias. These were typically, the year/quarter of observation and the county of residence. We are not overly concerned about this increase in bias: (a) all of our observation period is after the crisis affected in Hungary, and since 2008 there was no major change in the labour market prospects; and (b) while county of residence is a significant determinant of employment prospects, we deem that matching on „finer" measures of labour market conditions (unemployment rate, income tax base per head) is more important.

matching, radius matching, kernel matching, stratification matching and exact matching)[19] are shown in Appendix 2.

Results of the regressions

We now turn to the regression adjustment and the inverse-probability weighted regression adjustment results. Most coefficients of the control variables had the expected sign and significance. The mother's age increases the probability of employment with a decreasing return (the coefficient on the square age is negative), and higher education level induces higher chances at the labour market as well. Women living in less deprived regions (higher income tax base per person and lower unemployment rate) are also more likely to be employed. The father's employment increases the chance of the mother being employed; one explanation can be assortative mating.

To save space, we do not report all the estimated coefficients but only the average marginal effects for our treatment variable. These are shown in Table 3. The results of the both the regression adjustment and the inverse-probability weighted regression adjustment are in line with the estimates based on propensity score matching. They show that after accounting for observable characteristics, the average treatment effect on the treated is -43 to -41 percentage points. We can also see that while adjusting for propensity scores slightly decreased the estimated treatment effects, these results are not statistically significantly different from those based on the regression adjustment.

	Average treatment effect on the treated	Standard error	t-statistic	p-value	Lower bound of 95% confidence interval	Upper bound of 95% confidence interval
Regression adjustment	-0.431	0.011	-40.630	0.000	-0.452	-0.410
IPW-regression adjustment, model 1	-0.414	0.010	-41.093	0.000	-0.434	-0.394
IPW-regression adjustment, model 2	-0.415	0.012	-35.781	0.000	-0.437	-0.392
IPW-regression adjustment, model 3	-0.414	0.011	-36.546	0.000	-0.436	-0.392

3*. Table: Treatment effect on the treated, regression-based methods*

[19] We also experimented with radius matching using several different callipers and kernel matching with different bandwidths. These results are not shown in the Appendix as they were very similar to the ones presented.

Heterogeneous effects by subgroups

We examined whether the estimated treatment effects differed by various subgroups of the population. Analysis by the mother's education level, the father's employment status and age of the youngest child was carried out as we expected different treatment effects for lower educated mothers, for those who had a working spouse and for those who had a child below schooling age (age below 7). We present the average treatment effects on the treated over these subgroups in Tables 4-6 . Since all IPRWA estimates were basically the same, we only show here the results for the inverse probabilities from model 1, without sampling weights. Standard errors were calculated using bootstrap. We also calculated the treatment effect for these subgroups based on propensity score methods, but since the results were very similar to those of the regression-based methods, we do not discuss these (please find a selection of the results in Appendix 4).

	Mother's education level	TOT	standard error	z-statistic	p-value	Lower bound of 95% confidence interval	Upper bound of 95% confidence interval
Regression adjustment	Max. 8 years	-0.322	0.017	-18.491	0.000	-0.356	-0.288
	lower secondary school	-0.453	0.016	-28.708	0.000	-0.484	-0.422
	upper secondary school	-0.502	0.017	-30.284	0.000	-0.534	-0.469
	university degree	-0.420	0.038	-10.984	0.000	-0.495	-0.345
IPW-regression adjustment	Max. 8 years	-0.305	0.017	-18.226	0.000	-0.338	-0.272
	lower secondary school	-0.430	0.017	-24.874	0.000	-0.464	-0.396
	upper secondary school	-0.486	0.017	-28.134	0.000	-0.520	-0.452
	university degree	-0.421	0.039	-10.897	0.000	-0.497	-0.346

4. *Table: Treatment effect on the treated, regression-based methods, by mother's education level*

Our estimation results show that there are large differences across education groups in the reduction in employment probability due to the presence of long-term ill children. The treatment effect for mothers with low education (no more than primary school) is only two-thirds as large as for mothers with higher education levels. Among mothers with more than elementary school education, those who have finished upper secondary school the reduction in employment probability is about five percentage points more than among the two other education groups. The result that the effect of the presence of long-term ill children on low-educated mothers' employment probability is smaller is somewhat puzzling. Since these mothers' potential wages (opportunity cost of caring for children) are the lowest, we would expect to see a larger increase in their home production time, and hence a larger reduction in employment probability. There might be several factors that counteract this effect. First, the household income of low-educated mothers is likely much lower (due to positive assortative mating and the low employment probability of their spouses), and the (relative) reduction in long-term income due to presence of long-term ill children is larger, and, as a consequence we will see a larger increase in mothers' labour supply. Second, there might also be differences in benefit claiming behaviour across education groups. If higher-educated families only claim benefits in case their child has serious health problems (due to social stigma attached to welfare), then the estimated effects will be larger for these groups as a result of the higher care needs.

	Father's labour market status	TOT	standard error	z-statistic	p-value	Lower bound of 95% confidence interval	Upper bound of 95% confidence interval
Regression adjustment	not employed	-0.245	0.017	-14.382	0.000	-0.278	-0.211
	employed	-0.486	0.010	-48.338	0.000	-0.506	-0.466
IPW-regression adjustment	not employed	-0.212	0.017	-12.549	0.000	-0.246	-0.179
	employed	-0.474	0.010	-45.837	0.000	-0.494	-0.453

5*. Table: Treatment effect on the treated, regression-based methods, by partner's employment status*

We can also see large differences across mothers by their partners' employment status, as we estimate the treatment effect of raising a long-term ill child to be twice as large for mothers whose partners are employed than for mothers with a non-employed partner. This is consistent with both the notion of specialisation in the market or the household, and with the fact that the (relative) reduction due to the ill-health of the child in household income in households with non-employed fathers is larger. Finally, our results show that having long-term children of kindergarten age has a slightly more negative effect on their mothers' employment probability than that of having older long-term ill children, but this difference is only marginally significant.

