The effect of cash transfer programs on educational mobility - an agent based approach*

Florian Wendelspiess Chávez Juárez**

November 18, 2013

Abstract

In this paper I develop a model to reproduce the phenomenon of high intergenerational correlations in education, particularly in Latin America. The model is based on empirical evidence and implemented through agent based modeling techniques. The effect of conditional cash transfer programs on educational mobility is then analyzed. The results suggest that conditional cash transfer program can substantially increase intergenerational mobility in education. I find that adapting the subsidies to the poverty level is more important than conditioning them to participation. Moreover, using parental education as eligibility criteria can increase the efficiency of a program in increasing educational mobility.

Keywords: educational mobility, conditional cash transfer program, agent based modeling, inequality, schooling decision

JEL-Classification: C63, D63, H52, I24, I3, J62

---

*A previous version of this paper was distributed under the title 'Education Mobility - An Agent Based Approach'. I would like to thank Christelle Garrouste, Luna Bellani, Tobias Müller, Marcelo Olarreaga and the members of the economics department of the *Universidad Iberoamericana* in Mexico City for their substantial comments on earlier versions of this paper. I am also grateful for comments made by the participants of the 1st Conference of the Economic Development Network (Mexico City), the INFER conference 2013 (Orléans, France), the AFSE meeting 2013 (Aix-en-Provence, France), the 12th Journées Louis-André Gérard-Varet (Aix-en-Provence, France), the 2013 ECINEQ conference (Bari, Italy), the NIP meeting in Mexico City and the LACEA conference in Mexico City. Comments on this version are very much appreciated; please send them by email to florian@chavezjuarez.com

**University of Geneva, Department of Economics, Uni-Mail 5211, 40 Bd. du Pont d'Arve, 1211 Geneva - Switzerland.

1 Introduction

Following the economic crisis of 2008, economic and social inequalities moved back into the focus of public interest and nourished social movements all over the world. In the U.S. for instance, the *99-percent*-movement pointed to the large and increasing disparities between the very rich people and the main body of the population. Stiglitz (2012) argues that not only the state of inequality gives rise to anger, but the fact that social mobility is very low, giving little hope to those at the bottom of the social ladder. Education is probably the key to move up the social ladder as it directly affects life time earnings of individuals (Ashenfelter and Krueger, 1994; Psacharopoulos and Patrinos, 2004; Machin, 2009). To promote education among the most disadvantaged people in the society is a main goal of many recently introduced conditional cash transfer program. It is therefore important to discuss the role of such government programs in promoting educational mobility. In this article I try to do so by developing a model of educational mobility. The model has two main goals. First, I try to reproduce the actual situation in countries with low educational mobility. In a second step, I use the model in order to see whether (conditional) cash transfer programs can increase educational mobility and what properties of these programs matter most.

The model is based on a large body of empirical literature, which aimed at measuring the phenomenon and identifying the causes[1]. Among the identified channels of intergenerational transmission of education are the biological transmission of ability (Björklund et al., 2010; Black et al., 2009; Anger and Heineck, 2010) and budget constraints of the parents when investing in the education of their children (Carneiro and Heckman, 2002; Attanasio and Kaufmann, 2009; Plug and Vijverberg, 2005).

In addition to these two main channels, I also consider several empirical regularities that might be important for educational mobility. Among these regularities, I consider for instance assortative mating and education dependent fertility. Their inclusion in the model can be dissuasive when looking for analytical solutions of a model. For this reason, I make use of techniques from agent based models (ABM), which I consider to be particularly well suited in this context.

One of the key advantages of ABM is their flexibility. This flexibility is particularly important in our context for two reasons. First, the modeler has a lot of flexibility in adding the aforementioned empirical regularities. In this sense, the model is able to consider heterogeneity for instance in ability, education and the number of children per family. All these heterogeneities can be implemented endogenously and through the dynamic aspect of the model they are transmitted to the next generation based on findings of the empirical literature. Thus, this modeling approach allows me to get closer to the empirical evidence without losing the advantages of a theoretical model[2]. Second, the ABM approach allows the modeler to simulate different policy

[1] A complete discussion of the literature can be found in surveys such as Black and Devereux (2011), Björklund and Jäntti (2009) and Piketty (2000)

[2] Obviously, there might be other modeling approaches able to achieve the same goals and complement the

measures in a very flexible way. For instance, it is easily possible to change eligibility rules of program participants without having to change elements of the model and recomputed the optimal behavior. Despite the advantages of agent based models just described, it has to be mentioned that the model I develop in this paper is using some ABM techniques but should not be seen as a full agent based model. For instance, agent based models normally include some endogenous emergence, which is not the case in this model.

The results suggest that conditioning the help on the private education investment of the families and adapting it to the level of poverty makes it more efficient in terms of cost and benefits as opposed to an unconditioned program. Surprisingly, the conditioning on the poverty level seems to be more important than the conditioning on private investment in education. A second result is that the more a government wants to invest, the larger the base of subsidy receiving families should be. Otherwise an overcompensation of the poorest occurs and reduces the effectiveness of the program. In the model a reduction of the intergenerational education correlation from about 0.55 to about 0.30 can be achieved with a proportional tax rate of approximately 10%. Finally, the model results suggest that using parental education as eligibility criterion for conditional cash transfer programs might increase their efficiency in terms of reducing intergenerational mobility in education as compared to criteria based on income.

The paper makes three contributions to the economic literature on educational mobility. First, the model allows me to analyze the potential impact of cash transfer programs on educational mobility. The analysis of different policy schemes aims at identifying the theoretically most efficient ways to reduce intergenerational correlations in education and increase social mobility. Second, by the use of agent based modeling techniques, I can consolidate the theoretical modeling approach and the numerous empirical findings. This should allow me to model intergenerational mobility in education closer to the data. Finally and more generally, the paper should allow me to demonstrate the usefulness of agent based modeling techniques in the literature on social mobility as a complement to the traditional tools[3].

The remainder of the paper is organized as follows. In section 2, I present some recent literature on educational mobility. Section 3 describes the model with a special focus on how empirical evidence is taken into account. In section 4 I first present the baseline model to show how it reproduces the status quo and then I analyze the effect of different policy schemes. Section 5 concludes the paper and provides some recommendations and an outlook on possible extensions.

---

empirical and theoretical models.

[3] Along with contributions from authors like Blake LeBaron, Leigh Tetsfatsion, Scott Page, Joshua Epstein and many others, this paper should also help to show that ABM can be a very appealing complement to traditional methods in economics in general. This was recently highlighted by Foley and Farmer (2009).

2 Educational mobility - some empirical evidence

Educational mobility and related phenomena have been important topics in empirical research. In this section I will briefly discuss some of the contributions in a non-exhaustive way, focusing on the most relevant findings for the model developed later in this paper[4]. Among these relevant findings - which the model aims at reproducing - I discuss especially the measures of educational mobility and two main generating mechanisms: the genetic transmission of ability and the intergenerational linkage through the economic situation of individuals.

2.1 Estimation of educational mobility

The measurement of educational mobility can be done in different ways. A first possibility is the use of transition matrices and a second is the correlation between the years of education of the parents and the children. The simplicity of the correlation - allowing us to summarize educational mobility in a single numerical value - makes it very useful for this analysis. I will therefore focus on this metric throughout the paper, even though other measures could also be used.

A further advantage of this simple metric is the availability of many estimates in the literature. Intergenerational correlations in education have been estimated for a large set of countries. For instance, Hertz et al. (2007) use it to rank 42 countries. Their ranking of countries shows an impressive pattern: the 7 Latin American countries taken into account take the first 7 ranks. This suggests that the educational mobility is particularly low in Latin America. Among these seven countries, Peru is with 0.66 the country showing the highest correlation, followed by Ecuador, Panama and Chile. For the remainder of their ranking, no clear geographical pattern can be identified. Dahan and Gaviria (2001) compare only Latin American Countries (LAC) and find that Mexico has the second lowest intergenerational mobility level behind El Salvador. The substantially higher correlations for Latin American Countries make them particularly interesting to study. I focus on Mexico to calibrate the model because it is a country with low educational mobility and a large availability of high quality data.

Table 1: Intergenerational education correlation estimates in the literature

Country	**Period**	**Correlation**	**Source**
Different LAC	1994-2004	0.55-0.66	Hertz et al. (2007)
Mexico	2005	0.55	de Hoyos et al. (2010)
Mexico	2005	0.40-0.45	Wendelspiess Chávez Juárez (2012)

Notes: The values of Wendelspiess Chávez Juárez (2012) include children at the age of tertiary schooling and represent therefore lower bound estimates for the whole population.

Looking at the Mexican case more closely, several estimates of the intergenerational correlations

[4] For a more extensive review of the empirical literature, see Black and Devereux (2011) or Wendelspiess Chávez Juárez (2012) and their references.

in education can be found in the literature. Table 1 displays some of them. The estimates differ across studies, but in general, they lie between 0.5 and 0.6 for most countries. This range of values is used as stylized fact that the baseline model aims at reproducing. In order to reproduce these figures I include several mechanisms which were identified in the literature on educational mobility. The most important of these mechanisms are discussed in the next subsection.

2.2 Mechanisms generating low educational mobility

The estimates of pure correlations as displayed in Table 1 are useful to quantify the phenomenon, but not enough to understand the mechanisms. The literature identified mainly three channels of intergenerational transmission: the biological transmission of intelligence and ability, the economic channel and the direct education-to-education channel. I will focus on the former two in the construction of the model and discuss them hereafter in some more details[5]. Besides the two main mechanisms of transmission, I include in this discussion as well two phenomena potentially helping to explain a part of low educational mobility: assortative mating and education dependent fertility. Both phenomena are not directly considered to be channels of transmission, but as I will outline, they might play an important role and should not be neglected.

