Child Care Provision: Semiparametric Evidence from a Randomized Experiment in Mexico.

Pierre Dubois[*] and Marta Rubio-Codina[†]

September 2011[‡]

Abstract

We estimate the impact of the Mexican conditional cash transfer program on the time mothers and older sisters spend taking care of children younger than 3. We exploit the random allocation of the program and heterogeneity in household composition for identification, and apply the methodology in Lewbel (2000). We find that mothers in treatment households substitute their first teenage daughter in the provision of child care. As a result, daughters devote more time to schooling and less time to taking care of their younger siblings. Overall, total household hours to child care increase. These findings indicate that Progresa not only fosters human capital accumulation through keeping teenage girls in school but also through more and arguably better (mother provided) child care.

Keywords: Child Care Provision, Substitution Effects, Semiparametric Estimation, Conditional Cash Transfer Programs.

JEL Codes: D10, J13, J22, I00.

1 Introduction

In recent years, the economic case for public investment in Early Childhood Development (ECD) has become increasingly forceful. Carneiro and Heckman (2003) argue that investments in the early years have higher returns than investments later on in life because of the existence of dynamic complementarities in the learning process: early learning fosters and facilitates later learning. Moreover, remediating early disadvantages later may be prohibitively costly (Cunha and Heckman 2007). In both developed and developing countries, low ECD outcomes—often linked to poor family environments—are associated with inadequate school readiness and lower school performance (Currie 2001; Maluccio et al. 2009; Walker et al. 2005); lower earning capacity (Currie and Thomas 2001; Hoddinott et al. 2008); higher criminality (Carneiro and Heckman 2003; Garces et al. 2002); and lower levels of social integration (Walker et al. 2006). Children growing up with nutritional and psychosocial deficits are also less likely to provide adequate care and resources for their own children, thus contributing to economic inequality (Behrman et al. 2009; Grantham-McGregor et al. 2007; Heckman and Masterov 2007).

The relevance of stimulation and the home environment on ECD is well established empirically (see Walker et al 2007 for a review). However, caregivers may fail to provide adequate care and stimulation if they lack sufficient time, energy, knowledge and money. Because these resources are often scarce in impoverished rural environments, there is scope for Conditional Cash Transfer (CCT) programs—such as the *Progresa* program in Mexico—to improve the circumstances in which children from beneficiary families begin their lives. Even if CCT programs are not specifically designed as ECD interventions per se, the monetary incentives they provide are likely to affect child rearing practices within a household through changes in the intra-household allocation of time to various

activities generated by income and substitution effects. Moreover, the additional benefits they provide—in the form of nutritional supplements, health monitoring, and educational talks covering best health, hygiene and nutrition practices (the so-called "pláticas")—can also affect child care provision through raising awareness and increasing knowledge.

In this paper, we investigate whether *Progresa* affects child care provision through one specific pathway—namely, a re-allocation of time given over to child care amongst household members. We exploit time use data on the randomized *Progresa* evaluation sample to semiparametrically identify the impact of the program on participation and on the extent of participation in child care activities for mothers and sisters (ages 12 to 17) of under 3 year old children. We focus the analysis on mothers and their older daughters as they are the two main child care providers in the household. Moreover, since transfers increase with grade and are larger for girls than boys at secondary school, daughters of secondary school age make the family eligible for receiving the largest transfers, conditional on their school attendance. This strengthens the case for larger economic incentives to enhance substitution effects in the allocation of time devoted to child care between mothers and their older daughters. Estimates support the existence of such substitution effect. We find a 14% increase in mother provided child care in treatment households with teenagers 12 to 17 and children less than 3 years old. In turn, these older daughters reduce their child care participation by 36% and increase their school participation by 10%. Overall, total household time to child care increases, which implies net increases in child care quantity.

The contribution of this study is twofold. Methodologically, we exploit the experimental nature of the *Progresa* evaluation data to obtain a semiparametric estimate of treatment on time allocation. We apply the Lewbel (2000) estimator for qualitative response models to binary and ordered data, and argue that it offers a consistent estimate of

the program impacts on intra-household time allocation decisions. Cogneau and Maurin (2001) and Goux and Maurin (2005) are, to the best of our knowledge, the only existing empirical applications of this method to date. In our case, in addition to the interest in avoiding parametric assumptions, the use of this method is motivated by the seek for robustness in the estimation of limited dependent variables and by the presence of non-classical measurement errors.

In terms of findings, the analysis provides evidence that Progresa increases human capital accumulation both by keeping teenage girls in school and through more and "better" (mother provided) child care, according to the literature on biology, psychology and economics of education. For example, Case and Paxson (2001) and Case et al. (2001) provide empirical evidence of the important role the biological mother—as opposed to the stepmother—plays in the adequate investment in children's health and education. In the case of *Progresa*, beneficiary mothers may also increase their knowledge on parenting through the interaction with medical staff at the health centers and by attending the educational talks. Hence, linking benefits to school attendance can simultaneously improve the quantity and quality of child care provided in the household.

The interest of economists in child care arrangements initially revolved around the responsiveness of female labor supply and child care demand to job related child care subsidizing policies (Heckman 1974; Michalopoulos et al. 1992; Averett et al. 1997). Since Blau and Robins (1988), a number of studies have addressed family labor supply, fertility and child care provision decisions within an intra-household time allocation framework (Mueller 1984; Tiefenthaler 1997). Following the expansion of CCT interventions worldwide, an increasingly extensive literature has developed around the impacts of these programs on child health and nutrition—see Lagarde et al. (2007) for a review—and

more recently, on ECD (Gertler and Fernald 2004; Fernald et al. 2008; Paxson and Schady 2010; Macours et al. 2008). This paper contributes to both literatures by shedding some light on one of the mechanisms—namely, changes in household time allocation—through which CCT programs can affect child care provision and in turn, ECD.

The remainder of the paper is organized as follows. The next section describes the *Progresa* program, the experimental design and the data. In Section 3, we discuss the potential mechanisms through which the program can affect time allocated to child care. In Section 4, we present the Lewbel (2000) semiparametric estimator and discuss identification. Results are in Section 5 and Section 6 concludes.

2   The Rural *Progresa* Program and the Experiment

2.1   The *Progresa* Program

The Mexican government established the *Progresa* CCT program in 1997 in order to break the inter-generational transmission of poverty by alleviating current poverty while investing in the human capital of the next generation.[1] The program provides financial incentives (cash) to parents to invest in the health, nutrition and education of their children. *Progresa* is one of the largest CCT interventions in the world, with approximately 4.5 billion US dollars distributed to some 5.8 million beneficiary households in 2010.[2]

When *Progresa* began rolling out in 1997, program eligibility was determined in two stages (Skoufias et al. 2001). First, underserved communities were identified based on the

---

[1] *Progresa* was renamed *Oportunidades* during the Fox administration in 2000. Because the data used here are from the rural *Progresa* evaluation (1997 to 2000), we refer to the program under its original name throughout the text.

[2] http://www.oportunidades.gob.mx/Portal/wb/Web/poblacion_objetivo_y_montos_asignados _para_el_ejer

proportion of households living in very poor conditions as defined using data from the 1995 census (*Conteo de Población y Vivienda*). Second, low-income households within those communities were chosen by means of a proxy means test constructed using basic socio-economic data collected in the *Encuesta de Características Socioeconómicas de los Hogares* (ENCASEH). This process designated 52% of the households in selected communities as eligible for benefits.[3] All eligible households living in treatment communities were offered *Progresa* and over 90% enrolled. Once enrolled, households received benefits for a three-year period conditional on meeting the program requirements with the possibility of being recertified.

Cash transfers from *Progresa* are given to the female head of the household and are conditional on children attending school, family members obtaining preventive medical care, and female heads attending the "pláticas".[4] They come bimonthly in two forms. The first, received by all beneficiary households, is a fix cash stipend of 90 pesos per month (in 1997 prices) intended for families to spend on more and better nutrition. It is complemented with nutritional supplements and immunization directed to 0 to 2 year olds, and to pregnant and lactating women. The second is an educational grant given to each child younger than 18 and enrolled in school between the third grade of primary school and the third grade (last) of secondary school conditional on attending school a minimum of 85% of the time and on not repeating a grade more than twice. The

---

[3]Subsequently, the Government decided that a subset of households had been unduly excluded and expanded the eligibility criteria to include a set of slightly wealthier households in a process called "densification" (Hoddinott and Skoufias 2004).

[4]This represented an innovation in the Mexican social protection system, which used to give transfers to the head of household (typically a male) up until *Progresa*. However, if a female head of household was not present or able to receive payments, an alternate household member was designated for receipt of payment. Compliance with conditions was verified through the clinics and schools. About 1% of households were denied the cash transfer for noncompliance.

educational scholarship varies by grade and gender. It rises substantially after graduation from primary school and is higher for girls than boys during secondary school. The rates vary from 60 pesos per month for children enrolled in third grade of primary to 225 pesos per month for females enrolled in the third year of secondary school. Hence, households with more female children enrolled in higher grades are eligible for larger transfers compared to similar households with male children enrolled, or children enrolled in lower grades. Beneficiary children also receive money for school supplies once a year during secondary school, and twice a year during primary school.

In order to prevent individual migration into the household, only children who were living in the household at the time of incorporation are eligible for the school transfers. Children born into the household will be eligible for future educational transfers once they reach 9 years old and enter the third grade of primary school. Finally, total transfers for any given household are capped at a pre-determined upper limit of 550 pesos per month. The cap implies that a household cannot receive an unlimited amount by increasing the number of children enrolled in school. It also implicitly implies that enrolling more than three older children (per household) in school will not increase the total amount of transfers received by the household. There appears to have been no effect of the program on fertility rates or family structure (Stecklov et al. 2007).

On average, cash transfers from *Progresa* represent over 20% of total household income. Skoufias (2005) discusses the program at length and provides a review of its impacts.

2.2   Experimental Design and Data

The Mexican Government committed to a rigorous evaluation of the program using a controlled-randomized design. Given budgetary and logistical constraints, the Govern-

ment could not enroll all eligible families in the country simultaneously and had to phase in the enrollment of entire communities over time instead. As part of the program's national scale up, the Government randomly assigned communities in the first seven states to be phased in, to treatment (320 communities) and control (186 communities) groups.

Eligible households in treatment communities began receiving benefits in March/April of 1998, while eligible households in control communities were incorporated in November/December of 1999. In order to minimize anticipation effects, households in control communities were not informed that *Progresa* would provide benefits to them until two months before incorporation. Behrman and Todd (1999) confirm that the original randomization balanced the control and treatment communities, and Attanasio et al. (2011) explicitly test but find no evidence of anticipation effects amongst control households.

The data used in this paper comes from the *Progresa* rural evaluation surveys, the *Encuestas de Evaluación de los Hogares Rurales* (ENCEL), and the ENCASEH baseline data. The ENCEL interviewed all households in the 506 evaluation communities every six months between 1998 and 2000, and again in November 2003. The May 1999 survey collected additional data on time use for all household members older than eight.[5] By then, treatment households had enjoyed benefits for over a year, while no control household had yet received transfers. This allows us to obtain an estimate of the average treatment effect on time devoted to child care.

