GOSPODARKA NARODOWA

2(310)2022, 1–19
DOI: 10.33119/GN/149193
gnpje.sgh.waw.pl
SGH

The Polish Journal of Economics

Istnieje od / Published since 1931

Filip Premik
FAME|GRAPE, Poland

Evaluating Poland's Family 500+ Child Support Programme*

Szacowanie bezpośrednich skutków programu Rodzina 500+

Abstract

I investigate the immediate effects of the introduction of a large-scale child benefit programme on the labour supply of household members in Poland. Due to non-random eligibility and the universal character of the programme, standard evaluation estimators may be inconsistent. In order to address this issue, I propose an approach that combines difference-in-difference (DID) propensity score based methods with the covariate balancing propensity score (CBPS) approach developed by Imai and Ratkovic [2014]. The DID estimators exploit the time dimension to uncover the causal effect of interest. The CBPS method is expected to significantly reduce the bias resulting from systematic differences between treated and untreated subpopulations. I also account for potential heterogeneity among households by focusing on comparisons between locally defined subpopulations of individuals, which jointly provide a comprehensive view on the overall impact. I find that on average previously employed mothers maintain their labour supply although there are heterogeneous weak responses depending on the age of the youngest children. Additionally, mothers who did not work before the introduction of the programme are even less likely to do so having received the benefit. The fathers' labour supply remained mostly unaffected by the programme, with the exception of previously unemployed fathers, who tend to work more often having received the benefits. This finding may suggest that the programme strengthens the traditional division of household roles, with the male being the main earner.

Keywords:
child benefits, labour supply, programme evaluation, difference-in-difference estimation, covariate balancing propensity score

JEL classification codes:
C21, C23, I38, J22

Article history:
submitted: July 31, 2021
revised: January 6, 2022
accepted: April 13, 2022

Streszczenie

W artykule zbadano bezpośredni wpływ wprowadzenia programu świadczeń Rodzina 500+ na podaż pracy gospodarstw domowych w Polsce. Ze względu na uniwersalny charakter programu standardowe metody ewaluacji programów mogą zwracać niezgodne oszacowania. W celu zachowania pożądanych własności statystycznych w artykule zaproponowano nowe podejście łączące strategię estymacji *difference-in-difference* (DID) z wagami bilansującymi rozkłady zmiennych towarzyszących (tzw. CBPS, Imai, Ratkovic [2014]). Podejście to rozwiązuje potencjalne problemy

Słowa kluczowe:
świadczenia na dzieci, podaż pracy, ewaluacja programów rządowych, estymacja DID, bilansowanie rozkładów zmiennych towarzyszących

Kody klasyfikacji JEL:
C21, C23, I38, J22

Historia artykułu:
nadesłany: 31 lipca 2021 r.
poprawiony: 6 stycznia 2022 r.
zaakceptowany: 13 kwietnia 2022 r.

* The author gratefully acknowledges the support of the Polish National Science Centre (grant No. 2016/23/N/HS4/03637).

z brakiem spełnienia równoległych trendów, a także zmniejsza skalę obciążenia wynikającego z różnic pomiędzy jednostkami pobierającymi i niepobierającymi świadczeń. Wyniki wskazują na zaniedbywalny wpływ programu na podaż pracy w kwartałach następujących bezpośrednio po wprowadzeniu programu.

Introduction

The effects of unconditional non-equivalent income support instruments on the labour supply are at the core of labour economics. The interest in child-care programmes and labour supply has been growing since the seminal paper of **Heckman** [**1974**]. Autonomous income in the form of a subsidy pushes the budget constraint upwards, expanding the feasible set of choices between consumption and leisure. Depending on preferences, one can expect various reactions of the labour supply.

Income support instruments are under scrutiny because of a possible employment discouraging effect: the utility flow generated by a social transfer might exceed the utility from working even if the earned income was higher than the transfer as work induces some non-zero disutility [**Besley, Coate, 1992**; **Spencer, 2003**]. The direction of the impact of social transfers may be ambiguous due to nonlinearities in the budget set inherited from the relationship between the wage rate and hours worked [**Burtless, Hausman, 1978**]; a stigma effect, [**Moffitt, 1983**; **Hoynes, 1996**]; or tied wage-hours contracts [**Averett, Hotchkiss, 1997**]. Child benefits as a social transfer programme may also affect the labour supply by limiting barriers in women's access to the labour market driven by the costs of care [**Kimmel, 1998**], which in turn may even lead to an increase in the labour supply.

Given these theoretical premises, the actual effects of the income support programme on the labour supply remain an empirical question. The biggest challenge lies in the causal identification of the data: individuals adjust the labour supply and consumption for a variety of reasons. Verifying whether a certain income support instrument has been the cause and labour supply adjustment has been the effect is at the heart of much of modern microeconometrics and labour economics [**Angrist, Pischke, 2010**]. Researchers use a range of tools related to programme evaluation to answer this sort of causal questions. **Blundell et al.** [**2004**] perform difference-in-difference propensity score matching, investigating the effects of a mandatory job search programme. **Luna** [**2011**] investigates the introduction of an unconditional child benefit in Spain on family well-being using a regression discontinuity design finding that mothers eligible for the programme stayed longer out of the labour market after giving birth. **Koebel and Schirle** [**2016**] use the difference-in-difference strategy to measure the effects of the Canadian universal child care benefit and show that the child care benefit programme had a negative effect on legally married women and increased the labour supply of single mothers.

In the case of child support programmes, an additional empirical difficulty comes from the fact that eligibility is frequently based on the number of children. That makes the ineligible households (i.e. those who cannot receive the transfer) a poorly defined control group, as one may expect they may be substantially different from the eligible households [**Graham, Beller, 1989**]. In order to tackle this challenge, I propose an approach to estimate the effects of the child support instrument on household labour supply combining the covariate balancing propensity score approach (CBPS, **Imai, Ratkovic, 2014**] with difference-in-difference (DID) estimators by **Abadie** [**2005**] and **Heckman, Ihimura, Todd** [**1997**]. The DID estimators exploit the time dimension to uncover the causal effect of interest. The CBPS method is expected to significantly reduce the bias resulting from systematic differences between the treated and untreated subpopulations. I also compare the results to doubly robust estimators by **Sant'Anna and Zhao** [**2020**].

In this paper, I investigate the immediate effects of the introduction of a large-scale child benefit programme in Poland on the labour supply of household members. Programme eligibility is not random because it is driven by the number of children below 18 within a household. Additionally, the universal character of the programme makes it difficult to define appropriate control groups as those eligible include various types of households at different stages of the life cycle. My study indicates that the introduction of a large-scale child benefit programme has a minor impact on the women's labour supply and almost no effect on the men's labour supply. All of these effects are *immediate*, i.e. they concern the months directly after the introduction of the programme. I find that on average previously employed mothers maintain their labour supply although there are heterogeneous weak responses depending on the age of the youngest children. Additionally, mothers who did not work before the introduction of the programme are even less likely to do so having received the benefit. The fathers' labour supply remained mostly unaffected by the programme, with the exception of previously unemployed fathers, who tend to work more often having received the benefits. This finding may suggest that the programme strengthens the traditional division of household roles, with the male being the main earner.

The paper is organised as follows. In section 2, I discuss some simple theoretical models of labour supply and explain why their predictions for the movement in labour supply as a response to an increase in non-labour income are ex-ante ambiguous. In section 3, I provide a brief description of the child benefit programme and related literature. Section 4 introduces the estimator to account for potential incomparability between the treated and control groups and discusses other estimators used as robustness checks. In section 5, I describe the data set. In section 6, I present and discuss the empirical results. Section 7 concludes.

Theoretical considerations

Economic theory does not predict the direction of change in the labour supply in response to the exogenous non-labour income shock. This section revises different approaches to this comparative statics exercise and illustrates the issue using simple examples.

