

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

PIERRE DUBOIS AND MARTA RUBIO-CODINA

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

Pierre Dubois\* and Marta Rubio-Codina†

April 2010‡

## Abstract

We estimate semiparametrically the impact of the Mexican conditional cash transfer program *Oportunidades* on the time mothers and older sisters spend taking care of children aged under 3, using the randomization of a program placement and the methodology in Lewbel (2000). Results support the existence of substitution effects: mothers in treatment households are more likely to substitute for their older daughters' time to child care. As a result, daughters devote more time to schooling and less taking care of their younger siblings. Overall, total household time allocated to child care increases. These findings indicate that *Oportunidades* 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.

---

\*Toulouse School of Economics (GREMAQ, INRA, IDEI): pierre.dubois@tse-fr.eu. 21 allée de Brienne, F-31000 Toulouse

†Institute for Fiscal Studies: marta\_r@ifs.org.uk. IFS, 7 Ridgmount Street, London, WC1E 7AE.

‡We would like to thank Márton Csillag, Paul Gertler, Sylvie Lambert, Arthur Lewbel, Thierry Magnac, seminar participants at IPEA, and participants at the NEUDC, ESEM, LAMES and RES meetings for useful comments and suggestions on earlier versions of this paper. All errors are our own.

# 1 Introduction

In recent years, the economic case for public investment in Early Child Development (ECD) has become increasingly forceful. Carneiro and Heckman (2003) argue that returns to investments at early ages are higher than returns to investments later on in life on two accounts. First, beneficiaries will have a longer time to enjoy the rewards from these investments. Second, there are dynamic complementarities as early childhood learning fosters and facilitates later learning. Therefore, redirecting funds towards the early years is likely to entail Pareto improving efficiency gains.

The existing empirical evidence supports the notion that ECD is an important predictor of success throughout life. Both in developed and developing countries, poor ECD outcomes –usually linked to poor family environments– are associated with inadequate school readiness and poor school performance (Currie 2001); a lower earning capacity (Currie and Thomas 2001); higher criminality (Carneiro and Heckman 2003; Garces et al. 2000) and lower levels of social integration. Children suffering from low ECD also tend to have more children and provide poorer care, which contributes to the intergenerational transmission of poverty and to economic inequality. A recent review by Grantham-McGregor and coauthors (2007) estimates that there are at least 200 million children younger than 5 in the developing world that fail to reach their potential in motor, cognitive and socio-emotional development because of poverty, malnutrition, poor health and deficient care.<sup>1</sup>

While the relevance of stimulation and the home environment on ECD is well established empirically (Grantham-McGregor et al. 1997; Paxson and Schady 2007), caregivers

---

<sup>1</sup>See Gertler and Fernald (2004) and Fernald et al. (2006) for a description of the high prevalence of low ECD outcomes amongst children in rural and semi-urban Mexico.

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 large scope for conditional cash transfer (CCT) programs –such as the *Oportunidades* 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 and additional benefits they provide –in the form of nutritional supplements, health checkups, and educational talks– are likely to affect child rearing practices within a household.

In this paper, we investigate whether *Oportunidades* affects child care provision through altering the allocation of time given over to child care amongst household members. We exploit time use data on the randomized *Oportunidades* 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 of under 3 year old children. We focus the analysis on mothers and their older daughters as they are found to be the two main child care providers in the household. Moreover and given the program design, the older daughter is the household member likely to contribute the most to the total transfer amount received by the household, conditional on her attending school. This strengthens the case for greater economic incentives to enhance substitution effects in the allocation of time devoted to child care between mothers and their older daughters. Both OLS, probit and semiparametric estimates support the existence of such a 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, older daughters –ages 12 to 17– reduce their child care participation by 40%, and increase their participation in school activities by 9%. 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 *Oportunidades* 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 randomization of treatment offers a unique setting in which to implement this recently developed estimator.<sup>2</sup> In terms of findings, the analysis provides evidence that *Oportunidades* increases human capital accumulation both by keeping teenage girls in school and through more and arguably "better" –mother provided– child care. Hence, linking benefits to school attendance can simultaneously solve two related inefficiencies. First, an inefficiency in the levels of school attendance and domestic work provided by older daughters. Second, an inefficiency in the levels of child care provided in the household.<sup>3</sup>

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 intrahousehold 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

---

<sup>2</sup>There are relatively few empirical applications of the Lewbel (2000) estimator. These include the works of Cogneau and Maurin (2001) and Goux and Maurin (2005).

<sup>3</sup>Of course this is under the assumption that the pre-program levels of schooling and child care provision were indeed inefficient, or in other words households were constraint in their choices.

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 rural *Oportunidades* program. In Section 3, we discuss the potential mechanisms at play behind the observed effects. Section 4 describes the data and the sample of analysis. In section 5, we present the Lewbel (2000) semiparametric estimator and discuss identification. Results are presented in section 6 and section 7 concludes.

## 2 The Rural *Oportunidades* Program

*Oportunidades* began in 1997 under the name *Progresa* as a national anti-poverty program. Over its first three years, the program extended benefits to almost all eligible families living in rural areas. Starting in 2001, it expanded to urban areas and currently covers around 5 million families all over Mexico.<sup>4</sup> In rural areas, *Oportunidades* determined household eligibility in two stages. First, underserved communities were identified based on the proportion of households living in poverty as defined in the 1995 population census. Second, low-income households within these communities were chosen using a proxy means test. Pre-intervention data to construct the index was collected on all households in eligible communities through the Survey of Household Socioeconomic Characteristics (*Encuesta Socioeconómica de Hogares*, ENCASEH). This process designated 52% of households in selected communities as eligible for benefits.<sup>5</sup>

---

<sup>4</sup>[www.oportunidades.gob.mx](http://www.oportunidades.gob.mx). See Skoufias (2005) for a review of the program design and impacts.

<sup>5</sup>Soon afterwards, a set of slightly wealthier households was included as eligible in a process called “densification”. Many of these households however were not incorporated into the program until later because of administrative delays. In order to not attribute treatment effects to untreated households, we

All eligible households living in treatment localities were offered *Oportunidades* and over 90% enrolled. Once enrolled, households received benefits for a three-year period with the possibility of being recertified if all household members obtained the prescribed preventive medical care, children attended school and mothers participated in educational talks –called "pláticas"– on health, hygiene and nutrition.