2.2.1 The biological transmission of ability

The biological transmission channel refers to the genetic transmission of ability. This transmission makes the ability endowments of both generations alike. Assuming that more able people are getting more educated, the years of education between two generations will be highly correlated as a consequence. The inherited part of total ability is substantial according to recent studies. Anger and Heineck (2010) use German panel data and find that a 1 point increase in the IQ score of the parents is associated with a 0.45 to 0.5 point increase in the comparable scores of the children. Björklund et al. (2010) use Swedish data and find a father-son IQ-correlation of 0.346 and around 0.5 between brothers. Finally, Black et al. (2009) find a father-son IQ-correlation of 0.38 using data from Norway. Using a sample of 7'576 children and young adults from the Mexican Family Life Survey (MxFLS), I find a value of 0.319 for the correlation with the father and 0.366 for the mother, which is in line with the values found for Europe.

2.2.2 The economic channel

Under *economic channel* I understand all arguments saying that the economic situation of the family is determined by the education of parents and determines itself the child's education. This link is due to families' budget constraints when investing in the education of their children. The problem of budget constraints can be related but is not limited to credit constraints. Some authors argue that credit constraints play an important role (Attanasio and Kaufmann, 2009;

[5] A completer discussion of the three channels can be found in Wendelspiess Chávez Juárez (2012) and its references.

Winter, 2007). However, an increasing number of authors argue that the long run economic situation of a family, rather than short run credit constraints, is determinant for the economic channel (Alfonso, 2009; Carneiro and Heckman, 2002). According to a prior study, this holds also in the Mexican case, where the long run economic situation matters more than the short run consumption in explaining the schooling outcome of young people (Wendelspiess Chávez Juárez, 2012). In the model I introduce the economic channel through the budget constraint that will determine families' investment in the education of their children.

2.2.3 Assortative mating

Assortative mating describes the fact that the process of finding a partner is not independent of socioeconomic characteristics such as education. Typically there is a strong correlation in education and cognitive ability between spouses. The analysis of the exact process of partner matching goes beyond the scope of this paper[6]. Mare (1991) and Behrman and Rosenzweig (2002) discuss and consider assortative mating in the context of intergenerational transmission of education and argue that it has a potentially important role.

Assortative mating makes it more likely to have parents with similar education and ability levels than would prevail under a completely randomized matching process. A child will therefore have either a double-benefit from highly educated parents or a double-disadvantage in the opposite case. Under assortative mating, the spread of education is likely to be constant over time, while under random matching a regression to the mean would be expected.

Mare (1991) reports increasingly positive assortative mating from 1930 to 1980 for the U.S. and hypothesizes about consequences on inequality and its intergenerational transmission. As I will show later in more details, the data from Mexico confirms the hypothesis of assortative mating impressively. The spouse IQ correlation attains values of around 0.4 and the same metric applied to years of education goes up to almost 0.7. The importance of assortative mating for intergenerational mobility is also highlighted by Ermisch et al. (2006), who use German and British data and find that on average about 40-50% of the covariance between parents and own family permanent income can be attributed to assortative mating. The authors explain their finding by very strong spouse correlations in education.

2.2.4 Education dependent fertility

Education does not only have an effect on the partner finding process, but also affects fertility. Highly educated couples tend to have fewer children than less educated couples (Drèze and Murthi, 2001). Particularly the education level of the mother is important in explaining different fertility rates (Breierova and Duflo, 2004; Cleland and Rodríguez, 1988). Having on average more

[6] For more details, see for instance Thiessen and Gregg (1980) who discuss and document the mating process with respect to different dimensions such as intelligence, anthropometric measures and personality traits. Wolf and Figueredo (2011) see positive assortative mating as conservative bet-hedging strategy in reproduction and offer a deeper understanding of the reasons behind the observed phenomenon.

children among less educated parents directly affects the intergenerational transmission. Less educated parents tend to be poorer and given the larger number of children, they have fewer resources per child to invest in education. Like for assortative mating, I will show in a later section the figures for Mexico that confirm the presence of education dependent fertility clearly.

2.2.5 Relevant evidence for the model

The evidence presented in this section represents the cornerstones for the model I develop in this paper. The first goal of the paper is to model the mechanism as suggested by the empirical evidence in order to reproduce the intergenerational correlations observed in Latin American Countries. Table 2 summarizes the most important stylized facts that the model aims at reproducing and satisfying.

Table 2: Summary of relevant stylized facts

Description	**Value**
Intergenerational correlation in education	[0.45 - 0.65]
Parent-child IQ correlation	[0.31 - 0.38]
Spouse IQ correlation	≈ 0.41
Spouse education correlation	≈ 0.665
Correlation between education and fertility	< 0

The figures are relatively crude measures of the phenomenon, but they allow me to easily compare the model to the real world. However, to build the model, some more detailed information is needed. In the next section I will introduce the model along with a more detailed discussion of how these empirical regularities are taken into account.

3 The model

Based on the empirical evidence and findings of the literature explained in the previous section, I develop in this section a theoretical framework aiming at reproducing the observed intergenerational correlations in years of education. First, I provide a general overview of all events taking place in each period. Afterwards I introduce the different mechanisms in detail and explain how they were calibrated.

3.1 General overview of the model

In the model, people live for two periods corresponding to the childhood and adulthood. During childhood individuals are getting educated and during adulthood they participate in the labor market and raise their children. Families are composed of two adults and an endogenous number n_c of children. In each period parents decide how much to invest in the education of children

and how much to consume. Wages and education are endogenous and defined through equations relating them to ability and family investment in education.
To initiate the model, a set of N individuals is created. Table 3.1 displays the four main steps taking place in each period.

Table 3: Overview of all steps taking place in each period

Step	Description	Details in section
1	The adults of the previous period leave the model and former children become adults. They start **searching a partner** and in case of finding one, a new family is created.	3.2
2	Each family has a positive probability of **getting children**, where the number of children depends on the education level of the mother.	3.3
3	Each **family earns a salary** according to equation 4 and pays taxes.	3.4
4	Based on all private information of the family, the parents decide how much to **invest in the education of children** and how much to consume. Children receive education according to the investment.	3.5

Once these steps are completed, the next period begins and the sequence starts again. I will now present each of these steps in detail and highlight how and to what extent empirical evidence was used in the modeling process[7].

3.2 Step 1: Partner search with assortative mating

Modeling the partner matching process in accordance with the empirical evidence on assortative mating can be done in several ways. For this model, I chose a very simple solution, where individuals with too large differences in terms of education and ability cannot form a couple. The threshold was defined by consulting the Mexican Life Family Survey, which shows that for 95% of the couples the IQ difference does not exceed 30 points and in the case of education 95% of all partners are within 1.5 standard deviations[8]. Using this simple exclusion rule, potential candidates for a partner match have to satisfy

$$\delta_{IQ} \equiv |IQ_{woman} - IQ_{man}| \leq 30 \quad \wedge \quad \delta_{educ} \equiv |educ_{woman} - educ_{man}| \leq 1.5\sigma_{educ} \qquad (1)$$

This double condition is a very simple way to account for assortative mating. Much more sophisticated algorithms could be imagined. However, as I will show in the result section, this simple implementation produces results that are very close to the spouse IQ-correlation of 0.408 and the spouse education correlation of 0.665 reported in Wendelspiess Chávez Juárez (2012).

[7]Additionally, an overview of all parameters in the baseline model can be found in appendix A.
[8]The two conditions are overlapping for a relatively large part of the population. However, using only one of them yields less satisfactory results than using them together.

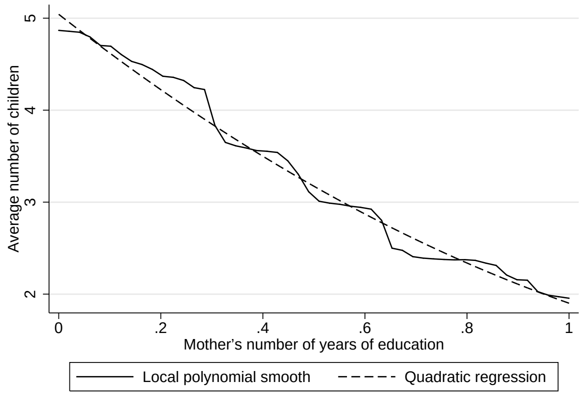
3.3 Step 2: Getting children

Once couples are formed, they have a positive probability of getting children. Two main processes are used to define the number of children and their ability level.

Education dependent fertility

The expected number of children depends on the education level of the mother. Figure 1 displays

Figure 1: Number of children as function of maternal education



Source: Based on data from the Mexican Census 2010

the local polynomial smooth and the quadratic OLS fit of the number of children in function of the education of the mother based on data from the 2010 Mexican Census. For simplicity, education is normalized to the interval of 0 (no education) and 1 (highest education in the population). I use only woman between 40 and 50 years old. Younger woman are excluded because they might still get more children, thus the observed number of children is probably not equal to the final number. The older women were excluded because they have substantially higher average number of children, which does no longer reflect the current situation in Mexico. Additionally I dropped extreme values of women with more than 25 children. The results are quite robust to small changes in these sample selection thresholds. The parametric version is used to calibrate the model. This choice is based on the good fit of the parametric model as compared to the non-parametric and to the easier implementation in the model.

Table 4: Number of children as a function of maternal education

Number of children (dep. var.)	Coef.	Std.Err
Mother's education	-4.3354***	(0.0089)
Mother's education (squared)	1.1941***	(0.0081)
Constant	5.0427***	(0.0022)
N	6544865	
Adj. R^2	.183	

Source: Author's calculation using the Mexican Census 2010.
Notes: Mother's education is normalized to the interval of 0 (no education) to 1 (highest education). Standard errors in parenthesis. Significance levels at 10% (*), 5% (**) and 1% (***).

