Girls' school attendance: A Dynamic discrete choice structural approach*

Hugo Reis†
Bank of Portugal

July 2014

Abstract

This paper studies the determinants of girls' schooling choices in a rural area of Pakistan where drop-out rates are considerably high for older girls. The paper estimates a dynamic structural model that allows for the analyses of interrelations between girls' schooling and mothers' labour market participation decision. The model explicitly takes into account the role of home production as an important economic factor. Moreover, the paper simulates the impact of different policies to increase girls' educational attendance. Our results suggest that monetary incentives, such as the conditional cash transfer (CCT), are a good mechanism to increase girls' school enrollment, but not necessarily the most cost effective. From a policy perspective, a school building program and the availability of free daycare centers would induce a greater impact on educational participation at similar cost. The impact of the gender CCT program on older girls' secondary school enrollment rate was only 34% as large as the school building program, while the same impact stood at 91% of the mandatory child care program. In terms of welfare, the impact of the gender CCT program is much closer to the impact of the school building program and becomes bigger when compared with the mandatory child care program.

Keywords: Girls' schooling, Dynamic discrete choice models, Development economics, Structural model
JEL Classification:I25,I28

*I am very grateful to Pedro Carneiro and Orazio Attanasio for their support and encouragement. I'm particularly thankful to Mónica Costa Dias for discussions and insights. I also have benefited from insightful comments made by Petra Todd, Lars Nesheim, Michael Keane, Aureo de Paula, Pierre-André Chiappori, Jishnu Das, Dan Black and Sankar Mukhopadhyay and participants at the 13th IZA/SOLE Transatlantic Meeting of Labor Economists. Hugo Reis acknowledges the support of Fundação para a Ciência e Tecnologia, World Bank and Banco de Portugal.

†Address: Economics and Research Department, Banco de Portugal, Av. Almirante Reis 71, 1150-012 Lisboa, Portugal. email: hfreis@bportugal.pt

1 Introduction

It is widely acknowledged that human capital plays a critical role in economic growth in developing countries, which enhanced political efforts towards educational participation. In particular, it is emphasised the role of women's education (Behrman et al. (1999)) and the role of human capital of future generations (Becker (1994)).

Although lower income countries have made significant improvements in terms of school enrolment, with a net primary enrollment rate close to 90% and a net secondary enrollment rate above 60% in 2012, Pakistan is still far behind reaching only 72% and 36% respectively. In Pakistan and in most developing countries women's literacy rates are dramatically low. Women are situated clearly at the bottom or lower end of the educational system in comparison to their male counterparts, particularly in rural and sub-urban areas.

Therefore, it is important to understand if investing in girls' education is efficient or why should people invest in girls' education in rural Pakistan. For instance, there is no market failure if households know private returns, if they are not altruists and if they are not credit constrained. In the opposite scenario, there is an efficiency argument to be in favour of more educational participation (higher school enrollment). Finally, disregarding efficiency, government may have simply different preferences from households.

Independently of the efficiency analysis, education can play a vital role in enhancing the status of women. An educated mother tends to have more influence in household decisions and may allow her to obtain more resources for her children, providing better education and health care[1]. Moreover, an educated girl is more likely to be self-confident, to take more informed decisions, to participate more in the formal labour market, to earn more income, to marry later, and to have smaller families. These benefits transmit across generations, as well as to the communities.

The aim of this paper is to examine the determinants of girls' schooling decision in a developing/lower income country. This study develops and estimates a dynamic structural model in which parents (mother) make labour and consumption decisions along with children's (girls') schooling decisions. The benefits of schooling are assumed to be uncertain,

[1]Mothers' education is highly related to their children's educational attainment. There is an extensive empirical literature on the strong association between parental schooling and children's outcomes focusing on developed and developing countries surveyed in Behrman (1997).

while the costs of schooling change (probably increase) as the child grows. Such a model allows for analysis of interrelations between girls' schooling choices and mothers' labour participation decision. The model allows for simple dynamic interactions in parental-decision making. For example, it can be expected that a girl whose mother participates in the labour market would be more valuable at home, replacing her mothers' housework. A decision model in which parents decide simultaneously labour market status and children's schooling emphasizes the relationship between the two decisions.

To increase school attendance rates among poor children in developing countries, policy makers have implemented several measures, namely conditional cash transfer (CCT) programs[2]. To explicitly address intra-household disparities in human capital investment some countries have implemented gender-targeted CCT (e.g. Female School Stipend Program in Punjab, Pakistan). Despite these efforts, educational participation is far from targets proposed by several international institutions, in particular for girls in Pakistan[3].

One of the main advantages of having a structural model is the possibility of performing policy experiments. In this framework, it is possible to analyse the effect of policies intended to increase school participation on girls' schooling and also the impact on mothers' labour market participation. Poverty, cultural constraints, and an inadequate supply of government schools for girls in some rural areas are three possible factors explaining the low school enrollment and the persistent disadvantage of rural girls in Pakistan. All of these features will be addressed by the model.

Overall, the contribution of this paper is twofold. First, it addresses important policy issues, explaining why drop-out rates for older girls are so high and comparing alternative measures to increase girls' educational participation in developing countries[4]. Second, the paper also contributes to a growing literature that addresses empirical questions using discrete choice dynamic programming models of individual behaviour.

[2]For more details regarding these programs see a policy research report of the World Bank by Fiszbein and Schady (2009).

[3]Concerning education, Pakistan is far behind the Millennium Development Goals (MDG). According to the MDG (2010) report, the Pakistani net primary enrollment ratio was around 60 percent, while the MDG target for 2010 was close to 80 percent. Another important indicator, the percentage of literacy rate is also behind. For girls, the target was around 65 percent in 2010 but the real figure is below 50 percent.

[4]This paper does not however explain whether educational investment it is efficient. Nevertheless, the positive estimate regarding the terminal value function parameter provide some evidence against the lack of knowledge about private returns.

These models are attractive as structural parameters have a clear interpretation and are useful tools for the evaluation of counterfactual policies (Aguirregabiria and Mira (2010), Keane et al. (2011) and Todd and Wolpin (2010)). The schooling-labour decision model developed in this paper builds on several literature, including dynamic models of occupational choices as in Keane and Wolpin (1997), dynamic model of employment-child care decisions of women as in Bernal (2008), and more closely related are the work of Attanasio et al. (2012) and Todd and Wolpin (2006) in which a dynamic schooling behavioural model is used to evaluate the impact of monetary incentives provided to families to increase their children's school attendance.

We extend the literature to a lower-income developing country using detailed information on the school choices decision in rural Punjab province, the largest state in Pakistan, using the Learning and Educational Achievement in Punjab Schools (LEAPS) dataset. In this paper, the mother makes a time allocation decision for herself and for each child age 6 to 15 in a similar way to Bernal (2008) and Todd and Wolpin (2006)[5]. However, we depart from their approach as we allow for the mother's labour supply decision and allow girls to drop out from school to work at home (modelling explicitly an home production function)[6] . The household receives utility in each period from their children's current school attendance and from the mothers' working decision. The model also allows the household to receive utility from consumption and from a home production, which depends on children's age and the number of very young children.The household's budget constraint includes the parents' exogenous income, the wage income of the mother, and educational expenditure for the children that go to school. There is an extra utility cost to attending school (at secondary level) that depends on the availability of school in the village. Households differ in their preferences for the choice variables according to children's age, mothers' age and education, and according to discrete "type" regarding tastes for the mutually exclusive decisions. The parameters of the model are estimated by simulated

[5]In Bernal (2008) the mother makes child care and labour market participation decisions. In Todd and Wolpin (2006), married couples are assumed to make sequential decisions over a finite horizon about the time allocation of all of their children age 6 through 15, including their school attendance and labour market participation, and about the timing and spacing of births. Parents' labour supply decisions are not included in their model.

[6]In contrast to Todd and Wolpin (2006), and despite being a potentially interesting feature, we are not considering a childbirth decision in the model.

maximum likelihood.

Among other factors, the results highlight the important role played by home production to explain the high drop-out rates among older girls, in particular at the secondary level, capturing in part a marriage premium effect. The estimated model fits the actual data reasonably well. In particular, it replicates the observed distribution of schooling choices for each particular child's age, i.e., the key decision analysed in the paper. In addition, the model simulates relatively well the results from the gender targeted CCT program developed in Punjab.

Policy simulations suggest that monetary incentives, as the conditional cash transfer case, are a good mechanism to keep and increase girls' school enrollment but not necessarily the most cost effective. This mechanism seems less effective in rural areas as older girls leave school to help in housework, instead of working for a salary. For girls in secondary school, the enrollment rate increases from less than 1 percentage point to around 5 percentage points, when the grant level is around 4 and 30 percent of average income, respectively. From a policy perspective, reducing transportation costs in secondary school through a school building program and making the use of free daycare centers mandatory are good alternatives at a similar cost. In particular, the impact of the CCT program on older girls' school enrollment rate is only 34% and 91% as large as the school building program and the mandatory free daycare centers, respectively. A combination of the different measures may be a good option as they may act throughout different mechanisms. The effect of these measures on mothers' labour market participation turns out to be negligible. Nevertheless, for older girls, if the mother's wages increase by 10 and 20 percent, with an employment/wage subsidy, mothers' labour market participation rises by 1 and 1.8 percentage points, respectively.

In terms of welfare, as the CCT scheme is an anti-poverty program and not specifically an educational program, the difference between schemes is much smaller. In particular, for the older girls, the utility gain of the school building program is equivalent to an increase in income of around 3 percent. The CCT cost comparable scheme effect is 83% as large as the school building program. Interestingly, given the mandatory nature of the program, the child care services scheme would lead to a negligible effect in terms of welfare, while the impact of the CCT cost comparable scheme is equivalent to an increase in income of around 2 percent.

The paper is organized as follows. In Section 2 we describe the structure of the model and estimation method. Section 3 describes the LEAPS data on which we estimate the model and highlights the overall patterns in the data. Section 4 presents the estimates of the model and evaluates its ability to fit the data. Section 5 shows the results from several policy experiments and Section 6 concludes.

2 Behaviour Model and Estimation

This section presents a structural model of mothers' sequential decisions about work and childrens' schooling. In each discrete time period t, the mother makes a time allocation decision for herself and for each girl age 6 to 15[7]. In particular, the mother decides whether to send a daughter to school or to let the girl remain at home (after age 12 this implies working at home). At age 16 children are assumed to become independent and make their own schooling and work decisions. In contrast to the mothers' labour supply decision, fathers' income is exogenous and the household cannot save or borrow. Educational expenditure on children is also subtracted from parental income to determine consumption. Specifically, we allow for two mother working options {work, no work} and two child schooling choices {enrolled, not enrolled} with a total of 4 possible options in a mother's choice set. The choice set is denoted as: $J = \{h_t, s_t : h_t = \{0, 1\} \text{ and } s_t = \{0, 1\}\}$, where h_t is an indicator for whether or not the mother works in the market, and s_t is an indicator for whether or not the girl is going to school in period t.

2.1 Utility Function

The current-period utility function given choice of option (h_t, s_t) is given by:

$$U_t(h_t, s_t; \Omega_t) = \frac{1}{\gamma} C_t^{\gamma} + \alpha_1 s_t + \alpha_2 h_t + \epsilon_t^{(h_t, s_t)} \tag{1}$$

where C_t is total consumption, α_1 and α_2 are, respectively, the utility/disutility from

[7] Note that siblings from the same mother have the same wage shock. In addition, as a robustness check, we estimate the model only for the older girls and we obtain similar results.

sending a child to school and from working, and $\epsilon_t^{(h_t,s_t)}$ is an alternative-specific preference shock. The utility function has the common CRRA form in consumption (γ).

2.1.1 Consumption

Total consumption is a composite of goods and services purchased in the market, c_t^M, and goods and services produced in the home, c_t^H. In particular,

$$C_t = [\theta c_t^{M\rho} + (1-\theta) c_t^{H\rho}]^{\frac{1}{\rho}} \tag{2}$$

and the parameter ρ controls the household's willingness to substitute between c_t^M and c_t^H (the larger is ρ , the greater is this willingness).