	Presence of a child under 7	TOT	standard error	z-statistic	p-value	Lower bound of 95% confidence interval	Upper bound of 95% confidence interval
Regression adjustment	no	-0.412	0.017	-24.529	0.000	-0.445	-0.379
	yes	-0.442	0.013	-34.190	0.000	-0.467	-0.417
IPW-regression adjustment	no	-0.393	0.017	-22.501	0.000	-0.427	-0.359
	yes	-0.426	0.014	-30.141	0.000	-0.454	-0.398

6*. Table: Treatment effect on the treated, regression-based methods, by children's age group*

Sensitivity analysis

The results we have presented estimate the causal effect of claiming extended child care allowance under the crucial assumption of unconfoundedness. The purpose of sensitivity analysis is to ask whether inferences about the effect of raising long-term ill children may be altered by the presence of unobservables affecting both employment probability and claiming extended child care allowance.

In order to perform the sensitivity analysis, we used Becker and Caliendo's (2007) 'mhbounds' program in Stata, which follows the paper of (Aakvik 2001). The basic idea is that if there is a dummy (0/1) unobserved variable affecting both the probability of receiving treatment and the outcome, then a Mantel-Haenszel type test statistic for testing whether the treatment has a significant impact on the outcome can be adjusted for the presence of such unobservables.[20] This adjustment depends on our assumptions about the strength of the association between the unobservable variable and selection into treatment, and the sign of the association between the unobservable variable and the outcome. The logic of the sensitivity analysis is then to look at alternative – plausible – scenarios for these associations, and see whether the test statistic, having adjusted for selection on the unobservable variable produces significant treatment effects.

In our case, the main concern is that unobservables that positively influence selection into treatment are negatively associated with employment outcomes; for example, women with unhealthy lifestyles might be more likely to have children affected by disabilities but might also have lower employment probabilities due to their own poor health. Indeed, when we performed our sensitivity analysis we found that if we assume that the estimated treatment effect is biased downward, it is no longer significant, but only in case where the influence of the unobservable variable is very strong. In other words, as long as the odds of receiving extended child care allowance for two mothers with the same observable variables does not differ by more than a factor of 3 due to an omitted characteristic, the estimated treatment effect is statistically significant.[21] We consider this as evidence that our results are robust deviations from the unconfoundedness assumption.

[20] A further assumption is that selection into treatment follows a logistic distribution.

[21] The results are available from the authors upon request.

6 Conclusion

Our results suggest that the presence of long-term ill children poses a very important burden on the employment of their mothers in Hungary. To illustrate the magnitude, the estimated effect of claiming extended child care allowance for a long-term ill or disabled child on their mothers' employment probability is comparable in size to the difference in employment probability between mothers who only finished elementary education and those who attended college. This conclusion is robust to alternative methods based on the ignorability assumption, and do not seem to be sensitive to "hidden bias" from the presence of unobservables.

Our study however has two weaknesses. First, that we do not observe child health directly, only through the receipt of extended childcare benefits, and second, we have no data on family income. Hence, we are unable to separate out the effect of increased caring needs of long-term ill children from the income effect of receiving childcare benefits. This lack data on is an important omission as (i) we expect to see heterogeneous effects on parental labour supply across children with different health conditions, and (ii) access to health-related variables and family income would be useful to include as more detailed controls in the matching procedures. While to our knowledge, no large-sample dataset exists that would include all such variables, but we consider it fruitful to re-estimate our models using the Hungarian Household Budget Survey as a future robustness analysis, which has detailed information on incomes and expenditures on health-related products.

References

Aakvik, Arild. 2001. “Bounding a Matching Estimator: The Case of a Norwegian Training Program.” *Oxford Bulletin of Economics and Statistics* 63 (1): 115–43. doi:10.1111/1468-0084.00211.

Bass, László. 2004a. “Szüljön Másikat? Súlyosan-Halmozottan Fogyatékos Gyermeket Nevelő Családok Életkörülményei Magyarországon [Should She Just Give Birth to an Other One? Report on the Circumstances of Families with Severely Disabled Children in Hungary.].” In *Jelentés a Súlyosan-Halmozottan Fogyatékos Embereket Nevelő Családok Életkörülményeiről*, 54–91. Fogyatékosságtudományi Tanulmányok VI. Budapest: Kézenfogva Alapítvány.

———. 2004b. “A Fogyatékos Gyermek Ellátása.” *Kapocs* III. (5 (14)).

Becker, Sascha O., and Marco Caliendo. 2007. “Mhbounds - Sensitivity Analysis for Average Treatment Effects”. SSRN Scholarly Paper ID 958699. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=958699.

Becker, Sascha O., and Andrea Ichino. 2002. “Estimation of Average Treatment Effects Based on Propensity Scores.” *Stata Journal* 2 (4): 358–77.

Busch, Susan H., and Colleen L. Barry. 2007. “Mental Health Disorders in Childhood: Assessing the Burden on Families.” *Health Affairs* 26 (4): 1088–95.