Table 4 displays the full estimation results of the parametric fit, based on an OLS estimation. While the average number of children for very little educated mothers is around 5, the highest educated women have on average only slightly less than two children.

In the model I use the conditional expectation obtained from the parametric fit as the argument of a Poisson distribution. This allows me to reproduce a certain heterogeneity in the number of children for a given level of education. Thus, the actual number of children a couple has is given by

$$n_c(e_m) \sim \mathcal{P}(E[n_c|e_m]) \tag{2}$$

where $E[n_c|e_m]$ is the expected number of children conditioned on the level of education of the mother using the above mentioned parametric model. In this model the number of children is not endogenously determined by the utility maximization, as this does not seem to be appropriate for the Mexican case. Moreover, the focus of this paper lies on possible policy measures, thus families would have to anticipate the subsidies before getting children. This is beyond the scope of this paper.

Ability transmission through genes

Besides the endogenous number of children, their ability level is also endogenous in the model and depends directly on the ability level of the parents. To approach as much as possible the actual transmission of cognitive ability, I make use of the Mexican Family Life Survey (MxFLS), where a short cognitive ability test is included. Limiting the observations to all children where I have data on parental ability allows me to regress the ability score of the child on the same measure of the parents. For the purpose of readability, I scaled the ability measure to the IQ-scale[9], which is a normal distribution with mean 100 and standard deviation 15.

Table 5: Intergenerational transmission of ability

Child's ability (Dep. var)	Coef.	Std.Err
Father's ability (IQ scale)	0.2059***	(0.0115)
Mother's ability (IQ scale)	0.2857***	(0.0116)
Constant	51.8772***	(1.2600)
N	7576	
Adj. R^2	.169	

Source: Author's calculation, based on data from the Mexican Family Life Survey (MxFLS)
Notes: Standard errors in parenthesis. Significance levels at 10% (*), 5% (**) and 1% (***).

Table 5 reports the results of an OLS regression of ability on parental ability. The ability of the child is significantly influenced by the ability measures of both parents. However, the coefficient for the mother is significantly (F=17.05) higher than the one of the father. The unexplained

[9]Note that the ability measure of the MxFLS is not a complete IQ test. Therefore it should only be considered as a proxy measure for the IQ.

part of children's ability follows a normal distribution $\mathcal{N}(0, 13.53)$. Thus, in the model the transmission is simulated as follows:

$$IQ_{child} = 51.8772 + 0.2059 IQ_{father} + 0.2857 IQ_{mother} + \mathcal{N}(0, 13.53) \tag{3}$$

This implementation generates an IQ correlation with the father of approximately 0.33, which is very close to the correlation of 0.346 found by Björklund et al. (2010) and 0.38 computed by Black et al. (2009). The aim of this estimation is not to distinguish the biological transmission from the environmental effects, as it is controversially discussed in the literature. The unique aim is to reproduce the total intergenerational link in ability without going into the details of a decomposition in different sub-processes.

3.4 Step 3: Earning a salary: the wage equation

The wage w of each adult depends on his or her education e. I assume the following general form:

$$w = \beta_0 + \beta_1 e^{\beta_2} + \epsilon_w \tag{4}$$

where the β's are parameters and ϵ_w is a random disturbance term accounting for unobserved heterogeneity in the agents. For the sake of readability, the base wage β_0 is normalized to 1. This normalization shapes the interpretation of w to a multiple of the base wage obtained without education. The parameter β_2 is assumed to be equal to 2, which corresponds to a quadratic relationship between education and wages. This value is based on a non-linear least squared estimation using the 2010 Mexican Census[10]. The estimated parameter is 2.027, which is not significantly different from 2. Finally, the last parameter β_1 is required to keep the education-wage relationship stationary over time. I use the value of 3.8, which appears to be the value generating a stationary solution. This value has no direct economic meaning. The random disturbance term is drawn from a normal distribution with mean 0 and standard deviation 2. Thus, the calibrated wage equation is given by:

$$w = 1 + 3.8e^2 + \epsilon_w \tag{5}$$

Like for all calibrated equations I ran also for this equation several robustness checks. Some of these tests are discussed in the appendix.

[10]For the analysis, the top 1% wages were excluded. They represent very extreme and implausible values. The model includes 3,340,646 observations and has an R^2 of 0.621

3.5 Step 4: Investment in education of children

The education investment decision is the crucial decision taken by families in each period. The education production function is assumed to be common knowledge. The family faces a trade-off between current consumption and investment in the education of children. First, I introduce the education production function and afterwards I come back to the optimization process. Education depends on the investment i in education and the ability level a of the child. I assume the general form of the education production function:

$$e(a,i) = \gamma_1 i^{\gamma_2} a^{\gamma_3} + \epsilon_e \quad (6)$$

where the γ's are parameters and ϵ_e is a random disturbance term to account for non-observed heterogeneity. γ_2 is the parameter shaping investment to education. The data suggests that education is increasing and slightly concave with respect to investment. The parameter used in the model is 0.8. The ability is assumed to have a unit mean and γ_3 can be seen as a sensitivity of the educational outcome to ability. I assume the value of 1 for the ability parameter. Finally, the parameter γ_1 simply ensures that the result is scaled correctly to ensure the model to work. In this case it takes the value of 0.45.

The proposed functional form ensures that investment in education and ability are imperfect substitutes. A very smart child who receives absolutely no support for education is likely to drop out of the schooling system quite quickly. On the other hand, even with massive investments in education, a child with a very low level of intelligence will not completely succeed in school.

The fact that education will be zero when families do not invest at all might be a surprising assumption. I justify this choice by the definition of education and investment in this case. First, education can be seen as formal education that adds to some basic knowledge children would acquire in any case. By looking at the wage equation we can see that without education there is still a positive wage. Second, investment should be seen in the larger sense and not be limited to paying schools fees. In this respect, even very poor families can invest in the education, e.g. by the simple fact of sending their children to school instead of letting them work at home.

The optimization problem of the family

Knowing the education production function, the family has to optimize investment in education taking into account that it reduces resources for current consumption. We can express total consumption as:

$$C = W - I - T(W) + S(i, n_c, W) \quad (7)$$

where the capital letters C, W and I refer to the family level of consumption, wages and investment respectively. $T(W)$ is the amount of taxes the family has to pay in function of their labor income and $S(i, n_c, W)$ is the amount of subsidies the family receives from the government.

In the absence of a government, the last two terms are simply equal to zero.
Families derive utility from their current consumption level and the expected education level of the children[11]. I assume intergenerational altruism given that investment in education has no direct return for the older generation. We can therefore assume for instance the following form of the utility function:

$$U(C, E[e_c]) = C^{\alpha} \prod_{i=1}^{n_c} E[e_c]^{1-\alpha} = C^{\alpha} \left(E[e_c]^{n_c}\right)^{1-\alpha} \tag{8}$$

where $E[e_c]$ is the expected education level of the children and C the current consumption level, both a function of investment in education. In the case of a single child family the utility function is a simple Cobb-Douglas function, while the relative importance of education becomes more important with the number of children. This formulation ensures that the total investment in education increases with the number of children, while the investment per child is decreasing.
As current consumption and expected education are determined by the investment in education per child i, we can express the utility as a function of i.

$$U(i) = [W - i \times n_c - T(W) + S(i, n_c, W)]^{\alpha} \left[E[e_c(a, i)]^{n_c}\right]^{1-\alpha} \tag{9}$$

The family then maximizes this utility function with respect to i, where the solution depends on the number of children, the income of the family and the form of $T(W)$ and $S(i, n_c, W)$.

3.6 Implementation as agent based model

Before turning to the result of the model, I would like to highlight the potential value added by the use of agent based modeling techniques in this specific context. It is important to mention that the way I model the phenomenon is by far not the only possibility. More standard economic model with analytical solutions or other computer based models with numerical optimization could be used as well. However, the main advantage of the approach taken in this paper is its flexibility in various aspects.
First, the approach allows me to include multiple heterogeneities of agents, which can even be interrelated. For instance, the heterogeneous fertility rate is closely related to the non-random partner search process. Second, the model is flexible in terms of parameterization and can be easily adapted to other countries based on different parameters or even different equations. Third, the possibility of simulating different policy measures is an appealing advantage. Clearly, other models can also simulate different policy schemes. However, this approach is flexible in the sense that no analytical solutions are required and therefore highly complex policy schemes could be analyzed without any problem.

[11] An alternative way would be that families derive utility from the expected future earnings of children. Since future earnings are monotonically increasing in education, it will give the same results, once the functional form of the utility function is adapted coherently.

Agent Based Models can be positioned in between the empirical and the theoretical modeling approach[12]. It allows us to go beyond the availability of data to explore possible policy interventions and perform a *what-if* analysis. On the other hand this way of theoretical modeling requires fewer simplifying assumptions (e.g. on the distribution of heterogeneities) than standard models.

4 Results

The results of the model are reported in two steps: first I discuss the baseline model without any policy intervention and then I move on to the analysis of different policy measures. The goal of the baseline model is to show how the model fits the empirical evidence and how the different components of the model contribute to it. The section on policy interventions represents then the active use of the model to predict hypothetical responses of the model to different policy measures. These policy measures are inspired by actual conditional cash transfer projects, such as *Oportunidades* in Mexico or *Bolsa Familia* in Brazil. In addition to these results, I present several robustness checks on the baseline model and the policy intervention analysis in appendix B.