Consumption in the market, c_t^M, is given by the budget constraint

$$c_t^M = y_t + w_t h_t - f_t s_t - \psi v_t s_t \tag{3}$$

The direct costs of attending school (f_t) are the costs of buying books as well as clothing items such as shoes[8]. The model allows for transport costs, a parameter to be estimated (ψ) that captures the effect of a village without an elementary/secondary school. If there is no school in a village it can become very costly to send a girl to school and this may be an important mechanism to explain the high drop-out rates among older girls. v_t is a categorical variable that takes value 1 if the village has no elementary or secondary school. y_t represents father's income and w_t the wage income earned by the mother if she decides to work.

The mother's wage affects the children's schooling decision in two opposite ways. On one hand, if w_t increases or the mother accepts a job there is an income effect. On the other hand, if the mother accepts a job or a higher wage this may lead older girls to stay at home to home-produce.

In addition to the budget constraint, in the spirit of Greenwood et al. (1993), mother

[8] If a child is going to private school it should also include the tuition fee. However, the results do not change if we add this cost to the model. In the sample, around 30 percent of the students go to private school but the private schools are relatively affordable (see Carneiro et al. (2013) for details).

has a home production function:

$$c_t^H = (\delta_0 + \delta_1 n_{05,t})[(1 - h_t) + (\eta_0 age^c_{1213,t} + \eta_1 age^c_{1415,t})(1 - s_t)] \quad (4)$$

The home production function yields consumption of the home goods and services (c_t^H) as a function of the time spent in home work of both mother $(1 - h_t)$ and children $(1 - s_t)$, and household characteristics, in particular the number of children less than five years old ($n_{05,t}$). This is another important mechanism that may lead girls to drop out; their value at home may be greater if there are very young children in the household. The number of young children enters as a joint production $(\delta_0 + \delta_1 n_{05,t})$, implying that home goods are produced at the same time and younger children do not change productivity of other home goods. We assume that home production is only technology. In addition, working at home is an option only after age 12 and we allow the value of having older girls at home to vary with age (η_0 and η_1). This captures the different effect of girls aged 12-13 vs. 14-15[9]. These terms try to capture the increasing number of hours working at home by age observed in the data[10]. $age^c_{1213,t}$ and $age^c_{1415,t}$ are dummy variables that take value 1 when children are 12 to 13 and 14 to 15 years old, respectively. This specification of the home production function implies that c_t^H is more/less valuable when the family has very young children if δ_1 is positive/negative[11].

2.1.2 Wage

One key factor is the mother's wage and we do not observe wages for mothers who are not working. We thus specify a standard Mincer type wage equation, in which the wage of a woman is determined by her age (age_t^m), education ($educ_t^m$), and local labour market conditions (z_{tv}) in village v according to

$$\ln w_t^m = \varphi_0 + \varphi_1 age_t^m + \varphi_2 educ_t^m + \varphi_3 z_{tv} + \xi_t^m \quad (5)$$

[9] Different specifications were used, in particular a quadratic form for children's age. This specification of the home production was determined in part using model fit criteria.

[10] Mothers spent on average around 9 hours per day working at home. Girls 12 and 13 years old working at home spent, on average, around 70 percent of the mothers time, while the older ones (between 14 and 15) spent more or less the same amount of time as mothers.

[11] Despite being a potentially interesting feature, we are not considering fertility decisions in the model, which in our setup does not influence the main results.

where φ_0 represents the log price of human capital, which we assume to be the same across villages, and $\xi_t^m \sim N(0, \sigma_w^2)$ represents a wage shock[12].

Following Attanasio et al. (2012), we estimate this wage equation separately from the rest of the model using a Heckman selection model. We then use predictions from this equation in place of actual wages and use $\widehat{\sigma}_w^2$ ($\xi_t^m \sim N(0, \widehat{\sigma}_w^2)$) to introduce the wage shock in the model.

2.1.3 Heterogeneity

Finally, we allow for observed and unobserved heterogeneity in mother's tastes for work (α_2)

$$\alpha_{2,k} = \alpha_{21} educ_t^m + \alpha_2 age_t^m + \overline{\alpha_{2,k}} \tag{6}$$

and in their tastes for children's schooling (α_1):

$$\alpha_{1,k} = \alpha_{11} educ_t^m + \alpha_{12} age_t^m + \alpha_{13} age_t^c + \alpha_{14} (age_t^c)^2 + \overline{\alpha_{1,k}} \tag{7}$$

where $\overline{\alpha_{1,k}}$ and $\overline{\alpha_{2,k}}$ represent the unobservable heterogeneity of the children and mother, respectively. We assume that there are two different types (k=low and high). Associated type proportions are denoted by $\Pi = \{\pi_l^m, \pi_h^m, \pi_l^c, \pi_h^c\}$, which are parameters to be estimated (Heckman and Singer (1984)).

Finally, we allow the term $\epsilon_t^{(h_t, s_t)}$ to be correlated across alternatives, capturing the fact that some alternatives are more similar than others[13]. We assume to have a joint normal distribution and are serially uncorrelated.

2.2 Value Function

At every period, the agent chooses his action $(h, s) \in J$ to maximize expected utility at t, the value function, which is given by

$$V_t(\Omega_t) = \max_{(h,s)_{t,\dots T}} E\left\{ \sum_{\tau=t}^{T} \beta^{\tau-t} U(h_\tau, s_\tau; \Omega_\tau) + \beta^{T+1-t} V_{T+1}(\Omega_{T+1}) \mid \Omega_t \right\} \tag{8}$$

[12] In the model, siblings have the same wage shock.

[13] It is expected that the alternative of going to school and mother working is correlated with going to school and mother not working but completely unrelated with not going to school and mother not working.

where $\Omega_t = \{f_t, v_t, y_t, educ_t^m, educ_t^c, age_t^m, age_t^c, n_{05,t}, z_t, \epsilon_t, \xi_t^w, \pi\}$ represents the state space (all the relevant factors affecting current or future utility) at time t. The parameter β represents the discount factor[14].

Household maximizes the expected utility subject to

Budget Constraint

$$c_t^M + f_t s_t + \psi v_t s_t = y_t + w_t h_t$$

Home production equation

$$c_t^H = h((1 - h_t), (1 - s_t), n_{05,t}, age_t^c)$$

Mother's Time Constraint

$$h_t \in \{0, 1\}$$

Child's Time Constraint

$$s_t \in \{0, 1\}$$

Mother's wage equation:

$$\ln w_t^m = \varphi_0 + \varphi_1 age_t^m + \varphi_2 educ_t^m + \varphi_3 z_{tv} + \xi_t^m$$

Children may not be successful in completing the grade, originating an important source of uncertainty to the model.

Child Education law of motion:

$$\begin{aligned} educ_{t+1}^c &= educ_t^c + 1 \quad if \quad s_t \neq 0 \quad and \quad progress \\ &= educ_t^c \qquad if \quad s_t = 0 \quad or \quad fail \end{aligned}$$

In a dynamic programming framework, the value function can be written as the maximum over alternative-specific value function, $V_t(h_t, s_t; \Omega_t)$,

[14] Since we do not model savings and borrowing, this will reflect liquidity constraints or other factors that lead the households to disregard more or less the future.

$$V_t(\Omega_t) = \max_{\substack{h_t \in \{0,1\} \\ s_t \in \{0,1\}}} [V_t(h_t, s_t; \Omega_t)] \tag{9}$$

The choice specific value function can be written as the expected discounted value that satisfies the Bellman equation:

$$V_t(h_t, s_t; \Omega_t) = U_t(h_t, s_t; \Omega_t) + \beta E \left[\max_{\substack{h_{t+1} \in \{0,1\} \\ s_{t+1} \in \{0,1\}}} V_{t+1}(h_{t+1}, s_{t+1}; \Omega_{t+1}) \mid \Omega_t, h_t, s_t \right] \tag{10}$$

If a grade is not completed successfully, we assume that the level of education does not increase. This may be important since failure may discourage school attendance. We assume that the probability of failing to complete a grade is exogenous and does not depend on the willingness to continue schooling. However, we allow this probability to vary with the grade in question, with the age of the individual, and with the mother's education. Moreover, we assume it is known to the individual[15]. In this case the choice-specific value function becomes:

$$\begin{aligned} V_t(h_t, s_t; \Omega_t) = U_t(h_t, s_t; \Omega_t) + \beta E\{ & pV_{t+1}(h_{t+1}, s_{t+1}; educ^c_{t+1} = educ^c_t + 1, \widetilde{\Omega_{t+1}}) \\ & +(1-p)V_{t+1}(h_{t+1}, s_{t+1}; educ^c_{t+1} = educ^c_t, \widetilde{\Omega_{t+1}})\} \end{aligned}$$

where

$$\begin{aligned} p &= p(educ^c_{t+1} = educ^c_t + 1 \mid s_t \neq 0 \;\; and \;\; progress) \\ 1-p &= p(educ^c_{t+1} = educ^c_t \mid s_t = 0 \;\; or \;\; fail) \\ \widetilde{\Omega} &= \Omega \setminus \{educ^c\} \end{aligned}$$

At $t = 11$ (girl who is 16 years old) the children schooling decision is taken by the child and not by the mother. We assume a terminal value function that is a function of children's level of education:

$$V_{T+1}(\Omega_{T+1}) = V_{T+1}(educ^c_{T+1})$$

[15]We estimate the probability of progressing for each grade as the ratio of individuals who passed to the next grade compared to the year before at a particular age and mother's education (see Table 13 in appendix).

Specifically, following Eckstein and Lifshitz (2011), we use a linear approximation

$$V_{T+1}(educ^c_{T+1}) = \phi educ^c_{T+1} \tag{11}$$

This parameter is capturing mainly beliefs about returns to education and altruism. Using different specifications as in Bernal (2008), Attanasio et al. (2012) or del Boca and Flinn (2012) does not seem to be critical to the main results[16].

2.3 Likelihood

Having solved the dynamic optimization problem, we are able to get the likelihood function. The probability that a mother chooses alternative $(h, s) = j$ at time t from her choice set J is given by:

$$\Pr((h_t, s_t) = j \mid \Omega_t) = \Pr[V^j_t(\Omega_t) \geq V^q_t(\Omega_t), \forall_{q \neq j \in J}]$$

An individual contribution to the likelihood is:

$$L_{it} = \sum_{\pi \in \Pi} \{ \prod_{t=t_i}^{T} \prod_{j \in J} \Pr((h_t, s_t) = j \mid \Omega_t)^{1[(h_t,s_t)=j]} \} \pi \tag{12}$$

where t_i is the first observation available in the data for each individual i.

2.3.1 Initial Condition

We assume that the initial conditions are exogenous conditional on type except for the children's level of education $(educ^c_t)$. The presence of stock of education generates an important initial conditions problem because we do not observe the entire history of schooling for the children in the sample, with the exception of those who are 6 years old, with information of education level. We cannot assume that the random variable in equation (7) is independent of past school decisions as reflected in the current level of schooling $(educ^c_t)$. To solve this problem we simulate life from 6 years of age until first observation is available

[16] Bernal (2008)): $V_{T+1}(educ^c_{T+1}) = U_{T+1}(h_t, s_t, \Omega_t) + f(educ^c_{T+1})$,
Attanasio et al. (2012): $V_{T+1}(educ^c_{T+1}) = \frac{\delta_1}{1+\exp(-\delta_2 educ^c_{T+1})}$ and
del Boca and Flinn (2012): $V_{T+1}(educ^c_{T+1}) = \delta educ^c_{T+1}(1-\beta)^{-1}$

to obtain the education level simulated for each individual and each type[17]. Then, conditional on each type and age, we get the probability of each girl with a certain education level ($f(educ_t^c|\pi_i)$, i.e., we construct the distribution of initial condition ($\widehat{f}(educ_t^c|\pi_i)$[18].

Then we incorporate this correction in the likelihood[19]:

$$L_{it} = \sum_{\pi\in\Pi}\{\prod_{t=t_i}^{T}\prod_{j\in J}\Pr((h_t,s_t)=j\mid\Omega_t)^{1[(h_t,s_t)=j]}f(educ_t^c|\pi)\}\pi \quad (13)$$

where t_i is the first observation available in the data for each individual i.

Evaluation of the likelihood itself requires the calculation of three-variate integrals. We use a GHK recursive probability simulator (Keane (1994)) of the choice probabilities and form a simulated maximum likelihood estimator[20].