Caliendo, Marco, and Sabine Kopeinig. 2008. “Some Practical Guidance for the Implementation of Propensity Score Matching.” *Journal of Economic Surveys* 22 (1): 31–72. doi:10.1111/j.1467-6419.2007.00527.x.

Corman, Hope, Kelly Noonan, and Nancy E. Reichman. 2005. “Mother’s Labor Supply in Fragile Families: The Role of Child Health.” *Eastern Economic Journal* 31 (4): 601–16.

DuGoff, Eva H., Megan Schuler, and Elizabeth A. Stuart. 2014. “Generalizing Observational Study Results: Applying Propensity Score Methods to Complex Surveys.” *Health Services Research* 49 (1): 284–303. doi:10.1111/1475-6773.12090.

Dunkelberg, Annalena, and C. Katharina Spieß. 2007. “The Impact of Child and Maternal Health Indicators on Female Labor Force Participation after Childbirth: Evidence for Germany”. Discussion Papers of DIW Berlin 686. DIW Berlin, German Institute for Economic Research. http://ideas.repec.org/p/diw/diwwpp/dp686.html.

Gupta, Prachi, Upasak Das, and Ashish Singh. 2013. “Child Disability and Maternal Work Participation: New Evidence from India”. Economics Discussion Paper 2013-6. Kiel Institute for the World Economy. http://ideas.repec.org/p/zbw/ifwedp/20136.html.

Hirano, Keisuke, and Guido W. Imbens. 2001. “Estimation of Causal Effects Using Propensity Score Weighting: An Application to Data on Right Heart Catheterization.” *Health Services and Outcomes Research Methodology* 2 (3-4): 259–78. doi:10.1023/A:1020371312283.

Imbens, Guido, Alberto Abadie, David Drukker, and Jane Herr. 2004. “Implementing Matching Estimators for Average Treatment Effects in Stata.” *The STATA Journal* 4 (3): 290–311.

Köllő, János, and Mónika Bálint. 2007. “Gyermeknevelési Támogatások a Munkaerő-Felvételben.” In mimeo. http://www.econ.core.hu/file/download/balintkollo.doc.

Lee, Sunhwa, Gi-Taik Oh, Heidi Hartmann, and Barbara Gault. 2004. *The Impact of Disabilities on Mothers’ Work Participation-Examining Differences Between Single and Married Mothers*. Citeseer.

http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.186.9338&rep=rep1&type=pdf.
Leuven, Edwin, and Barbara Sianesi. 2014. *PSMATCH2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing*. Statistical Software Components. Boston College Department of Economics. http://ideas.repec.org/c/boc/bocode/s432001.html.
Nagy, Beáta. 2009. "A Munkavállalás És Gyermekvállalás Paradoxona – Bevezető Gondolatok." *Szociológiai Szemle* 3: 81–91.
Porterfield, Shirley L. 2002. "Work Choices of Mothers in Families with Children with Disabilities." *Journal of Marriage and Family* 64 (4): 972–81. doi:10.1111/j.1741-3737.2002.00972.x.
Powers, Elizabeth T. 2001. "New Estimates of the Impact of Child Disability on Maternal Employment." *The American Economic Review* 91 (2): 135–39.
Rosenbaum, Paul R., and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70 (1): 41–55. doi:10.1093/biomet/70.1.41.
———. 1985. "Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score." *The American Statistician* 39 (1): 33–38. doi:10.1080/00031305.1985.10479383.
Salkever, David S. 1982a. "Children's Health Problems and Maternal Work Status." *The Journal of Human Resources* 17 (1): 94. doi:10.2307/145526.
———. 1982b. "Children's Health Problems: Implications for Parental Labor Supply and Earnings". NBER Chapters. National Bureau of Economic Research, Inc. http://ideas.repec.org/h/nbr/nberch/6551.html.
Scharle, Ágota. 2007. "The Effect of Welfare Provisions on Female Labour Supply in Central and Eastern Europe." *Journal of Comparative Policy Analysis* 9 (June): 157–174(18). doi:doi:10.1080/13876980701311596.
Szabó-Morvai, Ágnes. 2013. "Who Benefits from Child Benefits? The Labor Supply Effects of Maternal Cash Benefit." https://iweb.cerge-ei.cz/pdf/gdn/RRCXII_25_paper_01.pdf.
Wasi, Nada, Bernard van den Berg, and Thomas C. Buchmueller. 2012. "Heterogeneous Effects of Child Disability on Maternal Labor Supply: Evidence from the 2000 US Census." *Labour Economics* 19 (1): 139–54. doi:10.1016/j.labeco.2011.09.008.
Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, Mass.: MIT Press.
Yamauchi, Chikako. 2012. "Children's Health and Parental Labour Supply." *Economic Record* 88 (281): 195–213. doi:10.1111/j.1475-4932.2012.00794.