4.1 Baseline model and sequential implementation

As discussed earlier, different processes are potentially affecting the outcome of the model. In order to show how these different elements play a role, the model is implemented first with all elements activated and then excluding always one element. This partial exclusion of processes allows us to understand their relative importance in the model. The three elements that are individually excluded from the model are assortative mating, the biological transmission of ability and the education dependent fertility. In addition to the individual exclusion of elements, I also excluded the first two elements together in order to see if there are interaction effects that are substantially different from the sum of individual effects. Moreover, in appendix B.1 I present the changes of the results when excluding the stochastic terms.

Table 6 summarizes the main statistics of the full model and the relative importance of the different elements. The different statistics of the model are displayed in columns, while rows refer to different settings by including and excluding some elements.

The first panel of values (row [1]) refers to the baseline model, where all elements are activated. In the first three columns the correlation in education are displayed, starting with the spouse correlation, followed by the correlation of the child with the mother and the father respectively (column 2 and 3). The spouse education correlation of 0.673 is close to the value observed in the data and results from the partner eligibility rules discussed in section 3.2. The education

[12] For a general discussion on the usefulness of agent based modeling see for instance Foley and Farmer (2009) or Miller and Page (2007)

Table 6: Average statistics of full model and loss when excluding components

		Edu. corr. partner (1)	**Edu. corr. mother (2)**	**Edu. corr. father (3)**	IQ corr. father (4)	IQ corr. mother (5)	IQ corr. partner (6)
	Full model						
[1]	Correlation	0.673	**0.616**	**0.563**	0.327	0.373	0.450
	Standard deviation	(0.023)	**(0.050)**	**(0.053)**	(0.023)	(0.022)	(0.030)
	Excluded element:						
[2]	Assortative mating	-100.5% **	**-22.5% ****	**-34.9% ****	-37.7% **	-23.9% **	-100.4% **
[3]	Biological ability transmission	-1.7% **	**-10.1% ****	**-10.6% ****	-100.3% **	-100.3% **	n.s.
[4]	Educ. dependent fertility	n.s.	**-1.7% ****	**6.4% ****	n.s.	n.s.	0.9% *
[5]	Combination of [2] and [3]	-100.3% **	**-29.7% ****	**-41.6% ****	-99.3% **	-99.0% **	-101.2% **

Notes: The first panel called 'Full model' displays the average values of the statistics in the full model. The second panel shows the percentage change in these statistics when excluding the element indicated in the first column. Only significant losses are reported. Significance levels: *=5%, **=1%, n.s. = not significant at 5%

correlation of the child with the parents is the fruit of all processes involved in the full model and attains as well values close to the observed figures. The IQ correlations between the different agents are directly based on the calibration. First, the correlation between spouses reported in column (6) follows from the assortative mating restriction (equation 1). The IQ correlations of the child with the parents are resulting from equation (3) and reproduce the empirical facts very well.

The second and larger panel in Table 6 reports the same figures but when the element indicated in the first column is excluded. Note that only changes that are significant at the 5% level are reported.

Row [2] displays the generated correlations when assortative mating is not considered in the model, thus when all individuals are potential partners. By construction, the spouse correlations go to zero, since the matching process becomes completely random. More importantly, we can observe a sharp decrease in the intergenerational correlations in education and ability. The decrease is particularly important for the correlation with the father, but even in the case of the mother the correlation falls by more than 20%. This finding is due to *regression to the mean* and illustrates that the exclusion of assortative mating would be very problematic.

Row [3] shows what would happen if the transmission of ability trough genes would be excluded from the model. In this case the intergenerational IQ correlations are by definition driven to zero, since ability becomes a purely random characteristic. The intergenerational education correlations are reduced as well by approximately 10% of their initial values, which is substantial.

In contrast, the spouse education correlation is also affected significantly, but the change remains very small and economically rather irrelevant.
In row [4], the correlations of the model when excluding the education dependent fertility rate feature are reported. Generally, the effects are relatively small, as they do not exceed a change of more than 6.4% of the initial values. The education correlation with the mother goes slightly down, while the one of the father increases by about 0.035. Comparing this to the correlations of the full model, we can see that when ignoring this empirical regularity, the simulated correlations for both parents are much closer to each other. This would be in contradiction to the data, which suggests that the correlation with the mother is higher. Hence, even though the changes are not very high, the inclusion might be justified by the difference in correlation with each of the parents.
Finally row [5] presents the figures when the assortative mating and the biological transmission of ability are excluded. The reduction in the education correlations is slightly less than the sum of the two individual exclusions. Nevertheless, there does not seem to be a major interaction effect between the two elements.
Overall, the baseline model succeeds at reproducing the main stylized facts appropriately and the exclusion of some elements of the model substantially affects the main statistics in some cases. Nevertheless, it would be possible to reproduce the stylized facts without these elements by the mean of a re-calibration. Such an approach would, however, suffer from the fact that some empirical evidence is left aside. Such a reduced model would - as a consequence - put too much weight on the core processes of the model, especially on the optimization process. As a result of that overweighting, the effect of policy interventions is likely to be overestimated too. For this reason, all additional features such as assortative mating or the transmission of ability through genes are kept in the model in the following section where I introduce the government.

4.2 Policy measures

Based on the baseline model just described, I now present how the key statistics react to the introduction of different policy schemes. All measures I am introducing should be understood as additional measures compared to the status quo. In a first step, I will discuss different policy schemes and identify the most efficient programs. I will then focus on the two most efficient and discuss them in more details.
Table 7 summarizes the main characteristics of the 6 different policy schemes I analyze in a first step. Common to all schemes is that all people below an eligibility threshold are entitled to receive financial assistance. Besides this common feature, the schemes differ in three dimensions. First, the amount of financial assistance can be a function of private investment or not. Second, the financial assistance can be the same for all eligible households or a function of the distance to the eligibility threshold. Finally, the eligibility criterion can be based on income or the highest parental education. In real conditional cash transfer programs, households are normally

chosen according to income or some asset holdings. However, with respect to the topic of intergenerational transmission of education, it seems also interesting to see whether targeting particularly on families with little educated parents might increase the efficiency of the program.

Table 7: Description of the four subsidy schemes

	Type 1	**Type 2**	**Type 3**	**Type 4**	**Type 5**	**Type 6**
Eligibility criterion based on	income	income	income	income	education	education
Conditional on private investment?	Yes	Yes	No	No	Yes	Yes
Depending on poverty level?	No	Yes	No	Yes	No	Yes
Eligible if	$W < \omega$	$W < \omega$	$W < \omega$	$W < \omega$	$e_p < \omega_e$	$e_p < \omega_e$
Subsidy $S(i, n, W)$	ηin	$\left(1 - \frac{W}{\omega}\right) \eta in$	$\eta \omega n$	$(\omega - W)\eta$	ηin	$\left(1 - \frac{e_p}{\omega_e}\right) \eta in$
Income tax $T(W)$	τW	τW	τW	τW	τW	τW

Types 1 to 4 are all directly inspired from conditional cash transfer programs, where the eligibility threshold is defined in the income dimension. For these schemes, all people below the threshold ω are eligible to receive subsidies. The threshold is defined as a multiple of the median income[13], analogous to the relative poverty line. For the last two types I use the highest parental education as eligibility criterion. ω_e is a certain education level and all households with the highest education level below this value are eligible.

Let us now discuss each scheme in detail. The first scheme (Type 1) is a simple reimbursement of a proportion η of private investment in education. Under this scheme, the subsidy is conditional on the private investment, but insensitive to the poverty level of the family. This sensitivity to the poverty level is taken into account in Type 2, where the poorest get reimbursed a higher proportion of their private investment in education. The amount of private investment is no longer determinant for the amount of subsidies under Type 3 and 4. Type 3 corresponds to a simple lump sum transfer to each child, while in Type 4 this subsidy is again dependent on the level of poverty. Type 5 and 6 correspond to Type 1 and 2 respectively, with the difference that eligibility is based on the highest parental education[14].

It could be argued that conditioning subsidies on private investment is not suitable in the context of very poor families as they might not be able to invest at all. This is true when considering private investment in the narrow financial sense. However, I argue that the private investment should be seen as a broader concept. For instance, school assistance instead of child labor is a sort of investment in terms of opportunity cost.

For all schemes, the corresponding tax scheme is a proportional tax on the sum of all wages of the family. In order to ensure that the amount of taxes corresponds to the amount of subsidies,

[13]This is $\omega = \theta \tilde{W}$, where $\tilde{W}$ is the median income

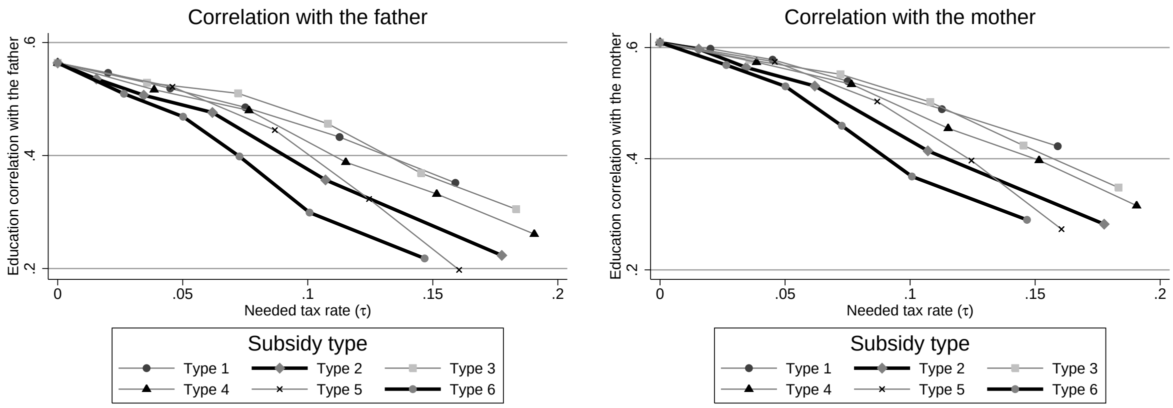
[14]Note that we could also apply the same logic to Type 3 and 4 and create two schemes more. However, simulations showed that they were not efficient and therefore I chose not to present them.

the tax rate is adapted endogenously. The tax rate in period t is set in order to cover the expenses on subsidies of period *t-1*. As a consequence, each generation pays as adult for the subsidies received as child, making each generation paying for their own subsidies[15]. This post-paid implementation is justified by the endogenous nature of the program size. The total amount of subsidies depends on the decision made by families, which depends itself on the amount of taxes. By paying the subsidies of the previous period, we can fix the amount of taxes and subsequently families can optimize their investment in education. This way, the model ensures that all subsidies are paid by taxes and the government does not incur debt for the financing.

