Intergenerational transmission of education - Uncovering the mechanisms behind high intergenerational correlations

Florian Wendelspiess Chávez Juárez

Intergenerational transmission of education - Uncovering the mechanisms behind high intergenerational correlations[*]

Florian Wendelspiess Chávez Juárez[†]

Department of Economics, University of Geneva

Abstract

Education is a main player in social mobility, however, it remains unclear through which channel the high observed intergenerational correlations of educational attainment are produced. The literature proposed very different mechanisms that could drive these correlations. This paper uses a structural equation model and data from Mexico to estimate the relative importance of the main channels simultaneously. The results sustain all of the most commonly proposed mechanisms and underline the importance to look at them simultaneously. The economic situation of a family has a large direct effect on children's schooling attainment, even when controlling for all other possible channels. Parental education has direct and indirect effects through the economic situation. Together, these two findings clearly reject the hypothesis that the educational correlations are only the fruit of a transmission of ability from one generation to the next and sustaining the idea that some inequality of opportunity is present.

Keywords: intergenerational transmission of education, social mobility, IQ transmission, inequality

JEL classification: D31, I21, I24, J62

[†] **Contact details:** Department of Economics, University of Geneva, Switzerland. florian.wendelspiess@unige.ch.

Intergenerational transmission of education - uncovering the mechanisms behind high intergenerational correlations[*]

Florian Wendelspiess Chávez Juárez[**]

October 26, 2011

Abstract

Education is a main player in social mobility, however, it remains unclear through which channel the high observed intergenerational correlations of educational attainment are produced. The literature proposed very different mechanisms that could drive these correlations. This paper uses a structural equation model and data from Mexico to estimate the relative importance of the main channels simultaneously. The results sustain all of the most commonly proposed mechanisms and underline the importance to look at them simultaneously. The economic situation of a family has a large direct effect on children's schooling attainment, even when controlling for all other possible channels. Parental education has direct and indirect effects through the economic situation. Together, these two findings clearly reject the hypothesis that the educational correlations are only the fruit of a transmission of ability from one generation to the next and sustaining the idea that some inequality of opportunity is present.

Keywords: intergenerational transmission of education, social mobility, IQ transmission, inequality

JEL-Classification: D31, I21, I24, J62

[**]Department of Economics, University of Geneva, Switzerland. florian.wendelspiess@unige.ch

1 Introduction

Economic inequality is a widespread phenomenon in societies and an important element of human development. A vast literature on inequality exists and in recent years, the dynamics of inequality became more and more analyzed. While a simple analysis of inequality in a static or cross-sectional framework provides information about the status quo, the analysis of intergenerational mobility and the transmission of inequality proposes to look more closely to the mechanisms behind inequality. Understanding the sources of inequality is key to the design of efficient inequality reducing measures, as not all kinds of inequality can be addressed with the same policy interventions.

In recent years, a rapidly growing literature deals with the numerous aspects of social mobility. While many researches focus on the intergenerational elasticities of income, other studies focus more on the processes and mechanisms behind[1]. A core element in the analysis is education, since it is generally easier to estimate than income and, more importantly, it provides insights in the mechanisms driving intergenerational earnings correlations.

The starting point is the observation of high correlations between the educational attainment of parents and their offspring. The existing literature focuses on three main mechanisms to explain these figures. First, there is a biological transmission of ability, oftentimes measured by the IQ, which explains a part of the relationship (Anger and Heineck, 2010; van Leeuwen et al., 2008; Björklund et al., 2010; Black et al., 2009). Secondly, it is generally assumed that returns to education of children with highly educated parents are higher (Black and Devereux, 2010), being a direct effect of parental education on their children's educational choice. This channel of direct education-to-education effects might also include preferences for education, aspirations and many other factors. Finally, it is oftentimes argued that poor families face credit constraints since they cannot borrow against future earnings of their offspring, which provides a third possible channel (see for instance Black and Devereux (2010); Attanasio and Kaufmann (2009); Stinebrickner and Stinebrickner (2007); Carneiro and Heckman (2002); Alfonso (2009)).