Monetary benefits represent on average over 20% of total household income. They are given bimonthly to the female head of the household in two forms. The first is a nutritional grant given to *all* 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; and regular health checkups. 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. The scholarship rises substantially after graduation from primary school and is higher for girls than boys during secondary school. It is received conditional on children attending a minimum of 85% of school days and on not repeating a grade more than twice.

### **3 *Oportunidades* 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. Even if the traditional division of labor between genders is less and less patent in western economies, 

---

will exclude "densified" households from the analysis.

it continues to be firmly established in rural Mexico (INEGI 2000; Parker and Skoufias 2000). As an illustration, Figure A plots participation rates in child care by age and sex in the *Oportunidades* evaluation sample. Child care is clearly a female activity: while participation rates oscillate between 40 and 60% for adult women; men participation rates are around 8%. A peak is observed for both men and women in their twenties. Note also that female participation increases sharply from the age of 12.

These patterns in the allocation of time devoted to child care suggest a twofold analysis. First, we frame the analysis within the "*separate spheres*" bargaining model of the household developed by Lundberg and Pollack (1993). This model of distribution within marriage differs from standard bargaining models (Manser and Brown 1980; McElroy and Horney 1981) in that the threat point is not divorce—a rather implausible choice in rural Mexico—but a noncooperative equilibrium within marriage that reflects traditional gender roles. The noncooperative equilibrium corresponds to a utility-maximizing strategy subject to individual resource constraints within each spouse's domain (sphere), taking the other spouse's strategy as given. In the current context, child care provision belongs to the wife's domain.<sup>6</sup> Second, we focus on the allocation of time to child care by the mother and her older daughter—aged 12 to 17—as they are the only two agents in charge of the provision of household public good child care.<sup>7</sup>

Hence, 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

---

<sup>6</sup>The separate spheres bargaining model is consistent with the view that allowances targeted to different spouses have different effects on the distribution of household resources—one of the premises on which *Oportunidades* is based. It also allows for both spouses to enjoy household public goods.

<sup>7</sup>We assume that market child care services are unavailable in these disadvantaged rural communities.



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. We assume that the noncooperative equilibrium allocations reached by the mother are the household's final equilibrium.<sup>8</sup> This means that the implications of the unitary household model apply within the mother's optimization sphere. In this stylized household, the *Oportunidades* intervention amounts to:

- (i) an increase in the mother's nonlabor income given the nutritional grant.
- (ii) the provision of a minimum level of maternal care, 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 living 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 of the intervention components described above 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 through an income effect. However, because child care requires maternal time as an input factor, increases in income might increase leisure, and reduce child care time.

---

<sup>8</sup>This may happen if transaction costs are high enough. As a matter of fact, the assumption is supported by the data: household members outside the sphere (males) do not alter their (initially very low) contribution to household child care time.

- (ii) direct increases in the total quantity of child care.
- (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 time and her first daughter's are substitute inputs in the production of child care.

In all treatment households, maternal time allocated to child care will be affected by the nutritional grant (income effect) and compliance with the program requirements (preventive health visits and attendance at "pláticas"). However, cross-substitution effects in child care time (effect *iii.c*) will only arise amongst those mothers whose daughters are eligible to receive the educational grant. In the empirical exercise, we will exploit both the random allocation of benefits 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 Experimental Design and Data

We benefit from the fact that the Mexican Government was committed to a rigorous evaluation of the impact of *Oportunidades* using a controlled-randomized evaluation design. Given budgetary and logistical constraints, the Government could not enrol all eligible families in the country simultaneously and decided to phase in the enrolment of entire

communities over time instead. As part of this process, the Government randomly chose 320 treatment and 186 control communities in seven states for a total of 506 experimental communities.<sup>9</sup> Eligible households in treatment communities began receiving benefits in April of 1998; whereas eligible households in control communities were not incorporated until November of 1999. In order to minimize anticipation effects, households in control communities were not informed that *Oportunidades* would provide benefits to them until two months before incorporation.<sup>10</sup>

In the empirical analysis, we use the *Oportunidades* rural evaluation data, or Evaluation Surveys (*Encuestas de Evaluación*, ENCEL), and the 1997 ENCASEH baseline data. The ENCEL household surveys collected data twice a year on a host of topics: demographics, assets, labor, consumption, etc. The May 1999 survey round collected additional data on time use for all household members older than eight.<sup>11</sup> 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.<sup>12</sup>

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 infor-

---

<sup>9</sup>Behrman and Todd (1999) test for statistical differences in the distributions of a large set of observable characteristics between treatment and control households to assess the validity of the randomization.

<sup>10</sup>Attanasio et al. (2005) find no evidence of anticipation effects among control households.

<sup>11</sup>Parker and Skoufias (2000) and Rubio-Codina (2006) provide further details on these data.

<sup>12</sup>Time use data was also collected as part of the November 2000 ENCEL round. As all eligible households were already receiving benefits by then, we use this cross-section to test for differential effects resulting from different lengths of exposure to the program benefits. We find increases in maternal time to child care for mothers that have been in the program for longer (i.e. households in the original treatment group). Results are available upon request.

mation on time allocation 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 dropping out of the sample 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 daily activities during the month prior to the interview.<sup>13</sup>

The final sample consists of 4,036 mothers with children younger than 3 (see Table 1). This represents about one third (34%) of all eligible households, as classified according to the original classification scheme. Approximately, 26% of these women are also mothers of a 12 to 17 teenage girl. Note that the proportion of mothers with different offspring compositions is very similar in treatment and control households and remains similar to the original (randomized) distribution –60% treatment and 40% control. This suggests that the potential for sample selection and sample composition biases is negligible.

Panel I in Table 2 shows summary statistics of the dependent variables for treatments and controls in May 1999. In the last column, we report a test of the equality of raw means. 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 is not statistically significant. For first daughters aged 12 to 17, there is

---

<sup>13</sup>These represent, respectively, 6.6 and 1.6% of the households in the estimation sample and are balanced across treatment and control groups. Parametric estimates are robust to keeping these households in the estimation.

a 6.8 points significant reduction in child care participation and a 7.9 points significant increase in school participation given treatment. There are no significant differences in the amount of time devoted to any of the activities considered conditional on participation.