Comparing different subsidy schemes

Figure 2 displays the reduction in the intergenerational education correlation with the father (left side) and the mother (right side) in function of the needed tax rate for the different subsidy schemes. Each point in the graph corresponds to a specific setting of the simulation. The points

Figure 2: Comparison of different subsidy schemes



are vertically not on the same values because the changing parameter when moving to the right is η. The reported τ on the horizontal axis is endogenously determined.

First and foremost, all schemes reduce education correlations significantly. Hence, such a policy intervention should increase educational mobility. The two best performing schemes are the two with double conditioning for both income and education as eligibility condition. They are highlighted in the graph with thicker lines. For all levels of tax rate, these two schemes yield higher reductions of the intergenerational correlations as compared to the other schemes. Additionally, the scheme where the eligibility is based on parental education (Type 6) is more efficient than the one on income.

When focusing only on the schemes based on income as eligibility criterion, we can observe that the second most efficient scheme is Type 4. Type 4 corresponds to the aid being dependent on the poverty level but not on private investment. Hence, it seems to be more important to adapt

[15]Since this way of implementing the model introduces a problem at the beginning of the simulation, the data used in the study always excludes the first periods and makes use of the model data once it is stabilized.

the subsidy to the poverty level rather than conditioning on private investment. This result of the model is in line with recent findings on conditional cash transfer programs. For instance, Banerjee and Duflo (2011, p.80) report that the main benefit of these programs in increasing school attendance stems from the income effect and not necessarily from the conditioning on sending children to school.

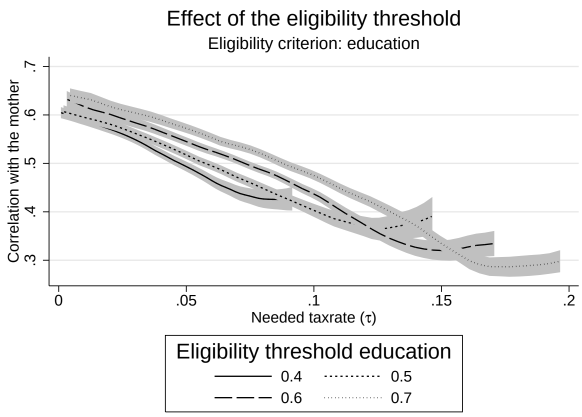
The role of the eligibility threshold

For the remainder of this section, I will focus on the two double conditioned schemes (Type 2 and 6). An important question is what eligibility threshold a program should set. By setting the threshold very low, only few people will benefit from the program, but the per capita subsidies are higher. Moving the threshold up the distribution of income, more and more people will benefit at the cost of lower per capita subsidies.

I performed the same simulation as above for all subsidy schemes and varied the eligibility threshold. The basic intuition that the threshold should be adapted according to the size of the program seems to be confirmed. Whenever the size of the program increases, the number of people eligible to receive financial assistance should increase as well.

Figure 3 displays the reduction of the education correlation with the mother for different levels of the eligibility threshold under scheme 6.

Figure 3: Reduction in function of tax rate for different levels of the threshold



We can observe that for relatively small programs with consequently small needed tax rates the most efficient threshold is the lowest. At a given size of the program, the reduction becomes smaller and eventually stops. Thus, for larger programs, higher thresholds become more efficient. For a tax rate of about 10%, the threshold where everybody below 0.5 of standardized education is eligible seems to be the most efficient. To fully understand the mechanisms here, it might be useful to look closer to the relationship between the education of parents and children.

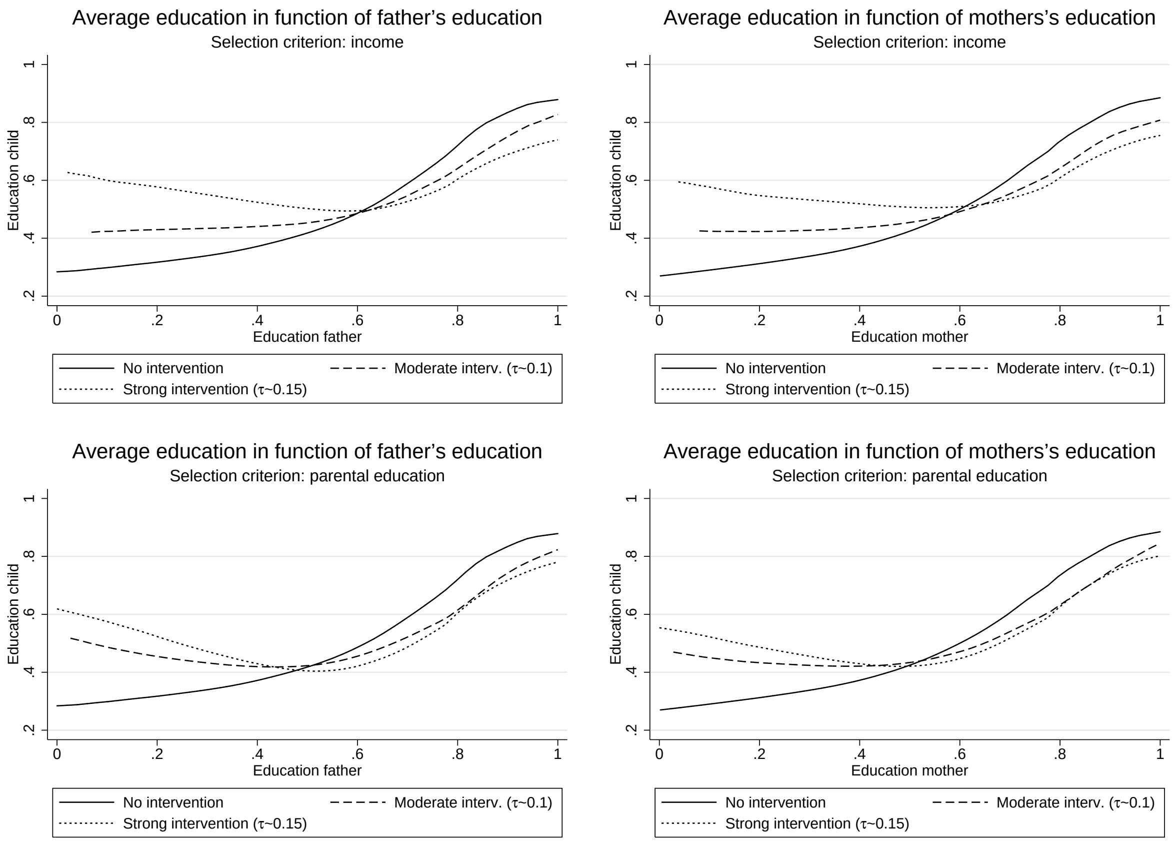
A closer look at the link between parental and offspring's education

To simplify the discussion, let us focus on three situations: no, moderate and strong intervention.

They correspond to a tax rate of 0% and approximately 10% and 15% respectively. Furthermore, let us take only the case where all people below the median income are eligible to receive subsidies.

Figure 4 shows the average education of children in function of the parental education using a local polynomial smooth. The graphs show the situation for the correlation with the father (left) and the mother (right) and for income (above) and education (below) as eligibility criterion. This allows us to better understand the effects of the program.

Figure 4: Effect of the subsidy on the education correlation with the parents

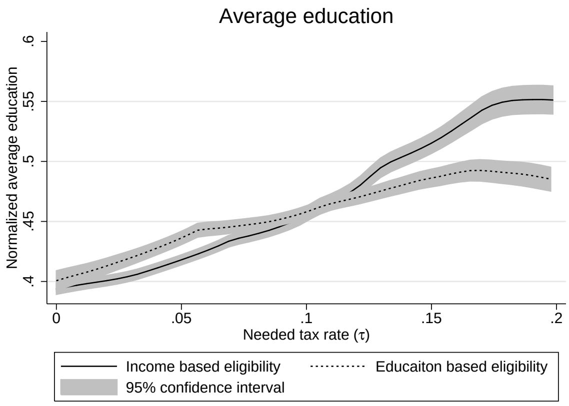


In the case of no state intervention, we observe an increasing and slightly S-shaped function, starting at values of 0.3 and going up to 0.85 approximately. For a moderate intervention ($\tau \approx 0.1$), the curve seems to turn around the midpoint, making the children of very little educated parents better off, while those of well-educated lose slightly. The increase for the children of parents with little education is due to a net transfer to them, while the reduction at the upper tail of the distribution is due to the income tax and therefore to fewer resources. Note that for a tax rate of about 10% and parental education as eligibility criterion we observe a slight overcompensation of the worst off. This overcompensation is generalized on all four graphs when considering the strong intervention. In the extreme case of a high tax rate ($\tau \approx 0.15$) a U-shaped function is observed, suggesting that the children of the very little and very highly educated parents get higher education as compared to those with parents of middle education. Two things should be learned from this case. First, it is easy to see that the group of eligible people should be enlarged in the case of a large intervention. This would allow the children of parents with a middle high education to benefit as well. Second, the beneficial situation for the least favored children is always accompanied by a decrease in average education for the best of students, which is straightforward from the fact that the policy measure has a purely redistributive character. To see if the beneficial impact at the bottom of the distribution offsets the loss at the top, I will now turn to the analysis of average education and income distributions.