2.4 Identification Issues

The influential works of Rust (1994) and Magnac and Thesmar (2002) have obtained negative results on the nonparametric identification of dynamic discrete structural models[21].

[17] A girl who is 7 years old has 0 years of education (if not enrolled or failed) or 1 year of education (if pass). At 8 years old, a girl can have 0, 1, or 2 years of education. The same logic is applied for the subsequent ages.

[18] For example, $educ_5^c$ (first observation available) implying that $educ_1^c, educ_2^c, educ_3^c$, and $educ_4^c$ are not observed. With many draws we simulate life from 1 to 5 obtaining $educ_5^c$ simulated for each type and obtain $f(educ_5^c|\pi)$.

[19] Let $t = t_i, T$

$$\Pr((h_t,s_t)=j\mid\Omega_t)^{1[(h_t,s_t)=j]}f(educ_t^c|educ_{t-1,...,}^c educ_{t_i}^c,\pi)f(educ_{t-1}^c|educ_{t-2,...,}^c educ_{t_i}^c,\pi)...f(educ_{t_i}^c|\pi)$$

In this case, $f(educ_t^c|educ_{t-1,...,}^c educ_{t_i}^c,\pi)$ and $f(educ_{t-1}^c|educ_{t-2,...,}^c educ_{t_i}^c,\pi)$ are deterministic and equal to 1. Then, the correction to the contribution to the likelihood becomes only the term $f(educ_{t_i}^c|\pi)$. This term is simulated as discussed above.

[20] Geweke et al. (1994) present a detailed discussion and in a Monte Carlo study of alternative approaches to simulation based inference concluded that classical methods based on GHK outperformed classical methods based on kernel smoothed probability simulators. Train (2009) also provides a good discussion of GHK probability simulator.
The GHK probability simulator was calculated conditional on the wage shock (ξ_t^m).

$$\widehat{P_j} = \frac{1}{D}\sum_{d=1}^{D}\widehat{P_{GHK}}((h_t,s_t)=j|\xi_t^m)$$

[21] Magnac and Thesmar (2002) characterize the degree of non-identification and show that the model is identified subject to ad hoc assumptions on distributions of unobservables and functional forms of agents' preferences. In particular, their identification result indicates that parametric specifications on the distributions of unobservables are indispensable for identifying the deep structural parameters of interest, which consequently motivates the maximum likelihood estimation. More recently, only under very specific conditions Aguirregabiria (2010) provides a nonparametric approach to evaluating the behavioral and welfare effects of counterfactual policies using a dynamic structural model.

Therefore, it is unfortunately not possible to provide a rigorous proof of identification for the parameters of the model[22]. Nevertheless, we'll provide some intuitive arguments regarding how the variation in the data may help in the parameters identification. In particular, despite not estimating the model by GMM we'll think and discuss identification in terms of moments. In addition, note that the identification comes also from the functional form assumed in the model.

We start the identification discussion on the home production parameters, as being one of the main focus of this paper. In this case, for each period, the covariance between mothers' working choices and girls' schooling decision ($Cov(h_t, s_t)$) conditional on the number of young children ($n_{05,t}$), children aged 12 and 13 ($age^c_{1213,t}$), children aged 14 and 15 ($age^c_{1415,t}$), helps in the identification of δ_0, δ_1, η_0 and η_1, the parameters of equation (4). In addition, the covariance between wages and mothers' working choices ($Cov(w_t, h_t)$) and between wages and girls' schooling decision ($Cov(w_t, s_t)$) are also an important source of variation that aid the identification of home production parameters.

Another important specification of the model is the heterogeneity in preferences which characterizes mainly the cultural factors. Regarding preferences towards mothers' working decision, the covariance between mothers' working decision and mothers' education ($Cov(h_t, educ^m_t$) and the covariance between mothers' decision and mothers' age ($Cov(h_t, age^m_t$) are important sources to provide identification of α_{21} and α_{22} in equation (6), respectively. Concerning preferences towards girls' schooling decision, the parameters in equation (7), we can apply a similar argument. The covariance between girls' schooling decision and mothers' education ($Cov(s_t, educ^m_t)$) and the covariance between girls' schooling decision and mothers' age ($Cov(s_t, age^m_t)$) helps in the identification of α_{11} and α_{12}, respectively. Furthermore, the relation between girls' schooling decision and girls' age helps in the identification of α_{13} and α_{14}, in particular we can think of the covariance between girls' schooling decision and girls' age ($Cov(s_t, age^c_t)$) and girls' age squared ($Cov(s_t, (age^c_t)^2)$). Another important source of variation in the data comes from the father's income (y_t) as the income effect becomes more relevant in this case. Thus, moments like the covariance between the father's income and mothers' working decision ($Cov(y_t, h_t)$) and the covari-

[22] We have no problems in terms of convergence and the results are robust to different starting points. This is a necessary condition but naturally far from being a sufficient one.

ance of father's income and girls' schooling decision ($Cov(y_t, s_t)$) helps in the identification of the preferences parameters. In addition, for the preferences towards girls' schooling decision an additional source of variation comes from the direct costs of attending school (f_t).

The parameters regarding unobservable heterogeneity of the children and mother ($\overline{\alpha_{1,k}}$ and $\overline{\alpha_{2,k}}$) and the type distribution $\Pi = \{\pi_l^m, \pi_h^m, \pi_l^c, \pi_h^c\}$ are also identified by each period choices. Heckman and Singer (1984) discuss the identification of unobserved heterogeneity related to duration models [23].

The variation in the variable not having a school in the village (v_t) conditional on other variables provide an important source of identification of the parameter capturing the cost of not having a school in the village ψ. In particular, the covariance between not having a school in the village and girls' schooling decision ($Cov(v_t, s_t)$) helps in the identification of this parameter.

Furthermore, moments relating mothers working choices (h_t) with girls' final education at age 16 ($educ^c_{T+1}$) help the identification of the terminal value function parameter (ϕ).

In this framework, since the household cannot save or borrow it works as we observe consumption. Therefore, consumption variation and moments relating choices with consumption help in the identification of the remaining parameters of the utility function.

Finally, we need to take the utility differences and to normalize for scale and level of utility to provide identification of the covariance matrix parameters. In particular, we take the difference against the alternative girls in school and mother not working choice and normalize the difference $\epsilon_t^{(1,1)} - \epsilon_t^{(0,1)}$ to 1. Train (2009) discusses the identification of the variance-covariance matrix using the GHK simulation with maximum likelihood.

3 Data

The paper uses the Learning and Education Achievement in Punjab School (LEAPS) project data set[24]. The LEAPS data is collected from 112 villages in the Punjab province,

[23]The idea is to allow for agents to differ in permanent ways unobserved to the econometrician and estimate the distribution of types to fit the persistence of various choices and outcomes of the agents. Naturally, if the number of types were allowed to approach the number of observations and allowed to vary over time then the data could be fit perfectly.

[24]The project details are available at www.leapsproject.org.

the largest state in Pakistan, located in the three districts of Attock (North), Faisalabad (Center), and Rahim Yar Khan (South). Villages were chosen randomly from a list of those with at least one private school according to the 2000 census of private schools. This allows for the study of the differences between private and public schools in the same village. The baseline survey in 2004 covered 823 schools (government and private) and around 1800 households (with almost 6000 children, 48 percent of which are girls). Table 1 presents the girls' age distribution between 6 and 15, which represents around 75 percent of all girls less than 15 years old.

Table 1: Girls' age distribution

Age	Observations	Percentage
6	158	8.0
7	191	9.7
8	227	11.5
9	195	9.9
10	245	12.4
11	180	9.1
12	225	11.4
13	188	9.6
14	219	11.1
15	141	7.2
	1969	100.0

Note: This table shows the girls' age distribution between 6 and 15 years of age, which represents around 75 percent of all girls less than 16 years old.

Table 2 presents mean characteristics of the variables used in the model estimation. The mean age of the girls in the sample is 10.5 years old with a level of education only slightly above 2 years. On average, the school enrollment rate is below 75 percent (around 80 percent between 6 and 11 years old and about 55 percent from 12 to 15) and only around 11 percent of the mothers work in the labour market. In addition, the mothers' mean age is 38 years old (around 40 percent are more than 40 years old) and 75 percent of the mothers have no education. Households have on average 1 young child (less than 5 years old), and around 30 percent of the households have more than one. Regarding school availability, only around 60 percent of the sample have an elementary/secondary school in the village[25]. Father's income is, on average, about 7110 rupees (approximately

[25]For only secondary schools the availability drops to around 40 percent.

83 U.S. dollars). Among those mothers who worked, the average wage was about 1200 rupees (around 14 U.S. dollars)[26].

Table 2: Sample summary statistics

	Obs	Mean	Std.Dev	Min	Max
Girls' school attendance^d	1969	0.72	0.45	0	1
Mother's working status^d	1969	0.11	0.32	0	1
Girls' age	1969	10.5	2.73	6	15
Girls' education	1969	2.28	2.23	0	8
Mother's education^d - (with at least some education)	1969	0.25	0.43	0	1
Mother's age^d - (more than 40 years old)	1969	0.40	0.49	0	1
Mothers with at least one young child^d	1969	0.62	0.49	0	1
Mothers with more than one young child^d	1969	0.29	0.46	0	1
Girls in villages with elementary/secondary school^d	1969	0.62	0.49	0	1
Female Local Labour Market Conditions					
Good (low village unemployment rate)^d	1969	0.36	0.48	0	1
Medium (medium village unemployment rate)^d	1969	0.28	0.45	0	1
Bad (high village unemployment rate)^d	1969	0.35	0.48	0	1
Father's income	1969	7.11	8.51	0.34	209.73

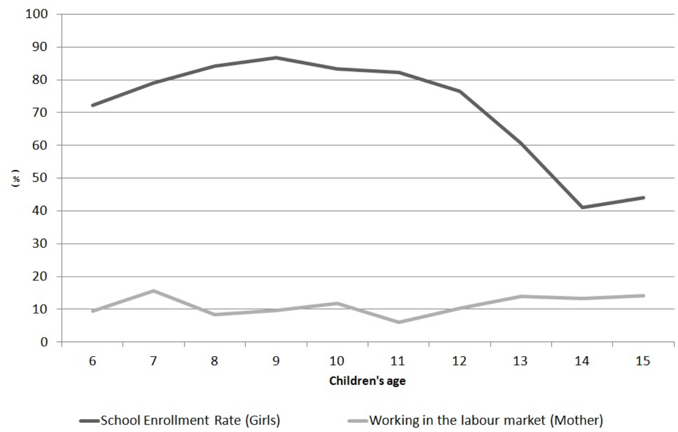
Notes: d - Categorical variables were used in order to obtain a more reasonable state space size. Budget constraint variables are expressed in rupees divided by 1000. For example, the monthly average parents' income is equal to 7110 rupees (≈ 83 U.S. dollars).
Female Local Labour Market conditions is measured as the village female unemployment rate. Using the household census data we split the 125 villages in 3 categories according to their female unemployment rate (up to 25th percentile, 25th to 75th percentile, and above 75th percentile). More specifically, the 25th percentile corresponds to an unemployment rate just below 6 percent and the 75th percentile just above 30 percent.

Figure 1 displays mother employment and child schooling choices of women in the sample with children between 6 and 15 years old. Enrollment rates decrease dramatically after 12 years old, reaching levels around 40 percent for the 14 and 15 year old girls. In general, mother labour market participation is low, without a clear girls' age pattern.

Figure 2 shows the 4 mutually exclusive choices used in the model. Until 12 years old, around 75 percent of the mothers stayed at home and sent their girls to school, while for 15 percent both mother and children stay at home. Furthermore, 7 percent work and send their girls to school, and the remaining 3 percent work but the girls stay at home. When girls grow (between 13 and 15) it changes completely, the percentage of girls not in school with mother working in the market increased to around 8 percent, and in the case where mother stays at home increased, as well, to around 40 percent.