In Panel II we report summary statistics on maternal and household characteristics in May 1999; and on baseline household and community characteristics for the sample of mothers under analysis. Mothers of children younger than 3 are around 30 and have 3.5 years of education on average. Less than 2% are the head of the household and between 6% (control) and 9% (treatment) work for a wage. 41% of the mothers in the sample are indigenous. On average, they have between 1 and 2 children younger than 3, 3 children younger than 7, and 1 child ages 8 to 17. The test of equality of means shows no significant differences for any of the variables reported, as is expected from random assignment. More importantly, there is not sufficient evidence to reject the hypothesis of (statistically) equal offspring composition between mothers in treatment and control groups during the intervention years. The only exception is the number of sons ages 12 to 17, which is significantly larger for treatment mothers.<sup>14</sup>

## 5 Semiparametric Estimation and Identification

Our objective 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 exploit the exogenous variation introduced by the random assignment of households to treatment and control communities, and the

---

<sup>14</sup>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.

heterogeneity in household composition for this purpose.

## 5.1 Empirical Specification

In order to estimate the overall program impact on child care provision, we specify the number of hours individual  $i$  spends on child care as:

$$\tilde{y}_i = \alpha_0 v_i + \alpha_1 T_i + \sum_{r=1}^R \beta_r x_{ri} + \varepsilon_i \quad \forall i = 1, \dots, N \quad (1)$$

where  $\tilde{y}_i$  is the number of hours allocated 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* accumulated since taking up the program; and the  $x_{ri}$  are  $R$  individual, household and community characteristics (listed below). Thus,  $\alpha_1$  identifies the average effect of treatment conditional on total transfers accumulated  $v_i$ . In other words, it identifies the counterfactual effect of being treated by the conditional cash transfer program given some transfers  $v_i$  received by the household. Notably, the total average effect of the program on a given individual also depends on the amount of transfer received by the household. Moreover, given the design of the program, the amount of transfers is a function of characteristics  $x_{ri}$ . Remark also that, in such a specification, interaction terms between  $v_i$  and  $T_i$  are already taken into account since  $v_i T_i = v_i$ , by construction. This will prove important in the semiparametric method whose validity requires an exogeneity and partial independence assumption which would not be true if some omitted interaction term appears in the error term  $\varepsilon_i$ .

Since the true number of hours is not directly observed,  $\tilde{y}_i$  is a latent variable. For simplicity, we model the observed variable  $y_i$  in two different ways: (i) as a participation dummy that equals 1 if  $i$  spends a positive number of hours on child care ( $\tilde{y}_i > 0$ ) and 0 otherwise ( $\tilde{y}_i \leq 0$ ); and (ii) as an interval indicator that can attain up to  $k$  different

values depending on the extent of time  $i$  devoted to child care (up to one hour, between one and two hours, etc.).<sup>15</sup>

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:

$$\tilde{y}_i = \alpha_0 v_i + \alpha_1 T_i + \alpha_2 S_i^j + \alpha_3 T_i S_i^j + \sum_{r=1}^R \beta_r x_{ri} + \varepsilon_i \quad \forall i = 1, \dots, N \quad (2)$$

where  $S_i^j$  equals 1 if the situation  $j$  is true. We consider two possible situations: (i) mother  $i$  has offspring 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 care will 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,  $\alpha_3$ , identifies this cross-substitution effect. Given the child care participation patterns shown in Figure A –i.e. females older than 12 are the main child care providers in these communities– we expect any cross-substitution effect to take place between mothers and their 12 to 17 year old daughters.<sup>16</sup> In this specification, the coefficient on the treatment dummy,  $\alpha_1$ , is the remaining effect of the program on the mother's time to child care. It is composed by a combination of the impacts coming from the nutritional grant (income effect) and from compliance with the program requirements (preventive health visits and attendance at

---

<sup>15</sup>Even if time (in minutes) devoted to child care is a priori a continuous variable, the empirical distribution has a discrete support on specific values, mainly hours and half hours.

<sup>16</sup>Moreover, because girls enrolled in secondary school receive the largest transfer amounts, teenage girls are the household members that most significantly contribute to total household transfers (conditional on secondary school enrolment).

"pláticas") that affect all mothers in treatment households –regardless of their offspring composition.

In the event of cross-substitution effects between mothers and first daughters, one should also expect reductions in the older daughter’s participation and in the time she spends caring for her younger siblings. To test this, we estimate equation (1) on the (extent of) child care participation of first daughters –aged 12 to 17– of those mothers in the sample described above.<sup>17</sup> We additionally estimate equation (1) for first daughters’ school participation and for their extent of schooling and leisure time in order to obtain a broader picture of their time allocation. We also estimate equations (1) and (2) on the extent of leisure time for mothers.

## 5.2 Semiparametric Identification

Parametric identification of the parameters in equations (1) and (2) is possible using a standard probit model –when  $y_i$  is dichotomous– and an ordered probit model –when  $y_i$  is polychotomous– 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 (i.e.  $\alpha_1$  and  $\alpha_3$  are different); and (ii) Oportunidades 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

---

<sup>17</sup>Recall that the number of daughters aged 12 to 17 is balanced between mothers in the treatment and control groups (see Table 2). Note also that over 70% of the mothers in the estimation sample have only one daughter 12 to 17 years old.



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 are available upon request. Consequently, and in order to avoid the problems of specifying a parametric model, 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 allows them to suffer from conditional heteroscedasticity of unknown form. Khan and Tamer (2009) provide an alternative distribution free estimator of such single index models. However, 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 A2 below), which is at the basis of the method. Moreover, this assumption, has the advantage of guaranteeing that the error term has a conditional mean which does not depend on interaction effects between partially independent regressors and other regressors, thus avoiding some form of misspecification error in the single index model.

Let us denote  $y$ ,  $v$  and  $\varepsilon$  the column vectors of  $y_i$  and  $v_i$  and  $\varepsilon_i$ , respectively.  $X$  is the  $n * (3 + R)$  vector of right hand side variables except  $v_i$  also called the "special regressor"; and  $\beta$  the column vector  $[\alpha_1, \alpha_2, \alpha_3, \beta_1, \dots, \beta_R]$ .  $I(\cdot)$  is an indicator function that equals one if  $\cdot$  is true and zero otherwise. Lewbel (2000) considers the binary choice model,

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

and the following assumptions, whose empirical validity will be discussed in the next subsection:

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

A.2: Partial Independence: the conditional distribution of  $\varepsilon$  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  $\beta$  can be estimated (with root N consistency) by an ordinary least squares regression of  $y^*$  on  $X$ , where

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

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  $v$  follows a normal distribution as parametrized in Lewbel (2006).<sup>18</sup>

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

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

where  $\alpha_0 = -\infty$  and  $\alpha_K = +\infty$ . In this case, the transformation of the dependent variable is written as:<sup>19</sup>

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

---

<sup>18</sup>A description of the implementation of the Lewbel estimator is available upon request.