Effects on average education and consumption

Figure 5 shows the average education in function of the tax rate used to finance the subsidies for both eligibility criteria.

Figure 5: Effect of the subsidy on the average education

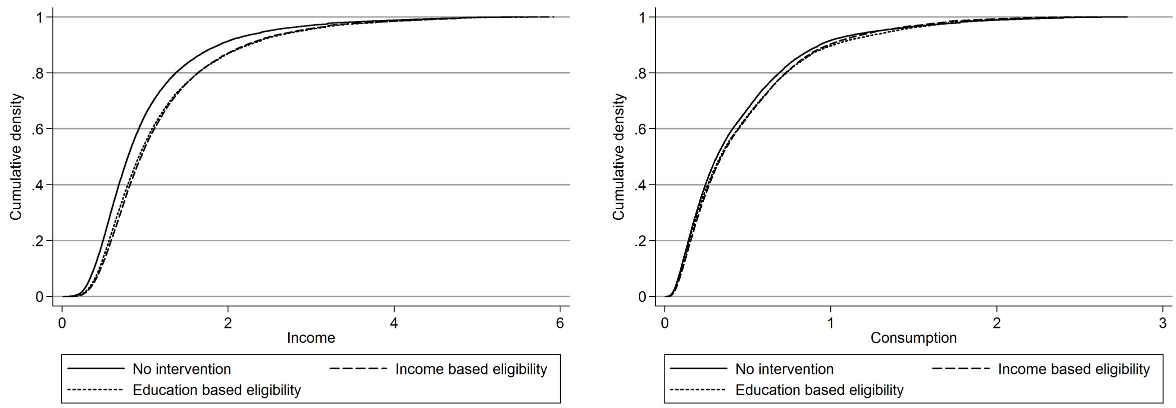


The relationship between average education and the tax rate is clearly positive, suggesting that the negative effect of the best off is more than compensated by the beneficial effect for the lower tail of the distribution. For relatively small programs the increase in average education is higher when using parental education as eligibility criterion. In contrast, for larger programs the scheme with income as eligibility criterion increases average education more than the one based on parental education. There are two possible explanations.

First, it can be that the threshold of eligibility is no longer efficient for large programs and that this inefficiency is more pronounced when using the education eligibility criterion. This might explain some part of the story, but there is also another reason. A program targeting on families with low education focuses essentially on reducing the correlation in education. However, by focusing on income as an eligibility criterion we target those families with fewer resources to invest. Thus, it is reasonable to see a better performance of the education eligibility in terms of correlation reduction while the focus on income has stronger effect of average education.

Now, one might argue that not the educational level is important, but what happens at the distribution of income. Figure 6 displays the cumulative distribution function of consumption per capita and income per capita for the moderate intervention case and both eligibility criteria.

Figure 6: Effect of the subsidy the distribution of income and consumption



The pattern of the effect is similar in both graphs. However, the effects are more pronounced in the case of the income distribution. For income, we can observe a clear first order dominance of the distributions where an intervention takes place. There is no difference due to the eligibility criterion. For consumption the distributions are very close to each other. One could now argue that what matters in the end is consumption and since there is basically no change, the intervention is not useful. Two arguments contradict this conclusion. First, as we have seen before, overall education increased and it is widely acknowledged that education has also a lot of positive non-monetary effects for the society. Second, social mobility is increased under the policy intervention. Therefore it is a fairer situation than the *laissez-faire* scheme. The pie is not getting bigger under this policy intervention, but the distribution of its slices is fairer.

5 Conclusions

In this paper I develop a model to reproduce the high intergenerational correlations in education particularly observed in Latin America. I consider two main generating processes: the biological transmission of ability and economic constraints for the investment in education. The use of agent based modeling techniques allows me to take into account several empirical regularities like the education dependent fertility or assortative mating.

Using data from Mexico to calibrate the model allows me to reproduce several key statistics as they are observed in the data. Among the reproduced statistics are the intergenerational correlations in education and ability and the distributions of education and income. Assortative mating seems to matter substantially, as the intergenerational education correlations drop sharply when not including it in the model.

The model is then used to simulate the effect of different policy schemes on educational mobility. Different subsidy schemes, all paid by a proportional tax, are simulated and their impact on educational mobility compared. The schemes differ essentially in two core characteristics:

subsidies are conditional on private investment or not and sensitive to the level of poverty or not. Additionally, two different eligibility criteria were used. First, I used the economic situation measured by parental income and second I used the highest parental education. Income as eligibility criterion is directly related to the current conditional cash transfer programs where income and family assets are mainly used to select eligible household. In contrast, the use of the highest parental education targets more directly the intergenerational transmission of education. A total of 6 different policy schemes are simulated. Several lessons can be drawn from this simulation exercise. First, subsidies to families with children, conditioned or unconditioned, are likely to increase social mobility and to have positive effects on the average education level. The gains for subsidy receiver more than offset the losses for the net contributors at the top of the income distribution. This finding holds for all policy schemes.

Second, subsidies that are conditioned on private investment and sensitive to the poverty level of the family seem to be more efficient than others. The sensitivity to the poverty level seems to be more important than the conditioning on private investment.

Third, for a given size of the government program the eligibility threshold based on the highest parental education yields higher reductions in intergenerational correlations in education as compared to income as criterion. In contrast, for bigger programs the average education increases more under a scheme where income is the eligibility criterion. This finding can be explained by the slightly different target population. When we focus on those families with little educated parents, we directly address the problem of high intergenerational correlations in education. Therefore is comes as no surprise that such a criterion increases educational mobility more for a given size of the program. In contrast, targeting on the poorest according to the income criterion increases average education more because we focus mostly on those not able to invest much in education, independently of their own education. Thus, the gains in terms of average education are likely to be higher when focusing on income as criterion, which is confirmed by the simulation exercise.

A last lesson from the simulation exercise is that the share of people eligible to receive subsidies should change in function of the size of the program. Larger programs should have a higher eligibility threshold than smaller programs.

Three policy recommendation can be derived from these findings. First, one might want to consider the introduction of poverty sensitive subsidies in conditional cash transfer programs instead of giving the same amount to all participants. Such a differentiated subsidy scheme is likely to make the program more efficient. However, they might also introduce additional administrative costs and jealousy among receivers, which could in turn reduce the efficiency gains.

The second recommendation concerns the eligibility threshold. This parameter turns out to be relatively important in optimizing the program. For a given size of the program, the eligibility threshold should be chosen accordingly and in case of changes in the program size its value should be adapted.

Finally, the exercise showed that whenever the goal is to increase educational mobility it might be more efficient to use parental education as eligibility criterion. However, this might also be a useful approach for programs aiming more at reducing poverty. It could, however, be argued that it would be easy for families to underreport the educational attainment. While this is true, it also applies to income or to a certain extent to family assets. The advantage of education as criterion is that there is no learning about how the program selects households. Therefore, households will not be able to hide the assets that matter more for the reevaluation. In practice, a combination of several indicators is most likely to be used and therefore the argument might simply be to give some more weight on parental education.

From a methodological point of view, this paper aims at underlying the usefulness of agent based modeling techniques in the economic literature. ABM techniques allowed me to model a rather complex phenomenon and easily include a series of empirical evidences. Including several empirical regularities at once using traditional modeling tools is likely to be intractable. Hence, the use of ABM techniques might be a way to consolidate theoretical models with empirical evidence.

This study is only a first step in modeling educational mobility with ABM techniques. I see mainly two ways for future research. It could be very interesting to complement this model of overlapping generations by a model that focuses more on the short run dynamics. The current model does not enable us to study such effect because the time unit is a generation. A second direction of future development could include the inclusion of even more empirical evidences and the calibration of the model to a very specific context, e.g. an existing social assistance program.

References

Alfonso, Mariana, "Credit Constraints and the Demand for Higher Education in Latin America," Inter-American Development Bank - Education Division - SCL, Working Paper #3 2009.

Anger, Silke and Guido Heineck, "Do smart parents raise smart children? The intergenerational transmission of cognitive abilities," *Journal of Population Economics*, 2010, *23*, pp. 1255–1282.

Ashenfelter, Orley and Alan Krueger, "Estimates of the Economic Return to Schooling from a New Sample of Twins," *American Economic Review*, 1994, *84* (5), pp. 1157–1173.

Attanasio, Orazio and Katja Kaufmann, "Educational Choices, Subjective Expectations, and Credit Constraints," NBER Working Paper 15087 2009.

Banerjee, Abhijit V. and Esther Duflo, *Poor Economics - A radical rethinking of the Way we fight global Poverty*, Public Affairs. ISBN 978-1-58648-798-0, 2011.

Behrman, Jere R. and Mark R. Rosenzweig, "Does Increasing Women's Scholing Raise the Schooling of the Next Generation?," *American Economic Review*, 2002, *92* (1), 323–334.

Björklund, Anders and Markus Jäntti, "Intergenerational income mobility and the role of family background," in Wiemer Salverda, Brian Nolan, and Timothy M. Smeeding, eds., *The Oxford Handbook of Economic Inequality*, Oxford University Press, 2009, chapter 20.