Non-convexities in the budget set

One of the explanations for why the labour supply may not necessarily decrease in response to an exogenous non-labour income shock is related to the kinks in the budget set, i.e. the points at which the boundary of the budget set is not differentiable; for reference, see, e.g., **Heckman [1974]**, **Burtless, Hausman [1978]** and **Moffitt [2002]**. Consider a simple static model of consumption and leisure choice:

$$max_{c,l} U(c,l)$$

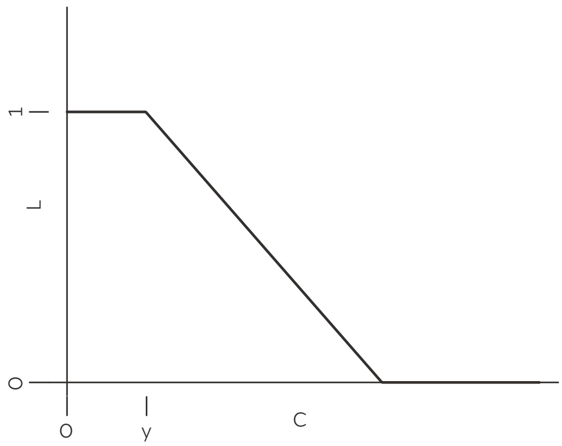
$$s.t. pc = y + w(1-l)$$

where c denotes consumption good, l – leisure (normalised to closed unit interval), p is price of consumption good, and w is the wage rate. Assume that U is increasing in both arguments. Figure 1 presents the budget set of the consumer.

Having obtained non-labour income y, a consumer may sustain non-negative consumption even if they spend the whole time endowment on leisure. Then the amount of feasible consumption increases with a decrease in labour. Consider an exogenous change in non-labour income y. **Moffitt [1990]** shows that unless the consumer's choice initially was on the kink of the budget set, the direction of change while considering comparative statics with respect to y is ambiguous.

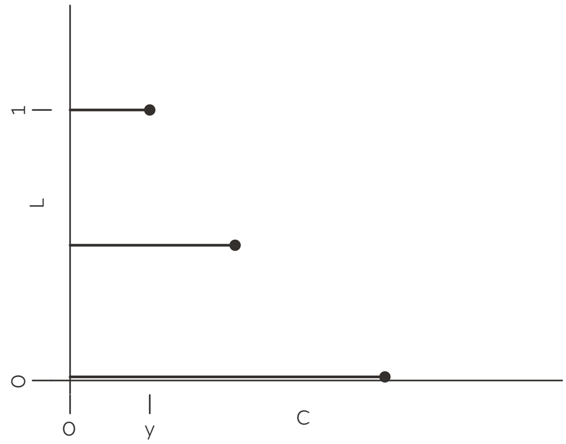
Moreover, the consumer's budget set is likely to be discontinuous. As the data shows, the vast majority of jobs observed are full-time jobs. This indicates that the choice of the amount of leisure feasible to the agent is discrete. This leads not only to kink-type points as discussed above, but also to discontinuities in the budget set. Therefore, depending on the shape of preferences, the labour supply response to a shock in y has an ambiguous direction.

Figure 1a. Kinked budget set



Source: Author's own elaboration.

Figure 1b. Consumer's budget set



Source: Author's own elaboration.

Preference shifters

Another explanation for the *a priori* ambiguous direction of changes in the labour supply in response to non-labour income shocks are preference shifters. Consider a simplified version of the life-cycle consumption-leisure choice model. Suppose an economy contains a positive measure of *a priori* identical consumers. Assume a finite horizon, no uncertainty or discounting. Suppose that all consumers enjoy perfect foresight.

Focus the attention on consumer i. They solve:

$$max_{c,l} \sum_{t=1}^{T} \frac{c_t^{1-\sigma}}{1-\sigma} - \frac{\chi}{1+\frac{1}{\epsilon}} \left(1-l_t\right)^{1+\frac{1}{\epsilon}}$$

$$s.t.\ c_t + a_{t+1} = w_t\left(1-l_t\right) + y$$

$$c_t \geq 0, a_{t+1} \geq -a, a_0 = 0$$

where y is the endowment of exogenous, non-labour income, a_t is the standard Arrow-securities, w_t is wage, all in period t. The parameters of interest are Frisch elasticity of labour ϵ, relative risk aversion coefficient γ and individual disutility from work χ.

One may easily show that the optimal solution to the problem satisfies:

$$c_t \equiv c_i \quad \forall_t$$

$$c - \frac{1}{T}\left(\sum_t \left(\frac{w_t}{\chi + i}\right)^{\epsilon}\right) c^{-\gamma\epsilon} - Ty = 0$$

$$1 - l_t = \left(\frac{w_t}{\chi}\right)^{\epsilon} c^{-\gamma\epsilon}$$

By the implicit function theorem, c is increasing in endowment y. This result is consistent with economic intuition. However, once one plugs it into the equation for within-period labour supply, the direction of the effect depends on the risk-aversion parameter and Frisch elasticity.

An additional insight comes if one allows χ to vary between consumers. This is especially appealing in the context of the child benefit programme as the opportunity cost of working an additional hour may differ between households that need to provide care services to their children and those who do not have kids. In a simple form, this heterogeneity may depend on the endowment of non-income labour: $\chi_i = \psi\left(y, \theta_i\right)$ for some function ψ and household i specific parameter vector θ_i. In this case, childless households have less disutility from working. At the same time, leisure (which implicitly also accounts for the time spent on child care services) is less costly once y is relatively high. The heterogeneity enters the labour supply equation, which in turn implies that the effects of an increase in the non-labour endowment are not necessarily monotone for households.

Home production

Household labour supply decisions may also be affected by home production technology. Consider a version of the **Ghez and Becker** [**1975**] household time use model. In this type of models, each type of commodity that enters a household's final consumption is produced through a technology that combines the input of market goods and the time spent by the household, with usually constant returns to scale. Let the number of commodities be J. The home-production function of commodity j in period t is given by:

$$c_{jt} = F^j\left(\xi_{jt} x_{jt}, \nu_{jt} h_{jt}\right)$$

where x_{jt} and h_{jt} denote the input of market goods and time respectively to produce commodity j, and ξ_{jt} and ν_{jt} are potential technology shifters for the input of goods and time. Then, the life-cycle utility function of a household with discount factor β and period utility function $U\left(c_1,\ldots,c_j\right)$ is given by:

$$\sum_{t=0}^{T} \beta^t U\left(c_{1t},\ldots c_{Jt}\right)$$

Each household maximises its expected life-cycle utility subject to the sequential budget constraint and time constraint:

$$\sum_{j=1}^{J} p_{jt} x_{jt} + a_{t+1} = w_t n_t + \left(1 + r_t\right) a_t + y_t$$

$$n_t + \sum_{j=1}^{J} h_{jt} = 1$$

where p_{jt} and w_t denote the prices of input goods and labour respectively, a_t and r_t are bonds and the interest rate respectively, and y_t denotes non-labor income. The household's endowment of time is normalised to 1, and n_t and h_{jt} denote the time spent on market production and home production of commodity j respectively. One may note that this general model formulation nests the standard household problem, where the utility is a function of just consumption goods and leisure. That is, households maximise the expected lifetime utility:

$$\sum_{t=0}^{T} \beta^t U\left(c_{1t}, c_{2t}\right)$$

subject to the budget constraint and $n_t + 1_t = 1$, with $c_{1t} = x_t$ and $c_{2t} = 1_t$.

To simplify, consider three commodities shaping household utility: home production good, childcare, and leisure. The household allocates its time endowment between labour for market production n_t, time spent on home production h_{1t} and childcare h_{2t}. That is, the household's optimisation problem becomes $E_0 \sum_{t=0}^{T} \beta^t U\left(c_{1t}, c_{2t}, 1_t\right)$ subject to $p_{1t}x_{1t} + p_{2t}x_{2t} + a_{t+1} = w_t n_t + \left(1 + r_t\right) a_t + Y_t$ and $n_t + h_{1t} + h_{2t} + 1_t = 1$, where $c_{jt} = F^j\left(\xi_{jt} x_{jt}, \nu_{jt} h_{jt}\right)$ is home production function F^1 and childcare production function F^2 respectively.