We construct a dataset of mothers older than 18 years of age living in eligible households in May 1999. We then match each mother to the characteristics and time allocation of her older daughter younger than 18 and still living in the household. We use time

---

[5]Parker and Skoufias (2000) and Rubio-Codina (2010) provide further details on these data. Rubio-Codina (2010) estimates the impacts of the program on intra-household time allocation but does not specifically focus on time to child care.

use data to construct our main dependent variables. Specifically, the question on time devoted to child care reads: "how many minutes did household member i devote yesterday to the care of small children, the elderly or the sick?". To narrow the scope of the question to the care of young children, we restrict the sample of analysis to mothers of children younger than 3 living in households where there are no elderly or sick members that might require care. Given the data available, this implies excluding households with: (i) elders older than 65 that did not engage in any paid or unpaid work activity during the week before the interview; and (ii) members older than 6 that reported being unable to perform regular activities during the month prior to the interview.[6]

The final sample consists of 4,036 mothers (2,571 treatment and 1,465 control) with children younger than 3. This represents 34% of all eligible households originally classified as poor.[7] Approximately 37% of these women are also mothers of a teenager—976 in treatment households and 536 in control households—and 24%—636 treatments and 343 controls—have a 12 to 17 year old daughter. Note that the proportion of mothers with different offspring compositions is balanced in treatment and control households and remains similar to the randomized distribution: 60% treatment and 40% control. This suggests that the potential for sample selection and sample composition biases is negligible.

---

[6]Although the program has improved self-reported health status for children and adults (Gertler and Boyce 2001) as well as children's nutritional status (Rivera et al. 2004; Behrman and Hoddinott 2005), we argue that dropping these households does not bias our estimates because: (i) they only represent 7% and 2% of the households in the estimation sample, respectively; (ii) these proportions are balanced between treatment and control groups; and (iii) parametric estimates are robust to keeping these households in.

[7]We do not use "densified" households—i.e. the set of wealthier and older households that were deemed eligible later—because their process of incorporation in the program is less well documented and many suffered substantial administrative delays in the receipt of benefits (Hoddinott and Skoufias 2004). Because households were categorized as "densified" in both treatment and control communities and under the same criteria, excluding them does not compromise the internal consistency of our estimates.

In Table I we provide further evidence that our sample of analysis is balanced by comparing the means of a set of maternal characteristics in May 1999 (Panel I), and baseline household and community characteristics (Panels II and III) between treatments and controls. The test of equality of means (reported in the last column) shows no statistically significant differences in any of the variables reported, which include offspring composition and baseline household demographics. The only exception is the number of sons ages 12 to 17, which is significantly larger for treatment mothers.

Table I also presents descriptive statistics on the analysis sample. On average, mothers of children younger than 3 are around 30 years old and have 3.5 years of education. Less than 2% are the head of the household, 41% report speaking an indigenous language, and between 6% (control) and 9% (treatment) work for a wage. On average, they have between 1 and 2 children younger than 3, 1 child ages 4 to 7, and 1 child 8 to 17.

3 *Progresa* and Child Care Provision

According to traditional household models, family utility is maximized when household members allocate their time to the production of those commodities in which they have a comparative advantage (Becker 1973). Women's believed comparative advantage in home time would thus explain part of the gender gap in market work participation and female specialization in household activities, including the care of children. Such traditional division of labor between genders is still well established in rural Mexico (INEGI 2002; Parker and Skoufias 2000). In the *Progresa* evaluation sample, 60.2% of prime age (18 to 54) men report working for a wage, as opposed to only 4.8% of prime age women. On the other hand, the proportion of prime age males engaged in domestic chores is significantly lower than that of prime age women. For example, only 2.3% of men versus

84.8% of women report having spent some time cleaning the house during the day before the interview. Similarly, 1.2% of men report having done some cooking in contrast to 88.6% of women.

Child care is also a predominantly female activity. As an illustration, Figure A plots participation rates in child care by age and sex amongst individuals in the *Progresa* evaluation sample. As shown, while participation rates oscillate between 40 and 60% for prime age women, men participation rates are around 8%. A peak is observed for both men and women in their early twenties. Note also that female participation in child care increases sharply from the age of 12. Indeed, while 7.5% of girls 8 to 11 engage in child care activities, this figure increases to an average rate of 14.1% amongst teenage girls (ages 12 to 17). For all boys ages 8 to 17, the average participation rate in child care is constant, at approximately 4%.

These patterns in the allocation of time devoted to child care could be framed into different types of household models, such as the "separate spheres" bargaining model (Lundberg and Pollak 1993), or collective models (Chiappori 1988; Browning et al. 2010), or even a standard unitary household model. Each specific set up would allow modeling the possible effects of *Progresa* on the intra-household allocation of time using interior or corner solutions to account for the fact that child care is a female (wife/mother) activity. The objective of this paper is not to determine which model fits the data better but rather to estimate reduced form solutions (compatible with different theoretical set ups) and identify the impacts of the intervention on time allocated to child care.

We focus on the allocation of time to child care by the mother and her first 12 to 17 years old daughter because: (i) they are the two main caregivers in the household, (ii) households with teenage girls enrolled in secondary school are eligible to receive the largest

transfer amounts, and (iii) the transfer is given directly to the household female head, who is likely to be the mother. All of these suggest that the impacts of the program on child care provision are likely to be largest amongst mothers and their older daughters.[8]

We consider that the mother allocates her time between child care and leisure, and her first daughter's time between child care, schooling, and leisure. She chooses the optimal levels to maximize her utility function—which is a function of total time to child care, her and her daughter's leisure, and her daughter's schooling—subject to a budget constraint, and to her and her daughter's time constraints. In this stylized household, the *Progresa* intervention amounts to:

(i) an increase in the mother's non-labor income given the nutritional grant.

(ii) the provision of a minimum level of maternal care, through increased awareness, knowledge and access, given the required attendance to the "pláticas", preventive health visits and the nutritional supplements.

(iii) a reduction in the price of schooling given the educational grant that the 12 to 17 year old daughter receives conditional on attendance. This implies that time in child care is more expensive relative to time in school for daughters in treatment households.

The budget constraint in treatment households thus integrates the change in the price of schooling and the unconditional nutritional grant. Assuming interior solutions, each one of these intervention components result in:

(i) an ambiguous effect on total maternal child care. If child care is assumed a normal good, maternal child care provision increases with income controlled by the mother

---

[8]We assume that market child care services are unavailable in these disadvantaged rural communities.

through an income effect. However, because child care requires maternal time as an input factor, increases in maternal income might increase her leisure, and hence reduce her time devoted to child care.

(ii) direct increases in the total quantity of child care provided.

(iii.a) increases in the daughter's time to school through the own-substitution effect because of the reduction in the price of schooling.

(iii.b) reductions in the daughter's time allocated to child care through the own-substitution effect, assuming the daughter's child care and schooling times are substitutes.

(iii.c) increases in maternal child care time through the cross-substitution effect, given her and her first daughter's time are substitute inputs in the production of child care.

In all treatment households, maternal time allocated to child care can be affected by the nutritional grant (income effect) and by the compliance with the program requirements (preventive health visits and attendance at "pláticas"). However, cross-substitution effects in child care time (effect iii.c) would only arise amongst those mothers whose daughters are eligible to receive the educational grant. In the empirical exercise, we will exploit the exogenous variation introduced by the random assignment of households (or rather communities) to treatment and control groups, and heterogeneity in mother's offspring to disentangle the cross-substitution effect from the composite of the other two effects (income effect and compliance with program requirements). This composite effect will be a residual in our empirical specification and its components unidentifiable.

4 Estimation and Identification

The central question explored in this paper is to identify empirically whether there have been changes in time allocated to child care in the household as a result of the intervention and understand the underlying mechanisms at play. In this section, we present some preliminary evidence, lay out the empirical specification and discuss identification.

4.1 Preliminary Evidence and Empirical Specification

We start by comparing the means of the dependent variables of interest in the treatment and control groups in May 1999.[9] Results in Table II show that, conditioning on having a teenager aged 12 to 17, mothers in the treatment group have higher participation rates in child care than mothers in the control group although the difference, of 4.9 percentage points, is not statistically significant. For first daughters aged 12 to 17, there is a 6.9 points significant reduction in child care participation and a 7.9 points significant increase in school participation given treatment. The comparison of raw means shows no significant differences in the amount of time devoted to any of the activities considered, conditional on participation.

We next estimate the effect of Progresa on hours to child care provided by mothers and their first teenage daughters, controlling for maternal, household and community characteristics.[10] Table III presents OLS and Tobit (left-censoring at zero) estimates.

---

[9]We cannot test the exogeneity of treatment by comparing baseline time allocation patterns between treatments and controls due to lack of data. Randomization should however guarantee that they were not statistically different.

[10]We include these covariates in the regression to reduce residual variance with respect to the simple comparison of raw means shown in Table II. They are: maternal age, age squared, years of education, ethnicity, head of the household status and whether she is a paid worker; first daughter's education (in years); baseline and contemporary household demographic composition; baseline assets (dirt floor,

Some would argue that OLS is preferable to Tobit on the basis that the zeros in time use data represent infrequencies rather than censoring of observations, given the short period of reference of the data collected (Stewart 2009). As shown in the last row of Table III, 49% of the mothers in control households and 79% of their first 12 to 17 year old daughters report zero time to child care over the day before the interview. As such, we consider the infrequency argument to hold only partially and present results on both OLS and Tobit for comparison purposes. Indeed, there are no dramatic differences (in terms of signs of effects) between both approaches although Tobit offers more precise estimates. Results show that maternal time to child care has increased, even if not significantly (columns 1 and 2), while that provided by the first daughter has decreased (last two columns). The latter effect is significant in Tobit specifications. As shown in columns 3 to 6, the coefficient on the interaction of the treatment dummy with having children ages 12 to 17 is positive and significant (Tobit model), suggesting that mothers substitute their older children in the provision of child care given treatment. The negative coefficient on the presence of teenage offspring in the household further supports this hypothesis.

These findings, consistent with the descriptive statistics on means reported in Table II, are not very significant statistically. Note, however, that these results correspond to average program effects across various treatments. As explained in Section 2, eligible households in treated villages can receive different amounts of transfers depending on family structure and, to a smaller extent, on when they took up the program (variation occurs amongst treated villages because of administrative—i.e. random—delays). Thus, we next estimate the effects of the program controlling for the heterogenous cash transfers

electricity and farm size); and community characteristics (male agricultural wage in the community, distance to large urban center, distance to secondary school, presence of pre-school and presence of junior high school imparted via TV, or "telesecundaria"in the community).

potentially received by a household since the first transfer payment received. This amounts to estimating the "intention to treat" effect but allowing for heterogeneity in treatment.

Denoting v_i the total transfer amount the household has potentially received since taking up the program, we estimate the treatment effects on child care provision using:

$$\tilde{y}_i = \alpha_0 v_i + \alpha_1 T_i + \sum_{r=1}^{R} \gamma_r x_{ri} + \varepsilon_i \tag{1}$$

where $\tilde{y}_i$ is the number of hours individual i allocates to child care, T_i is a binary variable equal to 1 if i lives in an original treatment community and 0 otherwise; v_i is the total transfer amount the household has potentially received since taking up the program; and the $\{x_{ri}\}_{r=1,..,R}$ are R individual, household and community characteristics listed above.

While all poor households in treated villages are eligible for benefits ($T_i = 1$), the program rules are such that the cash benefits they are eligible to get differ notably according to the education levels of children, and because of delays in program take-up across treatment villages. Thus, the total average effect of the program on i is the average of $\alpha_0 v_i + \alpha_1 T_i$. Note that interaction terms between v_i and T_i are implicitly included since, by construction, $v_i T_i = v_i$. This implies that any omitted interaction terms cannot appear in the error term ε_i.