Recent literature surveys such as Black and Devereux (2010), Björklund and Jäntti (2009) or Piketty (2000) coherently argue that more empirical research must be done to understand the mechanisms behind educational mobility and social mobility in general. Black and Devereux (2010, p.69) conclude that *"[...] there is still much work to do to pin down which family background factors are most important"* and Björklund and Jäntti (2009, p.516) argue that *"a major challenge for future research is to find out what in the family other than income is important for the future of children"*.

The present study is a step in the proposed direction, incorporating simultaneously the three channels previously outlined. Using structural equation models, it is possible to distinguish the relative importance of the different channels and shed some more light on the actual mechanisms behind the observed correlations. The empirical study is based on rich data from Mexico including cognitive ability scores for both, parents and children. This country is particularly interesting to study, since educational correlations between generations are found to be especially high in Latin American Countries (LAC). To the extent of my knowledge, this is the first empirical study able to distinguish the relative importance of the channels and to estimate direct and indirect effects of parental education. It is found that even when controlling for parental education and ability, the economic situation of the family has the largest direct effect on the schooling outcome of children, suggesting an important source of inequality of opportunity one might be able to reduce with sound policy interventions.

---

[1] Recent literature surveys can be found in Black and Devereux (2010); Björklund and Jäntti (2009); Piketty (2000)

2 Theory and empirical evidence on education mobility

The transmission of educational attainment from one generation to the next is a complex mechanism and can only be described in a very unsatisfactory way by intergenerational correlations of educational outcome. Digging deeper in the mechanisms of transmission, the literature found mainly three channels of transmission, which can in turn be subdivided again. The first channel is the biological transmission of ability, the second refers to the dependence of schooling outcome on the economic situation of the parents and the third deals with the direct education-to-education effects. In this section, some theoretical and empirical contributions to the understanding of these channels are presented.

2.1 Ability transmission through genes

The direct transmission of abilities, which needn't be limited to the simple IQ transmission, represents a more biological explanation of the phenomenon. For instance, Becker and Tomes (1979) use the term *endowments acquired from parents* to describe this direct transmission. They provide a theory of intergenerational transmission based on rational choices through a human capital theory approach, where the ability level of the offspring is a key determinant of the decision. Their model was consequently extended by Loury (1981) and Solon (2004) and serves as a benchmark in many analyses.

More empirically, much work has been done to determine the importance of this channel. Numerous studies tried to estimate the importance of IQ transmission, including Anger and Heineck (2010), van Leeuwen et al. (2008), Björklund et al. (2010) and Black et al. (2009). For instance, Anger and Heineck (2010) use German panel-data with two ultra short IQ-tests to estimate the parent-offspring relation. They find that a 1-point increase in the score of parents results in a 0.45-point increase in the coding speed (inherent ability) and 0.50-point increase in word fluency scores. The estimated coefficients remain stable at the inclusion and exclusion of control variables. Björklund et al. (2010) use Swedish data from military IQ tests and official registers. They estimate intergenerational and sibling correlations to estimate the impact of father's IQ and all other family background variables on the IQ. The estimated IQ-correlations are all highly significant and attain values such as 0.346 for father-son, 0.510 for siblings and 0.65 for twins. According to the authors, their estimations represent rather a lower bound of the true values. Black et al. (2009) find a similar father-son IQ-correlation (0.38) in a comparable study with Norwegian data.

van Leeuwen et al. (2008) go even further in the analysis of IQ-transmission by dividing it further up. They find a spousal IQ-correlation of about 0.33 suggesting a relatively high degree of assortative mating. No evidence for cultural transmission of the IQ was found and no indication that intelligent parents provide children with intelligence promoting circumstances. Individual differences in intelligence were found to be largely accounted for by genetic differences.