The technology shifters ξ_{jt} and ν_{jt} and non-labor income y_t may introduce uncertainty to the model. Denote the state in period t by s_t and the history of states up to period t $s^t = \left(s_0, s_1, \ldots, s_t\right)$. Let $\lambda_t\left(s^t\right)\beta^t$ be the multiplier of the household budget constraint and $\pi\left(s^t\right)$ be the unconditional probability of the realisation of s^t. Then, the first-order conditions with respect to $x_{jt}\left(s^t\right)$, $h_{jt}\left(s^t\right)$, $n_t\left(s^t\right)$, and a_{t+1} are given by:

$$\left[x_{jt}\left(s^t\right)\right] \qquad \pi\left(s^t\right)\frac{\partial U}{\partial c_{jt}\left(s^t\right)}\frac{\partial F^j}{\partial x_{jt}\left(s^t\right)} = \lambda_t\left(s^t\right)p_{jt}\left(s^t\right)$$

$$\left[h_{jt}\left(s^t\right)\right] \qquad \frac{\partial U}{\partial c_{jt}\left(s^t\right)}\frac{\partial F^j}{\partial h_{jt}\left(s^t\right)} = \frac{\partial U}{\partial \mathrm{l}_t\left(s^t\right)}$$

$$\left[n_t\left(s^t\right)\right] \qquad \pi\left(s^t\right)\frac{\partial U}{\partial \mathrm{l}_t\left(s^t\right)} = \lambda_t\left(s^t\right)w_t\left(s^t\right)$$

$$\left[a_{t+1}\left(s^t\right)\right] \qquad \lambda_t\left(s^t\right) = \left(1+r_{t+1}\left(s^{t+1}\right)\right)\beta\lambda_{t+1}\left(s^{t+1}\right)$$

After combining the first three first-order conditions, the intra-production substitution between input goods x_{jt} and time h_{jt} is characterised by

$$\frac{\partial F^j}{\partial x_{jt}\left(s^t\right)} \Big/ \frac{\partial F^j}{\partial h_{jt}\left(s^t\right)} = p_{jt}\left(s^t\right)/\,w_t\left(s^t\right)$$

and the cross-production substitution is characterised by

$$\frac{\dfrac{\partial F^1}{\partial x_{1t}\left(s^t\right)} \Big/ \dfrac{\partial F^1}{\partial h_{1t}\left(s^t\right)}}{\dfrac{\partial F^2}{\partial x_{2t}\left(s^t\right)} \Big/ \dfrac{\partial F^2}{\partial h_{2t}\left(s^t\right)}} = \frac{p_{1t}\left(s^t\right)}{p_{2t}\left(s^t\right)}$$

Therefore, the elasticities of substitution of home production technology F^j and utility function U determine the optimal allocation of input goods and time use.

These elasticities are likely to vary with household characteristics such as the number or age of children. If non-labour income increases, the relative cost of market good inputs decreases relative to the time input. For example, with a larger budget for the current period, the household may want to allocate more time spent with children. However, it may also be the case that market care (preschool, extracurricular classes and activities, etc.) is significantly more efficient in terms of care production technology, so that the additional income is spread as inputs for other types of home production.

First and second earner time allocation

Labour supply decisions within the household are likely to be made simultaneously by each of the members[1]. **Blundell et al. [2016]** find that in households with children the response to permanent and transitory income shocks differs between the male and the female[2]. In response to a transitory wage shock, the elasticity of the husband's labour supply does not depend on the presence of children, while with young kids in the household the woman's labour supply elasticity is significantly larger. In response to a permanent increase in the wage, men allocate more time in the market labour to the cost of both leisure and childcare, whereas women tend to spend less time working in the market but more time working at home with children. However, when a woman receives a permanent wage increase and reduces her time spent in childcare significantly, it is the man who reduces his hours, increases leisure and tends to spend slightly more time providing childcare.

I account for this issue in the empirical part, considering separate effects for males and females within a household.

1 **Connelly [1992]** provides conditions under which single parent optimisation is consistent with household optimisation.

2 In more general terms: first and second earner. Data for Poland suggests that the mother is much more likely to be the second earner in the presence of children in the household.

The child benefit programme

The *Rodzina 500 Plus* (Family 500 Plus) programme is one of the most significant policy interventions affecting the household sector in Poland in recent years. Introduced in the second quarter of 2016, this large-scale programme, which costs about 2% of the GDP a year, provides a monthly non-equivalent benefit of roughly 20% of the net average monthly salary to each family with two children, and another 20% of the net average monthly salary for each third and next child. Additionally, families below a certain income level are entitled to obtain the benefit also for the first-born child. In order to obtain the benefit, eligible households must register with local authorities. This kind of simplicity strongly stimulates participation among those eligible. The authorities predict that around 2.7 million families bringing up to 3.7 million children are or will be enrolled in the programme. The main goal of the programme is to improve the financial well-being of families bringing up children and to stimulate fertility.

The programme was announced in the first quarter of 2016. The first payments arrived to the treated households in April 2016. Due to the short time between the announcement and implementation of the programme, households were unlikely to adjust their behaviour to the expected arrival of benefits. The lack of anticipation effects facilitates statistical inference.

It is not clear whether households perceived the benefit as a long-term increase in non-labour income. If they expected the programme to be terminated in the reasonably short horizon or believed that the eligibility requirements might change, then they were likely to treat the increase in non-labour income as transitory. Therefore, the labour supply would not adjust. Moreover, even if a household perceived the programme as permanent, the process of labour supply adjustment may take several periods due to labour market frictions, existing contracts, etc. Hence, adjustments in the aggregate labour supply are likely be observed in data with some delay. This paper focuses on the immediate effects of the programme using data up to the fourth quarter of 2016.

The *Rodzina 500 Plus* programme has already been studied in the literature. **Magda et al.** [**2018**] investigate in a similar fashion the effects on the aggregate labour supply and find treatment effects implying a drop of 2 to 3 percentage points in the female labour force supply in response to the programme. **Myck** [**2016**] and **Myck and Trzcinski** [**2019**] estimate a discrete choice model of female labour supply using data on the pre-treatment period to predict the effects of introducing the programme suggesting a small or insignificant drop in the labour force supply, mostly affecting mothers in families with one or two children. **Paradowski, Wolszczak-Derlacz, Sierminska** [**2020**] applies the DID framework to show a substantial reduction in poverty and inequality indices among Polish households after the introduction of the programme. **Premik** [**2021**] investigates the effects of the programme in the longer term, emphasising the importance of discouragement in job searching activities among previously unemployed females.

Estimation

Estimating the effects of the non-equivalent child benefit programme on the labour supply falls into the category of programme evaluations. However, eligibility for the programme is not random as long as we believe that the number of children within a household is not random, but rather a result of household optimisation, as in **Rosenzweig and Schultz** [**1985**]. This implies that families that do not obtain the benefit may differ systematically from the treated subpopulation. Moreover, the number of children is also likely to affect labour supply measures; **Connelly** [**1992**], **Nakamura and Nakamura** [**1992**], **Black et al.** [**2013**].