In a second specification, we interact the household treatment status with dummies controlling for the mothers' offspring composition to capture heterogeneous responses across mothers living in different household environments:

$$\tilde{y}_i = \alpha_0 v_i + \alpha_1 T_i + \alpha_2 S_i^j + \alpha_3 T_i S_i^j + \sum_{r=1}^{R} \gamma_r x_{ri} + \varepsilon_i \tag{2}$$

where S_i^j equals 1 if the situation j is true. We consider two possible situations: (i) mother i has daughters or sons aged 12 to 17, and (ii) mother i has daughters aged 12 to 17. As noted in Section 3, cross-substitution effects in maternal time allocated to child

16

care only arise amongst those mothers who live in treatment households ($T_i = 1$) and have school aged children eligible to receive the educational grant ($S_i^j = 1$). We argue that the coefficient on the interaction of treatment and offspring composition, α_3, identifies this average cross-substitution effect. Given the child care participation patterns shown in Figure 1—i.e. females older than 12 are the main child care providers in the household—we expect any cross-substitution effect to take place between mothers and their 12 to 17 year old daughters.

In this specification, the coefficient on the treatment dummy, α_1, is the remaining effect of the program on the mother's time to child care. It combines the effects of the nutritional grant (income effect) and of the compliance with the program requirements (preventive health visits and attendance to the "pláticas") that affect all mothers in treated households—regardless of their offspring composition.[11]

In the event of cross-substitution effects between mothers and first daughters, we should also expect reductions in the amount of time the oldest daughter spends taking care of her younger siblings. To test this, we estimate equation (1) on child care participation of the first daughters—aged 12 to 17—of the mothers in the sample.[12] We additionally estimate equation (1) on their school participation and leisure time so as to obtain a broader picture of their time allocation. Similarly, we also estimate equations (1) and (2) on maternal leisure.

---

[11] These effects could be confounded with other factors—such as maternal education—also correlated with household composition. However, the random allocation of benefits and the fact that the subsample of treatment and control mothers is balanced (see Table I) dismisses such concern.

[12] As shown in Table I, the number of daughters aged 12 to 17 is balanced between mothers in the treatment and control groups. Over 70% of the mothers in the estimation sample have only one daughter 12 to 17 years old.

4.2 Identification and Semiparametric Estimation

If $\tilde{y}_i$ is observed, equations (1) and (2) can be estimated by OLS and offer consistent estimates of all parameters, provided that all right hand side variables are exogenous. Exogeneity of the treatment variable T_i is granted by the randomization; exogeneity of demographic characteristics S_i can be assumed. OLS is still consistent if ε_i includes some classical measurement error on the dependent variable—namely, an additive error independent of all explanatory variables. However, estimating (1) and (2) by OLS using hours of care as observed in the survey as dependent variables, we obtain insignificant parameters α_0 and α_1 (results not shown). Such imprecision is indicative of potential important measurement errors, which are common in time use data.

Indeed, the empirical distribution of the dependent variables of interest, time to child care, hinges to possible non-classical measurement error, such as that resulting from rounding time to an integer number of hours. As shown in Figure 2, the distribution of time (in hours) to child care provided by the mother and her first teenage daughter present a discrete support on specific values, mainly hours and half hours. In such cases, a measurement error ω_i, possibly correlated with explanatory variables, may affect the observed outcome: $\check{y}_i = \tilde{y}_i + \omega_i$. This generates an endogeneity bias in the estimation of (1) or (2) by OLS since ω_i will be correlated with some explanatory variables. The non-independent error ω_i is very problematic in this type of data. While it is not possible to test whether such non- classical measurement error ω_i exists, the empirical distribution of $\check{y}_i$, plotted in Figure 2, with mass points, bunching and multiple modes suggests so. For this reason and in order to get rid of non-classical measurement error, we choose to discretize the observed outcomes even if at the cost of "losing" some information. This amounts to defining an observed binary outcome: $y_i = I_{\{\tilde{y}_i > 0\}}$. Assuming that measure-

ment errors ω_i on the "continuous" variable $\tilde{y}_i$ are such that they do not change this discrete outcome—i.e. assuming that the non-classical measurement error on the number of hours does not set hours to zero when they are non zero and viceversa, or equivalently $I_{\{\tilde{y}_i>0\}} = I_{\{\check{y}_i>0\}}$—this discretization is sufficient to identify coefficients consistently. Then, one can use the binary outcome model $1_{\{\check{y}_i>0\}} = I_{\left\{\alpha_0 v_i+\alpha_1 T_i+\sum_{r=1}^{R} \gamma_r x_{ri}+\varepsilon_i>0\right\}}$ to identify parameters. We also take a less restrictive approach which categorizes total hours into a limited number of intervals defined sensitively to the underlying thresholds in the empirical distribution of hours and guaranteeing enough power in each cell.

Parametric identification of (1) and (2) is possible using standard probit and ordered probit models under the assumption that the error term of the latent variable follows a normal distribution and the normalization of a parameter ($\alpha_0 = 1$, for example). However, parametric identification relies too heavily on the chosen distribution of the error term. In our setting, additional problems arise given that the error term is likely heteroscedastic for two reasons: (i) we are testing for heterogeneous treatment effects across households with different demographic compositions; and (ii) *Progresa* sampled a large number of randomized communities (clusters), each consisting of relatively few poor correlated households. As such, maximum likelihood will only be valid as the number of observations in the cluster tends to infinity with the cluster unit fixed.

We have tested—and generally rejected—normality using Conditional Moments and other standard tests (results available upon request). Consequently, we estimate these discrete choice models semiparametrically. Semiparametric estimation has the advantage of not imposing any particular distribution (normal, logistic, etc.) on the latent variable errors and allow them to suffer from conditional heteroscedasticity of unknown form. We opt for the method of Lewbel (2000) because of the suitability of our data for such purpose.

Indeed, we can exploit the randomization of treatment in the data to justify the partial independence assumption (Assumption A.2 below), which is at the core of the method.

Let us denote y, v and ε the column vectors of y_i and v_i and ε_i, respectively. $X = (T, x_1, .., x_r)$ is the matrix of right hand side variables including T, but without v also called the "special regressor"; and we denote $\beta = [\alpha_1, \gamma_1, ..., \gamma_R]$. Normalizing $\alpha_0 = 1$, Lewbel (2000) considers the binary choice model,

$$y = I_{\{v+X\beta+\varepsilon>0\}} \tag{3}$$

and the following assumptions:

A.1: Continuity: the conditional distribution v given X is continuous.

A.2: Partial Independence: the conditional distribution of ε is independent of v given X,

$F_\varepsilon(\varepsilon|v, X) = F_\varepsilon(\varepsilon|X)$.

A.3: Large Support: the conditional distribution of v given X has support $[v_L, v_H]$ that

contains zero: $v_L \leq 0 \leq v_H$. The support of $-X'\beta - \varepsilon$ is a subset of $[v_L, v_H]$.

A.4: Uncorrelated errors: $E(\varepsilon X) = 0$, as in linear models.

Under assumptions A.1 to A.4, Lewbel (2000) shows that β can be estimated (with root N consistency) by an OLS regression of y^* on X, where

$$y^* = \frac{y - I_{\{v>0\}}}{f(v|X)} \tag{4}$$

and $f(v|X)$ denotes the conditional probability density function of v given X. If the distribution of v is unknown, a nonparametric first stage is needed to estimate it. Alternatively, an ordered data estimator can be used under more stringent conditions. For simplicity and precision in the estimation, we will assume that $f(v|X)$ is normally distributed as in

Lewbel (2006). Lewbel (2000) shows that the proposed methodology extends to ordered response models with K choices defined as:

$$y = \sum_{k=0}^{K-1} k I_{\{\alpha_k < v + X\beta + \varepsilon \leq \alpha_{k+1}\}} \tag{5}$$

where $\alpha_0 = -\infty$ and $\alpha_K = +\infty$. In this case, the transformation of the dependent variable is written as:[13]

$$y^* = \frac{\frac{y}{K-1} - I_{\{v>0\}}}{f(v|X)} \tag{6}$$

In our application, the chosen "special regressor" v_i is the potential transfer amount that the household should have accumulated at the time of analysis (May 1999) since it first received benefits. This potential transfer v_i is determined by the treatment rules in case of compliance. It depends on the household composition and children's school enrollment in 1997 (baseline). It corresponds to the value of cash transfers the household can have obtained since take up assuming it complies with the program rules like school attendance. As discussed, v_i and T_i differ since different treatment households are eligible for different transfer amounts depending on the gender and age of their school-aged children and on whether they are enrolled in school. Since v_i is predicted projecting forward baseline household composition and school enrolment, the required assumption for identification is that pre-program household composition and school enrolment are exogenous to program allocation. This is guaranteed by the randomization and validated by Behrman and Todd (1999) for the evaluation sample, and in Table I for our subsample

---

[13]Lewbel (2000) also proposes an extension to censored data which consists in applying the ordered data estimator repeatedly. More precisely, it involves: (i) defining a continuum of values for α_k (thresholds), (ii) obtaining a $\hat{\beta}_k$ for each threshold defined, and (iii) efficiently combining the $\hat{\beta}_k$'s. Estimating as many $\hat{\beta}'_k$'s as values takes the dependent variable is an unnecessary computational burden. For simplicity and given the empirical distribution of the dependent variable, we chose to categorize total hours into a limited number of intervals, as discussed earlier.

of analysis.

We next discuss why the proposed v_i satisfies the assumptions required for identification. First, the amount of accumulated potential transfers is a continuous variable (A.1) and includes zero in the support (A.3): it is positive for treatment households and zero for control households.

Concerning (A.2), and as defined above, $v_i = T_i^* f^*(D_{i97})$ where $f^*(.)$ is known and determined by the program design as a function of a subset of baseline observables $D_{i97} \subset X_{i97}$, and where t_i^* represents the length of time the household has been receiving benefits for in May 1999. We can rewrite v_i as $v_i = T_i f(D_i) u_i$ where T_i is the randomized treatment status and $f(.)$ is a known function of a subset of May 1999 observables D_i (household demographics, school attendance and grade attended) and $D_i \subset (x_{1i}, .., x_{ri})$. Thus, by construction, $u_i = \frac{t_i^* f^*(D_{i97})}{T_i f(D_i)}$ is an exogenous variable varying according to two random elements: (i) the administrative difficulties that delayed the reception of benefits amongst beneficiaries; and (ii) any departure in household demographics and children's school attendance in May 1999 from the situation predicted using baseline information. Note that $u_i \neq 0$ implies that v_i is a nondeterministic function of other regressors in $x_{1i}, .., x_{ri}$, as is required for identification.

The large support condition (A.3) requires that, given X, the support of $-X'\beta - \varepsilon$ is a subset of $[v_L, v_H]$ for all X. In our particular case, (A.3) is likely to be satisfied everywhere except when the treatment variable (which is amongst X) is equal to zero. In this case, v is equal to zero and thus the support of v is $[v_L, v_H]=\{0\}$ and cannot include the support of $-X'\beta - \varepsilon$. However, we are in a particular case in which identification is restored because the large support assumption is valid on the "treated only" $(T = 1)$ and we can assume that the distribution of ε is independent of T given all other X's.

Magnac and Maurin (2007) show that when the large support assumption is not valid for some X, the parameters β are only set-identified. This means that there are several vectors β and conditional distributions $F_\varepsilon(\varepsilon|v, T, X)$ that can be observationally equivalent. In our case, since the large support assumption is valid when treatment is one ($T = 1$), the identified set of coefficients β that is coherent with the probability $P(y_i = 1|v_i, T_i = 1, X)$ is a singleton (in some sense β can be identified from the treated only). Because of the randomization, we can assume that ε is independent of T and v given $x_1, .., x_r$, then $F_\varepsilon(\varepsilon|v, T, x_1, .., x_r) = F_\varepsilon(\varepsilon|x_1, .., x_r)$ which means that the distribution function of ε is independent of whether the household is being treated or not. Then, the method allows to nonparametrically identify the distribution of ε on the treated ($T = 1$) and thus all parameters.