<sup>19</sup>Lewbel (2000) also proposes an extension to censored data which consists of applying the ordered data estimator repeatedly. More precisely, it involves: (i) defining a continuum of values for  $\alpha_k$  (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, we

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  $\varepsilon$ . 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:

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

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 will thus devote some particular attention to the support condition in order to apply this method and have tested the stability and robustness of our estimator to variations in the trimming parameters.

### 5.3 The Special Regressor

Our 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. To 

---

chose to categorize total hours into a limited number of intervals defined sensitively to the underlying thresholds in the empirical distribution of hours, which as noted earlier suffers from considerable bunching around specific values.

construct  $v_i$ , we first take the household’s composition and children’s school enrollment in 1997 (baseline) and move it forward to May 1999. Next, we compute the transfer amount by applying the program benefit allocation rules since first take up, assuming no school drop out and no grade repetition.<sup>20</sup> In this subsection, we argue that the stock of the potential transfers received by the household is a valid special regressor as it satisfies the assumptions required for identification: continuity, monotonicity, partial independence and large support.

First, the amount of accumulated potential transfers is a continuous variable (A.1) and includes zero in the support (A.3). It is indeed positive for treatment households and zero for control households.

Second, randomization of treatment at the community level guarantees that the treatment dummy is independent of observables and unobservables,  $(\varepsilon_i, X_i) \perp T_i$ . In turn, this implies that  $\varepsilon_i \perp v_i | X_i$  (A.2) since the accumulated potential transfer is exogenous by construction. As defined above, the cash transfers are  $v_i = T_i^* f^*(D_{i97})$  where  $f^*(\cdot)$  is known, given by the design of the program and a function of a subset of baseline observables  $D_{i97} \subset X_{i97}$  and  $T_i^*$  represents the length of time the household has been receiving benefits by May 1999. We can also rewrite  $v_i$  as  $v_i = T_i f(D_i) u_i$  where  $T_i$  is the randomized treatment status and  $f(\cdot)$  is a known function of a subset of May 1999 observables  $D_i$  (household demographics, school attendance and grade attended) and  $D_i \subset X_i$ . Thus by construction  $u_i = \frac{T_i^* f^*(D_{i97})}{T_i f(D_i)}$  is an exogenous variable composed by two random elements:

---

<sup>20</sup>Note that  $v_i$  and  $T_i$  differ in that different treatment households are eligible for different transfer amounts depending on household size and composition (gender and age of children in school age) and enrolment status. Since  $v_i$  in May 1999 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.

(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 some of the other regressors in  $X_i$ , as is required for identification.

As noted, a testable implication of assumptions A.1 to A.4 and A.5 is the monotonicity of the conditional probability of success  $pr(y_i = 1|v_i, X_i)$  over the support of  $v_i$ . In this setting, monotonicity 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 "tested" its empirical validity (results available upon request).

## 6 Results and Discussion

We semiparametrically estimate the effect of *Oportunidades* on participation in child care, and on the extent of participation in child care and leisure for mothers of children under 3 and for the older daughter –aged 12 to 17– of these mothers. For the first daughter, we also estimate the program impact on participation and on the extent of participation in school. For comparison purposes, we first report OLS estimates (Models A), then probit –or ordered probit– results (Models B), and finally the semiparametric (Lewbel) estimates (Models C). As our interest concerns the application of the Lewbel method, we will focus the discussion on the semiparametric results.<sup>21</sup> In all sets of estimates, we

---

<sup>21</sup>Note that the OLS, probit (or ordered probit) and Lewbel estimates are not comparable in magnitude for various reasons. First, OLS estimates are likely biased given the binary and ordered nature of the

trim observations with values of the special regressor  $v_i$  in the top 7% of its distribution to conform to the symmetry condition (A.5). 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  $y^*$ . An analysis of the sensitivity of the Lewbel estimator to trimming is available upon request.

All regressions include the following explanatory variables: maternal age, age squared, years of education, ethnicity, head of the household status and whether she is a paid worker; the first daughter's years of education; baseline and contemporary household demographic composition; baseline assets (dirt floor, 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). Estimates are robust to the exclusion of these controls.

## 6.1 Maternal Time

Table 3 shows estimates of the impact of the program in May 1999. Panels I and II present results for mothers' participation in child care (binary outcome) and for their extent of participation (ordered polychotomous outcome). For mothers, the extent of participation in child care variable takes six different values:  $k \in \{0, 1, \dots, 5\} = \{\text{no time devoted to child care, up to one hour, between one and two hours, two and four hours, four to seven, seven to thirteen}\}$ .<sup>22</sup> Estimation results follow very similar patterns for both types of dependent variables. Second, probit and ordered probit impose a normalization to one of the variance of the errors for identification. Third, OLS, probit and ordered probit do not impose the normalization to one of the coefficient on the special regressor as in the Lewbel method.

<sup>22</sup>Results are robust to a redefinition of these categories, defined to ensure enough power in each cell.

outcomes.

Models 1A to 1C in Panel I 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, 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) *Oportunidades* attenuates the cross-substitution effect. The semiparametric coefficient on having children 12 to 17 interacted with treatment is 0.44 (Model 2C, Panel I) and the mean marginal effect is 7.15 percentage points.<sup>23</sup> Given an initial participation rate of 51.4%, this results in a 13.9 percentage increase in participation in child care for mothers with children 12 to 17. 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, Panel I). Rubio-Codina (2010) reports a similar cross-substitution effect between mothers and teenage daughters in time devoted to domestic activities amongst *Oportunidades* beneficiaries.