The universal character of the programme makes it difficult to define appropriate control groups even with additional reweighing coming from the CBPS approach. There are plenty of households on different stages of the life cycle and composition [**Graham, Beller, 1989**]. In order to improve the comparability of outcomes between households assigned to the treatment and control group, I focus on comparisons within locally defined subpopulations. Several approaches are considered. First, I compare households with at least two children and the youngest child under the age of six to those with one child under the age of six. The idea behind this identification scheme relies on the premise that children require the most attention when they are young, and this attention cannot be easily replaced through the market for caring services. At this stage of child development, mothers are more likely to be inactive (care leave after maternity leave), hence weaker disincentives for work may be sufficient to decrease their labour supply. The choice of the children's age threshold for this comparison is motivated by school duty. Second, I compare eligible households with a woman being the second earner to households with eligible households in which the woman is a single earner. Employment elasticities tend to be larger for single mothers; **Connelly and Kimmel** [**2003**], **Blundell et al.** [**2016a**]. Additionally, both single and married women obtain a sort of non-labour income: single mothers only in the form of benefit, full family mothers also in the form of the husband's earnings. However, single women are likely to be more constrained financially and because of providing child care services. Benefits may loosen the constraints and in this way enable single women to move to a new optimum, which is not necessarily a corner solution.

To address these issues and consistently estimate the effects of interest, I combine the covariate balancing propensity score method [**Imai, Ratkovic, 2014**] with the difference-in-difference estimator proposed by **Abadie** [**2005**] and **Heckman Ihimura, Todd** [**1997**]. The strategy based on the DID approach exploits the quasi-natural experimental character of the child benefit programme, which was introduced swiftly, not allowing for anticipation effects. The remaining differences between the treated and control households motivate applying CBPS estimates of propensity score to ensure proper comparisons between these subpopulations. The CBPS estimator employs the moment-based approach to the force balancing of the conditional distributions of the covariates. This is expected to significantly reduce the bias resulting from the systematic differences between the treated and untreated subpopulations. It may also smooth the time trends between the pre- and post-treatment periods conditionally on the covariates, further validating the DID approach.

I follow standard programme evaluation notation throughout the paper. I observe a household i in two periods indexed by $t \in \{0,1\}$. Treatment D_{it} is binary and is observed only in t=1, hence $\forall_i \, D_{i0} = 0$ and $D_{i1} \equiv D_i \in \{0,1\}$. Potential outcomes are denoted as Y_{it}^D, where $t \in \{0,1\}$, $D \in \{0,1\}$. For instance, Y_{i1}^1 is the value of the outcome variable in a world in which individual i receives treatment. The outcome variable observed by an econometrician is $Y_{it} = Y_{it}^0 (1 - D_{it}) + Y_{it}^1 D_{it}$. I consider a set of pre-treatment covariates $X_i \equiv \{x_{i0}\}$. The main estimand of interest is the average treatment effect on the treated (ATT):

$$\tau^{ATT} \equiv \tau = E\left[Y_{it}^1 - Y_{it}^0 | D_{it} = 1\right] = E\left[Y_{i1}^1 - Y_{i1}^0 | D_i = 1\right] \tag{1}$$

and its conditional-on-covariates versions. The conditional probability of treatment assignment is given by $P\left[D_i = 1 | X\right] = F_\beta(X_i)$, where β is a set of parameters. Let $F_\beta(X_i)$ denote the first derivative of $F_\beta(X_i)$ with respect to the parameters. If nothing else is mentioned, $F_\beta(X_i)$ is the logistic cumulative distribution function.

Covariate balancing propensity score

Since the influential paper of **Rosenbaum and Rubin** [**1983**], propensity score related methods have been continuously gaining attention in research, especially in the context of causal analysis; see, for example, **Dehejia and Wahba** [**1999**], **Caliendo and Kopeinig** [**2008**], and **Austin** [**2011**]. The propensity score as a conditional probability of treatment assignment serves mainly for matching or weighting. The true propensity score is rarely known, so it must be estimated by a researcher. The standard framework is limited to maximum likelihood binomial models. However, as it stems from the likelihood theory, misspecification of the propensity score model leads to inconsistent estimators of binomial model parameters and therefore to inconsistent estimates of the propensity score. That, in turn, might result in potentially heavily biased causal analysis **Zhao** [**2004**]. Usually researchers do not know the correct functional form of the data generating process, which pushes them towards experimenting with many probably incorrect specifications and choosing *the best* one. Estimating the conditional probability of being eligible to the child benefit programme is no different. In this paper, instead of estimating many probably misspecified propensity scores and checking their quality by a covariate balance check [**Dehejia, Wahba, 2002**], I employ the approach of **Imai and Ratkovic** [**2014**]. They exploit the dual nature of the propensity score, involving the conditional probability of the treatment assignment and the covariate balancing score. The goal is to find a vector of propensity scores $F_\beta(X_i)$, parametrized by β, which balances the distribution of covariates between the treated and the control groups. The β parameters are identified through the following moment condition:

$$E\left[D\tilde{X} - \frac{F_\beta(X_i)(1-D)\tilde{X}}{1-F_\beta(X_i)}\right] = 0 \tag{2}$$

One may set $\tilde{X} = X$ to balance the first moments, $\tilde{X} = X^2$ to balance the second moments, etc. Restrictions that impose a balance on a chosen moment for all the covariates are sufficient to ensure just-identification of the parameters β from the probability function $F_\beta(X_i)$. In general, imposing restrictions on more than one class of moments leads to an over-identified estimator, expected to be more biased in finite samples but asymptotically more efficient. One may add either first-order conditions from the binary ML model with $F_\beta(X_i)$ as the cumulative probability function or balancing conditions for (higher order) moments of covariates. Additionally, the over-identified version enables the researcher to perform a standard specification test of over-identifying restrictions. In the context of CBPS estimation, the specification test might be perceived as a verification of the reliability of unconfoundedness assumption [**Imai, Ratkovic, 2014**].

The sample moment conditions in the just-identified version of CBPS are:

$$\frac{1}{N}\sum_i \frac{N}{N_1} \cdot \frac{D - F_\beta(X_i)}{1-F_\beta(X_i)} \cdot X_i = 0 \tag{3}$$

The just-identified models are at the core interest in the paper as I utilise a relatively small sample. I consider the over-identified versions only to perform specification tests.

The CBPS method is preferred because it is robust for misspecification of the probability model, as opposed to ML estimators. Additionally, it balances the distribution of covariates, which is advantageous in the context of the paper due to the expected heterogeneity between the control and treated groups.

Abadie's difference-in-difference

Difference-in-difference estimators exploit time dimension in data to generate means of a natural experiment. Following **Abadie** [**2005**], I maintain a version of the parallel trends assumption:

$$E\left[Y_{i1}^0 - Y_{i0}^0 | X_i, D_i = 1\right] = E\left[Y_{i1}^0 - Y_{i0}^0 | X_i, D_i = 0\right] \tag{4}$$

which states that, in the absence of treatment, outcomes in both groups would behave in the same way. I also assume overlap:

$$P\left[D_i = 1 | X_i\right] < 1$$

which is a standard identification condition for selection on observables. These assumptions are enough to uncover ATT.

Abadie [2005] proposes an estimation procedure that relies on the least squares approximation to $\tau_{X_k} = E\left[Y_1^1 - Y_1^0 | D_i, X_k\right]$:

$$\tau_{\{X_k\}} = argmin_{\{\theta \in \Theta\}} E\left[\frac{F_\beta(X_i)(\mathrm{D_i} - F_\beta(X_i))}{F_\beta(X_i)(1-F_\beta(X_i))}\left((Y_{\{i1\}} - Y_{\{i0\}}) - g(X_k;\theta)\right)^2\right] \tag{5}$$

where $X_k \subseteq X$ and $g(X_k,\ \theta)$ is the approximating function. If \$$g(X_k;\ \theta) = \theta$ then $\tau x_k = \tau$. In the empirical part of my paper, I focus mainly on the *homogeneous*[3] average effect on the treated. Considering sample analogues, equation (2) implies the following first-order conditions:

$$\frac{1}{N}\sum_i -F_\beta(X_i)\left(\frac{\mathrm{D}_i - F_\beta(X_i)}{1-F_\beta(X_i)}\,\Delta Y_i - F_\beta(X_i)\tau\right) = 0 \tag{6}$$

This paper combines moment conditions from equations (3) and (6) in order to obtain a GMM representation for a one-step ATT estimator. It is appealing in many ways as it combines the advantages of both the CBPS approach and Abadie's DID method. First, it is a one-step method so there is no need to adjust the DID estimation for the stochastic nature of the previously estimated propensity score. Second, I express the whole estimation in terms of the GMM framework. It allows me to obtain the (efficient) covariance matrix in a simple fashion. Third, the one-step estimation is justified because the way the propensity score is defined produces perfect balancing properties by construction. Fourth, given the conditional independence assumption, it is robust for functional form misspecification, as neither **Imai and Ratkovic** [**2014**] nor **Abadie** [**2005**] assume any functional form for the true data generating process. In further discussions I refer to this estimator as ACBPS.