Assumptions (A.1) to (A.4) imply that the conditional probability of success, $pr(y = 1|v, X)$, increases monotonically and varies from 0 to 1 over the support $[v_L, v_H]$ of v. As this is admittedly very restrictive in empirical applications, Magnac and Maurin (2007) propose an alternative assumption to (A.3): a symmetry condition on the tails of the errors ε. Let $y_{v_L} = X'\beta + v_L + \varepsilon$ be the propensity of success for individuals with the smallest v, v_L; and $y_{v_H} = -(X'\beta + v_H + \varepsilon)$ the propensity of failure for individuals with the largest v, v_H. Then, the symmetry condition can be expressed as:

A.5: $E(X'y_{v_L} I_{\{y_{v_L} > 0\}}) = E(X'y_{v_H} I_{\{y_{v_H} > 0\}})$

A.5 requires that the propensity of success y_{v_L} (or $pr(y = 1|v_L, X)$) and the propensity of failure y_{v_H} (or $pr(y = 0|v_H, X)$) are identically distributed. If so, the Lewbel (2000) estimator is unbiased. If symmetry of the tails is not satisfied, it is always possible to choose conditional distributions for y_{v_L} and y_{v_H}—by trimming outliers in the distribution of v—in such a way that symmetry is more likely satisfied (Magnac and Maurin 2007).

Moreover, Khan and Tamer (2010) have shown that the rate of convergence of the Lewbel (2000) estimator can be slower than the parametric one and that numerical instability can happen depending on tail distributions. We have thus devoted some particular attention to the support condition and have tested the stability and robustness of our estimator to trimming parameters (results available upon request).

Finally, note that in this setting, monotonicity of the conditional probability of success $pr(y = 1|v, X)$ over the support of v amounts to assuming that: (i) the mother's child care time and the first daughter's schooling time are nondecreasing functions of the cumulative potential transfers v_i; and (ii) the first daughter's child care time is a nonincreasing function of v_i. We have discussed the theoretical validity of these assumptions in Section 3 and have "tested" its empirical validity (results available upon request).

5   Results and Discussion

We estimate the effect of *Progresa* on participation in child care, and on the extent of participation in child care and leisure for mothers of children under 3 (Tables IV and V) and for the older daughter—ages 12 to 17—of these mothers. We define leisure as 24 hours minus total time devoted to other (work and non-work) activities. It includes time spent sleeping, eating and socializing. For the first daughter, we also estimate the program impact on participation and on the extent of participation in school (Tables VI and VII). In each table we first report OLS estimates (Models A), then probit or ordered probit estimates depending on the nature of the dependent variable (Models B), and finally the semiparametric (Lewbel) estimates (Models C).

All regressions include the explanatory variables listed in Section 4 but estimates are robust to the exclusion of covariates. We also trim extremely low values of the conditional

probability density function of v_i as they imply, by construction, outlier observations of the transformed dependent variable y^*.

5.1 Maternal Time

Table IV presents estimates of the impact of the program on maternal participation in child care (binary outcome) and Table V on their extent of participation (ordered discrete outcome). The latter variable takes six different values: $k \in \{0, 1, ..., 5\} = \{$no time devoted to child care, up to one hour, between one and two hours, two and four hours, four to seven, seven to thirteen$\}$.[14]

Estimation results follow very similar patterns for both types of outcomes. Models 1A to 1C in Table IV show no significant effect of treatment on maternal participation in child care when all mothers are pooled together neither parametrically nor semiparametrically. However, when treatment is interacted with whether the mother has offspring aged 12 to 17 (Models 2A to 2C), the effect becomes positive and significant. Moreover, the coefficient on having 12 to 17 year old children alone is negative and significant. We interpret these findings as indicative of: (i) a cross-substitution effect between mothers and their older children in child care provision; and (ii) *Progresa* attenuates the cross-substitution effect. The semiparametric mean marginal effect of *Progresa* on child care participation for mothers with children younger than 3 and 12 to 17 teenagers is 7.15 percentage points (point estimate of 0.44 as shown in Model 2C).[15] This results in a 13.9

---

[14]Results are robust to a redefinition of these categories.

[15]For dichotomous outcomes, we compute the marginal effects on the estimated coefficients as: $M_{ij} = \frac{\triangle[1-\hat{G}(-v_i-x_i\hat{\beta})]}{\triangle x_j} = \frac{\triangle[1-\hat{G}(z_i\hat{\beta})]}{\triangle x_j} = \hat{G}(\hat{z}_i^0) - \hat{G}(\hat{z}_i^1)$, where M_{ij} is the effect of switching the jth binary variable, x_j, from 0 to 1 on the probability that y_i equals 1; and $\hat{z}_i^1$ is the value of the index $-v_i - x_i\hat{\beta}$ when the j-th binary variable is set to 1, and similarly for $\hat{z}_i^0$ when the value of j is set to 0. $\hat{G}(.)$ is the estimated cumulative distribution function of the probability of y_i given z_i. We nonparametrically

percentage increase in child care participation, given an initial participation rate of 51.4% amongst these women. We further interact treatment with a dummy equal to 1 if the mother has a daughter—as opposed to a child—aged 12 to 17 and find that the coefficient on this interaction is also positive and significant (Model 3C in Table IV).

Panel I in Table V reports similar effects on the extent of participation in child care. Mothers of children (daughters) 12 to 17 are more likely to increase the amount of time they spend taking care of their younger children given treatment (Model 2C (3C) in Panel I). Consistently, the analysis on the extent of maternal leisure in Panel II shows that these women enjoy less leisure given treatment. This reduction in their leisure time is partly explained by increases in time to child care, to other domestic activities and to compliance with the program requirements.

The qualitative evidence in Adato et al. (2000) endorses these findings. During the summer of 1999, the authors conducted focus groups with 230 beneficiary and non-beneficiary women to learn about their perceptions on the program. The authors report: "Another reason that women's time burden increases is because of the need to do work that was previously done by children who are now attending school, particularly secundaria. However, their mothers see this as worthwhile in order for their children to study. (...) Although some women said that the father also does some of this work, more often it was the mother." (Adato et al. 2000, p. xiii).

5.2 First Daughter's Time

Next, we turn our attention to the allocation of time by first daughters. Panels I in Tables VI and VII present the estimated program impact on their participation and extent of

---

estimate $\hat{G}(.)$ running the kernel regression of y_i on z_i. Results reported are computed using a Gaussian Kernel and 500 equally spaced points in the range of z_i and are robust to larger numbers of points.

participation to child care. The latter variable takes four different values: $k \in \{0, 1, 2, 3\} =$ {no time devoted to child care, up to one hour, between one and two hours, more than two hours}.

The semiparametric estimates in Table VI evidence a reduction of 6.3 percentage points in the first daughter's probability of engaging in child care activities, or a 35.8% decrease in participation (estimated coefficient of -0.031, Model C in Panel I). Significant reductions are also observed in their extent of participation (Model C in Panel I, Table VII). A plausible concern is that these reductions are in fact driven by an increase in school attendance of other siblings in primary school age. While this hypothesis is not testable given that almost all teenage girls with siblings younger than 3 also have siblings in primary school age (6 to 11), it is well-known that the program had little effect on primary school enrolment (Schultz 2004).

About 69% of girls in the sample living in households with children younger than 3 that engage in child care activities do not attend school: of these, 3% report that they are not in school because they have to help in the house.[16] Moreover, conditional on enrollment, another 4% report having to take care of their siblings as one of the reasons why they miss school. Other more frequent reasons are teacher absenteeism, illness and care of the sick.

These figures are somewhat indicative that child care and schooling are substitute activities for teenage girls. Not surprisingly, we find a significant increase in school participation (attendance and doing homework) for first daughters aged 12 to 17 (Model C in Panel II, Table VI). The estimated coefficient of 0.07 translates into a marginal effect of 4.1 percentage points, or a 9.8% increase in schooling. This is consistent with the positive

---

[16] Other, more common, reasons why they are not in school are the lack of money (47%), that they do not like it (16%), or that the school is too far (8%).

program impact on female secondary school enrolment reported in Schultz (2004) and Parker and Skoufias (2000). Lewbel estimates in Panels II and III in Table VII show that while the increase in the extent of time allocated to schooling activities by first daughters aged 12 to 17 is not significant, the reduction in the extent of leisure time is. This might suggest that these girls reduce (or stop) their contribution to child care in the household to take up school full time, as a result of the reduction in the price of schooling. Alternatively or in addition, the intervention may be affecting the allocation of teenage girls' time to other activities and in particular, to other household chores as documented in Rubio-Codina (2010).[17]

5.3   Quantity and Quality of Care

We start this subsection by investigating whether mothers (and possibly other household members) just or more than substitute for their daughters' child care provision. To this end, we estimate equation (2) on total hours devoted to child care at the household level. We define total household hours to child care as the sum of hours spent in child care by each household member and categorize the variable as $k = \{0, 1, ..., 6\} = \{$no time allocated to child care, up to one hour, one to two hours, two to three hours, three to five, five to seven, more than seven$\}$.

If mothers and their 12 to 17 year old daughters are the only two household members taking care of the very young in the household and if mothers fully substitute for the time to child care previously provided by their older daughters, then we should expect no effect on total household hours. On the other hand, if mothers (and possibly other

---

[17]As a robustness check, we re-estimate all previous regressions on the restricted sample of mums and older daughters whose time information refers to weekdays. The estimated coefficients (available upon request) are very similar—albeit less precisely estimated—to the ones reported here.

household members) increase the amount of time they devote to child care by more than the amount previously devoted by teenage girls, then α_3 in equation (2) should be positive and significant. A negative and significant α_3 would imply the opposite effect and as a consequence, a reduction in total household hours to child care in treatment households.

Estimates in Table VIII show a significant increase in the extent of total household hours to child care in treatment households with children under 3 and teenagers aged 12 to 17. This implies that *Progresa* fosters net increases in total child care provided within the household.

The next natural question is whether mothers alone are more than compensating for care time previously provided by their older daughter or whether other household members are also contributing. Table IX shows results from estimating equation (2) using the share of child care hours provided by household member m over total household hours as the dependent variable. We consider the following groups of household members: the mothers of 0 to 3 years old; other adult women; adult men; brothers aged 8 to 11 and 12 to 17; and sisters aged 8 to 11 and 12 to 17 of under 3 year old children.

Results are consistent with the premise that the mother is the household member providing a larger share of child care in substitution of the older daughter and overall. The coefficient on treatment interacted with girls (daughters) 12 to 17 on the share of mother's time is positive and almost significant at the 10%. The coefficient on treatment for the share of time devoted to child care for daughters aged 12 to 17 continues to be negative and significant. Surprisingly, daughters aged 8 to 11—who are the sisters of the 12 to 17 year old daughter—also seem to increase their share of child care provision in treatment households. Because of the low participation rates for this group (around 3%) we are inclined to think that this effect is driven by a few outliers.

Hence, and as long as we believe that the mother is more productive in the provision of child care, these findings are suggestive that the program involves gains not only in the quantity but also in the quality of the care given to the very young. The biology and psychology literature have repeatedly acknowledged the mother as the best child nurturer. Frequent breast-feeding and mother warmth are widely recognized as key care practices (UNICEF 2001). Variations in the quality of maternal care are also proven to produce lasting changes in stress reactivity, anxiety, and memory function in the offspring (Grantham-McGregor et al. 2007). The works of Case and Paxson (2001) and Case et al. (2001) are examples in the economic literature of the important role the biological mother—as opposed to the stepmother—plays in the adequate investment in the child's health and education. Moreover, beneficiary mothers increase their knowledge on parenting both through the interaction with medical staff at the health centers and by attending the "pláticas".