2.2 Credit constraints and SES

The second channel is the most directly related to the topic of inequality of opportunity, as the socioeconomic status of the family directly affects the schooling attainment of children. Stiglitz (1969) does not directly consider educational mobility and focuses especially on the transmission of wealth through capital accumulation and provides arguments for determinants tending to equalize or making wealth distribution more unequal in the long run. The idea that poorer families might face credit constraints making the optimal investment in the human capital of their offspring impossible can be found for example in Banerjee and Newman (1994) and Loury (1981).

Empirical research was not able to conclude on an exact estimation of its importance. While it seems that the impact of credit constraints is relatively modest in richer countries, some evidence was found that in developing countries the effect is larger.

Stinebrickner and Stinebrickner (2007) analyze the situation at a college in the U.S. using panel data of students. They find that a group of students is credit constrained in consumption during their stay at the college, however, many of them are not willing to borrow.

Carneiro and Heckman (2002) critically revise the literature on the question of credit constraints. They compute that using modern U.S. data, only about 8% of students really face short-term credit constraints. They argue that long-term effects, being the family environment during the whole schooling period of children, play a much bigger role.

Winter (2007) elaborates a computable general equilibrium model to evaluate the role of credit constraints in the decision whether to go or not to college. The model is calibrated for the U.S. economy and predicts observed patterns quite well. The findings contrast the results of few credit constrained students found by other studies and argues that econometric estimates such as in Carneiro and Heckman (2002) are downward-biased. In line with the results of other studies is the observation that the share of people financially constrained increased (dramatically) over the past decades.

Alfonso (2009) presents a study of 4 Latin American countries (Mexico, Chile, Colombia and Peru). It is evidenced that the effect of credit constraints disappears in regression analysis, once we control for long run family variables (parental education, family assets, etc). However, the relatively small effect of credit constraints increases from the oldest to the newest datasets used in the study.

Attanasio and Kaufmann (2009) use Mexican data to analyze the relationship between post-secondary school decisions and subjective expectations. Among other findings on the role of expectations, they show that credit constraints represent an important issue for poor Mexicans, in contrast to some literature coming from higher developed countries, where these effects do not seem to be as present.

2.3   Education to education transmission

Finally a third channel considered is the more direct way that highly educated parents encourage their children to continue school more often (Merton, 1953; Boudon, 1973, 1974). Sociologists argued that the encouragement to study might depend on students' social background or that highly educated parents simply have higher preferences for education and therefore send their children to school during a longer period. This third channel is - compared to the previous two - kind of a black box, which allows many different explanations.

2.4   Relevance of educational mobility, especially in LAC

The particular importance of the parent-child schooling correlations for Latin America can be shown very impressively by the finding of Hertz et al. (2007). They compare the correlation between parental and child education of 42 countries including 7 Latin American countries. These seven countries take the first seven places when ranking them according to their intergenerational education correlation. The country with the highest (0.66) correlation is Peru, followed by Ecuador, Panama and Chile. For the remainder of the list, no clear geographical order can be found, suggesting a very special pattern of Latin American Countries in this context.

López-Calva and Macías (2010) use recent data on social mobility from Mexico to estimate the interaction of child labor, education and future income. They find that parental education and the socioeconomic status of parents play a crucial role in the schooling success of children. Moreover, they detect that the effect of child

labor on future income and social mobility is negative and important. People having worked in early years, have up to 15% less income, which concerns educational mobility since it might be reasonable to assume that child labor is essentially a phenomenon of poor families. de Hoyos et al. (2010) use the same data as López-Calva and Macías (2010), but their focus relies more on the transmission of education from one generation to the next. They report correlations between the education in years of parents and children for the case of Mexico, finding the highest correlation of about 0.6 for the children cohorts born between 1942-1951, followed by a reduction of the correlation to about 0.5 for the cohort 1962-1971 and finally a new increase of the correlation to 0.55 for the youngest cohort, being children born between 1972 and 1981. According to the same authors, this recent increase even reaches higher values than the first cohort when using different data sources. They present the matrix of educational mobility displayed in table 1, where the education of the children is related to the education of parents.