Robustness checks

In the empirical part, I apply the just-identified ACBPS as an efficient GMM, where the parameters needed to estimate consistently the moment covariance matrix come from a pre-estimated logit and logit-based Abadie's DID. I compare the results with Heckman's difference-in-difference propensity score estimators using both logit and CBPS propensity score estimates. My method is likely to be advantageous compared to the conventional estimators in the presence of an improperly balanced control group. Additionally, ACBPS should be less biased in small samples as it skips the matching step.

Finally, I also present results using the standard OLS estimator. However, this method is expected to perform poorly due to the lack of a control for the imbalance between the treatment and control groups. We maintain an assumption of conditional independence, i.e. conditioning on the set of covariates described in the next section, there are no systematic differences in the probability of obtaining the benefit between the treated and control subsamples.

We compare the estimates from our estimators with other estimators for ATT to verify the robustness of the results. We investigate:

- Basic DID (as a baseline, **Card and Krueger** [**1994**]

$$\tau = \left(\overline{\{y\}}_{\{11\}} - \overline{\{y\}}_{\{01\}}\right) - \left(\overline{\{y\}}_{\{10\}} - \overline{\{y\}}_{\{00\}}\right)$$

- Regression-based DID controlling for covariates:

$$\Delta Y_i = \beta_0 + \tau D_i + X_{i\beta} + e_i$$

- Abadie's estimator in the original two-step formulation, with ML logit estimation for the propensity score.
- Heckman's estimator in the original formulation with ML logit estimation for the propensity score and kernel matching.
- **Sant'Anna and Zhao** [**2020**] doubly robust estimator in two versions – *drimp* being a DR estimator based on the inverse probability of tilting and weighted least squares, and *ipw* – inverse probability weighting DiD estimator in the spirit of **Abadie** [**2005**].

3 The quality of any non-constant approximation for the effects of interest is expected to be poor due to the limited number of observations in my sample.

The Stata codes for ACBPS and CBPS estimation are available on request. In the paper, we also make use of codes by Houngbedji [2016], Leuven and Sianesi [2003] and Sant'Anna and Zhao [2020].

Data

I utilise Labor Force Survey (LFS) data collected by the Central Statistical Office in Poland. It is a rotating panel in which households are interviewed in two consecutive quarters and then reinterviewed in two respective periods a year after. The time dimension allows me to apply the DID strategy.

This paper presents the estimates of the immediate effects of the programme. I focus on a balanced panel with $T_0 \in \{2015Q3, 2015Q4\}$ and $T_1 \in \{2016Q3, 2016Q4\}$. Note that 2016Q2 is treated as pre-intervention, as the programme was introduced in the middle of this period and a non-trivial share of households received their payments with a few months' delay.

I consider only families with children and drop households in which another child was born between T_0 and T_1. Depending on the definition of the control groups, I keep households with various characteristics in the sample. I distinguish five main groups of covariates that we control for in order to satisfy the unconfoundedness assumption. Table 1 summarises the choice of covariates used for conditioning in my analysis.

Table 1. Covariates controlled for in the model specification

(1)	Household	presence of non-parental household members, small city and village indicators, mother's age at birth of the youngest child
(2)	Mother demographics	age, work experience, education
(3)	Mother's job	indicators for public company, working as an employee, tenure contract, full-time job, binary variable for working in the field of education
(4)	Father demographics	age, age squared, work experience, education
(5)	Father's job	indicators for public company, working as an employee, tenure contract, full-time job, binary variable for working in the field of education

Source: Own tabulation based on BAEL data.

Results

This section presents estimated treatment effects concerning the impact of introducing non-equivalent child support on the labour supply within a household. I consider various definitions of control and treated subpopulations to provide a comprehensive assessment of the total effect. Wherever the sample size permits, I separate effects on households with respect to the age of the youngest child as it is a significant determinant of the labour supply, especially in the case of women [Jacobsen, Pearce III, Rosenbloom, 1999]. The set of conditioning covariates is described in the previous section. The results are presented in Tables 2–7.

The specification test discussed in the tables is the over-identifying restrictions test for the CBPS propensity score. It is valid for both Heckman's and Abadie's approaches, which are based on the same just-identified estimates of CBPS. Rejection of the null hypothesis might suggest violation of the conditional independence assumption indicating failure in accounting for observable factors.

Full families with children at the same age

I begin with comparisons between families in which both the mother and father are present. The treated group is defined as a subset of households that have at least two children and are thus eligible for participation in the child benefit programme, whereas families with only one child are classified as a control group. However, some of the control group households may also be eligible. The data gives no opportunity to verify their participation status as a household's income is not observed. However, one should not expect a high percentage of potentially treated households in the control group, as they contain at least both the father and mother, making their total income likely to exceed the eligibility threshold. I separate the effects for parents who worked in the pre-treatment period and those who did not.

First, I analyse the labour supply paths of parents who worked in the pre-treatment period. Table 2 presents the estimated effects for women, while Table 3 shows the data for men. The average effects on the pooled sample are insignificant for both men and woman. Most *local* effects for both women and men who worked in T_0 and have obtained a benefit indicate that they worked in T_1 with approximately the same probability and the same number of hours as those employed in T_0 with only one child. This finding is robust with respect to the choice of the specification. In the majority of the comparison groups, the parameters are not only statistically insignificant but also tend to zero numerically for most subgroups.

Table 2. Effects on mothers who worked in T0

	OLS		Abadie		Heckman		Sant'Anna and Zhao	
	raw	with x	cbps	logit	cbps	logit	drimp	abadie
Probability of working in T1								
	all households							
tau	0.001	−0.001	0.001	−0.015	−0.003	−0.002	−0.003	−0.001
	(0.009)	(0.009)	(0.010)	(0.013)	(0.010)	(0.010)	(0.008)	(0.008)
	N= 3512, N treated = 1667, Specification test for CBPS p-val = 0							
	youngest child younger than 6							
tau	−0.012	−0.030	−0.026	−0.155	−0.004	−0.005	−0.023	−0.020
	(0.016)	(0.018)	(0.016)	(0.224)	(0.009)	(0.010)	(0.013)	(0.015)
	1048, N treated = 648, Specification test for CBPS p-val = 0							
	youngest child between 6 and 11							
tau	−0.001	−0.016	−0.011	0.004	−0.001	0.001	−0.007	−0.007
	(0.014)	(0.016)	(0.015)	(0.020)	(0.010)	(0.010)	(0.013)	(0.013)
	N= 1232, N treated = 705, Specification test for CBPS p-val =.17							
	youngest child older than 11							
tau	0.033*	0.028	0.038	0.032	−0.003	−0.002	0.027	0.028
	(0.016)	(0.018)	(0.027)	(0.021)	(0.010)	(0.010)	(0.016)	(0.016)
	N= 1085, N treated = 247, Specification test for CBPS p-val = 1							
Hours worked								
	all households							
tau	0.493	0.381	0.409	0.341	0.429	0.448	0.465	0.496
	(0.399)	(0.449)	(0.480)	(0.458)	(0.472)	(0.469)	(0.379)	(0.387)
	N= 3512, N treated = 1667, Specification test for CBPS p-val = 0							
	youngest child younger than 6							
tau	−0.571	−1.325	−0.298	−0.578	0.405	0.434	−0.344	−0.010
	(0.703)	(0.810)	(0.776)	(0.943)	(0.444)	(0.457)	(0.642)	(0.764)
	N= 1195, N treated = 715, Specification test for CBPS p-val = 0							
	youngest child between 6 and 11							
tau	0.504	0.458	0.625	0.532	0.596	0.702	0.741	0.765
	(0.646)	(0.755)	(0.740)	(0.716)	(0.458)	(0.458)	(0.623)	(0.619)
	N= 1232, N treated = 705, Specification test for CBPS p-val =.17							
	youngest child older than 11							
tau	1.667*	1.756	2.122	1.844	0.390	0.452	1.670*	1.704*
	(0.799)	(0.958)	(1.317)	(0.967)	(0.462)	(0.464)	(0.809)	(0.811)
	N= 1085, N treated = 247, Specification test for CBPS p-val = 1							
I control for the following set of covariates: (1), (2), (3), (4), (5) and labour status of the father (see Table 1). Standard errors in parentheses. * p < 0:05, ** p < 0:01,*** p < 0:001								