On the other hand, one could argue that environmental factors (the education of the caregiver, for example) matter more than biological attachment. In the current context, this would imply that older daughters are better caregivers as they are, on average, more educated than their mothers. However, when we interact the treatment dummy with years of education of the mother we find stronger substitution effects amongst more educated mothers (results available upon request). This result confirms that the increase in quantity of care is, on average, likely to go hand in hand with increases in quality.[18]

---

[18]We are using more recent rounds of the *Progresa* evaluation data to investigate whether increases in maternal time to child care do improve child cognitive and non-cognitive development. Because we have to rely on a different identification strategy, we consider this issue to be beyond the scope of this paper.

6   Conclusion

This paper provides semiparametric estimates of the effect of *Progresa* on child care provision. We have first shown that child care is almost exclusively provided by females in the rural communities where the program operates and have argued that the intervention might lead to increases in the quantity of care provided within the household. The nutritional supplements, health checkups and "pláticas" result in direct increases in the quantity of child care mothers provide and young children receive. Moreover, the conditional-on-attendance education grants result in a reallocation of time to "better paying" activities given the change in the relative shadow values of household members' time it entails. As a consequence, increases in maternal care in substitution for care previously provided by her older daughter were expected. These increases can arguably result in more and better care if the mother is assumed to be a better caregiver.

We have applied the Lewbel (2000) semiparametric method and provided evidence in support of such mother-daughter substitution effect in the provision of child care. While mothers with 0 to 3 and 12 to 17 year old children are significantly less likely to participate in child care, this behavior is reversed given treatment. In addition, older daughters devote their freed up time to schooling. We also observe an increase in total household time given over to child care in these households. These findings suggest that—by linking benefits to school attendance—the *Progresa* program fosters human capital accumulation both through keeping teenage girls in school and through more and "better" child care. Increased maternal care is likely to lead to better development in the early ages and increased school readiness. Note that it would have been unfeasible to increase the levels of child care provided by directly conditioning the reception of benefits to maternal time allocated to child care, as it is not possible to monitor how much time mothers spend

with their children.

Interestingly, we have not found significant program impacts on child care provision amongst mothers of children younger than 12 nor on the pooled sample of mothers. This suggests two things: first, the program mainly alters the household allocation of time devoted to child care through the reduction in the price of schooling (educational grant). Second, the educational talks and preventive care do not influence maternal child care provision as much as desired. Gertler and Fernald (2004) point at the inadequate development of the "pláticas" as an explanation of the inexistent program impacts on cognitive development. An alternative explanation could be that mothers in control communities also attend these talks, which would confound the estimated treatment effect. However, it seems unlikely that households in control communities know about them, let alone travel to treatment communities to attend. In any event, and given the important role of ECD in long run individual and societal welfare, CCT programs could re-consider introducing more intense parental and community training activities oriented to promote child stimulation and early education.

References

[1] Adato, Michelle, Bénédicte de la Brière, Dubravka Mindek and Agnes Quisumbing (2000), "The Impact of Progresa on Women's Status and Intrahousehold Relations", IFPRI Report, Washington D.C.

[2] Attanasio, Orazio, Costas Meghir and Ana Santiago (2011), "Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate Progresa", forthcoming in *Review of Economic Studies*.

[3] Averett, Susan L., Elizabeth H. Peters and Donald M. Waldman (1997), "Tax Credits, Labor Supply, and Child Care", *Review of Economics and Statistics*, 79(1):125-135.

[4] Becker, Gary S. (1973), "A Theory of Marriage: Part I", *Journal of Political Economy,* 81(4): 813-846.

[5] Behrman, Jere R. and Petra E. Todd (1999), "Randomness in the experimental samples of PROGRESA (education, health, and nutrition program)", IFPRI Report, Washington D.C.

[6] Behrman, Jere R. and John Hoddinott (2005), "Programme Evaluation with Unobserved Heterogeneity and Selective Implementation: The Mexican PROGRESA Impact on Child Nutrition", *Oxford Bulletin of Economics and Statistics*, 67(4): 547-569.

[7] Behrman, Jere R., María C. Calderon, Samuel H. Preston, John Hoddinott, Reynaldo Martorell and Aryeh D. Stein (2009), "Nutritional Supplementation of Girls Influences the Growth of their Children: Prospective Study in Guatemala", *American Journal of Clinical Nutrition*, 90(5):1372-1379.

[8] Blau, David M. and Philip K. Robins (1988), "Child-Care Costs and Family Labor Supply", *Review of Economics and Statistics*, 70(3):374-381.

[9] Browning Martin, Pierre A. Chiappori and Valérie Lechene (2010), "Distributional effects in Household Models: Separate Spheres and Income Pooling", *Economic Journal*, 120(6): 786-799.

[10] Carneiro, Pedro and James J. Heckman (2003), "Human Capital Policy", in *Inequality in America: What Role for Human Capital Policies*, James Heckman and Alan Krueger, eds. MIT Press. Cambridge, MA.

[11] Case, Anne, I-Fen Lin and Sara McLanahan (2001), "Educational Attainment in Blended Families", *Evolution and Human Behavior*, 22(4): 269-289.

[12] Case, Anne and Christina Paxson (2001), "Mothers and Others: Who Invests in Children's Health?", *Journal of Health Economics*, 20(3):301-328.

[13] Chiappori, Pierre A. (1988), "Rational household labor supply", *Econometrica*, 56(1): 63-89

[14] Cogneau, Denis and Eric Maurin (2001), "Parental income and school attendance in a low-income country: a semi-parametric analysis", DIAL Working Paper, 2001/16.

[15] Cunha, Flavio and James J. Heckman (2007), "The Technology of Skill Formation", *American Economic Review*, 97(2):31-47.

[16] Currie, Janet (2001), "Early Childhood Education Programs", *The Journal of Economic Perspectives*, 15(2):213-238.

[17] Currie, Janet and Duncan Thomas (2001), "Early Test Scores, Socioeconomic Status and Future Outcomes", *Research in Labor Economics*, 20:103-132.

[18] Fernald, Lia C., Paul J. Gertler and Lynette M. Neufeld (2008), "Role of cash in conditional cash transfer programmes for child health, growth, and development: an analysis of Mexico's *Progresa*", *The Lancet*, 371(9615):828-37.

[19] Garces, Eliana, Duncan Thomas and Janet Currie (2002), "Longer-Term Effects of Head Start", *The American Economic Review*, 92(4):999-1012.

[20] Gertler, Paul J. and Simone Boyce (2001), "An Experiment in Incentive-Based Welfare: The Impact of PROGRESA on Health in Mexico", mimeo UC Berkeley.

[21] Gertler, Paul J. and Lia C. Fernald (2004), "The Medium Term Impact of *Oportunidades* on Child Development in Rural Areas", mimeo UC Berkeley.

[22] Goux, Dominique and Eric Maurin (2005), "The effect of overcrowded housing on children's performance at school", *Journal of Public Economics*, 89(5-6):797-819.

[23] Grantham-McGregor, Sally, Yin Bu Cheung, Santiago Cueto, Paul Glewwe, Linda Richter, Barbara Strupp and the International Child Development Steering Group (2007), "Developmental Potential in the First 5 Years for Children in Developing Countries", *The Lancet*, 369(9555):60-70.

[24] Heckman, James J. (1974), "Effects of Child-Care Programs on Women's Work Effort", *Journal of Political Economy*, 82(2):S136-S163.

[25] Heckman, James J. and Dimitri V. Masterov (2007), "The Productivity Argument for Investing in Young Children", *Review of Agricultural Economics*, 29(3):446-493.

[26] Hoddinott, John and Emmanuel Skoufias (2004), "The Impact of OPORTUNIDADES on Consumption", *Economic Development and Cultural Change*, 53(1):37-61.

[27] Hoddinott John, John A. Maluccio, Jere R. Behrman, Rafael Flores and Reynaldo Martorell (2008), "Effect of a Nutrition Intervention during Early Childhood on Economic Productivity in Guatemalan Adults", *The Lancet*, 371(9610):411-416.

[28] INEGI (2002), *Uso del Tiempo y Aportaciones en los Hogares Mexicanos*, Instituto Nacional de Estadística (INEGI), Aguascalientes.

[29] Khan, Shakeeb and Elie Tamer (2010), "Irregular Identification, Support Conditions, and Inverse Weight Estimation", *Econometrica*, 78(6):2021-2042.

[30] Lagarde, Mylene, Andy Haines and Natasha Palmer (2007), "Conditional Cash Transfers for Improving Uptake of Health Interventions in Low- and Middle-Income Countries: A Systematic Review", *Journal of the American Medical Association*, 298(16):1900-1910.

[31] Lewbel, Arthur (2000), "Semiparametric Qualitative Response Model Estimation With Unknown Heteroscedasticity or Instrumental Variables", *Journal of Econometrics*, 97(1): 145-177.

[32] Lewbel, Arthur (2006), "Simple Endogenous Binary Choice and Selection Panel Model Estimators", Boston College Working Papers in Economics No. 613.

[33] Lundberg, Shelly and Robert A. Pollak (1993), "Separate Spheres Bargaining and the Marriage Market", *Journal of Political Economy*, 101(6):988-1010.

[34] Macours, Karen, Norbert Schady and Renos Vakis (2011), "Cash Transfers, Behavioral Changes, and Cognitive Development in Early Childhood: Evidence from a Randomized Experiment", forthcoming in *American Economic Journal: Applied Economics*.

[35] Maluccio John A., John Hoddinott, Jere R. Behrman, Agnes Quisumbing, Reynaldo Martorell and Aryeh D. Stein (2009), "The Impact of Nutrition During Early Child-

hood on Education among Guatemalan Adults", *Economic Journal*, 119(537):734–763.

[36] Magnac, Thierry and Eric Maurin (2007), "Identification and Information in Monotone Binary Models", *Journal of Econometrics*, 127(1):76-104.

[37] Michalopoulos, Charles, Philip K. Robins and Irwin Garfinkel (1992), "A Structural Model of Labor Supply and Child Care Demand", *Journal of Human Resources*, 27(1):166-203.

[38] Mueller, Eva (1984), "The value and allocation of time in rural Botswana", *Journal of Development Economics*, 15(1-3):329-360.

[39] Parker, Susan and Emmanuel Skoufias (2000), "The Impact of Progresa on Work, Leisure, and Time Allocation", Final Report, IFPRI, Washington, D.C.

[40] Paxson, Christina and Norbert Schady (2010), "Does Money Matter? The Effects of Cash Transfers on Child Health and Development in Rural Ecuador", *Economic Development and Cultural Change*, 59(1):187-229.

[41] Rivera, Juan A., Daniela Sotes-Alvarez, Jean-Pierre Ilabricht, Teresa Shamah and Salvador Villalpando (2004), "Impact of the Mexican Program for Education, Health, and Nutrition (Progresa) on Rates of Growth and Anemia in Infants and Young Children", *Journal of the American Medical Association*, 291(21):2563-2570.

[42] Rubio-Codina, Marta (2010), "Intra-household Time Allocation in Rural Mexico: Evidence from a Randomized Experiment", *Research in Labor Economics*, 31: 219-257

[43] Schultz, Paul (2004), "School Subsidies for the Poor: Evaluating the Mexican PRO-GRESA Poverty Program", *Journal of Development Economics*, 74(1): 199-250.