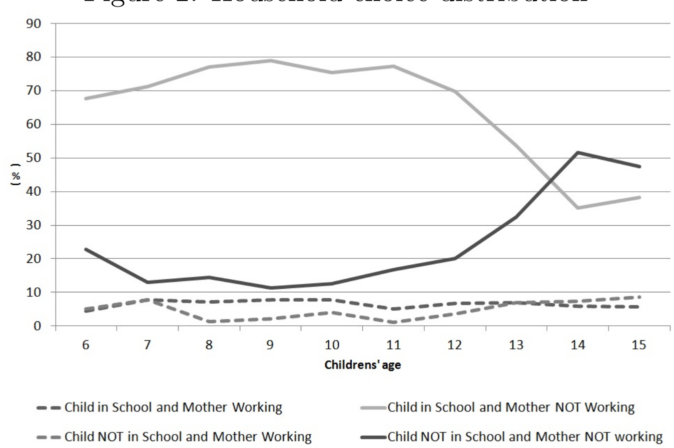
[26] 1 U.S. dollar ≈ 85.6 Pakistani rupees

Figure 1: Girls' school enrollment rate and mother working status



Note: This figure shows the girls' school enrollment rate and the percentage of mothers working in the labour market for girls between 6 and 15 years old.

Figure 2: Household choice distribution



Note: This figure shows the 4 mutually exclusive household choices distribution.

4 Results

This section presents estimates of the structural model presented in Section 2. The estimation procedure involves a maximization of the likelihood function given by equation (13). We first solve the dynamic programming problem for each individual conditional on a given type and then write the probability function. As usual, we consider the reasonableness of the parameter values and the within sample fit. The use of some discrete variables, as described in Section 3, allowed us to obtain a size of state space that does not make the problem computationally intractable[27].

4.1 Estimation

Estimates of the model structural parameters are presented in this sub-section. Table 3 reports the estimates of the parameters in the utility function, budget constraint, home production constraint, and observed heterogeneity. The second panel presents the unobservable heterogeneity parameters. Finally, the variance-covariance matrix is reported in Table 12 in the appendix. The wage equation parameters used in the estimation of the model are reported in Table 4. The probability of a girl passing grade at age t with a mother's education ($educ^m$) used in the estimation are reported in Table 13 in the appendix.

Overall the estimated parameters, especially the most important from the policy perspective have the expected signs and are statistically significant [28]. The discount factor was estimated to be 0.91 and it was obtained from a grid search over several values.

The results indicate significant heterogeneity among mother's types regarding both work and child education. In particular, women are very different in their tastes for children's education. Although one of the types derives utility from sending children to school ($\alpha_{1,h} = 4.0$), the other type derives disutility from doing so ($\alpha_{1,l} = -3.2$). Younger and more educated mothers have a clear preference for education in contrast to what happens with older and less educated ones. Mother's education impact seems more

[27]We discretize the continuous state variables and solve for the Emax functions only on the grid of discretized values, i.e., we interpolate between grid points (see Keane et al. (2011)).

[28]The estimated CRRA parameter is similar to the result presented in Ahmed et al. (2012) of 0.58 and lies within the ranges of 0.05 and 2.57 for developing countries in Cardenas and Carpenter (2008).

Table 3: Structural Parameters Estimates

Parameter	Estimate	Std. Error
Utility Function		
Consumption		
CRRA (γ)	0.70	0.0001
Proportion of market and home production (θ)	0.63	0.0009
Willingness to substitute between c_t^M and c_t^H (ρ)	0.95	0.0001
Budget constraint		
Transport cost - Secondary School Not Available in the village (ψ)	2.29	0.0016
Home Production		
Other goods/technological factor (δ_0)	4.81	0.0006
Children from 0 to 5 years old (δ_1)	4.31	0.0003
child contribution - 12 and 13 years old (η_0)	0.32	0.0001
child contribution - 14 and 15 years old (η_1)	2.90	0.0001
Observed heterogeneity in taste for child education		
Mother's education (α_{11})	2.47	0.0009
Mother's Age (α_{12})	-1.08	0.0007
Children's age (α_{13})	1.04	0.0002
Children's age squared (α_{14})	-0.12	0.0002
Observed heterogeneity in taste for work		
Mother's education on taste for work (α_{31})	-0.37	0.0001
Mother's Age on taste for work (α_{32})	0.21	0.0009
Terminal value function - child education in T+1 (ϕ)	1.96	0.0003
Unobservable heterogeneity		
Children (taste for child education)		
α_{1l} (type I - low)	-3.15	0.0001
α_{1h} (type II - high)	3.98	0.0007
Type proportions		
π_l^c - type I	0.35	0.0023
π_h^c -type II	0.65	-
Mother (taste for work)		
α_{3l} (type I - low)	-2.41	0.0001
α_{3h} (type II - high)	0.54	0.0006
Type proportions		
π_l^m - type I	0.54	0.0026
π_h^m - type II	0.46	-
Log likelihood	5709.9	

Notes: This table shows the estimated structural parameters and asymptotic standard errors.

important than age effect (2.5 and 1.1, respectively).

Interestingly, utility is concave regarding children's age, explaining to some extent the hump-shaped form of the enrolment rate for girls (Figure 3). The utility from sending a girl to school increases up the age of 9/10 years decreasing dramatically for the subsequent ages. The results suggest that parents derive very low utility if girls are 13 years old and disutility from sending girls to school when they reach 14 or 15 years old. Interestingly, 6 year old girls seem to be considered still too young to go to school, as the utility that parents derive from sending them to school is as lower as the one derived from sending a 13 year old girl. These results can be explained to some extent with cultural factors and constraints, such as seclusion, toward girls that limit parents' willingness to send their girls to school. In addition, the most traditional attitudes toward girls' schooling tend to be held by the least educated, which derive less utility from educating girls as seen before leading to a perpetuation of girls disadvantage. Nevertheless, the main results are driven by the economic factors, as shown in sub-section 4.3 - robustness check.

Figure 3: Preferences by age

Note: This figure shows the girls' age profile of mother's preferences in taste for child education in utility terms.

In what concerns tastes for work, one of the types dislikes work while the other obtains very low utility from work. Mother's education increases the disutility derived from work while age has the opposite effect. Nevertheless, both have a very modest impact on utility.

Among other economic factors, the results suggest that home production and, in particular, the presence of young children in the household is an important mechanism explaining the high drop out rates among girls aged 12-15 as δ_1 is positive[29]. If this is the case, it is a rational efficient response to the economic environment.

The child's contribution to home production implies a significant difference between girls aged 12-13 and older girls. The coefficient associated with the younger ones is around 0.3, while for those aged 14-15 it is significantly higher (2.9). These results should be compared to the mothers' term which is normalized to 1, the case where girls spend the same time working at home as mothers, or are as productive as mothers[30]. Data suggests that girls aged 14 to 15 spend around the same time at home as mothers. Therefore, the high value estimated by the model can be partly explained if girls are more productive than mothers. One possible explanation for the magnitude may be related to the incentives regarding marriage market premium. If this is the case, home production may also be capturing the fact that girls are being trained to the marriage[31]. In addition, this parameter may be capturing the effect of other household members that are not taken directly into account in the model. Results for girls aged 12-13 are in line with the lower amount of time spent at home observed in the data.

Finally, the estimated cost of going to school if there is no secondary school available in the village is significantly high (around 2290 rupees $\approx$ 27 U.S. dollars), which is around 6 times the average educational expenditure for those in grade 5 or above and around 30 percent of household income. Given the high percentage of villages without elementary/secondary schools (more than 40 percent) this result suggests that this might be an important factor explaining the high drop-out rate among older girls. From a policy perspective this is something that cannot be ignored by the authorities.

As discussed in Section 2, we estimate the wage equation outside the model and then

[29] Another implication of this finding is that if the number of younger children increases, we have less market consumption. Despite not being intuitive, this result becomes reasonable given that c_t^H and c_t^M are highly substitutes (ρ =0.95).

[30] These coefficients being higher/lower than 1 can only be explained by two reasons: spending more/less time than mothers at home and/or being more/less productive than mothers.

[31] In the data there is some evidence that boys are not leaving the house so early, which may suggest that some girls may be leaving home to get married (TO DOUBLE CHECK).

use predictions from that equation in place of actual wages. We use an Heckman selection model to estimate the wage equation and test if selection is an issue or not. To obtain the Inverse Mills ratio to be used in the wage equation, we estimate a reduced form probit for labour market participation as a function of variables used in the structural model. These variables are mothers' age and education, number of children less than 1 year old, girls' school attendance, secondary school availability, and female unemployment rate in the village.

Table 4: Wage equation

Parameter	Estimate	Std.Error
Log wages equation		
Mother education	0.79	0.21
Mother age	0.08	0.12
Mother age squared / 100	-0.09	0.14
Female Local Labour Market Conditions		
Medium (medium village unemployment rate)	-0.63	0.21
Bad (high village unemployment rate)	-0.39	0.20
Inverse Mills ratio (λ)	-0.14	0.21
Constant	-1.13	2.50
First Step (Labour market participation)		
Mother education	0.07	0.12
Mother age	0.01	0.07
Mother age squared /100	0.01	0.08
Female Local Labour Market Conditions		
Medium (medium village unemployment rate)	0.19	0.12
Bad (high village unemployment rate)	0.03	0.11
Children from 0 to 5 years old	-0.24	0.12
Children in school	-0.18	0.10
Log income (father)	-0.53	0.07
Secondary school availability in the village	0.05	0.11
Constant	-1.22	1.34

Notes: This table shows the results of the wage equation estimated outside the model. The first block displays log wages equation (second step) while the second block displays mother's labour market participation decision. Female Local Labour Market conditions is measured as the village female unemployment rate. Using the household census data, we split the 125 villages into 3 categories according to their female unemployment rate (up to 25th percentile, 25th to 75th percentile, and above 75th percentile). More specifically, the 25th percentile corresponds to an unemployment rate just below 6 percent and the 75th percentile just above 30 percent.

The resulting estimated wage equation is presented in Table 4. The selection effect is not an issue as the Inverse Mills ratio coefficient is not significant, reflecting to some extent the fact that female labour is relatively homogenous given age and education[32]. Table 4 also shows that the education effect is significant and has the expected sign, while the age effect has the expected sign but is not significant. Local labour market conditions (female unemployment rate in the village) also seems to have some significant impact on wages. Nevertheless, the key determinant of wages at the individual level is education.

Note that, in this type of model, to get identification we need at least one independent variable that affects selection but not the outcome. In our case we have number of children less than 6 years old, father's income, if child is in school, and secondary school availability in the village[33]. All appear with the correct sign and are significant except the availability of secondary schools in the village. Therefore, a women with young children, with higher father income, and with children at school are less likely to work in the market.

4.2 Model Fit

Figure 4 depicts the fit of the model to the choice distributions, based on a simulation of 10000 individuals. It compares the model's prediction of the distribution of children (school, or home) and mother (work, or home) activity allocations at individual ages with the actual distribution.

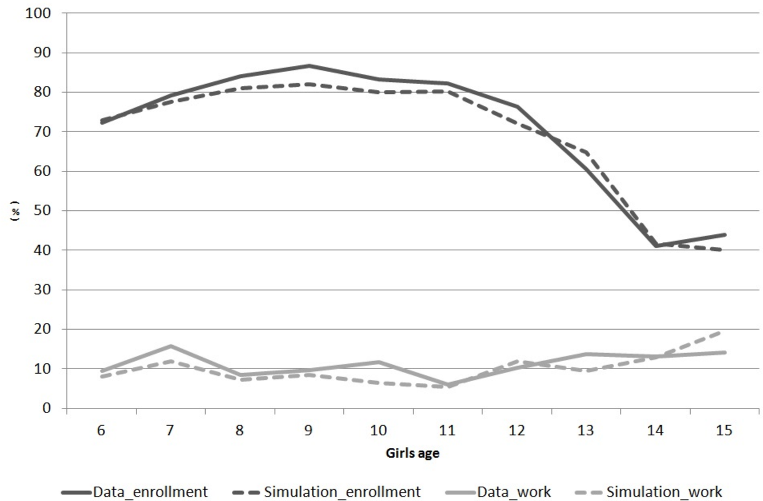
As can be seen, the model matches the data quite well, especially looking at the girl school enrollment decision. Despite not being the main focus of the paper, the mother working status presents a reasonable result but a clear lower fit with data[34]. More specifically, Figure 11 in appendix, shows that the within-sample fit of our model is particularly good in the case of the most chosen alternative, i.e., mother staying at home with children enrolled in school. The other alternatives present a reasonable match, with the exception of the first half of the distribution of the alternative of mother working and children staying at home. Furthermore, concerning girl school enrollment decision, the model also matches

[32] The Inverse Mills ratio has an unexpected sign, implying that those who go to the labour market tend to have lower wages. This result may be due to a strong income effect. Those who do not get a job are from a wealthier family.