_ , Karin Hederos Eriksson, and Markus Jäntti, "IQ and Family Background: Are Associations Strong or Weak?," *The B.E. Journal of Economic Analysis & Policy*, 2010, *10* (1).

Black, Sandra E. and Paul J. Devereux, "Chapter 16 - Recent Developments in Intergenerational Mobility," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 4, Part B, Elsevier, 2011, pp. 1487 – 1541.

_ , _ , and Kjell G. Salvances, "Like Father, Like Son? A Note on the Intergenerational Transmission of IQ Scores," *Economic Letters*, 2009, *105* (1), pp. 138–140.

Breierova, Lucia and Esther Duflo, "The Impact of Education on Fertility and Child Mortality: Do Fathers Really Matter Less Than Mother?," NBER Working Paper 10513: http://www.nber.org/papers/w10513 2004.

Carneiro, Pedro and James J. Heckman, "The Evidence on Credit Constraints in Post-Secondary Schooling," *The Economic Journal*, 2002, *112* (482), 705–734.

Cleland, John and Germán Rodríguez, "The Effect of Parental Education on Marital Fertility in Developing Countries," *Population Studies: A Journal of Demography*, 1988, *42* (3), pp. 419–442.

Dahan, Momi and Alejandro Gaviria, "Sibling Correlations and Intergenerational Mobility in Latin America," *Economic Development and Cultural Change*, 2001, *49* (3), pp. 537–554.

de Hoyos, Rafael, Juan M. Martínez de la Calle, and Miguel Székely, "Educación y movilidad social en México," In: Serrano, Julio and Florencia Torche (eds): Movilidad social en México - Población, desarrollo y crecimiento. Centro de Estudios Espinosa Yglesias. Mexico City. 2010.

Drèze, Jean and Mamta Murthi, "Fertility, Education and Development: Evidence from India," *Population and Development Review*, 2001, *27* (1), pp. 33–63.

Ermisch, John, Marco Francesconi, and Thomas Siedler, "Intergenerational Mobility and Marital Sorting," *The Economic Journal*, 2006, *116* (513), pp. 659–679.

Foley, Duncan and J. Doyne Farmer, "The economy needs agent-based modelling," *Nature*, 2009, *460* (6), pp. 685–686.

Hertz, Tom, Tamara Jayasundera, Patrizio Piraino, Sibel Selcuk, Nicole Smith, and Alina Verashchagina, "The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends," *The B.E. Journal of Economic Analysis & Policy*, 2007, *7* (2), 1–46.

Machin, Stephen, "Education and Inequality," in Wiemer Salverda, Brian Nolan, and Timothy M. Smeeding, eds., *The Oxford Handbook of Economic Inequality*, Oxford University Press, 2009, chapter 17.

Mare, Robert D., "Five Decades of Educational Assortative Mating," *American Sociological Review*, 1991, *56* (1), pp.15–32.

Miller, John H. and Scott E. Page, *Complex Adaptive Systems - An introduction to computational model of social life*, Princeton University Press, 2007.

Piketty, Thomas, "Theories of Persistent Inequality and Intergenerational Mobility," in A.B Atkinson and F. Bourguignon, eds., *Hanbook of Income Distribution, Volume 1*, 2000.

Plug, Erik and Wim Vijverberg, "Does Family Income Matter for Schooling Outcomes? Using Adoptees as a Natural Experiment," *The Economic Journal*, 2005, *115*, pp. 879–906.

Psacharopoulos, George and Harry Anthony Patrinos, "Returns to Investment in Education: A Further Update," *Education Economics*, 2004, *12* (2), pp. 111–134.

Stiglitz, Joseph E., *The Price of Inequality - How Today's divided Society endangers our Future*, W. W. Norton and Company Inc. New York, 2012.

Thiessen, Del and Barbara Gregg, "Human Assortative Mating and Genetic Equilibrium: An Evolutionary Perspective," *Ethology and Sociobiology*, 1980, *1* (2), pp. 111–140.

Wendelspiess Chávez Juárez, Florian, "Intergenerational transmission of education - uncovering the mechanisms behind high intergenerational correlations," Working paper available at SSRN: http://ssrn.com/abstract=1986022. 2012.

Winter, Christoph, "Accounting for the Changing Role of Family Income in Determining College Entry," European University Institute Working Paper ECO 2007/49 2007.

Wolf, Pedro S.A. and Aurelio José Figueredo, "Fecundity, Offspring Longevity, and Assortative Mating: Parametric Tradeoffs in Sexual and Life History Strategy," *Biodemography and Social Biology*, 2011, *57* (2), pp. 171–183.

A Overview of calibration for the baseline model

Table A.1 lists the full set of parameters along with a short description. A detailed description on how these values were chosen can be found in section 3.

Table A.1: Values of the parameters used in the baseline model

Parameter	Value	Explanation
α	0.5	Parameter in the utility function: relative importance of current consumption and education of children
β_0	1	Wage equation: Base income without education
β_1	3.8	Wage equation: Scaling factor for education
β_2	2.0	Wage equation: Quadratic specification
γ_1	0.45	Education production function: Scaling factor for the wage
γ_2	0.8	Education production function: Parameter to ensure concave relationship of education and investment
γ_3	1.0	Education production function: Parameter to reduce the impact of the ability of education
ϵ_w	$\mathcal{N}(0, 2)$	Random term in the wage equation
ϵ_e	$\mathcal{N}(0, \frac{\sigma_e}{4})$	Random part in education production function
$T(W)$	0	No taxes in the baseline model
$S(i, n, W)$	0	No subsidies in the baseline model
N	1000	Initial number of agents in the model
δ_{educ}	$1.5\sigma_e$	Maximal spouse difference in education (1.5 standard deviations)
δ_{IQ}	30	Maximal spouse difference in IQ (30 IQ points)

The results of the model using these values are reported in table 6

B Robustness checks

In this section I present several robustness checks of the model. In section B.1 I show how the baseline model behaves when excluding the stochastic terms. In section A I then discuss how the model and the policy measures are affected when changing some parameters of the model. I focus mostly on the parameters that are not directly based on empirical evidence. At the end of this appendix, I discuss all the tests in general.

B.1 Exclusion of stochastic terms

Table B.1 is an extension of Table 6 in the main body of the paper and displays the reduction of key statistics when excluding the stochastic elements from the model.

Table B.1: Average statistics of full model and loss when excluding components

		Edu. corr. partner (1)	**Edu. corr. mother** **(2)**	**Edu. corr. father** **(3)**	**IQ corr. father** (4)	**IQ corr. mother** (5)	**IQ corr. partner** (6)
	Full model						
[1]	Correlation	0.673	**0.616**	**0.563**	0.327	0.373	0.450
	Standard deviation	(0.023)	**(0.050)**	**(0.053)**	(0.023)	(0.022)	(0.030)
	Excluded element:						
[2]	Stochastic wages	-0.5% *	**3.0%** **	**3.6%** **	n.s.	n.s.	n.s.
[3]	Stochastic education	0.8% **	**1.4%** **	**1.9%** **	n.s.	n.s.	n.s.

Notes: The first panel called 'Full model' displays the average values of the statistics in the full model. The second panel shows the percentual change in these statistics when excluding the element indicated in the first column. Only significant losses are reported. Significance levels: *=5%, **=1%, n.s. = not significant at 5%

The two stochastic elements in the education production function and the wage equation have only very small effects on the education correlations, as reported in rows [2] and [3]. An explanation for this is that the stochastic values introduce some more extreme values, which are then driving the correlations upwards. However, the changes are very small and the exclusion of the stochastic terms does not alter the model substantially. Put differently, the results do not directly depend on the assumptions on the stochastic terms.

B.2 Robustness on parameters

With the complexity of the model, the number possible robustness checks that can be done increase sharply. In this section I present some robustness checks by changing main parameters, especially those I was not able to estimate directly. Besides the possible values to include in the robustness checks, there are also different ways to present them. I present here the graphs showing the relationship between education and the parental education. These graphs are likely to give a more precise impression than focusing on the correlations. Table B.2 displays the parameters I am changing in the robustness checks.

Equation	Parameter	Value in baseline model	Values in robustness check
Wage equation	β_0	1.0	0.9, 1.0, 1.1
	β_2	2.0	1.8, 1.9, 2.0, 2.1, 2.2
Educ equation	γ_2	0.80	0.75,0.80,0.85
	γ_3	0.40	0.35,0.40,0.45
Utility	α	0.50	0.45,0.50, 0.55

Table B.2: Robustness checks

The presented robustness checks are all performed using income as eligibility criterion.

B.2.1 Utility function

The main parameter of the utility function that might change the analysis is α, the relative importance of current consumption compared to the education of children. In the baseline model I assume a value of 0.5. Here I compare this baseline value to the values of 0.45 (giving more weight to the education of children) and 0.55 (more weight on current consumption). Figure B.1 displays the graphs for the two robustness checks and the baseline model.

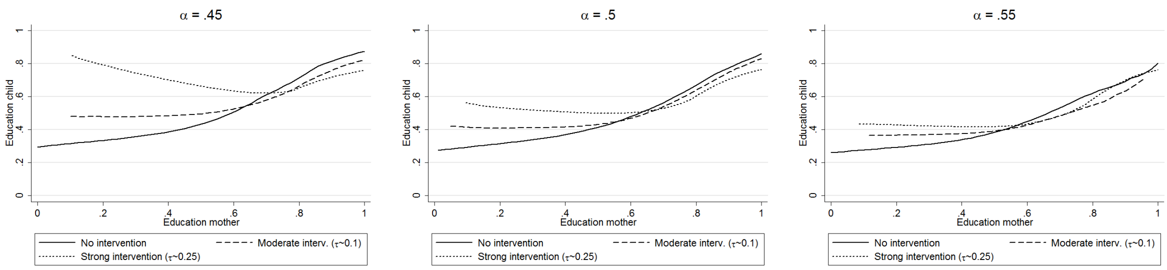


Figure B.1: Robustness checks for α

We can see that the qualitative effects of the program on the intergenerational link in education are the same. The education of children with little educated parents is increased under the program, while those at the top achieve slightly less education. However, the size of the effects changes in function of the parameter. When the weight on education in the utility function is higher, the program has more effect. This finding is reasonable in the sense that when parents care more about education, they will be more sensitive to subsidies on education.