Source: Own calculation based on BAEL data.

Table 3. Effects on fathers who worked in T0

	OLS		Abadie		Heckman		Sant'Anna and Zhao	
	raw	with x	cbps	logit	cbps	logit	drimp	abadie
Probability of working in T1								
	all households							
tau	0.003	0.002	−0.002	−0.016	0.001	0.001	0.002	0.002
	(0.006)	(0.006)	(0.007)	(0.009)	(0.007)	(0.007)	(0.005)	(0.005)
	N= 4020, N treated = 2487, Specification test for CBPS p-val = 0							
	youngest child younger than 6							
tau	−0.000	−0.002	0.011	−0.016	0.001	−0.001	0.001	0.008
	(0.010)	(0.011)	(0.014)	(0.009)	(0.006)	(0.007)	(0.010)	(0.014)
	N= 2126, N treated = 1294, Specification test for CBPS p-val = 0							
	youngest child between 6 and 11							
tau	−0.001	−0.005	−0.001	−0.016	−0.001	−0.002	0.002	0.004
	(0.010)	(0.011)	(0.012)	(0.009)	(0.007)	(0.007)	(0.009)	(0.010)
	N= 1567, N treated = 895, Specification test for CBPS p-val =.175							
	youngest child older than 11							
tau	0.020	0.011	−0.006	0.000	0.001	0.001	0.007	0.008
	(0.013)	(0.011)	(0.015)	(0.015)	(0.007)	(0.007)	(0.010)	(0.011)
	N= 1290, N treated = 298, Specification test for CBPS p-val = 1							
Hours worked								
	all households							
tau	0.227	0.498	0.295	0.365	0.411	0.355	0.381	0.376
	(0.303)	(0.371)	(0.423)	(0.397)	(0.394)	(0.394)	(0.327)	(0.335)
	N= 4020, N treated = 2487, Specification test for CBPS p-val = 0							
	youngest child younger than 6							
tau	−0.411	−0.387	0.286	−0.140	0.481	0.434	0.077	0.076
	(0.498)	(0.663)	(0.772)	(1.905)	(0.379)	(0.388)	(0.576)	(0.875)
	N= 2126, N treated = 1294, Specification test for CBPS p-val = 0							
	youngest child between 6 and 11							
tau	0.617	0.753	0.786	0.759	0.339	0.293	0.297	0.360
	(0.502)	(0.677)	(0.728)	(0.692)	(0.386)	(0.386)	(0.575)	(0.590)
	N= 1567, N treated = 895, Specification test for CBPS p-val =.175							
	youngest child older than 11							
tau	1.164	1.288	0.699	1.156	0.362	0.351	1.223	1.272
	(0.620)	(0.748)	(1.016)	(0.778)	(0.394)	(0.395)	(0.723)	(0.773)
	N= 1290, N treated = 298, Specification test for CBPS p-val = 1							
I control for the following set of covariates: (1), (2), (3), (4), (5) and labour status of the father (see Table 1). Standard errors in parentheses. * p < 0:05, ** p < 0:01,*** p < 0:001								

Source: Own calculation based on BAEL data.

Two comparisons require further comments. First, most estimates consistently suggest a small decrease in labour participation for treated mothers whose youngest child is below the age of six, despite low estimation precision. This might indicate weak regularity in quitting jobs by participating females with two or more kids. This regularity is consistent with economic theory, as additional money from non-equivalent transfer shifts the budget line so the allocation with no labour (and full-time maternal care) becomes available. The trade-off between work and care is likely to be stronger for mothers with more children. Similarly, mothers of two or more children whose youngest child is over 11 years old are more less likely to keep employment after receiving the benefit. In this case, the additional portion of income may be spent on extracurricular activities for older children. Teenagers do not require full-time maternal care. On the contrary, market services might outperform informal care by means of quality as the needs of teenagers are different than those of younger children. Therefore, given a non-equivalent transfer, mothers become able to afford additional after-school classes or other form of activities for a child gaining additional time to work.

Provided leisure plays a minor role in total utility, both the mother and the child might spend their time more productively. Nevertheless, it should be noted that in both cases the estimates do not pass the significance test at 5% confidence level. The low precision of the estimates may be indicative of heterogeneity in treatment effects, i.e. the existence of smaller subpopulations of individuals (being a subset of the comparison groups analysed in this paper) for which the effects are strong. In addition, the relatively small sample size is also likely to decrease the precision of the estimates.

Responses to the programme at the intensive margin are even weaker. Estimates obtained on the pooled sample of mothers working in T_0 imply the programme introduced no difference in patterns for the number of hours worked between the analysed groups. Estimates with respect to the age of the oldest child confirm this finding, being insignificant both statistically and economically.

The pooled sample results do not pass the CBPS specification, which may suggest the existence of unobserved factors deteriorating the covariate balance between the analysed groups. The situation improved when I considered subsamples with respect to the age of the youngest children. In this case, in most of the specifications, I cannot reject the null. This suggests that household behaviour differs significantly with respect to the age of the youngest child.

There is no evidence that the programme has changed patterns of employment between fathers with two or more kids and fathers with one child. Regardless of the comparison group, all estimates are statistically indistinguishable from zero and very low numerically.

Tables 4 and 5 present analogous effects of the programme on the subsample of parents who did not work in T_0. The results for females uncover an interesting pattern suggesting that mothers of more than one child who did not work in T_0, are less likely to find a job after the benefits are granted. This employment discouragement effect is particularly visible in terms of the extensive margin. Consistently across specifications, the programme discouraged employment search. Estimates from the pooled sample indicate a 3 percentage point drop in the probability of working among the treated in T_1. Estimates for subpopulations with respect to the age of the youngest child confirm this finding. Though not precisely estimated, the negative estimates in this section may suggest a stronger trend in job search discouragement that would strengthen with time as households internalise the long-term existence of the programme. Using data with longer time series, **Premik [2021]** confirms that supposition and describes its role in the overall drop of labour force participation in response to the programme. In turn, the response on the intensive margin was negligible. Nearly all the estimates are statistically insignificant. The greatest effect in numerical terms is a drop in the number of hours worked by less than 2 hours a week on average.