Similar effects are observed on the extent of participation in child care. Mothers of

---

<sup>23</sup>For dichotomous outcomes, we compute the marginal effects on the estimated coefficients as:  $M_{ij} = \frac{\Delta[1-\hat{G}(-v_i-x_i\hat{\beta})]}{\Delta x_j} = \frac{\Delta[1-\hat{G}(z_i\hat{\beta})]}{\Delta x_j} = \hat{G}(\hat{z}_i^0) - \hat{G}(\hat{z}_i^1)$ , where  $M_{ij}$  is the effect of switching the  $j^{th}$  binary variable,  $x_j$ , from 0 to 1 on the probability that  $y_i$  equals one; 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 zero.  $\hat{G}(\cdot)$  is the estimated cumulative distribution function of the probability of  $y_i$  given  $z_i$ . We nonparametrically estimate  $\hat{G}(\cdot)$  running the kernel regression of  $y_i$  on  $z_i$ . Results reported in the text were computed using a Gaussian Kernel and 500 equally spaced points in the range of  $z_i$  and are robust to larger numbers of points.

daughters 12 to 17 are more likely to increase the amount of time taking care of their younger children given treatment (Model 2C, Panel II). As before, the coefficient on treatment alone is negative and non statistically significant (Model 1C, Panel II). The analysis on the extent of maternal leisure in Panel III shows opposite effects: mothers with children younger than 3 and teenagers 12 to 17 enjoy less leisure given treatment. This reduction 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).

## 6.2 First Daughter's Time

Next, we turn our attention to the allocation of time by first daughters. Panels I and II in Table 4 present the estimated program impact on first daughters' participation in child care and on their extent of participation. The latter variable takes four different values:  $k \in \{0, 1, 2, 3\} = \{\text{no time devoted to child care, up to one hour, between one and two hours, more than two hours}\}$ .

Semiparametric estimates evidence a reduction of 7.1 percentage points in the first daughter's probability of participating in child care (estimated coefficient of -0.032, Model



C in Panel I) or a 40.1% decrease. Significant reductions are also observed in terms of the extent of participation (Model C in Panel II). A concern is that the observed reductions in teenage child care participation are in fact driven by an increase in school attendance of other siblings in primary school age. Unfortunately, this hypothesis is not testable as almost all teenage girls with siblings younger than 3 also have siblings in primary school age (6 to 11). It is well-known, however, that the program had little effect on primary school enrolment (Schultz 2004), which makes this an unlikely driver of the results found.

About 65% of the first daughters in the sample that engage in child care activities do not attend school. 3% of them report that they are not enrolled because they have to help in the house.<sup>24</sup> Moreover, conditional on school enrolment, another 3% 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, or care of the sick.

These figures are somewhat indicative that child care and schooling are substitute activities. Not surprisingly, we semiparametrically identify a significant increase in participation in schooling activities (attendance and homework) for first daughters aged 12 to 17 (Model C in Panel III). The estimated coefficient of 0.08 translates into a marginal effect of 3.9 percentage points or a 9.3% increase in schooling, which is consistent with the positive treatment impact on female secondary school enrolment reported in Schultz (2004) and Parker and Skoufias (2000). The increase in the extent of time allocated to schooling activities is, however, not significant. Moreover, we observe a significant reduction in the extent of leisure for first daughters (Model C in Panel V). This might suggest that these girls reduce (or stop) their contribution to child care and instead, take up schooling full time, as a result of the reduction in the price of schooling. Another

---

<sup>24</sup>Note, however, that a majority of teenage girls report not going to school because they do not have enough money (51%), they do not like it (18%) or the school is too far (8%).

possibility is that the intervention affects the allocation of teenage girls' time to other activities.<sup>25</sup>

### 6.3 Quantity and Quality of Care

In this subsection, we estimate equations (1) and (2) on total household hours devoted to child care. We define total household hours to child care as the sum of hours each household member reports spending in child care and categorize the variable as  $k = \{0, 1, \dots, 6\} = \{\text{no time allocated to child care, up to one hour, one to two hours, two to three hours, three to five, five to seven, more than seven}\}$ .

We should expect no effect on total household hours if mothers and their 12 to 17 year old daughters are the only two household members taking care of the very young and if mothers fully substitute for their daughters' time. Conversely, if mothers –and possibly other household members– increase time devoted to child care by more than the amount previously devoted by teenage girls,  $\alpha_3$  in equation (2) should be positive and significant. A negative and significant  $\alpha_3$  would imply a reduction in total household hours to child care in treatment households. Table 5 shows a significant increase in the extent of total household hours to child care in treatment households with children under 3 and teenagers ages 12 to 17. This implies that *Oportunidades* 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 6 shows results from estimating (1) and (2) using the share

---

<sup>25</sup>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.

of child care hours provided by household member  $m$  over total household hours as the dependent variable. We consider the following "sets" of household members: mothers of 0 to 3 year old; other adult women living in the household; 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 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 subpopulation (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 entails 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" –the educational talks covering best health, hygiene and nutrition practices.

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.<sup>26</sup>

## 7 Conclusion

This paper provides robust semiparametric evidence that does not rely on *ad hoc* parametric models of the effect of the *Oportunidades* program on child care provision.

We have first shown that child care provision is a female activity in rural Mexico. Next, we have argued that the intervention might lead to increases in the quantity of care provided within the household, using the set up provided by a separate spheres bargaining model and assuming the mother as the sole decision-maker. The nutritional supplements, health checkups and educational talks ("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

---

<sup>26</sup>We are currently using more recent rounds of the *Oportunidades* 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 somewhat different identification strategy, we consider this issue to be beyond the scope of this paper.

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 a mother-daughter substitution hypothesis. 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 *Oportunidades* 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 found no significant effect of treatment on child care provision amongst mothers with no children older than 12 or 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) and the resulting mother-daughter cross-substitution effect. Second, the educational talks and preventive care do not influence maternal child care provision as substantially 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. Consequently, and given the crucial 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 (2005), “Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate Progresa”, IFS Working Paper EWP05/01, London.
- [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] Blau, David M. and Philip K. Robins (1988), “Child-Care Costs and Family Labor Supply”, *Review of Economics and Statistics*, 70(3):374-381.
- [7] 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.
- [8] Case, Anne, I-Fen Lin and Sara McLanahan (2001) "Educational Attainment in Blended Families", *Evolution and Human Behavior*, 22(4): 269-289.
- [9] Case, Anne and Christina Paxson (2001), “Mothers and Others: Who Invests in Children’s Health?”, *Journal of Health Economics*, 20(3):301-328.
- [10] 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.
- [11] Currie, Janet (2001), “Early Childhood Education Programs”, *The Journal of Economic Perspectives*, 15(2):213-238.
- [12] Currie, Janet and Duncan Thomas (2001), “Early Test Scores, Socioeconomic Status and Future Outcomes”, *Research in Labor Economics*, 20:103-132.
- [13] Fernald, Lia C., Lynette M. Neufeld et al. (2006) “Parallel Deficits in Linear Growth and Mental Development in Low-Income Mexican Infants in the Second Year of Life", *Public Health Nutrition*, 9(2): 178-186.
- [14] 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 *Oportunidades*", *Lancet*, 371:828-37.