[33] The availability of secondary schools in the village increases the likelihood of a woman participating in the labour market. This means that more secondary schools would increase the girls' enrollment and then more mothers would need to stay at home. Nevertheless, this result is not statistically significant.

[34] In some situations the chi-square statistic associated with a test of the null that the predicted and actual distributions are the same is rejected.

Figure 4: Model Fit - Girls' school enrollment and mother working status



Notes: This figure shows the fit of the model to the mother's labour market participation (Data work vs. Simulation work) and girls' schooling (Data enrollment vs. Simulation enrollment), based on a simulation of 10000 individuals. It provides the fit of the model at each girls' age between 6 and 15 years old.

the data well if we look at households with and without young children (see Figure 12 in appendix).

Table 5 presents the fit of the one-period transition rates by different group of ages for girls' schooling decisions. Once girls leave school, they rarely return. In particular, only 5.5 percent of girls aged 11 to 15 who were at home (not enrolled) one year attended school the next year. The school-to-school transition also exhibits substantial persistence, with more than 90 percent of the girls aged 6 to 13 and around 83 percent of the girls aged 14 to 15 who attended school in one year also attending the following year.

As seen in Table 5, the model seems to capture the age group differences but it is not capturing schooling transition patterns very well, in particular, the not enrolled-to-enrolled transition rates. One concern is that lagged schooling status does not have an effect on current enrollment. In future research, a natural and important extension to the current model that deals with this problem is to include a cost of getting back to school after drop-out (a cost to re-entry to school). An additional source of dynamics is to make completion rate dependent on girls' effort and at the same time to depend on school quality. This is important as the model does not incorporate much dynamics as utility depends mainly on school attendance and not on cognitive achievement.

Table 5: One-period transition rates by age

	DATA			MODEL	
	6-10			**6-10**	
	Not Enrolled (t)	Enrolled (t)		Not Enrolled (t)	Enrolled (t)
Not Enrolled (t-1)	77.7	22.3	Not Enrolled (t-1)	47.1	52.9
Enrolled (t-1)	5.4	94.6	Enrolled (t-1)	13.2	86.8
	11-13			**11-13**	
	Not Enrolled (t)	Enrolled (t)		Not Enrolled (t)	Enrolled (t)
Not Enrolled (t-1)	94.3	5.7	Not Enrolled (t-1)	63.2	36.8
Enrolled (t-1)	9.0	91.1	Enrolled (t-1)	19.3	80.7
	14-15			**14-15**	
	Not Enrolled (t)	Enrolled (t)		Not Enrolled (t)	Enrolled (t)
Not Enrolled (t-1)	94.8	5.2	Not Enrolled (t-1)	72.7	27.3
Enrolled (t-1)	17.4	82.6	Enrolled (t-1)	31.7	68.3

Notes: This table shows the fit of the one-period transition rates (in percentage) by different group of ages for girls' schooling decisions.

4.3 Robustness Check

In this section we present two robustness checks on our previous results. We first consider a different model specification including father's income as endogenous to infer the reasonability in terms of identification of the conditional cash transfer experiment. We then check what happens to our results if we do not allow taste for school to vary with children's age.

Endogenous father's income - In this setup we re-estimate the model considering father's income as endogenous. In particular, it depends on wife age and education, male local labour market conditions, and an income shock:

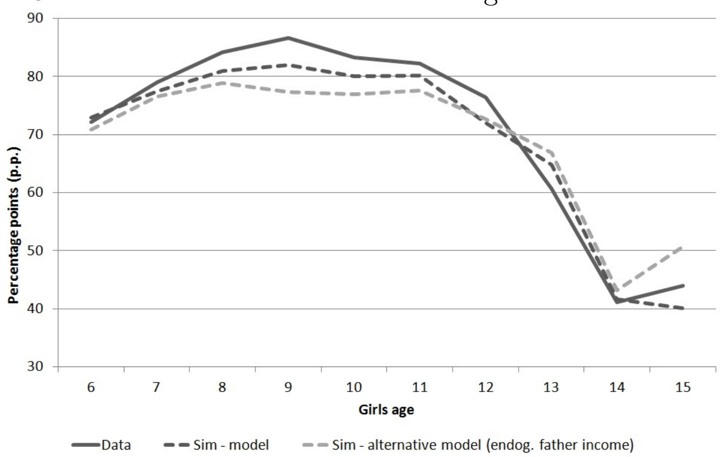
Father income wage equation:

$$y_t = g(age_t^m, educ_t^m, z_{tv}^{male}, \varsigma_t^y)$$

The estimation results are similar but the fit is not as good as in the exogenous case

(Figure 5)[35]. Nevertheless, the similarity with the exogenous case gives some assurance about the identification and interpretation of the cash transfer policy of our simulation.

Figure 5: Girls' school enrollment rate - Endogenous father's income case

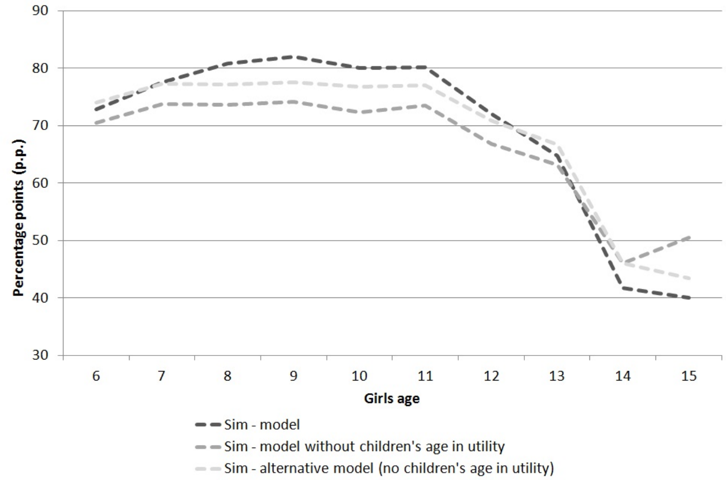


Notes: This figure compares the actual girls' school enrollment with the simulation of the model with and without father's income as endogenous. Alternative Model (endog. father income): The simulation of the alternative model implies the estimation of a new model with a different specification, including the father income (y_t) as endogenous. It depends on mother's age, mother's education, male local labour market conditions, and income shock.

Children's age in the utility function - To check if economic factors are driving the results, in particular, the dramatic decrease of girls' enrollment rate for older girls, we perform two exercises: 1) simulate the model imposing the coefficients of children's age equal to zero ($\alpha_{13} = 0$ and $\alpha_{14} = 0$), and 2) re-estimate the model with a new specification excluding children's age from the utility. Figure 6 shows that in both cases age naturally has some relevance but the main conclusions are still coming from the economic factors discussed before, in particular for older girls. These exercises, simulation, and re-estimation of the model without children's age in the utility function, confirm economic factors as the main determinants of the girls' enrolment rate. The result is particularly strong for older girls, but even the initial increase for younger girls is, to some extent, determined by economic factors and not entirely by the inclusion of age in a quadratic form in the utility function.

[35]The results are available upon request.

Figure 6: Girls' school enrollment rate - Children's age case



Notes: This figure compares the model simulation of the girls' school enrollment rate with two different model specifications: 1) Model without children's age in utility function: we simulate the model presented in the paper with the coefficients of age and age squared equal to zero ($\alpha_{13} = 0$ and $\alpha_{14} = 0$); 2) The simulation of the alternative model implies the estimation of a new model with a different specification excluding the terms α_{13} and α_{14} from the estimation.

5 Policy Experiments

One of the main advantages of having a structural model is the possibility of performing policy experiments. We now use the model to simulate school participation under different scenarios. We quantify the effect of demand-side policies such as (conditional) cash transfers, availability of daycare centers for young children, and employment/wage subsidy. We also present results of the effect of a supply-side policy on school attendance, i.e., the increase in the number of communities where a secondary school is available. We discuss also the impact of each policy on welfare.

In order to get credible results we need exogenous variation in the data related with the policy experiment. Therefore, we'll provide some intuitive arguments regarding the identification issues related to the effects of the policy experiments performed in the paper, i.e., policy relevant variation. The policy effect is estimated without direct policy variation, i.e., we did not need to observe households in both states of the world, for example, with and without the cash transfer program.

In the *employment subsidy (wage variation)* case, to get policy relevant exogenous variation requires the presence of one variable in the wage equation that does not directly affect the utility of sending a girl to school. For example, female local labour market conditions enter the wage equation but do not directly affect preferences. Therefore, we assume that variation in the female local labour market factors might cause exogenous variation in employment and schooling decisions while not directly affecting preferences. For this argument to be valid we require the following assumptions: these local labour market conditions cannot vary across villages due to supply changes, for example, a common shock to women (for example to tastes to work) cannot affect local female unemployment rate, and there is no systematic variation in women's unobserved heterogeneity across villages.

In the case of the *cash transfer* policy, the argument is similar but applied to father's income instead of mother's wage[36]. As before, in this case male local labour market conditions in the father's income equation provide us the required exogenous policy-relevant variation.

The *mandatory public(free) child care services* policy relies mainly on the model and on the credibility of the results of the home production equation. In this case, there is no obvious variation in the data providing exogenous policy-relevant variation.

Finally, in the *school building program*, we assume that, conditional on other variables, variation in the variable not having a school in the village might plausibly generate exogenous variation in girls' schooling decisions. In that case, this assumption is enough to provide policy-relevant exogenous variation.

5.1 School Subsidy

5.1.1 Gender Targeted Conditional Cash Transfer

Conditional Cash Transfer (CCT) programs targeting poor households are rapidly becoming a key policy instrument used by developing countries to reduce poverty and increase human capital investments. To explicitly address intra-household disparities in human capital investment some countries, like Pakistan, have implemented gender-targeted CCT.

We perform four alternative simulations with a gender CCT program: i) first we perform an exercise similar to the one implemented in Pakistan - Female School Stipend Program in Punjab; ii) then our approach follows closely the grant scheme implemented in Mexico (Oportunidades) with different grants for primary and secondary school[37]; iii)

[36] The endogeneity of father's income is addressed in sub-section 4.3

[37] See Attanasio et al. (2012) for more details on the Oportunidades grants.

a scheme with the same grant for primary and secondary school and iv) a grant scheme that would close the enrollment rate gender gap.

i) Grant Scheme similar to the Female School Stipend Program in Punjab

Under the Female School Stipend Program in Punjab eligible girls receive a stipend (of around 3 U.S dollars per month) conditional on her being enrolled in grade 6-8 in a government girls school in a target district and conditional on her maintaining average class attendance of at least 80 percent.

This program was evaluated by Chaudhury and Parajuli (2010)[38] and Hasan (2010) with an impact on school enrollment rate between 2 and 3 percentage points for girls aged 12-14 in rural areas. With a similar grant amount (of around 3 U.S dollars per month) we simulate this gender CCT program estimating an impact of almost 1 p.p. for girls in the same group age.

The difference in the results can be explained by the following reasons. First, the baseline is much lower than the one studied in this paper (29 and 50 percent, respectively). Second, the program was addressed to the less literate regions[39], where only one (Rahim Yar Khan) of our 3 districts would have been included in the original program[40]. In this case, their results are naturally more sensitive to an income effect. Finally, the program targeted only public schools and the high percentage of private schools in our sample may affect the results in a non-clear way. Therefore, it can be said that the model simulates relatively well the results from the Female School Stipend Program in Punjab.

ii) Grant Scheme a la Oportunidades

The simulation includes monthly grants for children of a family qualified as beneficiary[41]. To be given a grant, children need to be enrolled in school. A child who does not pass a

[38] This paper is discussed and analysed in the Conditional cash transfers policy research report by Fiszbein and Schady (2009).