Table 4. Effects on mothers who did not work in T0

	OLS		Abadie		Heckman		Sant'Anna and Zhao	
	raw	with x	cbps	logit	cbps	logit	drimp	abadie
Probability of working in T1								
	all households							
tau	−0.022	−0.007	−0.029	−0.052	−0.020	−0.034	−0.028	−0.036*
	(0.018)	(0.019)	(0.021)	(0.027)	(0.022)	(0.022)	(0.017)	(0.018)
	N= 2095, N treated = 1095, Specification test for CBPS p-val = 0							
	youngest child younger than 6							
tau	−0.052	−0.001	0.006	−0.237	−0.003	−0.017	−0.015	−0.008
	(0.027)	(0.028)	(0.027)	(0.790)	(0.021)	(0.021)	(0.022)	(0.025)
	N= 1156, N treated = 720, Specification test for CBPS p-val = 0							

	OLS		Abadie		Heckman		Sant'Anna and Zhao	
	raw	with x	cbps	logit	cbps	logit	drimp	abadie
	youngest child between 6 and 11							
tau	−0.060	−0.028	−0.056	−0.034	−0.001	−0.004	−0.026	−0.028
	(0.036)	(0.039)	(0.044)	(0.049)	(0.021)	(0.021)	(0.033)	(0.033)
	N= 513, N treated = 285, Specification test for CBPS p-val =.99							
	youngest child older than 11							
tau	−0.040	−0.015	0.080	0.029	−0.025	−0.024	−0.012	−0.011
	(0.039)	(0.042)	(0.082)	(0.047)	(0.022)	(0.022)	(0.035)	(0.036)
	N= 426, N treated = 90, Specification test for CBPS p-val =.862							
Hours worked								
	all households							
tau	−0.614	−0.518	−1.396	−1.468	−1.004	−1.499	−1.438*	−1.715**
	(0.657)	(0.717)	(0.769)	(0.806)	(0.805)	(0.816)	(0.622)	(0.660)
	N= 2095, N treated = 1095, Specification test for CBPS p-val = 0							
	youngest child younger than 6							
tau	− 2.172*	−0.470	−0.363	−0.167	−0.340	−0.905	−1.132	−0.870
	(1.000)	(1.047)	(0.979)	(1.278)	(0.772)	(0.797)	(0.823)	(0.944)
	N= 1156, N treated = 720, Specification test for CBPS p-val = 0							
	youngest child between 6 and 11							
tau	−1.177	−0.836	−1.507	−1.101	−0.315	−0.375	−0.791	−0.811
	(1.230)	(1.373)	(1.547)	(1.355)	(0.784)	(0.780)	(1.168)	(1.179)
	N= 513, N treated = 285, Specification test for CBPS p-val =.99							
	youngest child older than 11							
tau	−1.467	−0.707	0.483	−0.004	−1.058	−1.026	−0.892	−0.877
	(1.313)	(1.578)	(2.163)	(1.484)	(0.804)	(0.808)	(1.235)	(1.275)
	N= 426, N treated = 90, Specification test for CBPS p-val =.862							
I control for the following set of covariates: (1), (2), (3), (4), (5) and labour status of the father (see Table 1). Standard errors in parentheses. * p < 0:05, ** p < 0:01,*** p < 0:001								

Source: Own calculation based on BAEL data.

Table 5. Effects on fathers who did not work in T0

	OLS		Abadie		Heckman		Sant'Anna and Zhao	
	raw	with x	cbps	logit	cbps	logit	drimp	abadie
Probability of working in T1								
	all households							
tau	0.130**	0.028	0.057	0.000	0.058	0.046	0.026	−0.001
	(0.041)	(0.039)	(0.049)	(0.062)	(0.047)	(0.046)	(0.036)	(0.042)
	N= 1458, N treated = 590, Specification test for CBPS p-val =.127							
Hours worked								
	all households							
tau	3.490***	0.628	1.375	0.870	1.750	1.340	0.352	−0.801
	(0.794)	(1.648)	(2.145)	(2.082)	(1.976)	(1.970)	(1.537)	(1.817)
	N= 1458, N treated = 590, Specification test for CBPS p-val =.127							
I control for the following set of covariates: (1), (2), (3), (4), (5) and labour status of the father (see Table 1). Standard errors in parentheses. * p < 0:05, ** p < 0:01,*** p < 0:001								

Source: Own calculation based on BAEL data.

As depicted in Table 5, fathers of children who received the benefits seemed to increase their participation in the labour market after receiving the transfers. The positive effect is in particular postulated by methods based on the proper balancing of covariate distributions. The increased job finding rate among fathers compared with a decrease in the same indicator among mothers is consistent with the differentiation of household roles as primary and secondary earners within a household. Traditionally in Poland, it is males who are the main earners. The child benefit programme seems to have strengthened this trend. Nevertheless, again, the precision of estimation is rather poor, which may suggest heterogeneity in effects on lower levels of aggregation. Notably, no important adjustments pop up on the intensive margin.

Single participating mothers and ineligible full families

In this section, I compare the labour outcomes of single mothers with two or more children, who would receive the benefit in T_1, to married mothers with only one child, who by assumption will not receive a transfer. The rationale for such comparisons lies in larger employment elasticities of the former [**Connelly, 2003**]. Single mothers may react strongly for any disincentive to work. However, the data provide mixed evidence for this hypothesis. Among mothers who worked in T_0, depending on the estimation method, the estimates at the extensive margin vary by direction and magnitude of the implied effect. Notably, the CBPS-based estimation would support the hypothesis of an increased labour supply of single mothers. However, the specification test fails. Given the substantially different implications of other estimates, one should be cautious about reaching strong conclusions in this case.

Table 6. Effects on single mothers who worked in T0

	OLS		Abadie		Heckman		Sant'Anna and Zhao	
	raw	with x	cbps	logit	cbps	logit	drimp	abadie
Probability of working in T1								
	all households							
tau	−0.041	−0.010	0.421***	0.142***	−0.038	0.002	−0.029	−0.044*
	(0.023)	(0.022)	(0.033)	(0.031)	(0.032)	(0.035)	(0.023)	(0.022)
	N= 1610, N treated = 183, Specification test for CBPS p-val = 0							
Hours worked								
	all households							
tau	−0.874	0.376	−0.829	−0.560	−1.043	0.579	−0.925	−1.476
	(1.003)	(0.986)	(1.202)	(1.088)	(1.387)	(1.615)	(0.972)	(0.933)
	N= 1610, N treated = 183, Specification test for CBPS p-val = 0							
I control for the following set of covariates: (1), (2), (3), (4), (5) and labour status of the father (see Table 1). Standard errors in parentheses. * p < 0:05, ** p < 0:01,*** p < 0:001								

Source: Own calculation based on BAEL data.

Estimates at the intensive margin imply no adjustments in the hours worked. Meanwhile, single mothers who did not work in T_0 seem to be searching and getting a job more eagerly than married women and also work slightly longer having found a job. I interpret this as a premise that some single treated mothers worked in T_1 significantly longer than their untreated counterparts. The positive sign of the estimate suggests that the programme provides incentives to work for single mothers of older children. In such a case, the difference between the quality of informal and market care is not dramatic. Additional non-labour income allows those mothers to buy more market care and therefore work longer, whereas married women might have used informal care provided by other members of a household in both periods.

Due to the relatively small number of single-woman households, I am not able to separate the effects with respect to the age of the youngest child. Moreover, estimation on the subsamples of single males is infeasible due to lack of data.

Table 7. Effects on single mothers who did not work in T0

	OLS		Abadie		Heckman		Sant'Anna and Zhao	
	raw	with x	cbps	logit	cbps	logit	drimp	abadie
Probability of working in T1								
	all households							
tau	0.005	0.038	0.037	0.250***	0.064	0.051	0.031	0.064*
	(0.038)	(0.036)	(0.060)	(0.043)	(0.052)	(0.050)	(0.033)	(0.032)
	N= 890, N treated = 137, Specification test for CBPS p-val = 0							
Hours worked								
	all households							
tau	0.333	1.564	2.263	3.446*	2.659	2.217	1.371	2.635*
	(1.448)	(1.375)	(2.299)	(1.342)	(1.984)	(1.941)	(1.263)	(1.250)
	N= 890, N treated = 137, Specification test for CBPS p-val = 0							
I control for the following set of covariates: (1), (2), (3), (4), (5) and labour status of the father (see Table 1). Standard errors in parentheses. * p < 0:05, ** p < 0:01,*** p < 0:001								

Source: Own calculation based on BAEL data.