Appendix

Appendix 1

A1.a Probit regressions for the estimation of the propensity score (2-parent households)

	Model 1	Model 2	Model 3
Mother’s education level (reference: max. 8 years)			
lower secondary school	-0.078	-0.073	-0.081
	(0.071)	(0.071)	(0.073)
higher secondary school	-0.188	-0.178	-0.180
	(0.072)***	(0.071)**	(0.073)**
university degree	-0.529	-0.521	-0.507
	(0.097)***	(0.096)***	(0.098)***
Mother's age	-0.007	-0.008	
	(0.005)	(0.005)	
Child 4-5 years old (dummy)	0.090		
	(0.188)		
Child 6-7 years old (dummy)	0.050		
	(0.189)		
Child 8-9 years old (dummy)	-0.046		
	(0.187)		
Child 10-18 years old (dummy)	-0.389		
	(0.184)**		
No. of children	0.285		
	(0.181)		
No. of 18-64 year old adults in household	-0.008		
	(0.006)		
Base of income tax/no. of people in settlement	-0.000		
	(0.000)		
Father is employed (dummy)	-0.146	-0.145	-0.154
	(0.067)**	(0.067)**	(0.067)**
Only child, 4-6 years old (dummy)		0.132	0.131
		(0.197)	(0.199)
Only child, 7-9 years old (dummy)		0.074	0.074
		(0.198)	(0.200)
1 child 4-6 years old, 1 child 7-9 years old (dummy)		0.387	0.386
		(0.199)*	(0.201)*
1 child 7-9 years old, 1 child 10-18 years old (dummy)		-0.076	-0.065
		(0.199)	(0.201)
1 child 4-6 years old, 1 child 10-18 years old (dummy)		0.061	0.075
		(0.203)	(0.205)
2 children, 4-6 years old (dummy)		0.667	0.667
		(0.217)***	(0.219)***
Log of unemployment rate in micro-region		0.071	

(quarterly)			
		(0.079)	
Mother's age category dummies (reference: 23-26)			
27-31			0.094
			(0.117)
32-36			-0.019
			(0.117)
37-41			-0.116
			(0.126)
42-47			-0.070
			(0.135)
Other 18-64 old people in household (dummy)			-0.042
			(0.063)
Unemployment rate in micro-region (quarterly)			0.009
			(0.008)
Constant	-1.556	-1.399	-1.905
	(0.240)***	(0.337)***	(0.277)***
Region dummies	√	√	
County dummies			√
Year and quarter dummies separately		√	√
Year*quarter dummies in interaction	√		
Pseudo R2	0.0517	0.0462	0.0578
N	7,457	7,457	7,457

* p<0.1; ** p<0.05; *** p<0.01

Households with two parents

Probit regression coefficients; standard errors in parentheses

A1.b Probit regressions for the estimation of the propensity score (Single-mother households)

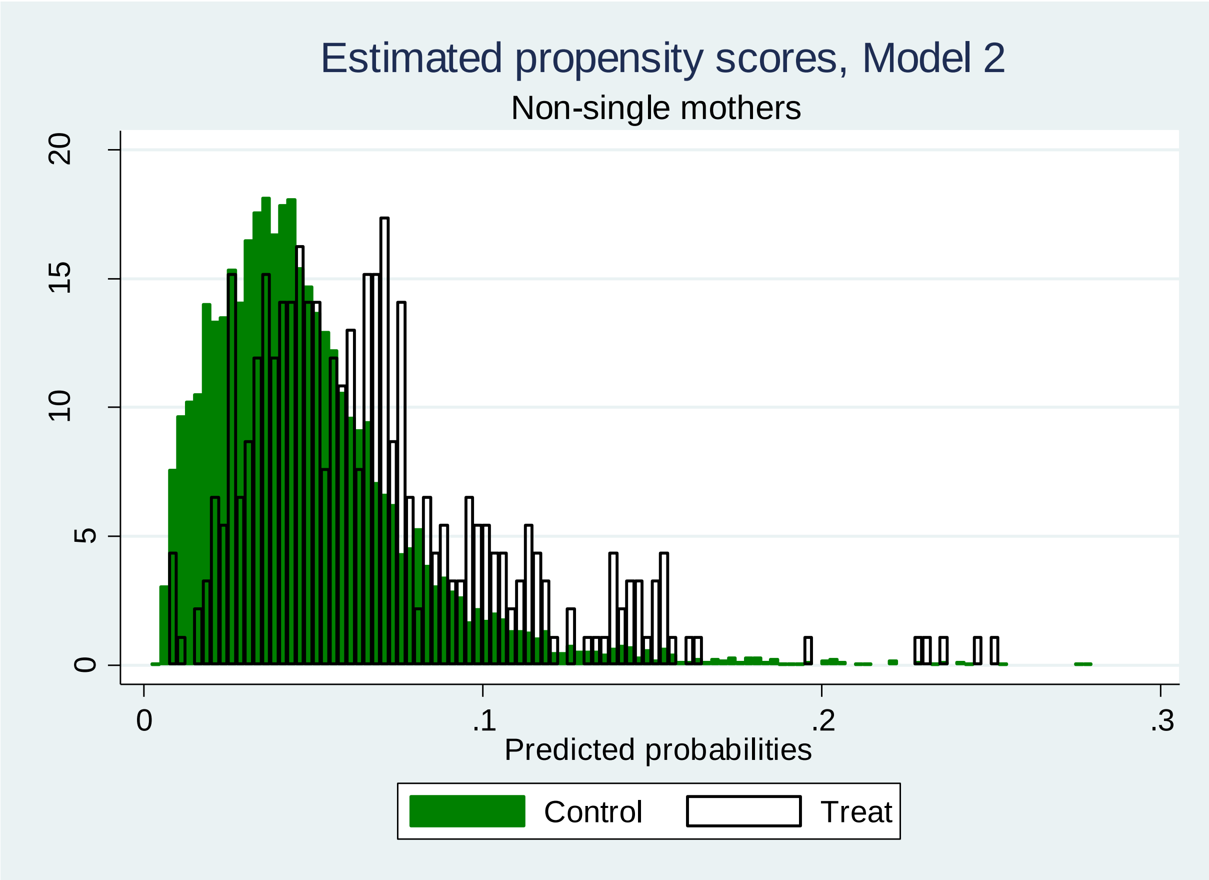
	Model 1	Model 2	Model 3
Mother’s education level (reference: max. 8 years)	-0.228	-0.189	-0.247
lower secondary school	-0.228	-0.189	-0.247
	(0.145)	(0.140)	(0.149)*
higher secondary school	-0.416	-0.358	-0.396
	(0.146)***	(0.142)**	(0.148)***
university degree	-0.798	-0.683	-0.717
	(0.229)***	(0.220)***	(0.230)***
Mother's age	0.004	0.004	
	(0.010)	(0.010)	
Child 4-5 years old (dummy)	3.743		
	(99.456)		
Child 6-7 years old (dummy)	3.658		
	(99.456)		
Child 8-9 years old (dummy)	3.555		
	(99.456)		
Child 10-18 years old (dummy)	3.858		
	(99.456)		
No. of children	-3.860		
	(99.456)		
No. of 18-64 year old adults in household	-0.019		
	(0.012)*		
Base of income tax/no. of people in settlement	0.000		
	(0.000)		
Only child, 4-6 years old (dummy)		-0.375	-0.328
		(0.331)	(0.346)
Only child, 7-9 years old (dummy)		-0.489	-0.483
		(0.332)	(0.347)
1 child 4-6 years old, 1 child 7-9 years old (dummy)		-0.858	-0.850
		(0.407)**	(0.421)**
1 child 7-9 years old, 1 child 10-18 years old (dummy)		-0.499	-0.499
		(0.341)	(0.358)
1 child 4-6 years old, 1 child 10-18 years old (dummy)		-0.429	-0.413
		(0.365)	(0.381)
Log of unemployment rate in micro-region (quarterly)		0.133	
		(0.170)	
Mother’s age category dummies (reference: 23-26)			
27-31			0.521
			(0.255)**
32-36			0.404
			(0.260)
37-41			0.333
			(0.271)