- [15] Garces, Eliana, Duncan Thomas and Janet Currie (2002), “Longer-Term Effects of Head Start”, *The American Economic Review*, 92(4):999-1012.
- [16] Gertler, Paul J. and Lia C. Fernald (2004), “The Medium Term Impact of *Oportunidades* on Child Development in Rural Areas”, mimeo UC Berkeley.
- [17] Goux, Dominique and Eric Maurin (2005), “The effect of overcrowded housing on children’s performance at school”, *Journal of Public Economics*, 89:797-819.
- [18] Grantham-McGregor, Sally, S. P. Walker, S. M. Chang and C.A. Powell (1997), “Effects of Early Childhood Supplementation With and Without Stimulation on Later Development in Stunted Jamaican Children”, *American Journal of Clinical Nutrition*, 66:247-253.
- [19] 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”, *Lancet*, 369:60-70.
- [20] Heckman, James J. (1974) “Effects of Child-Care Programs on Women’s Work Effort”, *Journal of Political Economy*, 82(2):S136-S163.
- [21] INEGI (2002), *Uso del Tiempo y Aportaciones en los Hogares Mexicanos*, Instituto Nacional de Estadística (INEGI), Aguascalientes.
- [22] Khan Shakeeb and Elie Tamer (2009), "Inference on Endogenously Censored Regression Models Using Conditional Moment Inequalities", *Journal of Econometrics*, 152, 104-119.



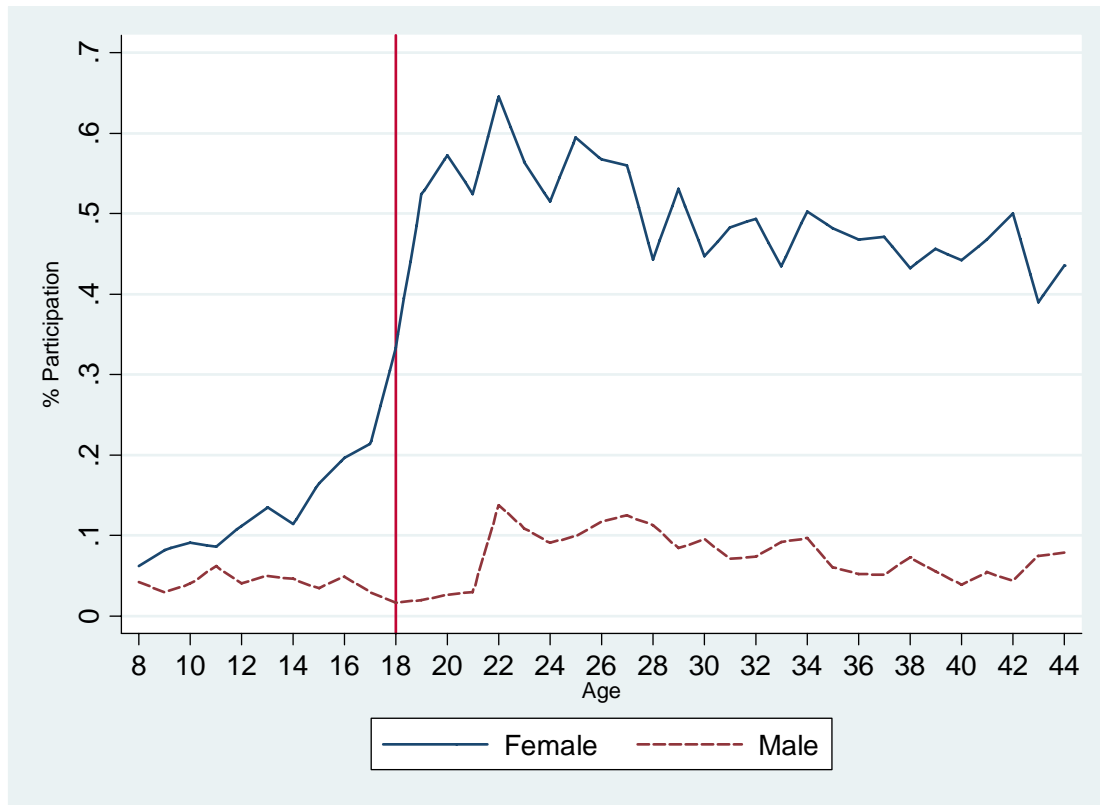
- [23] Khan Shakeeb and Elie Tamer (2010), "Irregular Identification, Support Conditions, and Inverse Weight Estimation" mimeo
- [24] Lagarde, Mylene, Andy Haines and Natasha Palmer (2007), "Conditional Cash Transfers for Improving Uptake of Health Interventions in Low- and Middle-Income Countries: A Systematic Review", *Journal of the American Medical Association*, 298:1900-1910.
- [25] Lewbel, Arthur (2000), "Semiparametric Qualitative Response Model Estimation With Unknown Heteroscedasticity or Instrumental Variables", *Journal of Econometrics*, 97(1): 145-177.
- [26] Lewbel, Arthur (2006), "Simple Endogenous Binary Choice and Selection Panel Model Estimators", Boston College Working Papers in Economics No. 613.
- [27] Lundberg, Shelly and Robert A. Pollak (1993), "Separate Spheres Bargaining and the Marriage Market", *Journal of Political Economy*, 101(6):988-1010.
- [28] Macours, Karen, Norbert Schady and Renos Vakis (2008), "Cash Transfers, Behavioral Changes, and the Cognitive Development of Young Children: Evidence from a Randomized Experiment", mimeo Johns Hopkins University and The World Bank.
- [29] Magnac, Thierry and Eric Maurin (2007), "Identification and Information in Monotone Binary Models", *Journal of Econometrics*, 127(1):76-104.
- [30] Manser, Marilyn and Murray Brown (1980), "Marriage and Household Decision-Making: A Bargaining Analysis", *International Economic Review*, 21(1): 31-44.