[39] Fifteen of Punjabs 34 districts were selected as program districts on the basis of average literacy rate of population 10 years old and older. The average literacy was estimated from Population Census 1998 data. The cut-off literacy rate was 40 percent: 15 selected districts below the cut-off were stipend-eligible and the remaining 19 above the cut-off were not.

[40] Rahim Yar Khan was not among the worst ones, having a literacy rate of around 33 percent (in an interval from 20 to 40 percent). Attock and Faisalabad presented higher literacy rates (49 and 52 percent, respectively).

[41] This excludes the ones above 95th percentile of the distribution.

grade is still eligible for the grant[42]. The grant increases with the years of schooling completed in a very similar way to the well known CCT program in Mexico (Oportunidades). In addition, using also other previous CCT schemes (Fiszbein and Schady (2009)), we built a lower and an upper bound for the impact on school attendance.

Table 6 shows the grant scheme used in both scenarios, one in which the secondary school grant is around 4 percent of the average income of the household (Lower bound scenario - CCT(L)) and another one where the same grant reaches 30 percent of the average income of the household (Upper bound scenario - CCT(U)).

Table 6: Grant Scheme a la Oportunidades (in U.S dollars per month)

Grade	CCT (U)	CCT(L)
Primary education		
grade 1 to 4	5.8	0.9
grade 5	11.7	1.8
Elementary and Secondary education		
grade 6 to 7	23.4	3.5
grade 8 or more	29.2	4.4

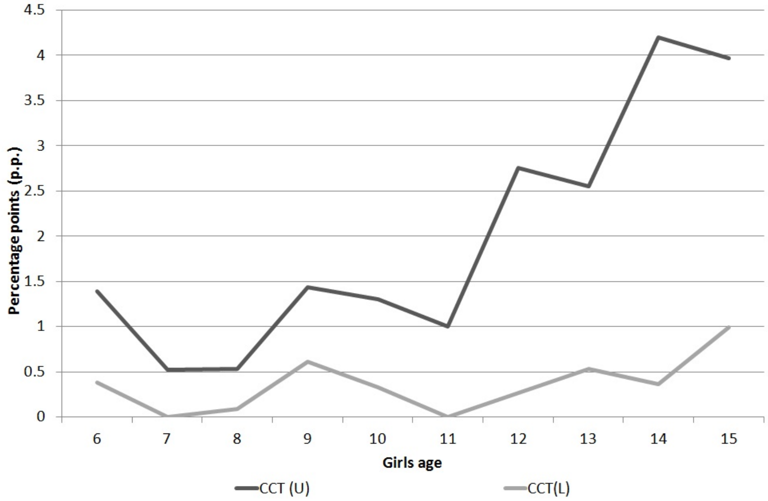
Notes: This table shows the grant scheme of the conditional cash transfer simulated in the model similar to the Oportunidades scheme. CCT (U) represents the upper bound case, where the secondary school grant represents around 30% of average income. CCT (L) represents the lower bound case, where the secondary school grant represents around 4% of average income.

Figure 7 shows that the CCT program would have a positive and increasing impact on school enrollment, especially after 12 years old using the upper bound scenario (CCT(U)). The average impact for girls between 12 and 15 is around 3.5 percentage points, while for younger girls is around 1 percentage point. Using the lower bound scenario (CCT(L)) the results become much smaller, reaching for older girls an impact of less than 1 percentage point and having a negligible effect on younger girls.

In general, the impact of this measure on mothers' labour market participation is modest (see Table 14 in appendix).

[42] In Oportunidades if the child fails the same grade twice, she/he loses eligibility. This is not taken into account in this simulation.

Figure 7: Conditional Cash Transfer - impact on girls' school enrollment rates



Notes: This figure shows the impact, in percentage points, on girls' school enrollment rate of the conditional cash transfer scheme presented in Table 6. CCT(U) represents the upper bound case, where the secondary school grant represents around 30% of average household income. CCT(L) represents the lower bound case, where the secondary school grant represents around 4% of average household income.

Interestingly, in terms of welfare, the impact of the CCT program is positive but decreasing in both scenarios. In particular, using the upper bound scenario, the utility gain for girls aged 11 to 13 is equivalent to an increase in income of 15.5 percent while for girls aged 14 to 15 it is equivalent to a 9.5 percent increase. The effects in the lower bound scenario are smaller and equivalent to an income positive variation of 2.3 and 1.3 percent, respectively.

iii) Same Grant for all grades

In this simulation, the grant does not increase with the years of schooling completed. Table 7 shows the grant scheme and Figure 8 shows the impact of this simulation. As before, we simulate two different scenarios, one in which the monthly grant is around 4 percent of the average income of the household, and another in which the same grant reaches 30 percent of the average income of the household.

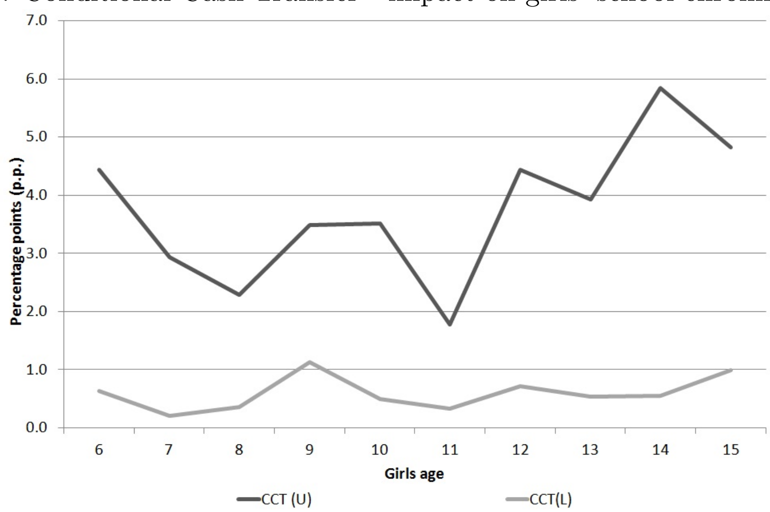
Interestingly, even with the same grant for all years of schooling completed, the CCT program would have a positive and increasing impact on school enrollment especially after 12 years old, using the Upper bound scenario (CCT(U)). The average impact for

Table 7: Same Grant for all grades (in U.S dollars per month)

Grade	CCT (U)	CCT(L)
All grades	23.4	3.5

Notes: This table shows the grant scheme of the conditional cash transfer simulated in the model, where the grant is the same for all grades. CCT (U) represents the upper bound case where the grant represents around 30% of average income. CCT (L) represents the lower bound case, where the grant represents around 4% of average income.

Figure 8: Conditional Cash Transfer - impact on girls' school enrollment rates



Notes: This figure shows the impact, in percentage points, on girls' school enrollment rate of the conditional cash transfer scheme presented in Table 7. CCT(U) represents the upper bound case, where the grant represents around 30% of average household income. CCT(L) represents the lower bound case, where the grant represents around 4% of average household income.

girls between 12 and 15 is around 5 percentage points, while for younger girls is around 3 percentage points. In the Lower bound scenario the impact on girls' enrollment rate increases slightly, to be on average around 0.6 percentage points and similar for younger and older girls.

As before, in terms of welfare, the impact of the program is positive but decreasing. The utility gain for girls aged 11 to 13 is equivalent to an increase of income by 21.5 percent using the upper bound scenario, while for the older girls it is equivalent to an increase of 12.2 percent. The effects in the lower bound scenario are smaller and equivalent to an increase in income of less than 3.5 percent.

iv) Which grant would close the enrollment rate gender gap?

Finally, we present the results for the simulation where the CCT grant would close the school enrollment rate gender gap. Table 8 shows the observed school enrollment gender gap in the sample and Table 9 shows the grant scheme required to close such large differences.

Table 8: School Enrollment Rate by age - Gender gap

	Age									
	6	7	8	9	10	11	12	13	14	15
Gender Gap	10.9	9.56	5.5	7.4	10.8	6.1	10.0	17.5	27.0	22.6

Notes: This table shows the school enrollment gender gap by age.

Table 9: Grant Scheme - Closing Gender Gap (in U.S dollars per month)

Grade	CCT	% of household average income
Primary		
grade 1 to 4	82	100.0
grade 5	94	114.3
Elementary/Secondary		
grade 6 to 7	152	185.7
grade 8 or more	245	300.0

Notes: This table shows the grant scheme of the conditional cash transfer simulated in the model which closes the gender gap in terms of school enrollment rates. It also shows the percentage of each grant in terms of the household average income.

For younger girls the average gender gap reaches on average 10 percentage points, while for girls aged 14 and 15, reaches an average of 25 percentage points. In the Primary school the level of the grant needs to be similar or slightly above the average household income. In the Secondary school, for grades 6 and 7 the grant has to be almost twice the household average income, and would need to be three times greater for the students in grade 8 or above. As we can see, it would be very costly to close the gender gap using only a program like the conditional cash transfer. To close the gap the grant for older girls needs to be around 10 times greater than the most generous CCT program that has already taken place, which reaches around 30 percent of the household average income. In terms of welfare, closing the school enrollment gender gap would provide an utility gain equivalent to an increase in income of around 100 percent.

5.1.2 Unconditional Cash Transfer

Table 10 compares the results of the conditional cash transfer (CCT) with an unconditional cash transfer (UCT) program. The impact of the UCT scheme is much smaller in both cases, lower and upper bound scenarios. Nevertheless, the higher the grant the bigger the difference between the two schemes. On average, for older girls, the impact of the UCT upper bound scenario is only 15 percent as large as the CCT scheme and 28 percent in the lower bound scenario. Furthermore, the CCT(L), which corresponds to a grant of around 4 percent of household income, yields results slightly better than the ones presented by UCT(U) with a grant 8 times larger.

Table 10: Unconditional vs. Conditional cash transfer - Same Grant for all grades

	Girls' age									
Policy measures	6	7	8	9	10	11	12	13	14	15
Upper Bound Scenario - (around 30 % of household average income)										
CCT (U)	4.4	2.9	2.3	3.5	3.5	1.8	4.4	3.9	5.8	4.8
UCT (U)	0.1	0.0	0.1	0.0	0.1	0.0	0.2	0.2	0.5	0.7
Lower Bound Scenario - (around 4 % of household average income)										
CCT (L)	0.6	0.2	0.4	1.1	0.5	0.3	0.7	0.5	0.5	1.0
UCT (L)	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.3	0.3

Notes: This table compares the impact, in percentage points, of the unconditional with the conditional cash transfer policy on girls' school enrollment rate. CCT(U) and UCT(U) - It represents the upper bound case for the Conditional and Unconditional cash Transfer, respectively, where the secondary school grant represents around 30% of average household income. CCT(L) and UCT(L) - It represents the lower bound case for the Conditional and Unconditional cash Transfer, respectively, where the secondary school grant represents around 4% of average household income.

These results are in line with the recent paper (Baird et al. (2011)) confirming big effects of the conditionality[43]. Therefore, this exercise confirms and quantifies from a policy perspective the widespread use of CCT schemes. In particular, when education is to be targeted, the need and importance of a conditional incentives scheme seems clear .

In terms of welfare, with this scheme (same grant for the UCT and CCT) we obtain as expected a higher utility gain for the UCT as the number of people receiving the grant is also higher. Nevertheless, a cost comparable version of the UCT and CCT would provide similar impacts in terms of welfare, except for the older girls where the UCT would provide a higher utility gain. In this case, the effect of the CCT scheme on older girls' would provide an utility gain equivalent to an increase in income by 12.2 percent, which is around 90 percent as large as the UCT version.

5.2 Availability of elementary/secondary schools

Access to the school has been an important factor inhibiting girls' enrollment, particularly in rural areas. Changes in government educational policy in recent years, as well as the rapid growth of low-fee private schools in rural areas, may be changing the educational opportunity structure for poor rural girls.