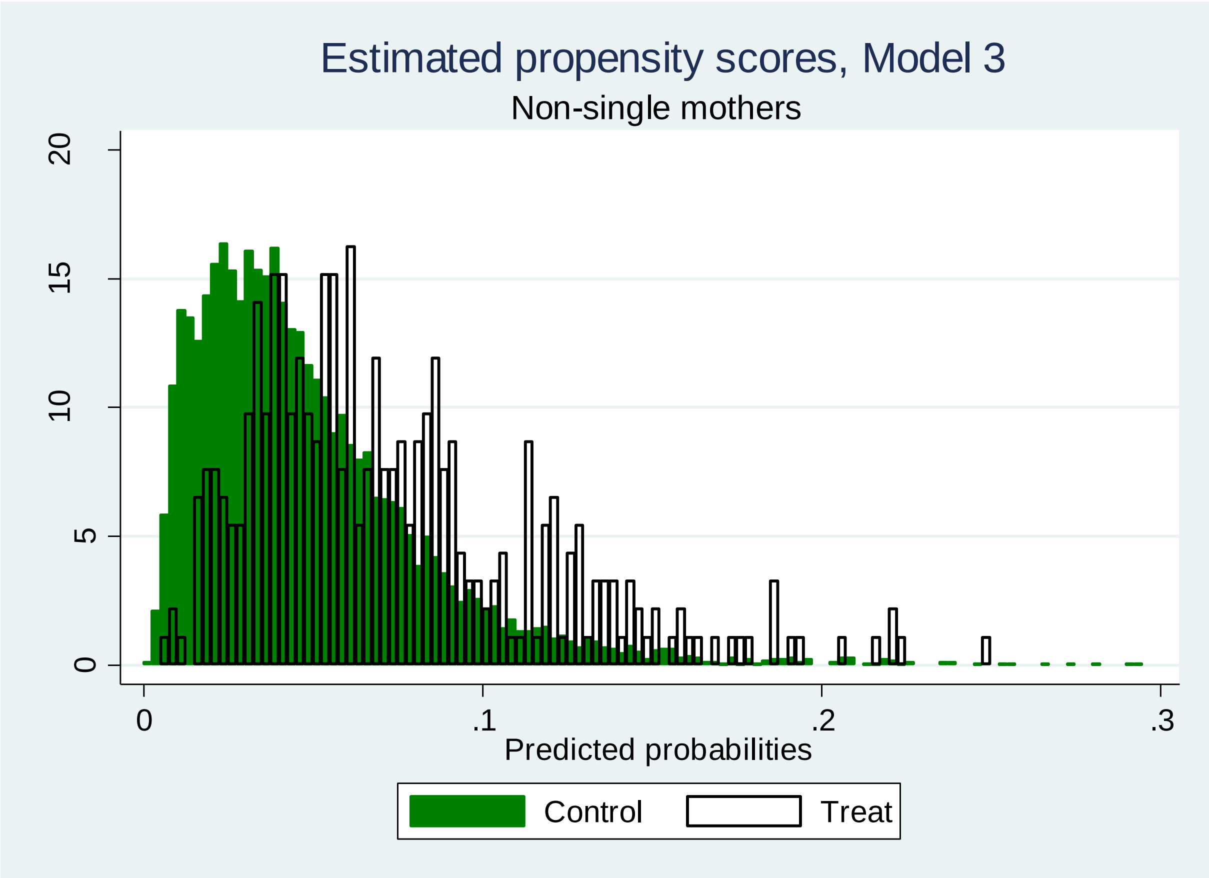
42-47			0.452
			(0.279)
Other 18-64 old people in household (dummy)			-0.126
			(0.118)
Unemployment rate in micro-region (quarterly)			0.015
			(0.017)
Constant	-1.141	-1.251	-1.754
	(0.491)**	(0.634)**	(0.516)***
Region dummies	√	√	
County dummies			√
Year and quarter dummies separately		√	√
Year*quarter dummies in interaction	√		
Pseudo R2	0.0871	0.0544	0.0875
N	1,585	1,621	1,621