[44] Skoufias, Emmanuel (2005) "PROGRESA and Its Impacts on the Welfare of Rural Households in Mexico", IFPRI Report 139, Washington D.C.

[45] Skoufias, Emmanuel, Benjamin Davis and Sergio de la Vega (2001), "Targeting the Poor in Mexico: An Evaluation of the Selection of Households into PROGRESA", *World Development*, 29(10):1768-1784.

[46] Stecklov, Guy, Paul Winters, Jessica Todd and Ferdinando Regalia (2007), "Unintended effects of poverty programmes on childbearing in less developed countries: Experimental evidence from Latin America." *Population Studies*, 61(2): 125-140.

[47] Stewart, Jay (2009), "Tobit or Not Tobit?", IZA Discussion Paper No. 4588, Bonn.

[48] Tiefenthaler, Jill (1997), "Fertility and Family Time Allocation in the Philippines", *Population and Development Review*, 23(2):377-397.

[49] UNICEF (2001), *The State of World's Children*: UNICEF, New York.

[50] Walker Susan P., Susan M. Chang-Lopez, Christine A. Powell and Sally M. Grantham-McGregor (2005), "Effects of early childhood psychosocial stimulation and nutritional supplementation on cognition and education in growth-stunted Jamaican children: prospective cohort study", *The Lancet*, 366(9499):1804-1807.

[51] Walker Susan P., Susan M. Chang, Christine A. Powell, Emily Simonoff and Sallly M. Grantham-McGregor (2006), "Effects of psychosocial stimulation and dietary supplementation in early childhood on psychosocial functioning in late adolescence: follow-up of randomized controlled trial", *British Medical Journal*, 333(7566): 472.

[52] Walker, Susan P., Theodore D. Wachs, Julie Meeks Gardner, Betsy Lozoff, Gail A. Wasserman, Ernesto Pollitt, Julie A. Carter and the International Child Development Steering Group (2007), "Child Development: Risk Factors for Adverse Outcomes in Developing Countries", *The Lancet*, 369(9556):145-157.

FIGURES and TABLES
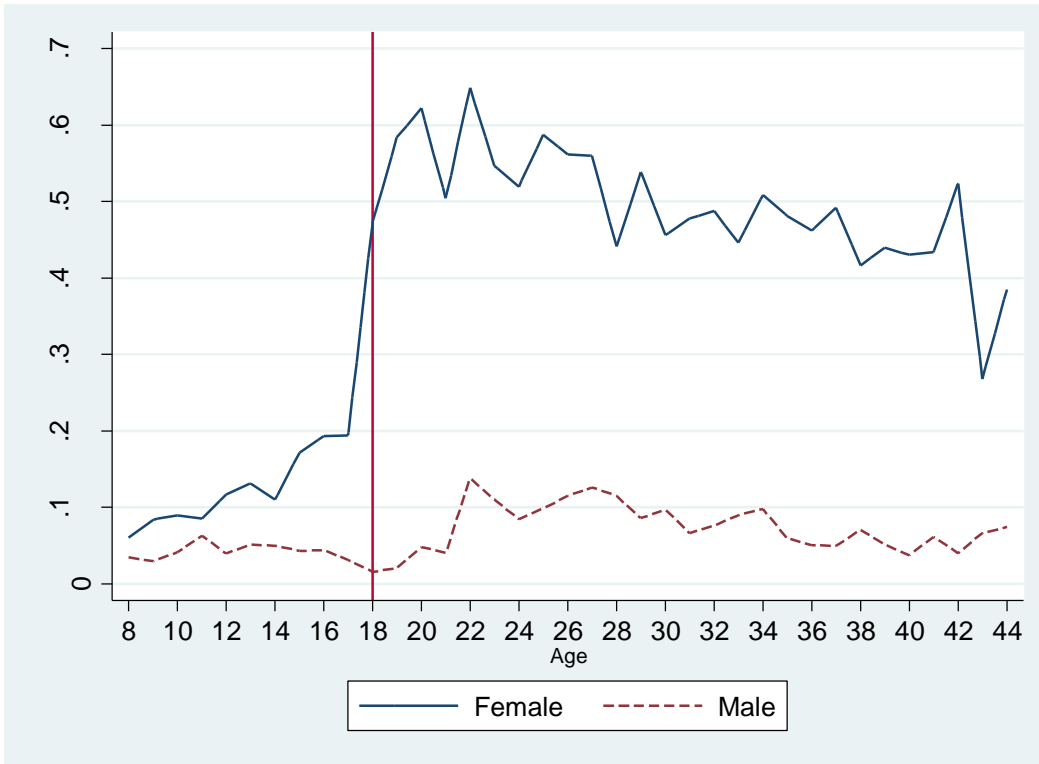### Figure 1: Child Care Participation Rates by Sex and Age

Figure 2: Distribution of Hours to Child Care by Mothers and First Daughters (Ages 12 to 17)
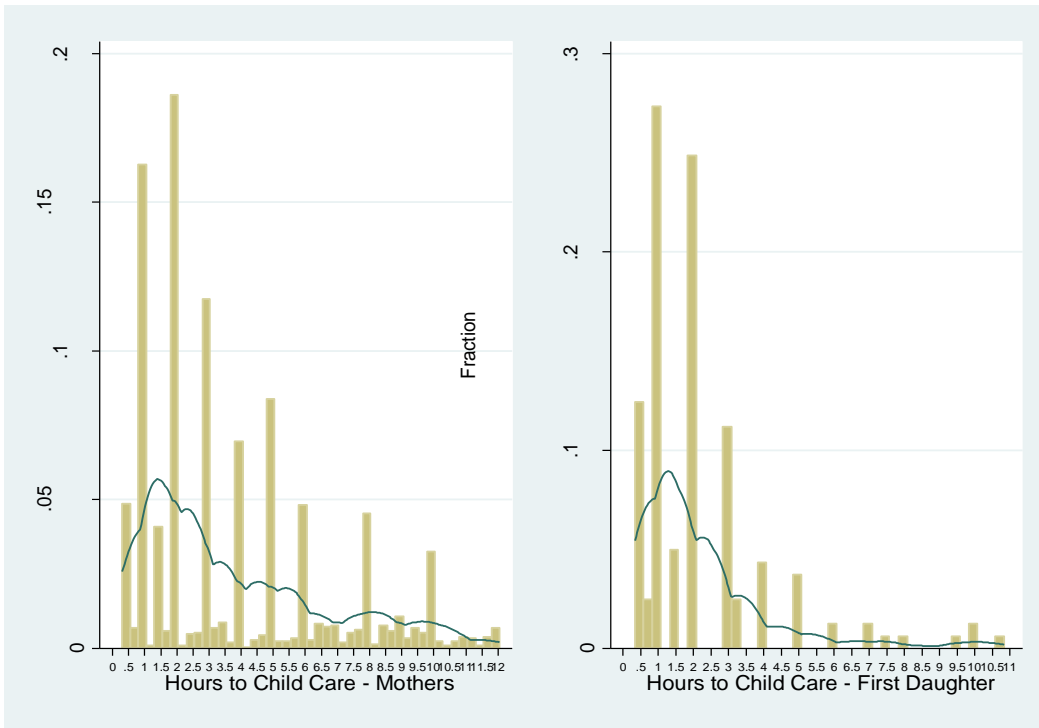
Table I: Balance between Treatment and Control Groups

	Treatment Group			Control Group			
	N	Mean	SD	N	Mean	SD	t-stat
I. *Current Maternal Characteristics* *(May 1999)*							
Age	2571	29.66	6.418	1465	29.57	6.451	0.412
Years of Education	2534	3.55	2.720	1447	3.51	2.781	0.224
Household Head =1	2537	1.89	0.136	1447	1.94	0.138	-0.094
Indigenous =1	2534	40.84	0.492	1445	43.04	0.495	-0.374
Working for a Wage =1	2534	8.72	0.282	1446	5.95	0.237	1.601
Number of Children 0 to 3	2571	1.25	0.465	1465	1.25	0.476	-0.382
Number of Children 4 to 7	2571	1.05	0.798	1465	1.05	0.814	0.266
Number of Sons 8 to 11	2571	0.42	0.624	1465	0.38	0.605	1.500
Number of Sons 12 to 17	2571	0.37	0.697	1465	0.32	0.615	2.390*
Number of Daughters 8 to 11	2571	0.40	0.604	1465	0.40	0.614	-0.233
Number of Daughters 12 to 17	2571	0.33	0.644	1465	0.32	0.631	0.588
Years of Education First 12 to 17 Year Old Daughter	2571	1.38	2.589	1465	1.29	2.512	0.918
II. *Baseline Household Characteristics* *(October1997)*							
Number of Children 0 to 3	2571	1.38	0.732	1465	1.41	0.740	-1.116
Number of Children 4 to 7	2571	1.09	0.851	1465	1.07	0.847	0.571
Number of Teenagers 8 to 17	2571	0.75	1.046	1465	0.72	1.002	0.837
Number of Adults 18 to 54	2571	2.26	0.779	1465	2.26	0.827	-0.284
Number of Adults Over 55	2571	0.14	0.415	1465	0.13	0.381	0.657
Electricity =1	2568	61.06	0.488	1464	62.09	0.485	-0.219
Dirtfloor =1	2563	71.91	0.450	1460	76.23	0.426	-1.503
Animal and Land (more than 3 ha) Ownership =1	2561	29.48	0.456	1463	31.99	0.467	-0.854
III. *Baseline Community Characteristics* *(October1997)*							
Pre-school =1	2514	91.45	0.280	1417	91.11	0.285	0.123
Junior Secondary Education Imparted via TV =1	2571	21.78	0.413	1465	25.26	0.435	-0.608
Health Center =1	2571	76.66	0.423	1465	82.87	0.377	-1.568
Distance to Closest Secondary School (Km)	2571	2.45	2.171	1465	2.67	2.782	-0.562
Minimum Distance to Large Urban Centre (Km)	2571	105.46	44.046	1465	101.52	47.534	0.716
Monthly Community Agricultural Male Wage (pesos)	2506	6.27	0.313	1395	6.29	0.304	-0.572

Notes: *Significant at 5%. Sample of mothers 18 to 44 with kids under 3 years old living in households, originally classified as poor, with no elder older than 65 not working and no sick individuals.

Table II: Comparison of Raw Means of Dependent Variables

	Treatment Group			Control Group			
	N	Mean	SD	N	Mean	SD	t-stat
Dependent Variables							
Participation in Child Care (All Mothers) =1	2571	50.80	0.500	1465	50.99	0.500	-0.075
Participation in Child Care (Mothers of 12 to 17 Teenagers) =1	973	47.28	0.500	536	42.35	0.495	1.421
Participation in Child Care (Mothers of 12 to 17 Teenage Girls) =1	636	49.37	0.500	343	44.90	0.498	1.107
Participation in Child Care (First Daughter) =1	626	14.22	0.350	337	21.07	0.408	-2.127*
Participation in School (First Daughter) =1	626	45.69	0.499	336	37.80	0.486	1.826+
Hours in Child Care (All Mothers)†	1306	3.76	2.862	747	3.74	2.854	0.088
Hours in Child Care (Mothers of 12 to 17 Teenagers)†	460	3.41	2.681	227	3.36	2.681	0.202
Hours in Child Care (Mothers of 12 to 17 Teenage Girls)†	314	3.32	2.703	154	3.39	2.631	-0.247
Hours in Child Care (First Daughter)†	89	2.31	2.247	71	2.06	1.497	0.846
Hours in School (First Daughter)†	286	6.15	1.586	127	6.00	1.471	0.969
Total Household (Hh) Hours in Child Care (All Hhs)†	1353	4.04	3.086	783	4.03	3.028	0.047
Total Hh Hours in Child Care (Hhs with 12 to 17 Teenagers)†	500	4.05	3.276	248	4.16	3.183	-0.420
Total Hh Hours in Child Care (Hhs with 12 to 17 Teenage Girls)†	345	4.03	3.435	170	4.38	3.241	-1.082

Notes: +Significant at 10%; *Significant at 5%. Sample of mothers 18 to 44 with kids under 3 years old living in households, originally classified as poor, with no elder older than 65 not working and no sick individuals.
†Conditional on being positive.