- [31] McElroy, Marjorie B. and Mary Jean Horney (1981), “Nash-Bargained Household Decisions: Toward a Generalization of the Theory of Demand”, *International Economic Review*, 22(2): 333-349.
- [32] 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.
- [33] Mueller, Eva (1984), “The value and allocation of time in rural Botswana”, *Journal of Development Economics*, 15(1-3):329-360.
- [34] Parker, Susan and Emmanuel Skoufias (2000), “The Impact of Progresa on Work, Leisure, and Time Allocation”, Final Report, IFPRI, Washington, D.C.
- [35] Paxson, Christina and Norbert Schady (2007), “Cognitive Development among Young Children in Ecuador: The Roles of Wealth, Health and Parenting”, *Journal of Human Resources*, 42(1):49-84.
- [36] Paxson, Christina and Norbert Schady (2010), “Does Money Matter? The Effects of Cash Transfers on Child Health and Development in Rural Ecuador”, forthcoming in *Economic Development and Cultural Change*
- [37] Rubio-Codina, Marta (2010), “Intra-household Time Allocation in Rural Mexico: Evidence from a Randomized Experiment”, forthcoming in *Research in Labor Economics*.
- [38] Schultz, Paul (2004), “School Subsidies for the Poor: Evaluating the Mexican PROGRESA Poverty Program”, *Journal of Development Economics*, 74(1): 199-250.

- [39] Skoufias, Emmanuel (2005) “PROGRESA and Its Impacts on the Welfare of Rural Households in Mexico”, IFPRI Report 139, Washington D.C.
- [40] Tiefenthaler, Jill (1997), “Fertility and Family Time Allocation in the Philippines”, *Population and Development Review*, 23(2):377-397.
- [41] UNICEF (2001), *The State of World’s Children*: UNICEF, New York.

**FIGURES and TABLES**

**Figure A: Child Care Participation Rates by Sex and Age (May 1999)**



Note: Households with children under 3 years old, no elder older than 65 not working and no sick individuals. Households deemed eligible for benefits according to the original classification scheme.

**Table 1: Sample of Mothers with Kids Younger than 3 Years Old**

	All	Treatment	Control
Mothers of Kids Younger than 3	4036	2571 (63.7%)	1465 (36%)
Mothers of Kids Younger than 3 & Teenagers 12 to 17	1509	973 (64.5%)	536 (35.5%)
Mothers of Kids Younger than 3 & Teenage Girls 12 to 17	979	636 (64.9%)	343 (35.1%)

Notes: Sample of mothers 18 to 44 living in households with no elder older than 65 not working and no sick individuals in May 1999. Households deemed eligible for benefits according to the original classification scheme.

**Table 2: Comparison of Raw Means between Treatment and Control Groups**

I. Dependent Variables (May 1999)	Treatment Group			Control Group			t-stat
	N	Mean	SD	N	Mean	SD	
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
<b>II. Covariates</b>							
<i>Current Maternal Characteristics (May 1999)</i>							
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
Hh 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
<i>Current Household Characteristics (May 1999)</i>							
Number of Kids 0 to 3	2571	1.25	0.465	1465	1.25	0.476	-0.382
Number of Kids 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 Daughters 12 to 17	2571	1.38	2.589	1465	1.29	2.512	0.918
<i>Baseline Household Characteristics (October 1997)</i>							
Number of Kids 0 to 3	2571	1.38	0.732	1465	1.41	0.740	-1.116
Number of Kids 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
<i>Baseline Community Characteristics (October 1997)</i>							
Pre-school =1	2514	91.45	0.280	1417	91.11	0.285	0.123
Junior High School 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 10%; \*Significant at 5%. Sample of mothers 18 to 44 living in households with no elder older than 65 not working and no sick individuals in May 1999. Households deemed eligible for benefits according to the original classification scheme.

**Table 3: Maternal Participation in Child Care and Maternal Extent of Participation in Child Care and Leisure**

		OLS	PROBIT	LEWBEL	OLS	PROBIT	LEWBEL	OLS	PROBIT	LEWBEL
		Mod 1A	Mod 1B	Mod 1C	Mod 2A	Mod 2B	Mod 2C	Mod 3A	Mod 3B	Mod 3C
<b>I. <u>Child Care Participation</u> =1</b> (Mean Dep Var = 0.51)	Special Regressor (v)	0.345 (8.853)	5.843 (22.617)	-	-5.314 (9.951)	-8.472 (25.386)	-	-2.533 (9.190)	-1.879 (23.580)	-
	Treatment (T) =1	0.014 (0.030)	0.020 (0.077)	0.007 (0.005)	-0.006 (0.030)	-0.033 (0.078)	-0.054 (0.053)	0.007 (0.030)	0.001 (0.077)	-0.003 (0.008)
	Children 12 to 17 =1				-0.166** (0.045)	-0.428** (0.120)	-0.600* (0.276)			
	T*Children 12 to 17 =1				0.095* (0.037)	0.246* (0.096)	0.435** (0.128)			
	Daughters 12 to 17 =1							-0.101 (0.071)	-0.259 (0.182)	-0.097 (0.085)
	T*Daughters 12 to 17 =1							0.062 (0.043)	0.164 (0.109)	0.257* (0.111)
	<hr/>									
<b>II. <u>Extent of Child Care Hours</u></b> (Mean Dep Var = 1.94)	Special Regressor (v)	6.700 (32.380)	11.828 (22.201)	-	-21.370 (36.543)	-6.995 (25.143)	-	1.093 (34.413)	6.041 (23.769)	-
	Treatment (T) =1	0.037 (0.107)	-0.000 (0.071)	-0.001 (0.002)	-0.001 (0.107)	-0.030 (0.070)	-0.059 (0.052)	0.030 (0.109)	-0.008 (0.071)	-0.007 (0.007)
	Children 12 to 17 =1				-0.344* (0.144)	-0.287** (0.105)	-0.369* (0.149)			
	T*Children 12 to 17 =1				0.297* (0.131)	0.217* (0.090)	0.306** (0.093)			
	Daughters 12 to 17 =1							-0.180 (0.232)	-0.145 (0.160)	-0.065 (0.056)
	T*Daughters 12 to 17 =1							0.094 (0.138)	0.098 (0.096)	0.142* (0.061)
<hr/>										
<b>III. <u>Extent of Leisure</u></b> (Mean Dep Var = 7.54)	Special Regressor (v)	13.184 (30.137)	-2.596 (19.251)	-	-21.415 (35.890)	-23.310 (22.717)	-	3.206 (32.887)	-7.842 (20.842)	-
	Treatment (T) =1	0.001 (0.104)	-0.035 (0.064)	-0.004 (0.003)	0.031 (0.105)	-0.016 (0.064)	0.056 (0.052)	0.014 (0.105)	-0.028 (0.064)	0.004 (0.007)
	Children 12 to 17 =1				0.250+ (0.139)	0.164+ (0.086)	0.493* (0.218)			
	T*Children 12 to 17 =1				-0.316* (0.126)	-0.191* (0.079)	-0.457** (0.118)			
	Daughters 12 to 17 =1							0.059 (0.217)	0.041 (0.135)	0.100 (0.067)
	T*Daughters 12 to 17 =1							-0.162 (0.141)	-0.086 (0.088)	-0.233** (0.078)