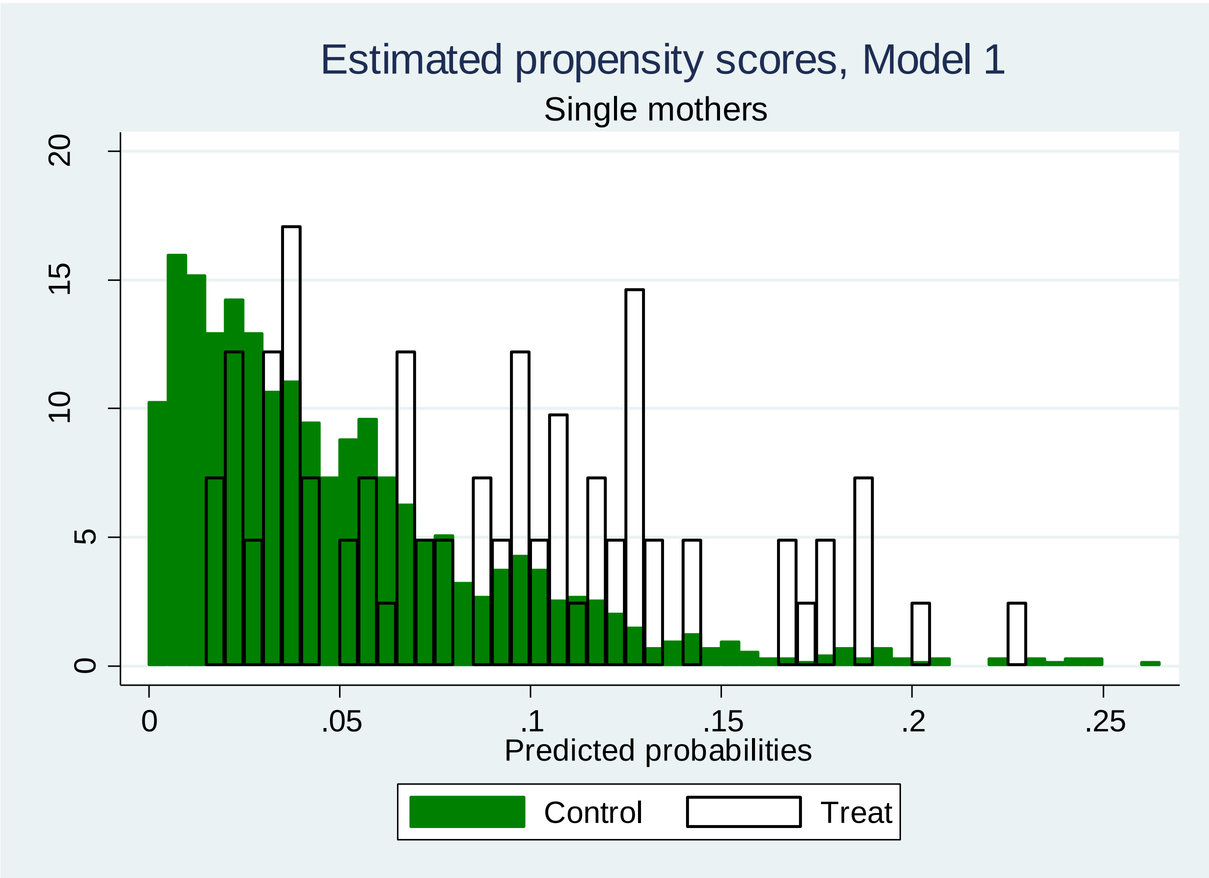
* *p*<0.1; ** *p*<0.05; *** *p*<0.01
Households with a single mother
Probit regression coefficients; standard errors in parentheses