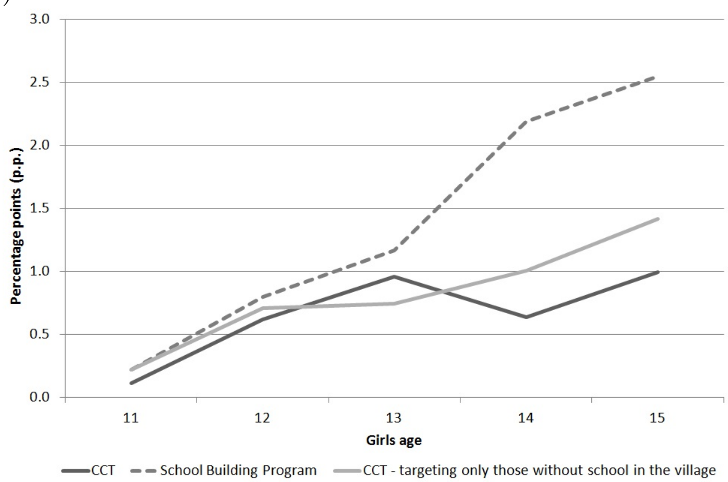
Nevertheless, not having an elementary/secondary school available in the village of residence implies transportation and time costs. Both costs decrease the utility of attending school. For example, if the government establishes a secondary school in one of those villages, a positive effect on school enrollment could be expected. In the next experiment we demonstrate the effects of a potential but ambitious school building program, in which all villages would have at least one school for girls (Figure 9). The impact is increasing with children's age and reaches a maximum of around 2.5 percentage points for the oldest girls.

The results are significant and bigger than the cost comparable CCT scenario, in particular, looking at older girls (14 and 15 years old) the impact on the girls' school enrollment rate is almost three times greater. The cost per student of this type of measure

[43]Baird et al. (2011) assess the role of conditionality in a cash transfer program using an experiment with adolescent girls in Malawi. In their paper, the impact of the UCT scheme on drop-out rate was only 43% as large as the impact of the CCT.

was calculated to be of around 465 rupees per month (≈ 5.4 U.S. dollars)[44]. This measure can be seen not only as an alternative but also as a complement from an education policy perspective that wishes to target high drop-out rates among girls in a more effective and direct way. Even if the CCT program was targeting only those in villages without school, the impact on girls' school enrollment would be 50 % as large as the school building program (see Figure 9)[45].

Figure 9: School Building Program vs. Conditional Cash Transfer (impact on girls' school enrollment rate)



Notes: This figure compares the impact, in percentage points, of the school building program with the conditional cash transfer program on girls' school enrollment rate. The two policies are cost comparable - the cost per student in an elementary/secondary school of around 465 rupees per month (≈ 5.4 U.S. dollars)). The cost per student per month is calculated in the following way: using an estimate of around $7000 per classroom (source: other school building programs) and assuming i) 20 years without major reforms, ii) on average 30 children per classroom, and iii) a student will stay on average 5 years.
In the case of targeting only those in villages without school, the grant is of around 1250 rupees (≈ 14.6 U.S. dollars).
1 US dollar ≈ 85.6 PAK rupees.

Notice that, on one hand, the results may be underestimated, since we are not taking into account the extra positive effect of building schools in villages where a school is already available and crowded. Furthermore, if we had used the same estimated total cost of the school building program (per year) instead of the cost per student, the effect on the CCT

[44]The cost per student per month is calculated in the following way: from other school building programs (from World Bank and other institutions) we use an estimate of around $7000 per classroom. Then, we assume i) 20 years without major reforms, ii) on average 30 children per classroom, and iii) a student will stay on average 5 years.

[45]In this case the grant is of around 1250 rupees (≈ 14.6 U.S. dollars).

scheme would have been even smaller. Finally, this effect relies on the assumption that the government builds a school in all villages initially without one.

Furthermore, as the CCT scheme is an anti-poverty program and not only an educational program it is also important to look at the impacts on welfare. As expected, in terms of welfare, the difference between the two schemes is much smaller. The utility gain of the school building program is equivalent to an increase in income of around 3 percent. The CCT cost comparable scheme effects are 83 percent as large as the school building program (compared to only 34 percent in terms of enrollment).

Interestingly, Table 14 in appendix shows that there is a negative impact on mothers' labour market participation for older girls, reaching a maximum impact of -0.65 percentage points when the girl is 14 years old. This seems to be a result of the substitutability between mothers and girls in working at home.

5.3 Mandatory Public/free Child care services

Suppose all the children aged 5 or less in the family are sent to a public (free) daycare center. Girls would no longer be needed at home to look after them and may go back to school. In the model, the utility of a girl going to school is negatively related with the number of children aged 5 or less in the family.

Introducing nursery into the model would change the home production function in the following way:

$$c_t^H = \delta_0 \Psi + \delta_1 n_{05,t}[(1 - \chi - h_t) + (\eta_0 age^c_{1213,t} + \eta_1 age^c_{1415,t})(1 - \chi - s_t) + \chi\tau] \qquad (14)$$

where χ represents the nursery number of hours, τ the nursery productivity factor and $\Psi = (1 - h_t) + (\eta_0 age^c_{1213,t} + \eta_1 age^c_{1415,t})(1 - s_t)$.

The time spent by the mother and the daughter at home would be smaller if the young children goes to the nursery, and it is reduced by χ. This work is replaced by the nursery and it has a productivity factor of τ.

The effect of a mandatory free child care service for all children aged 5 or less on girls' school enrollment can be approximated by simulating girls' choices after setting the time in the nursery by 8 hours (χ=0.333) and nursery being half as productive as the mother

(τ=0.5)[46].

Table 11: Public Child Care vs. Conditional Cash Transfer

	Girls' Age				
Policy measures	11	12	13	14	15
Impact on Girls' School Enrollment Rate					
Mandatory Free Child Care Services	-0.2	0.1	-0.1	0.7	0.6
CCT targeting only secondary school students	0.0	0.5	0.7	0.5	0.8

Notes: This table compares the impact, in percentage points, of a free child care service with a cost comparable conditional cash transfer program. The two policies are cost comparable presenting the same total cost in the sample. Total Cost in the sample: the cost of a contract teacher per child is of around 1200 Rupees per year (≈ 14 U.S. dollars) (the cost per month is of around 3000 rupees and dividing that by 20-30 children to emerge at something around 1200 rupees per year). In the sample there are 1195 children less than 6, which means a total cost in the sample of 1434000 rupees per year (119500 per month). In the secondary school there are around 280 girls, which gives a monthly cash transfer of around 420 rupees (≈ 4.9 U.S. dollars) per student in the secondary school.

Table 11 shows that the impact is significant only for older girls' school enrollment (aged 14-15) and on average of 0.67 percentage points. It should be noted that the effect of the mandatory free child care may be overestimated, as it is not taking into account the age at which daycare accepts children and most importantly because of the mandatory nature of the policy (there is no choice). Interestingly, the results in Table 11 show also that the impact of a CCT scheme targeting only secondary school students is only 91 percent as large as the impact of providing free child care services, when the two measures have a similar total cost [47]. If the CCT scheme wanted to target students in the primary school as well, the mandatory free child care service becomes much more effective, as more students will be eligible for the grant [48]. The effect on mother's labour

[46] Similar to conclusions after using different specifications, such as χ=0.25 ≡ 6 hours in the nursery and different productivity factor for the nursery (0.25 and 0.75).

[47] Total Cost in the sample: the cost of a contract teacher per child is of around 1200 Rupees per year (≈ 14 U.S. dollars) (the cost per month is of around 3000 rupees and dividing that by 20-30 children to emerge at something around 1200 rupees per year). In the sample there are 1195 children less than 6, which means a total cost in the sample of 1434000 rupees per year (119500 per month). In the secondary school there are around 280 girls, which gives a monthly cash transfer of around 420 rupees (≈ 4.9 U.S. dollars) per student in the secondary school.

[48] If we consider all girls aged 11 to 15 years old enrolled in school and not only the ones enrolled in the secondary school, the monthly cash transfer drops to around 200 Rupees per month (≈ 2.3 U.S. dollars). In this case, this measure becomes more effective, with an impact of around twice the one presented by the CCT.

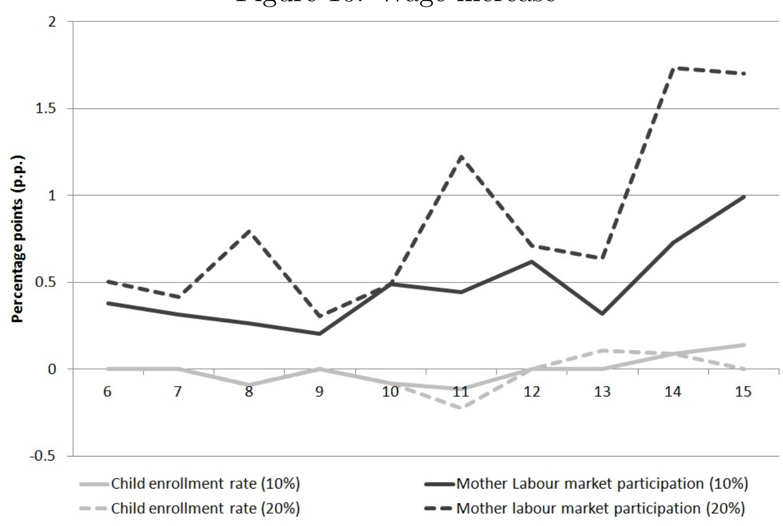
market participation is decreasing and turns out to be negligible for older girls (see Table 14 in appendix). Nevertheless, for younger girls the maximum impact reaches only 0.3 percentage points.

Interestingly, given the mandatory nature of the program, the child care services scheme would lead to a negligible but slightly negative effect in terms of welfare. This result is affected mainly by its mandatory nature but it also depends, in some extent, to the levels of the new parameters of the new home production specification, in particular of the productivity factor (τ). Finally, the CCT cost comparable program would provide an utility gain equivalent to an increase in income of around 2 percent.

5.4 Employment subsidy

Another interesting experiment is the impact of an employment subsidy (or wage subsidy)[49] on school enrolment and mother labour market participation (Figure 10). We present the results of an increase in wages by 10 and 20 percent.

Figure 10: Wage increase



Note: This figure shows the impact, in percentage points, on girls' school enrollment rate and mother's labour market participation rate of an increase in wages by 10 and 20 percent.

In both cases, the effect on girls' school enrollment is modest, more so for the youngest girls. To a large extent, this is explained by the high substitutability between home and

[49]This policy would be in the same spirit of the Working Tax Credit in the United Kingdom.

market production (see Table 3) observed in the model. Interestingly, for a mother with an older girl, we may end up with a small but negative impact. In this case, the substitution effect becomes more important than the income effect. As for the mother's labour market participation, the impact is positive and increasing in both scenarios. The impact of an increase in wages by 20 percent reaches a maximum of 1.8 percentage points, when the child is 15 years old (around 1 percentage point for the 10 percent increase).

5.5 The effect of a more educated generation

As mentioned above, the most traditional attitudes towards girls' schooling tend to be held by the older and less educated population. Therefore, there is a positive generation effect that is interesting to quantify and that should be taken into account. In this case, we simulate and quantify what happens to girls' school attendance if mothers' education distribution is different, i.e., in the way the model predicts. In one generation (10 to 15 years), the oldest population (50-65 years old) would be out and replaced by the generation analysed in the model. The figures are clear: among the oldest population, the percentage of educated mothers is around 5 percent, compared to the 60 to 65 percent in the youngest generation. In practical terms, the overall percentage of educated mothers rises from 25 to 30 percent. This simple exercise is an underestimation of a pure generation effect, as it does not take into account other preference changes that, of course, are also taking place. The impact on older girls is around 1.5 percentage points, which is relatively modest. Despite the other factors that are not taken into account in this exercise, this is just a confirmation that any cultural change takes time to have a clear and visible effect.

6 Concluding Remarks

In this paper we focus on the labour supply and children school decisions of women using and estimating a dynamic behavioural model in rural Pakistan. Nevertheless, our research is relevant not only for Pakistan but also for other low-income countries, in particular, from the perspective of how these countries may deal with high drop-out rates among older girls. The mechanisms and the consequences of different policies may provide good insight to other lower-income countries from a policy perspective.