Notes: (N = 3710). +significant at 10%, \*significant at 5%, \*\*significant at 1%. SE in parentheses and clustered at the community level. Observations in the top 7% of Fv and with  $f(v|x) > 1 \cdot 10^{-5}$  have been trimmed. 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 years 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, or "telesecundaria" in the community.

**Table 4: First Daughters (Ages 12 to 17) Time Use - Child Care, Schooling and Leisure**

		OLS	PROBIT	LEWBEL
		Model A	Model B	Model C
<b><i>I. Child Care Participation =1</i></b> (Mean Dep Var = 0.18)	Special Regressor (v)	-6.994 (12.140)	-40.453 (53.705)	- -
	Treatment (T) =1	-0.085 (0.052)	-0.379+ (0.220)	-0.032** (0.011)
<b><i>II. Extent of Child Care Hours</i></b> (Mean Dep Var = 0.40)	Special Regressor (v)	-4.755 (26.676)	-38.338 (53.429)	- -
	Treatment (T) =1	-0.132 (0.112)	-0.356+ (0.215)	-0.032** (0.011)
<b><i>III. School Participation =1</i></b> (Mean Dep Var = 0.42)	Special Regressor (v)	35.239+ (18.103)	100.092* (50.348)	- -
	Treatment (T) =1	-0.066 (0.077)	-0.185 (0.214)	0.008* (0.004)
<b><i>IV. Extent of Hours in School</i></b> (Mean Dep Var = 2.51)	Special Regressor (v)	87.903+ (46.025)	86.352* (43.990)	- -
	Treatment (T) =1	-0.137 (0.194)	-0.146 (0.188)	0.002 (0.002)
<b><i>V. Extent of Leisure</i></b> (Mean Dep Var = 9.04)	Special Regressor (v)	17.391 (62.764)	10.248 (37.904)	- -
	Treatment (T) =1	-0.180 (0.239)	-0.116 (0.143)	-0.013* (0.006)

Notes: (N = 758) +significant at 10%, \*significant at 5%, \*\*significant at 1%. SE in parentheses and clustered at the community level. Observations in the top 7% of Fv and with  $f(v|x) > 1 \cdot 10^{-4}$  have been trimmed. All regressions include covariates (see Table 3).

**Table 5: Total Household Hours in Child Care**

Extent Household Hours in Child Care	OLS	PROBIT	LEWBEL	OLS	PROBIT	LEWBEL	OLS	PROBIT	LEWBEL
	Model 1A	Model 1B	Model 1C	Model 2A	Model 2B	Model 2C	Model 3A	Model 3B	Model 3C
Special Regressor (v)	16.364 (39.153)	17.024 (21.341)	- (0.002)	-19.451 (44.630)	-3.856 (24.416)	- (0.052)	13.278 (41.990)	11.820 (23.157)	- (0.007)
Treatment (T) =1	-0.003 (0.129)	-0.029 (0.070)	-0.001 (0.002)	-0.039 (0.129)	-0.056 (0.069)	-0.059 (0.052)	-0.004 (0.130)	-0.035 (0.070)	-0.007 (0.007)
Children 12 to 17 =1				-0.290 (0.189)	-0.243* (0.110)	-0.342* (0.138)			
T*Children 12 to 17 =1				0.340* (0.165)	0.220* (0.090)	0.299** (0.091)			
Daughter 12 to 17 =1							-0.084 (0.297)	-0.107 (0.161)	-0.061 (0.054)
T*Daughter 12 to 17 =1							0.041 (0.177)	0.080 (0.097)	0.148* (0.062)

Notes: (N=3710) Mean Dep. Var. = 2.16 hours. \*significant at 5%, \*\*significant at 1%. SE in parentheses and clustered at the community level. Observations in the top 7% of Fv and with  $f(v|x) > 1 \cdot 10^{-5}$  have been trimmed. All regressions include covariates (see Table 3).



**Table 6: Share of Child Care Hours Provided by Different Household Members**

	OLS												
	Share Mums	Share Mums	Share Other Women	Share Other Women	Share Men	Share Men	Share Sons 8 to 11	Share Sons 8 to 11	Share Sons 12 to 17	Share Sons 12 to 17	Share Daughters 8 to 11	Share Daughters 8 to 11	Share Daughters 12 to 17
Treatment (T) =1	0.025 (0.021)	0.009 (0.024)	0.012 (0.024)	-0.013 (0.033)	-0.006 (0.006)	-0.008 (0.008)	-0.002 (0.005)	-0.007 (0.008)	0.008 (0.007)	0.005 (0.011)	-0.010 (0.011)	-0.026 (0.016)	-0.031* (0.015)
Girls 12 to 17 in the Household =1		-0.072* (0.033)		-0.051 (0.046)		0.002 (0.010)		-0.011 (0.011)		0.006 (0.015)		-0.048** (0.018)	
T*Girls 12 to 17 in the Household =1		0.051 (0.032)		0.058 (0.045)		0.006 (0.009)		0.011 (0.011)		0.005 (0.015)		0.041* (0.018)	
Observations	4036	4036	476	476	4100	4100	1382	1382	1047	1047	1381	1381	988
Mean Share Child Care Hours	0.44	0.44	0.10	0.10	0.03	0.03	0.02	0.02	0.02	0.02	0.04	0.04	0.08

Notes: \*significant at 5%, \*\*significant at 1%. SE in parentheses clustered at the community level. All regressions include covariates (see Table 3).  
Share Household Member m = Total Hours to Child Care of Household Member m / Total Household Hours to Child Care.