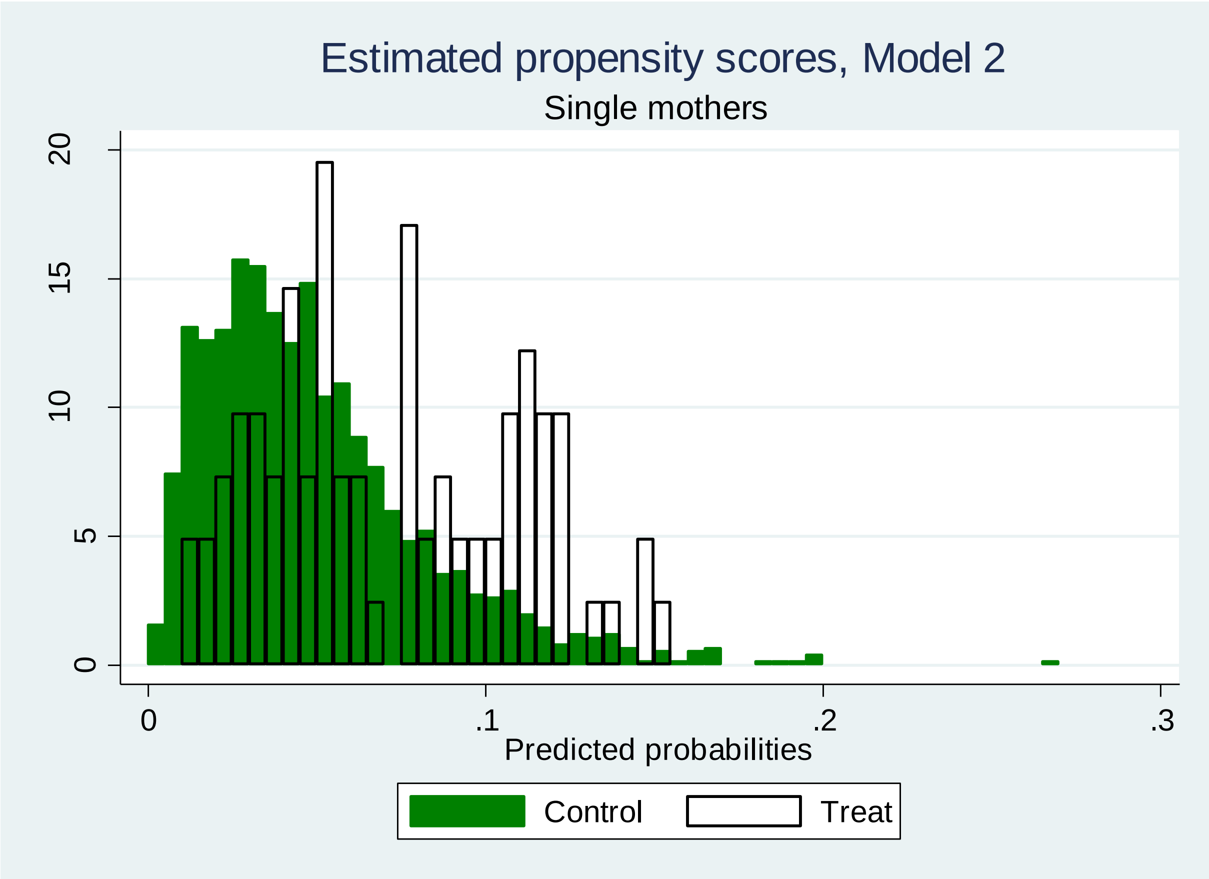
Appendix 2

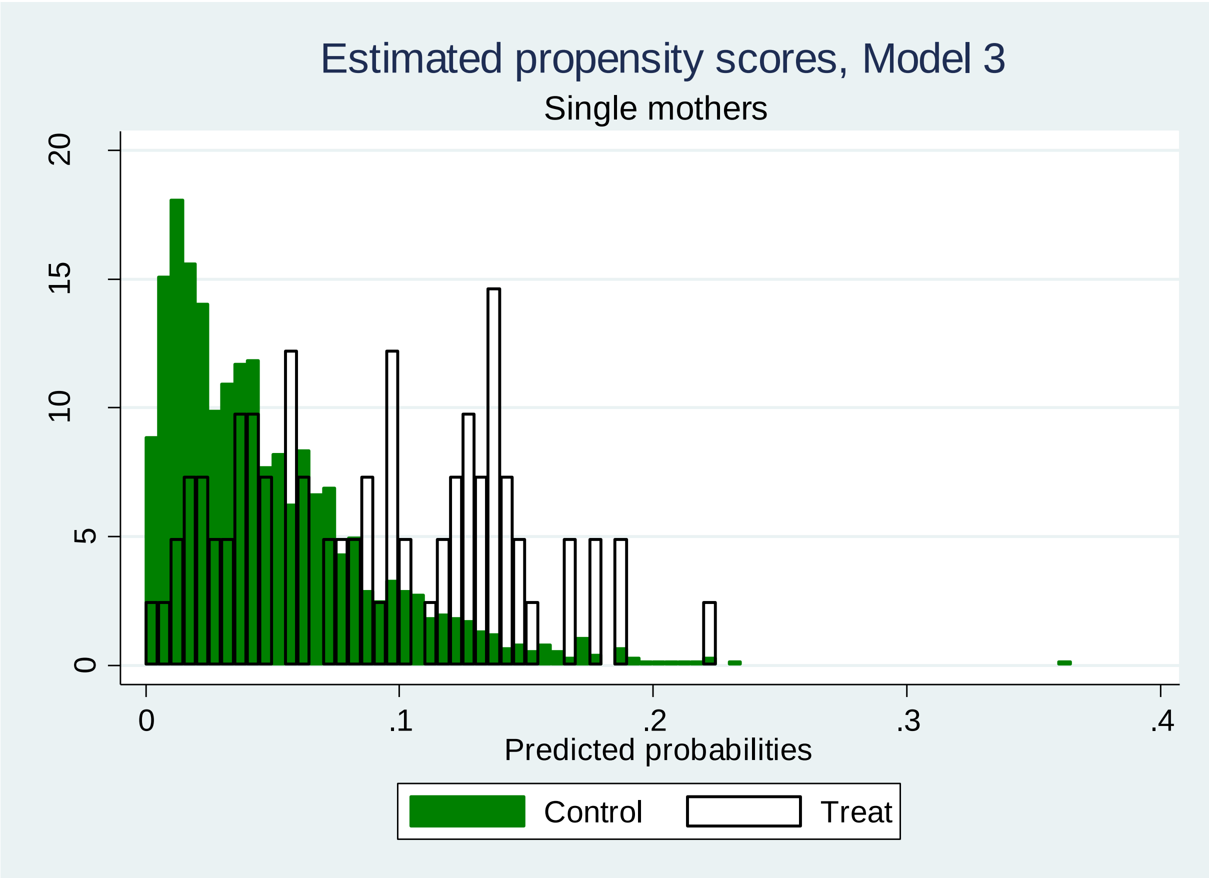
Distributions of the estimated propensity scores by treatment level




Estimated propensity scores, Model 3
Non-single mothers
0
5
10
15
20
0
.1
.2
.3
Predicted probabilities
Control
Treat


Estimated propensity scores, Model 1
Single mothers
0
5
10
15
20
0
.05
.1
.15
.2
.25
Predicted probabilities
Control
Treat


Estimated propensity scores, Model 2
Single mothers
0
5
10
15
20
0
.1
.2
.3
Predicted probabilities
Control
Treat


Estimated propensity scores, Model 3
Single mothers
0
5
10
15
20
0
.1
.2
.3
.4
Predicted probabilities
Control
Treat

Appendix 3

A3.a Matching results: Differences in means. Households with two parents.