This paper seeks to explain the high drop-out rates among girls, in particular, the role of home production for older girls, and the impact of monetary incentives and supply side education policies on girls' school enrollment. Previous studies have provided evidence that these incentives are important to increase school attendance but never combine mother labour supply with children school decision, in particular, the key role played by home production. Results in this paper show that mothers' working status affects the girls' utility of staying at home. It can be expected that a girl whose mother works in the labour market would be more valuable at home, replacing her mothers' housework. Also, the family composition, especially the number of young children has an important role for the value of both mother and girls at home. The model highlights the relevance of economic factors, through budget constraint and home production, in the decision of children attending school. The importance of cultural factors is also analysed and presented in the household preferences. Most importantly, estimation of the structural model allows us to explore the effects of counterfactual policy experiments. Our results may lead to important policy implications for developing countries. Simulations suggest that monetary incentives are a good mechanism to keep and increase girls' school enrollment, such as the conditional cash transfer case. Nevertheless, to obtain sizable effects, they seem not to be cost effective. These targets can also be achieved by reducing transportation costs in secondary school and to a less extent by making daycare centers available. This paper shows that in rural areas, CCT programs are clearly not the only way to target educational attainment, in particular when some basic school environments and resources are not attained. Nevertheless, in terms of welfare, as the CCT scheme is an anti-poverty program and not specifically an educational program, the difference between schemes is much smaller. Further, the effect of a more educated generation should not be ignored, but it seems clear that it will take time (three or four decades) to see a real and visible effect.

References

Aguirregabiria, V. (2010). Another look at the identification of dynamic discrete decision processes: An application to retirement behavior. *Journal of Business & Economic Statistics*, 28(2):201–218.

Aguirregabiria, V. and Mira, P. (2010). Dynamic discrete choice structural models: A survey. *Journal of Econometrics*, 156(1):38–67.

Ahmed, W., Haider, A., and Iqbal, J. (2012). Estimation of discount factor (beta) and coefficient of relative risk aversion (gamma) in selected countries. Mpra paper, University Library of Munich, Germany.

Attanasio, O. P., Meghir, C., and Santiago, A. (2012). Education choices in mexico: Using a structural model and a randomized experiment to evaluate progresa. *The Review of Economic Studies*, 79(1):37–66.

Baird, S., McIntosh, C., and Özler, B. (2011). Cash or condition? evidence from a cash transfer experiment. *The Quarterly Journal of Economics*, 126(4):1709–1753.

Becker, G. S. (1994). *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education (3rd Edition)*. Number beck94-1 in NBER Books. National Bureau of Economic Research, Inc.

Behrman, J. R. (1997). *Mother's Schooling and Child Education: A Survey.* University of Pennsylvania - Institute for Economic Research.

Behrman, J. R., Foster, A. D., Rosenzweig, M. R., and Vashishtha, P. (1999). Women's Schooling, Home Teaching, and Economic Growth. *Journal of Political Economy, University of Chicago Press*, 107(4):682–714.

Bernal, R. (2008). The effect of maternal employment and child care on children's cognitive development. *International Economic Review*, 49(4):1173–1209.

Cardenas, J. C. and Carpenter, J. (2008). Behavioural development economics: Lessons from field labs in the developing world. *Journal of Development Studies*, 44(3):311–338.

Carneiro, P., Das, J., and Reis, H. (2013). Parental valuation of school attributes in developing countries: Evidence from pakistan. mimeo, University College London.

Chaudhury, N. and Parajuli, D. (2010). Conditional cash transfers and female schooling: the impact of the female school stipend programme on public school enrolments in punjab, pakistan. *Applied Economics*, 42(28):3565–3583.

del Boca, D. and Flinn, C. (2012). Household behavior and the marriage market. Working papers, Human Capital and Economic Opportunity Working Group.

Eckstein, Z. and Lifshitz, O. (2011). Dynamic female labor supply. *Econometrica*, 79(6):1675–1726.

Fiszbein, A. and Schady, N. (2009). *Conditional Cash Transfers: Reducing Present and Future Poverty*. World Bank policy research report.

Geweke, J., Keane, M. P., and Runkle, D. (1994). Alternative computational approaches to inference in the multinomial probit model. *The Review of Economics and Statistics*, 76(4):609–32.

Greenwood, J., Rogerson, R., and Wright, R. (1993). Putting home economics into macroeconomics. *Quarterly Review*, (Sum):2–11.

Hasan, A. (2010). Gender-targeted conditional cash transfers : enrollment, spillover effects and instructional quality. Policy Research Working Paper Series 5257, The World Bank.

Heckman, J. and Singer, B. (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica*, 52(2):271–320.

Keane, M. P. (1994). A computationally practical simulation estimator for panel data. *Econometrica*, 62(1):95–116.

Keane, M. P., Todd, P. E., and Wolpin, K. I. (2011). *The Structural Estimation of Behavioral Models: Discrete Choice Dynamic Programming Methods and Applications*, volume 4 of *Handbook of Labor Economics*, chapter 4, pages 331–461. Elsevier.

Keane, M. P. and Wolpin, K. I. (1997). The career decisions of young men. *Journal of Political Economy*, 105(3):473–522.

Magnac, T. and Thesmar, D. (2002). Identifying dynamic discrete decision processes. *Econometrica*, 70(2):801–816.

MDG (2010). Pakistan millennium development goals - report 2010. Technical report, Centre for Povery Reduction and Social Policy Development.

Rust, J. (1994). *Structural estimation of markov decision processes*, volume 4 of *Handbook of Econometrics*, chapter 51, pages 3081–3143. Elsevier.

Todd, P. and Wolpin, K. (2006). Assessing the impact of a school subsidy program in mexico: Using a social experiment to validate a dynamic behavioral model of child schooling and fertility. *American Economic Review*, 96(5):1384–1417.

Todd, P. E. and Wolpin, K. I. (2010). Structural estimation and policy evaluation in developing countries. *Annual Review of Economics*, 2(1):21–50.

Train, K. (2009). *Discrete Choice Methods with Simulation*. Cambridge University Press.

7 Appendix

Table 12: Variance-Covariance Matrix

Parameter	Estimate	Standard Error
var($\epsilon_t^{(0,1)}$) - var($\epsilon_t^{(0,1)}$)	-	-
var($\epsilon_t^{(1,1)}$) - var($\epsilon_t^{(0,1)}$)	1	-
var($\epsilon_t^{(0,0)}$) - var($\epsilon_t^{(0,1)}$)	4.71	0.003
var($\epsilon_t^{(1,0)}$) - var($\epsilon_t^{(0,1)}$)	0.04	0.005

Note: This table shows the variance estimates and the asymptotic standard errors. The variance-covariance matrix was estimated using the GHK method. The results are in utility differences with the alternative school/not working.

Table 13: Probability of passing grade

Mother' s Education - No Education

	Girls' Age									
Girls' Education	6	7	8	9	10	11	12	13	14	15
0	0.36	0.46	0.67	0.85	0.86	0.83	0.40	0.67	0.67	0.00
1	0.42	0.51	0.72	0.64	0.80	0.71	0.78	0.75	0.25	0.00
2	0.00	0.70	0.62	0.77	0.79	0.84	0.69	0.52	0.76	0.50
3	0.00	0.00	0.63	0.48	0.65	0.80	0.66	0.59	0.86	0.65
4	0.00	0.00	0.00	0.50	0.77	0.73	0.76	0.78	0.80	0.71
5	0.00	0.00	0.00	0.00	0.20	0.80	0.41	0.41	0.38	0.20
6	0.00	0.00	0.00	0.00	0.00	0.71	0.74	0.60	0.78	0.61
7	0.00	0.00	0.00	0.00	0.00	0.00	0.60	0.62	0.73	0.58
8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50	0.69	0.73
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.40	0.67
10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50

Mother's Education - at least some education

	Girls' Age									
Girls' Education	6	7	8	9	10	11	12	13	14	15
0	0.52	0.74	0.86	0.78	0.67	0.67	0.40	0.67	0.67	0.00
1	0.38	0.63	0.67	0.76	0.86	0.68	0.75	0.75	0.50	0.00
2	0.00	0.50	0.82	0.77	0.86	0.89	0.87	0.76	0.50	0.00
3	0.00	0.00	0.72	0.59	0.67	0.83	0.74	0.79	0.95	0.00
4	0.00	0.00	0.00	0.82	0.74	0.78	0.83	0.68	0.82	0.88
5	0.00	0.00	0.00	0.00	0.50	0.63	0.42	0.63	0.36	0.14
6	0.00	0.00	0.00	0.00	0.00	0.57	0.79	0.78	0.95	0.43
7	0.00	0.00	0.00	0.00	0.00	0.00	0.80	0.60	0.78	0.73
8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.56	0.92	0.85
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.64
10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.67

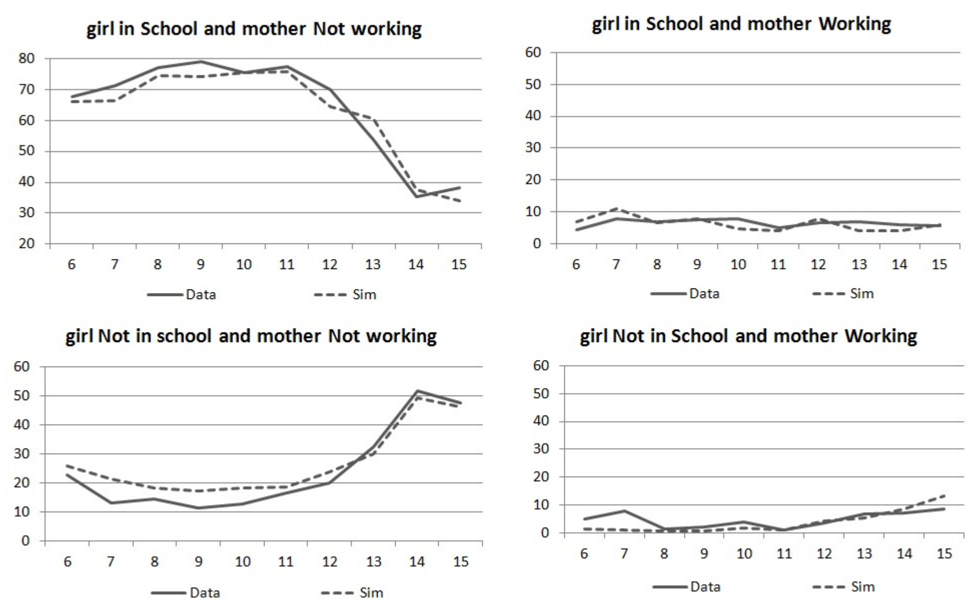
Notes: We estimate the probability of progressing for each grade as the ratio of individuals who passed to the next grade compared to the year before at a particular age, given their education and their mother's education. Mother's education is a binary variable, which can represent a mother with no education or a mother with at least some education.

Table 14: Impact on mother's labour market participation rate

	Girls' age	
Policy measures	6-10	11-15
CCT(U)	0.04	-0.02
School Building Program	0.06	-0.07
Free Child Care Services	0.06	-0.03

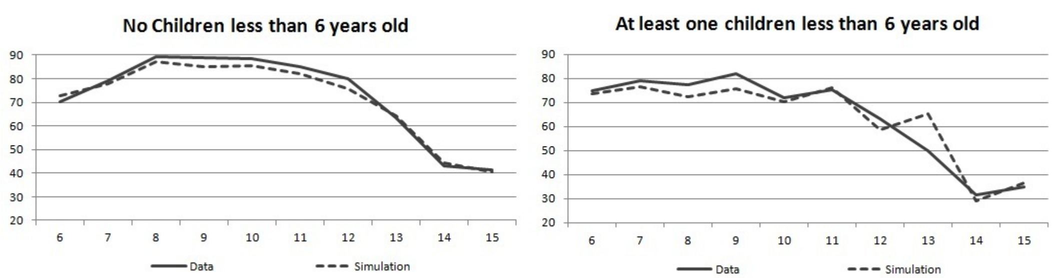
Notes: This table shows the impact, in percentage points, on mother's labour market participation rate of the Conditional cash transfer scheme (Upper bound scenario), the school building program and of the free child care services program. The table displays the average impact for younger girls (6 to 10 years old) and for older girls (11 to 15 years old).

Figure 11: Model fit - household choice distribution



Note: This figure shows the model fit for the 4 mutually exclusive household choices.

Figure 12: Model fit - Girls' school enrollment rate by number of young children



Note: This figure compares the model fit for the girls' school enrollment rate for households with and without children less than 6 years old.