B.2.2 Education production function

For the education production function two parameters are of interest for the robustness check. γ_2 is the sensitivity of education to ability and γ_3 shapes the investment to education. The last parameter γ_1 is not included in the robustness check as its only purpose is to keep the model stationary over several generations. Again, the quality of the results does not change

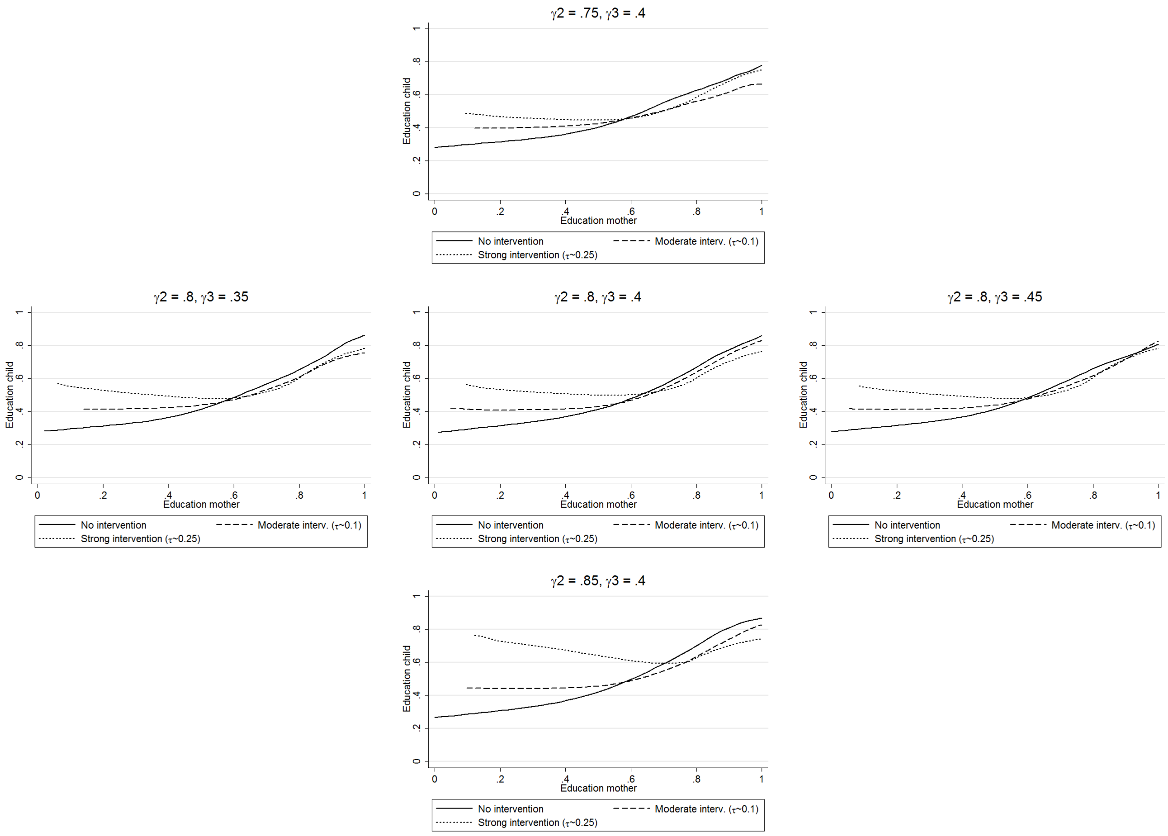


Figure B.2: Robustness checks for the education equation

with the two parameters γ_2 and γ_3. In this case, even the size of the effects remains very stable, suggesting that the model is not very sensitive to these two parameters.

B.2.3 Wage equation

For the wage equation I show the robustness check for two parameters of interest. First, I let the base income β_0 vary. Figure B.3 displays the graphs for different value of the base income β_0. The general pattern remains the same, suggesting an increased educational mobility under the

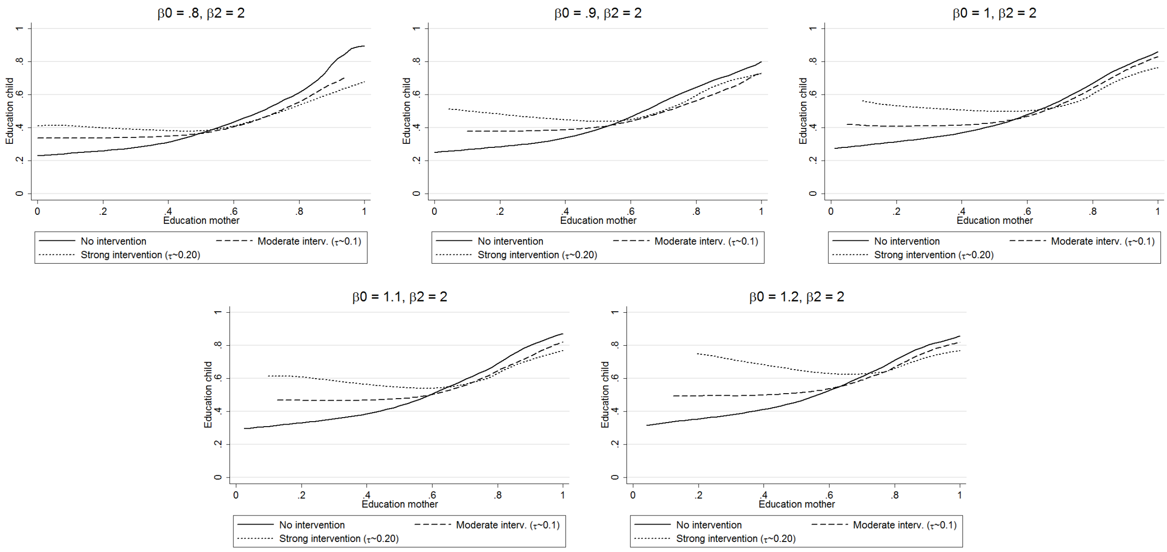


Figure B.3: Robustness checks for the wage equation

cash transfer program. The size of the program's effect, however, changes with the parameter, at least for large government interventions. For the case of the moderate intervention, the changes in the parameter do not affect the results.

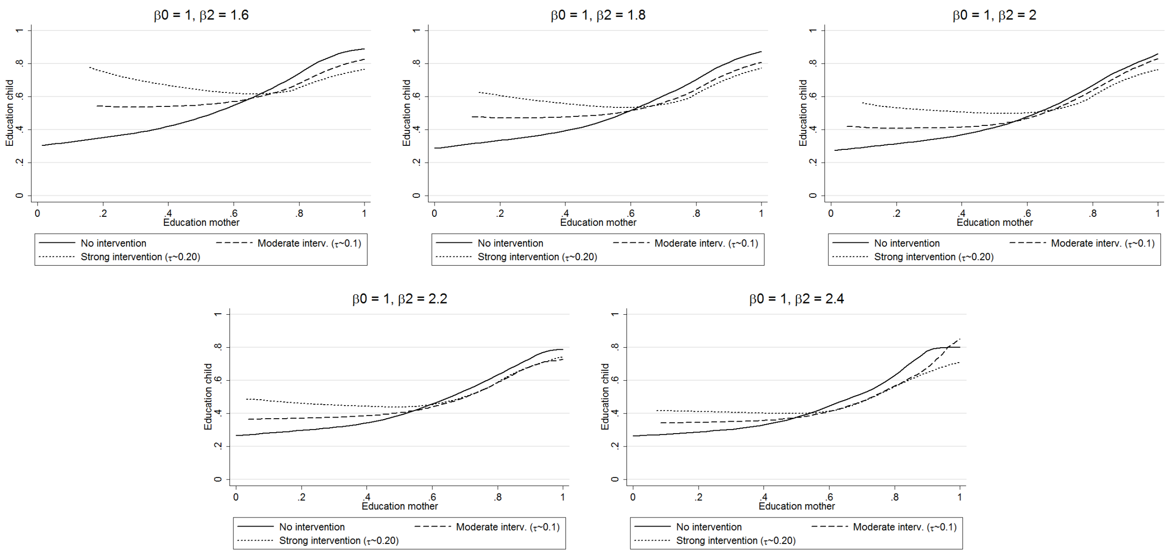


Figure B.4: Robustness checks for the wage equation

A very similar effect can be observed in Figure B.4 where I display the graphs for different values of β_2, the parameter shaping education to wages. Qualitatively the results do not change. However, the program seems to have smaller effects when the relationship between wage and education becomes more convex. This is reasonable given that the program affects mainly the lower tail of the distribution. In relative terms, the gains of the program are much less when the relationship is very convex.

B.3 General discussion on robustness checks

As a general conclusion on the robustness checks we can probably say that the qualitative results presented in section 4 and especially in section 4.2 are robust to changes in the parameter. The exact quantitative results, however, depend on the calibration. This need not be a problem of the model, as it reflects that the effects of a policy measure depend upon the context. It is reasonable to believe that in a different context (e.g. using data from another country), the parameters would be slightly different and the quantitative results as a consequence too. Note however, that the quantitative changes remain small when looking at the moderate intervention and become only substantial for the strong intervention simulated in this paper.