	Model 1			Model 2			Model 3		
	Difference in means	No. of treated	No. of controls	Difference in means	No. of treated	No. of controls	Difference in means	No. of treated	No. of controls
Without matching (full sample)	-0.504	369	7088						
st. error	0.023								
Nearest-neighbour, 1-to-1, with replacement	-0.436	369	354	-0.388	369	349	-0.439	369	344
st. error	0.034			0.035			0.034		
Nearest-neighbour, 1-to-1, without replacement	-0.434	369	369	-0.401	369	369	-0.428	369	369
st. error	0.033			0.033			0.033		
Nearest-neighbour, 1-to-5, with replacement	-0.415	369	1506	-0.432	369	1470	-0.437	369	1515
st. error	0.026			0.026			0.026		
Radius matching (0.05)	-0.463	369	7088	-0.465	369	7088	-0.458	369	7088
st. error	0.023			0.023			0.023		
Radius matching (0.01)	-0.425	369	7087	-0.427	369	7058	-0.425	369	7081
st. error	0.023			0.023			0.023		
Kernel matching (bandwith: 0.06)	-0.460	369	7088	-0.461	369	7088	-0.455	369	7088
st. error	0.023			0.023			0.023		
Stratification matching	-0.431	369	6880	-0.428	369	7062	-0.431	369	7028
st. error	0.023			0.023			0.023		

A3.b Matching results: Differences in means. Households with single mothers.

	Model 1			Model 2			Model 3		
	Difference in means	No. of treated	No. of controls	Difference in means	No. of treated	No. of controls	Difference in means	No. of treated	No. of controls
Without matching (full sample)	-0.552	82	1577						
st. error	0.040								
Nearest-neighbour, 1-to-1, with replacement	-0.488	82	78	-0.5	82	74	-0.415	82	75
st. error	0.068			0.070			0.071		
Nearest-neighbour, 1-to-1, without replacement	-0.500	82	82	-0.5	82	82	-0.415	82	82
st. error	0.066			0.066			0.067		
Nearest-neighbour, 1-to-5, with replacement	-0.490	82	328	-0.5	82	302	-0.480	82	314
st. error	0.048			0.049			0.048		
Radius matching (0.05)	-0.516	82	1503	-0.519	82	1538	-0.499	82	1538
st. error	0.041			0.040			0.041		
Radius matching (0.01)	-0.496	82	1405	-0.500	82	1523	-0.469	82	1535
st. error	0.042			0.041			0.043		
Kernel matching (bandwith: 0.06)	-0.514	82	1503	-0.517	82	1538	-0.497	82	1538
st. error	0.041			0.040			0.041		
Stratification matching	-0.499	82	1172	-0.514	82	1368	-0.488	82	1507
st. error	0.042			0.042					

Appendix 4

A4.a Differences in mean employment rate, by mother's education level.

Nearest-neighbour matching, 1-to-1 with replacement. Standard errors are bootstrapped.

Mother's education level	Model 1			Model 2			Model 3		
	Difference in means	No. of treated	No. of controls	Difference in means	No. of treated	No. of controls	Difference in means	No. of treated	No. of controls
max. 8 years	-0.265	98	92	-0.388	98	84	-0.337	98	94
st. error	0.068			0.077			0.065		
lower secondary school	-0.458	118	116	-0.398	118	107	-0.530	118	114
st. error	0.068			0.068			0.068		
higher secondary school	-0.500	120	115	-0.517	120	111	-0.517	120	114
st. error	0.064			0.062			0.059		
university degree	-0.424	33	33	-0.485	33	31	-0.394	33	32
st. error	0.091			0.096			0.105		

A4.b Differences in mean employment rate, by father's employment status.

Father's employment level	Model 1			Model 2			Model 3		
	Difference in means	No. of treated	No. of controls	Difference in means	No. of treated	No. of controls	Difference in means	No. of treated	No. of controls
not employed	-0.250	84	76	-0.190	84	73	-0.286	84	79
st. error	0.079			0.082			0.084		
employed	-0.495	285	273	-0.451	285	274	-0.493	285	274
st. error	0.040			0.043			0.040		

Nearest-neighbour matching, 1-to-1 with replacement. Standard errors are bootstrapped.

A4.c Differences in mean employment rate, by children's age group.

Presence of a child under 7	Model 1			Model 2			Model 3		
	Difference in means	No. of treated	No. of controls	Difference in means	No. of treated	No. of controls	Difference in means	No. of treated	No. of controls
no	-0.333	132	127	-0.405	132	127	-0.477	132	129
st. error	0.067			0.067			0.062		
yes	-0.451	237	219	-0.426	237	221	-0.451	237	227
st. error	0.041			0.047			0.047		

Nearest-neighbour matching, 1-to-1 with replacement. Standard errors are bootstrapped.