Table III: Parametric Evidence on Maternal and First Daughters (Ages 12 to 17) Time to Child Care

	Mother						First Daughter	
	OLS	**Tobit**	**OLS**	**Tobit**	**OLS**	**Tobit**	**OLS**	**Tobit**
Treatment =1	0.043	0.102	-0.054	-0.151	0.010	-0.009	-0.095	-0.838*
	(0.126)	(0.243)	(0.149)	(0.265)	(0.140)	(0.258)	(0.075)	(0.408)
Children 12 to 17 =1			-0.246	-0.865*				
			(0.200)	(0.423)				
Treatment * Children 12 to 17 =1			0.263	0.713*				
			(0.177)	(0.353)				
Daughters 12 to 17 =1					-0.074	-0.332		
					(0.320)	(0.643)		
Treatment * Daughters 12 to 17 =1					0.142	0.483		
					(0.190)	(0.389)		
Observations	4036	4036	4036	4036	4036	4036	971	971
Mean Dependent Variable (Controls)	1.909	1.909	1.909	1.909	1.909	1.909	0.365	0.365
Proportion Zeros (Controls)	0.49	0.49	0.49	0.49	0.49	0.49	0.79	0.79

Notes: *significant at 5%. SE clustered at the community level in parentheses. All regressions include the following covariates: maternal age, age squared, years of education, ethnicity, head of the household status and whether she is a paid worker; the first daughter's year of education; baseline and contemporary household demographic composition; baseline household assets (dirt floor, electricity and farm size); male agricultural wage in the community at baseline, distance to large urban center, distance to secondary school, presence of pre-school and presence of junior high school imparted via TV ("telesecundaria") in the community.

Table IV: Maternal Participation in Child Care

	Model 1A	Model 1B		Model 1C		Model 2A	Model 2B		Model 2C		Model 3A	Model 3B		Model 3C	
	OLS	PROBIT	Mean Marginal Effect	LEWBEL	Mean Marginal Effect	OLS	PROBIT	Mean Marginal Effect	LEWBEL	Mean Marginal Effect	OLS	PROBIT	Mean Marginal Effect	LEWBEL	Mean Marginal Effect
Child Care Participation =1 (Mean = 0.51)															
Special Regressor (v)	0.345	5.843	2.247	-		-5.314	-8.472	-3.248	-		-2.533	-1.879	-0.722	-	
	(8.853)	(22.617)	(8.696)	-		(9.951)	(25.386)	(9.731)	-		(9.190)	(23.580)	(9.057)	-	
Treatment =1	0.014	0.020	0.008	0.007	0.004	-0.006	-0.033	-0.013	-0.054	-0.001	0.007	0.001	0.000	-0.003	-0.001
	(0.030)	(0.077)	(0.030)	(0.005)		(0.030)	(0.078)	(0.030)	(0.053)		(0.030)	(0.077)	(0.030)	(0.008)	
Children 12 to 17 =1						-0.166**	-0.428**	-0.163**	-0.600*	-0.001					
						(0.045)	(0.120)	(0.044)	(0.276)						
Treatment * Children 12 to 17 =1						0.095*	0.246*	0.093**	0.435**	0.071**					
						(0.037)	(0.096)	(0.036)	(0.128)						
Daughters 12 to 17 =1											-0.101	-0.259	-0.099	-0.097	0.010
											(0.071)	(0.182)	(0.069)	(0.085)	
Treatment * Daughters 12 to 17 =1											0.062	0.164	0.063	0.257*	0.091*
											(0.043)	(0.109)	(0.041)	(0.111)	

Notes: (N = 3710).*significant at 5%, **significant at 1%. SE clustered at the community level in parentheses. Observations in the top 7% of Fv and with $f(v|x) > 1*10^{-5}$ have been trimmed. All regressions include the list of covariates in Table III.

Table V: Maternal Time to Child Care and Leisure

	Model 1A	Model 1B	Model 1C	Model 2A	Model 2B	Model 2C	Model 3A	Model 3B	Model 3C
	OLS	ORDERED PROBIT	LEWBEL	OLS	ORDERED PROBIT	LEWBEL	OLS	ORDERED PROBIT	LEWBEL
I. Extent of Child Care Hours (Mean =1.94)									
Special Regressor (v)	6.700	11.828	-	-21.370	-6.995	-	1.093	6.041	-
	(32.380)	(22.201)	-	(36.543)	(25.143)	-	(34.413)	(23.769)	-
Treatment =1	0.037	-0.000	-0.001	-0.001	-0.030	-0.059	0.030	-0.008	-0.007
	(0.107)	(0.071)	(0.002)	(0.107)	(0.070)	(0.052)	(0.109)	(0.071)	(0.007)
Children 12 to 17 =1				-0.344*	-0.287**	-0.369*			
				(0.144)	(0.105)	(0.149)			
Treatment * Children 12 to 17 =1				0.297*	0.217*	0.306**			
				(0.131)	(0.090)	(0.093)			
Daughters 12 to 17 =1							-0.180	-0.145	-0.065
							(0.232)	(0.160)	(0.056)
Treatment * Daughters 12 to 17 =1							0.094	0.098	0.142*
							(0.138)	(0.096)	(0.061)
II. Extent of Leisure (Mean = 7.54)									
Special Regressor (v)	13.184	-2.596	-	-21.415	-23.310	-	3.206	-7.842	-
	(30.137)	(19.251)	-	(35.890)	(22.717)	-	(32.887)	(20.842)	-
Treatment =1	0.001	-0.035	-0.004	0.031	-0.016	0.056	0.014	-0.028	0.004
	(0.104)	(0.064)	(0.003)	(0.105)	(0.064)	(0.052)	(0.105)	(0.064)	(0.007)
Children 12 to 17 =1				0.250+	0.164+	0.493*			
				(0.139)	(0.086)	(0.218)			
Treatment * Children 12 to 17 =1				-0.316*	-0.191*	-0.457**			
				(0.126)	(0.079)	(0.118)			
Daughters 12 to 17 =1							0.059	0.041	0.100
							(0.217)	(0.135)	(0.067)
Treatment * Daughters 12 to 17 =1							-0.162	-0.086	-0.233**
							(0.141)	(0.088)	(0.078)

Notes: (N = 3710). +significant at 10%, *significant at 5%, **significant at 1%. SE clustered at the community level in parentheses. Observations in the top 7% of Fv and with $f(v|x) > 1*10^{-5}$ have been trimmed. All regressions include the list of covariates in Table III.

Table VI: First Daughters (Ages 12 to 17) Participation in Child Care and Schooling

	Model A	Model B		Model C	
			Mean Marginal Effect		Mean Marginal Effect
	OLS	PROBIT		LEWBEL	
I. Child Care Participation =1 (Mean =0.18)					
Special Regressor (v)	-6.691	-39.487	-9.300	-	-
	(11.966)	(53.200)	(12.530)	-	-
Treatment =1	-0.081	-0.367+	-0.088	-0.031**	-0.063**
	(0.051)	(0.218)	(0.061)	(0.011)	-
II. School Participation =1 (Mean =0.42)					
Special Regressor (v)	36.248*	103.020*	36.748*	-	-
	(17.587)	(49.191)	(17.548)	-	-
Treatment =1	-0.069	-0.193	-0.068	0.007*	0.041*
	(0.075)	(0.208)	(0.072)	(0.004)	-

Notes: (N = 758) +significant at 10%, *significant at 5%, **significant at 1%. SE clustered at the community level in parentheses. Observations in the top 7% of Fv and with f(v|x) > 1*10-4 have been trimmed. All regressions include the list of covariates in Table III.

Table VII: First Daughters (Ages 12 to 17) Time to Child Care, Schooling and Leisure

	Model A	Model B	Model C
	OLS	ORDERED PROBIT	LEWBEL
I. Extent of Child Care Hours (Mean =0.40)			
Special Regressor (v)	-4.755	-38.338	-
	(26.676)	(53.429)	-
Treatment =1	-0.132	-0.356+	-0.032**
	(0.112)	(0.215)	(0.011)
II. Extent of Hours in School (Mean =2.51)			
Special Regressor (v)	87.903+	86.352*	-
	(46.025)	(43.990)	-
Treatment =1	-0.137	-0.146	0.002
	(0.194)	(0.188)	(0.002)
III. Extent of Leisure (Mean =9.04)			
Special Regressor (v)	17.391	10.248	-
	(62.764)	(37.904)	-
Treatment =1	-0.180	-0.116	-0.013*
	(0.239)	(0.143)	(0.006)

Notes: (N = 758) +significant at 10%, *significant at 5%, **significant at 1%. SE clustered at the community level in parentheses. Observations in the top 7% of Fv and with $f(v|x) > 1*10\text{-}4$ have been trimmed. All regressions include the list of covariates in Table III.

Table VIII: Total Household Hours to Child Care

Extent of Child Care Hours	Model 2A OLS	Model 2B PROBIT	Model 2C LEWBEL	Model 3A OLS	Model 3B PROBIT	Model 3C LEWBEL
Special Regressor (v)	-19.451	-3.856	-	13.278	11.820	-
	(44.630)	(24.416)	-	(41.990)	(23.157)	-
Treatment =1	-0.039	-0.056	-0.059	-0.004	-0.035	-0.007
	(0.129)	(0.069)	(0.052)	(0.130)	(0.070)	(0.007)
Children 12 to 17 =1	-0.290	-0.243*	-0.342*			
	(0.189)	(0.110)	(0.138)			
Treatment * Children 12 to 17 =1	0.340*	0.220*	0.299**			
	(0.165)	(0.090)	(0.091)			
Daughters 12 to 17 =1				-0.084	-0.107	-0.061
				(0.297)	(0.161)	(0.054)
Treatment * Daughters 12 to 17 =1				0.041	0.080	0.148*
				(0.177)	(0.097)	(0.062)

Notes: (N=3710). Mean Dep. Var. = 2.16 hours. *significant at 5%, **significant at 1%. SE clustered at the community level in parentheses. Observations in the top 7% of Fv and with $f(v|x) > 1*10^{-5}$ have been trimmed. All regressions include the list of covariates in Table III. Model labels relate to those in Tables IV and V.