Conclusion

In this paper, I provide evidence that the introduction of a large-scale child benefit programme had a minor immediate impact on the women's labour supply and almost no immediate effect on the men's labour supply. I find some variation in the mothers' responses to the benefits, though they vary by direction within samples defined by the age of the youngest child in the household and cancel out in the aggregate. More interestingly, the programme seems to have discouraged previously unemployed females from employment. Moreover, all changes in the labour supply attributed to the programme have occurred on the extensive margin. The results do not suggest any major changes in the hours worked. The effects on the fathers are less pronounced. Previously employed fathers maintained their jobs after receiving the benefit. In turn, previously unemployed fathers seem to have increased their labour supply in T_1. That may suggest the programme has confirmed and strengthened the traditional division of roles within a household in which the male is the main earner. Nevertheless, the estimates lack precision, which makes it difficult to draw strong conclusions.

There are some possible explanations for such outcomes. First, as mentioned before, I provide estimates of the immediate effects of the programme's introduction. Households might not have enough time to adjust their behaviour to the presence of additional non-labour income. The process of quitting or finding a job is not immediate and may last over a quarter or two. Second, one should not expect any reaction in the labour supply if households do not internalise the benefit as a long-term stable increase in income. Lastly, the precision of the estimators suffers from the rather small sample size.

References

Abadie A. [2005], Semiparametric difference-in-differences estimators, *The Review of Economic Studies*, 72: 1–19.

Angrist J.D., Pischke J.-S. [2010], The credibility revolution in empirical economics: How better research design is taking the con out of econometrics, *Journal of Economic Perspectives*, 24: 3–30.

Austin P.C. [2011], An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies, *Multivariate Behavioral Research*, 46: 399–424.

Averett S.L., Hotchkiss J.L. [1997], Female labor supply with a discontinuous, nonconvex budget constraint: incorporation of a part-time/full-time wage differential, *Review of Economics and Statistics*, 79: 461–470.

Besley T., Coate S. [1992], Workfare versus welfare: Incentive arguments for work requirements in poverty-alleviation programs, *The American Economic Review*, 82: 249–261.

Black D.A., Kolesnikova N., Sanders S.G., Taylor L.J. [2013], Are children normal?, *Review of Economics and Statistics*, 95: 21–33.

Blundell R., Dias M. C., Meghir C., Reenen J. [2004], Evaluating the employment impact of a mandatory job search program, *Journal of the European Economic Association*, 2: 569–606.

Blundell R., Costa Dias M., Meghir C., Shaw J. [2016a], Female labor supply, human capital, and welfare reform, *Econometrica*, 84: 1705–1753.

Blundell R., Pistaferri L., Saporta-Eksten I. [2016b], Consumption Inequality and Family Labor Supply, *American Economic Review*, 106: 387–435.

Burtless G., Hausman J. A. [1978], The effect of taxation on labor supply: Evaluating the Gary negative income tax experiment, *Journal of Political Economy*, 86: 1103–1130.

Caliendo M., Kopeinig S. [2008], Some practical guidance for the implementation of propensity score matching, *Journal of Economic Surveys*, 22: 31–72.

Card D., Krueger A. B. [1994], Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania, *American Economic Review*, 84: 772–793.

Connelly R. [1992], The effect of child care costs on married women's labor force participation, *The Review of Economics and Statistics,* 74: 83–90.

Connelly R., Kimmel J. [2003], Marital status and full-time/part-time work status in child care choices, *Applied Economics*, 35: 761–777.

Dehejia R. H., Wahba S. [1999], Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs, *Journal of the American Statistical Association*, 94: 1053–1062.

Dehejia R. H., Wahba S. [2002], Propensity score-matching methods for nonexperimental causal studies, *Review of Economics and Statistics*, 84: 151–161.

Ghez G., Becker G. S. [1975], *The Allocation of Time and Goods Over the Life Cycle*, NBER.

Graham J. W., Beller A. H. [1989], The effect of child support payments on the labor supply of female family heads: An econometric analysis, *Journal of Human Resources,* 24: 664–688.

Heckman J. J. [1974], Effects of child-care programs on women's work effort, *Journal of Political Economy*, 82: S136–S163.

Heckman J. J., Ichimura H., Todd P. E. [1997], Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme, *The Review of Economic Studies*, 64: 605–654.

Houngbedji K. [2016], *ABSDID: Stata module to estimate treatment effect with Abadie semiparametric DID estimator*, Statistical Software Components, Boston College Department of Economics.

Hoynes H. W. [1996], Welfare Transfers in Two-Parent Families: Labor Supply and Welfare Participation Under AFDC-UP, *Econometrica*, 64: 295–332.

Imai K., Ratkovic M. [2014], Covariate balancing propensity score, *Journal of the Royal Statistical Society*: Series B (Statistical Methodology), 76: 243–263.

Jacobsen J. P., Pearce III J. W., Rosenbloom J. L. [1999], The effects of childbearing on married women's labor supply and earnings: using twin births as a natural experiment, *Journal of Human Resources,* 34: 449–474.

Kimmel J. [1998], Child care costs as a barrier to employment for single and married mothers, *Review of Economics and Statistics*, 80: 287–299.

Koebel K., Schirle T. [2016], The Differential Impact of Universal Child Benefits on the Labour Supply of Married and Single Mothers, *Canadian Public Policy*, 42: 49–64.

Leuven E., Sianesi B. [2003], *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.

Luna L. G. [2011], *The effects of a universal child benefit*, mimeo.

Magda I., Kielczewska A., Brandt N. [2018), *The Effects of Large Universal Child Benefits on Female Labour Supply*, IZA Discussion Paper No. 11652.

Moffitt R. [1983], An economic model of welfare stigma, *The American Economic Review*, 73: 1023–1035.

Moffitt R. [1990], The Econometrics of Kinked Budget Constraints, *The Journal of Economic Perspectives*, 4: 119–139.

Moffitt R. [2002], Welfare programs and labor supply, *Handbook of Public Economics*, 4: 2393–2430.

Myck M. [2016], *Estimating Labour Supply Response to the Introduction of the Family 500+ Programme*, Centre For Economic Analysis (CenEA) Working Paper Series WP01/16.

Myck M., Trzcinski K. [2019], *From Partial to Full Universality: The Family 500+ Programme in Poland and its Labor Supply Implications*, ifo DICE Report, 17: 36–44.

Nakamura A., Nakamura M. [1992], The econometrics of female labor supply and children, *Econometric Reviews*, 11: 1–71.

Paradowski P. R., Wolszczak-Derlacz J., Sierminska E. [2020], *Inequality, poverty and child benefits: Evidence from a natural experiment*, mimeo, LIS Working Paper Series.

Premik F. [2021], *Estimating the effects of universal transfers: new ML approach and application to labor supply reaction to child benefits*, mimeo.

Rosenbaum P. R.,. Rubin D. B [1983], The central role of the propensity score in observational studies for causal effects, *Biometrika*, 70: 41–55.

Rosenzweig M. R., Schultz T. P. [1985], The Supply and Demand of Births and their Life-Cycle Consequences, *American Economic Review*, 75: 992–1015.

Sant'Anna P. H., Zhao J. [2020], Doubly robust difference-in-differences estimators, *Journal of Econometrics*, 219: 101–122.

Spencer D. [2003], Love's labor's lost? The disutility of work and work avoidance in the economic analysis of labor supply, *Review of Social Economy*, 61: 235–250.

Zhao Z. [2004], Using matching to estimate treatment effects: data requirements, matching metrics, and Monte Carlo evidence, *Review of Economics and Statistics*, 86: 91–107.