Table IX: Share of Child Care Hours Provided by Different Household Members

	OLS						
	Share Mother	**Share Other Women**	**Share Men**	**Share Sons 8 to 11**	**Share Sons 12 to 17**	**Share Daughters 8 to 11**	**Share Daughters 12 to 17**
Treatment (T) =1	0.009	-0.013	-0.008	-0.007	0.005	-0.026	-0.031*
	(0.024)	(0.033)	(0.008)	(0.008)	(0.011)	(0.016)	(0.015)
Girls 12 to 17 in the Household =1	-0.072*	-0.051	0.002	-0.011	0.006	-0.048**	
	(0.033)	(0.046)	(0.010)	(0.011)	(0.015)	(0.018)	
T*Girls 12 to 17 in the Household =1	0.051	0.058	0.006	0.011	0.005	0.041*	
	(0.032)	(0.045)	(0.009)	(0.011)	(0.015)	(0.018)	
Observations	4036	476	4100	1382	1047	1381	988
Mean Share Child Care Hours	0.44	0.10	0.03	0.02	0.02	0.04	0.08

Notes: *significant at 5%, **significant at 1%. SE clustered at the community level in parentheses. Share Household Member m =Total Hours to Child Care of Household Member m / Total Household Hours to Child Care. All regressions include the list of covariates in Table III.
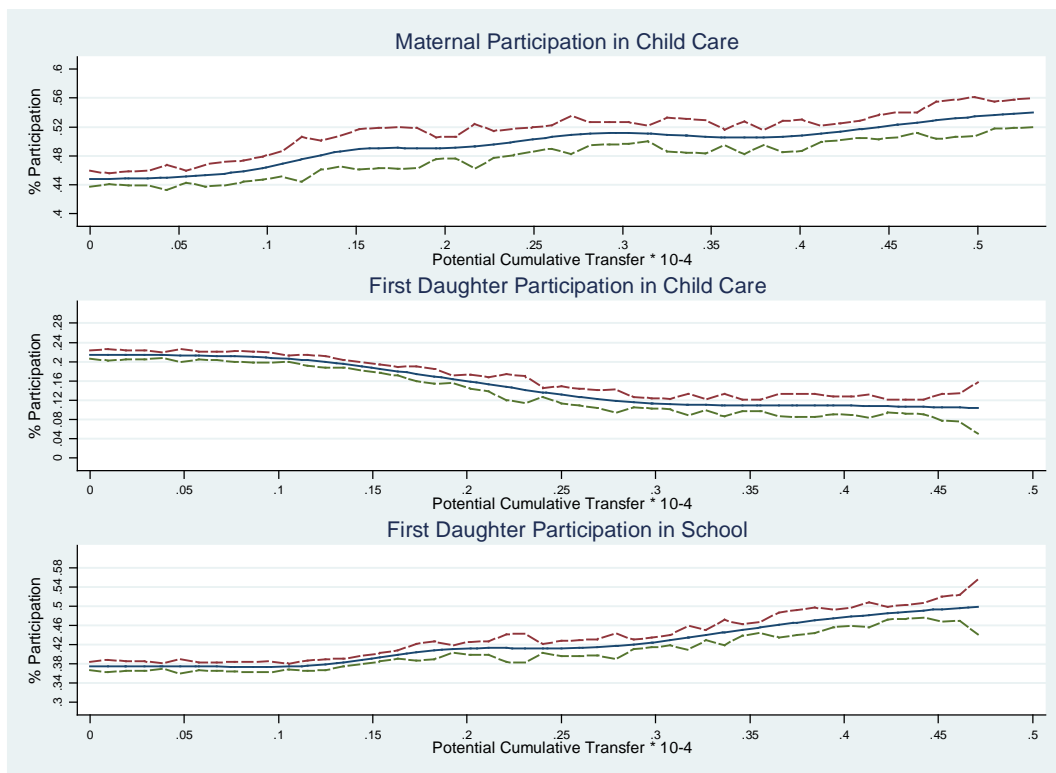
FIGURES and TABLES

Figure A1: Monotonicity

Table AI: "Regression Analogue" Test for Normality (Pagan and Vella, 1989)

	Model	Moment Restriction	T-Stat	Joint Test: Chi-Squared (2)
Mother's Participation in Child Care	Probit	$E(Pred^2*GR)$	-1.18	4.56
		$E(Pred^3*GR)$	-1.70	(0.1025)
Mother's Hours in Child Care	Tobit	$E(Pred^2*GR)$	-0.44	5.71
		$E(Pred^3*GR)$	-2.20	(0.0575)
Mothers' Leisure Hours	Tobit	$E(Pred^2*GR)$	-1.39	4.29
		$E(Pred^3*GR)$	-1.31	(0.1171)
First Daughter's Participation in Child Care	Probit	$E(Pred^2*GR)$	-0.54	0.29
		$E(Pred^3*GR)$	0.51	(0.8651)
First Daughter's Hours in Child Care	Tobit	$E(Pred^2*GR)$	-0.04	0.01
		$E(Pred^3*GR)$	0.01	(0.9948)
First Daughter's Participation in School	Probit	$E(Pred^2*GR)$	-0.76	0.58
		$E(Pred^3*GR)$	0.54	(0.7501)
First Daughter's Hours in School	Tobit	$E(Pred^2*GR)$	-1.82	3.33
		$E(Pred^3*GR)$	1.07	(0.1893)
First Daughter's Leisure Hours	Tobit	$E(Pred^2*GR)$	-0.01	0.57
		$E(Pred^3*GR)$	-0.03	(0.7530)

Notes: Pred are liniar predictions and GR are generalized residuals.

Table AII: Maternal Time to Child Care and Leisure. Sensitivity of Results to Trimming Outliers of Fv and f(v|x)

	All	**All & f(v|x) >1*10⁻⁵**	**98% Fv & f(v|x) > 1*10⁻⁵**	**95% Fv & f(v|x) >1*10⁻⁵**	**90% Fv & f(v|x) >1*10⁻⁵**
I. Child Care Participation =1					
	Mod 2C - 1	**Mod 2C - 2**	**Mod 2C - 3**	**Mod 2C - 4**	**Mod 2C - 5**
Treatment =1	-0.100	-0.029	-0.049	-0.027	-0.005
	(0.346)	(0.050)	(0.030)	(0.028)	(0.012)
Children 12 to 17 =1	-19.347	-0.430+	-0.475*	-0.464*	-0.392*
	(17.068)	(0.251)	(0.202)	(0.223)	(0.187)
Treatment*Children 12 to 17 =1	4.141	0.361**	0.410**	0.381**	0.384**
	(2.882)	(0.138)	(0.150)	(0.139)	(0.143)
Observations	4036	4031	3910	3789	3593
	Mod 3C - 1	**Mod 3C - 2**	**Mod 3C - 3**	**Mod 3C - 4**	**Mod 3C - 5**
Treatment =1	-0.017	-0.002	-0.007	0.002	0.011
	(0.196)	(0.018)	(0.008)	(0.009)	(0.011)
Daughters 12 to 17 =1	0.919	-0.131	-0.075	-0.111	-0.089
	(3.471)	(0.149)	(0.089)	(0.108)	(0.118)
Treatment*Daughters 12 to 17 =1	3.633	0.272*	0.197*	0.281*	0.442*
	(2.672)	(0.129)	(0.091)	(0.138)	(0.202)
Observations	4036	4033	3911	3790	3595
II. Extent of Child Care Hours					
	Mod 2C - 1	**Mod 2C - 2**	**Mod 2C - 3**	**Mod 2C - 4**	**Mod 2C - 5**
Treatment =1	-0.168	-0.057	-0.089*	-0.046	-0.012
	(0.167)	(0.050)	(0.040)	(0.029)	(0.010)
Children 12 to 17 =1	-7.471	-0.323	-0.318*	-0.337*	-0.283*
	(6.783)	(0.236)	(0.136)	(0.164)	(0.117)
Treatment*Children 12 to 17 =1	1.647	0.300*	0.296**	0.301**	0.262**
	(1.094)	(0.122)	(0.081)	(0.097)	(0.090)
Observations	4036	4031	3910	3789	3593
	Mod 3C - 1	**Mod 3C - 2**	**Mod 3C - 3**	**Mod 3C - 4**	**Mod 3C - 5**
Treatment =1	-0.048	-0.027	-0.030*	-0.007	0.000
	(0.077)	(0.023)	(0.014)	(0.008)	(0.006)
Daughters 12 to 17 =1	0.653	-0.125	-0.074	-0.090	-0.066
	(1.253)	(0.124)	(0.076)	(0.084)	(0.065)
Treatment*Daughters 12 to 17 =1	1.349	0.192*	0.140*	0.159*	0.232*
	(1.027)	(0.097)	(0.061)	(0.077)	(0.097)
Observations	4036	4033	3911	3790	3595
III. Extent of Leisure					
	Mod 2C - 1	**Mod 2C - 2**	**Mod 2C - 3**	**Mod 2C - 4**	**Mod 2C - 5**
Treatment =1	0.226	0.029	0.066*	0.032	0.008
	(0.305)	(0.037)	(0.034)	(0.023)	(0.009)
Children 12 to 17 =1	15.467	0.360+	0.473**	0.409*	0.328*
	(14.142)	(0.206)	(0.172)	(0.171)	(0.130)
Treatment*Children 12 to 17 =1	-3.143	-0.333**	-0.394**	-0.416**	-0.351**
	(2.267)	(0.114)	(0.103)	(0.111)	(0.102)
Observations	4036	4031	3910	3789	3593
	Mod 3C - 1	**Mod 3C - 2**	**Mod 3C - 3**	**Mod 3C - 4**	**Mod 3C - 5**
Treatment =1	0.087	0.022	0.021+	0.004	-0.007
	(0.152)	(0.024)	(0.011)	(0.009)	(0.008)
Daughters 12 to 17 =1	-1.573	0.138	0.118	0.140	0.087
	(2.579)	(0.127)	(0.095)	(0.100)	(0.087)
Treatment*Daughters 12 to 17 =1	-2.749	-0.246*	-0.249*	-0.281**	-0.386**
	(2.125)	(0.105)	(0.100)	(0.107)	(0.148)
Observations	4036	4033	3911	3790	3595

Notes: +significant at 10%, *significant at 5%, **significant at 1%. SE in parentheses. Outliers trimmed at different points of Fv and f(v|x) as indicated. Fv is the empirical distribution of v, and f(v|x) denotes the conditional probability density function of v given x, as estimated from the data (predicted). Models 2C and 3C labels relate to the labels in Table IV (Lewbel estimates). All regressions include the list of covariates in Table III.

Table AIII: First Daughters (Ages 12 to 17) Time to Child Care, Schooling and Leisure.
Sensitivity of Results to Trimming Outliers of Fv and f(v|x)

	All	All & f(v|x) >1*10^{-4}	98% Fv & f(v|x) > 1*10^{-4}	95% Fv & f(v|x) >1*10^{-4}	90% Fv & f(v|x) >1*10^{-4}
	Mod C - 1	Mod C - 2	Mod C - 3	Mod C - 4	Mod C - 5
I. Child Care Participation =1					
Treatment =1	-0.028*	-0.028*	-0.026**	-0.031**	-0.030**
	(0.011)	(0.011)	(0.009)	(0.011)	(0.009)
Observations	963	963	872	807	704
II. Extent of Child Care Hours					
Treatment =1	0.011	0.011	0.007	0.008+	0.009*
	(0.008)	(0.008)	(0.005)	(0.004)	(0.004)
Observations	962	962	871	807	702
III. School Participation =1					
Treatment =1	-0.029**	-0.029**	-0.026**	-0.031**	-0.030**
	(0.011)	(0.011)	(0.009)	(0.011)	(0.009)
Observations	963	963	872	807	704
IV. Extent of Hours in School					
Treatment =1	0.002	0.002	0.002	0.002	0.004
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Observations	961	961	870	807	702
V. Extent of Leisure					
Treatment =1	-0.010+	-0.010+	-0.009*	-0.012*	-0.013*
	(0.005)	(0.005)	(0.004)	(0.005)	(0.006)
Observations	973	973	880	815	710

Notes: +significant at 10%, *significant at 5%, **significant at 1%. SE in parentheses. Outliers trimmed at different points of Fv and f(v|x) as indicated. Fv is the empirical distribution of v, and f(v|x) denotes the conditional probability density function of v given x, as estimated from the data (predicted). Model C labels relate to the labels in Table V (Lewbel estimates). All regressions include the list of covariates in Table III.