Table 1: Educational transition matrix for Mexico using MSMS data

| Father	Education of the child					
	None	**Primary**	**Secondary**	**High School**	**College**	**Graduate**
None	20.1%	**56.3**%	15.7%	5.2%	2.6%	0.0%
Primary	2.8%	**37.3**%	29.5%	18.5%	10.9%	0.9%
Secondary	0.8%	6.4%	31.8%	**34.6**%	25.6%	0.8%
High School	0.1%	6.6%	15.0%	34.6%	**38.5**%	5.2%
College	0.5%	3.9%	9.0%	24.3%	**59.9**%	2.5%
Graduate	0.0%	0.4%	0.0%	4.6%	**81.4**%	13.6%

Source: de Hoyos et al. (2010)

There is a very strong relation between the parental education and the children's schooling outcome. One can observe a slight shifting to more education, which corresponds to an overall increase in terms of education from the first to the second generation, being a positive absolute social mobility. However, the two educational distributions are strongly correlated, indicating a low degree of relative social mobility.

3 Model

3.1 Conceptual model of educational mobility

Based on the literature presented in section 2, a relatively simple conceptual framework of the transmission of education from one generation to the next is developed. The main objective of the framework is to look behind the intergenerational correlation of education and to quantify the relative importance of transmission channels. This integrated vision of the whole transmission procedure comes obviously at the cost of not being able to consider each of the channels in all its details. However, this was already largely done by the numerous studies in the literature and a few of them were mentioned in the previous section. Thus, the objective here is to put the pieces together and to investigate the relative importance. This is important because analyzing a specific channel and finding significant effects does not tell us much about the relative importance of the analyzed channel with respect to other possible channels. A channel might show very significant effects and being at the same time relatively irrelevant for the whole transmission. Wendelspiess Chávez Juárez (2011) elaborates a conceptual framework for the analysis of social mobility in general, based on ideas of the complex system theory. Figure 1 presents a strongly simplified version of the model where some factors are regrouped and only the part of educational mobility is retained.

Figure 1: Simplified conceptual system
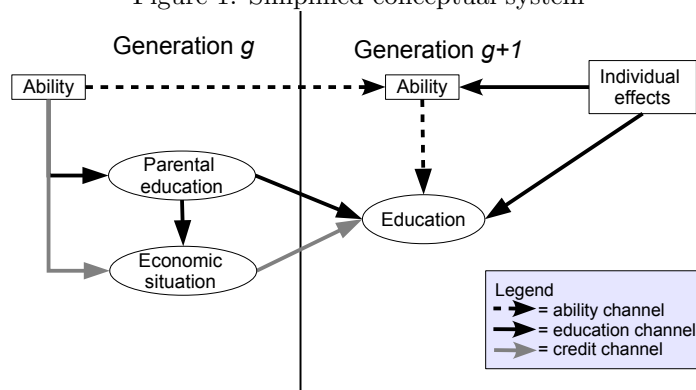
This simple illustration makes use of the three channels identified by the literature. First, there is a direct link between abilities of parents and children, presented with the dotted lines. Ability of the parents influences the ability of children, which in turn increases their propensity for education. The economic channel is represented by gray arrows using the regrouping term of *economic situation* since it might include short-term credit constraints and long run effects of assets among others. Finally, the third channel illustrated with solid black arrows represents the direct education-to-education transmission, which is based on many different hypotheses as explained earlier.

The interesting feature of presenting this framework as a system is that there are, without any doubt, very important interactions between the channels. Even though the theory normally does not exclude such interactions, it might be worthy to explicitly model them, giving us the possibility to distinguish between direct and indirect effects. Parents' ability plays a role in all channels, since it is very likely that the parental education is already the fruit of their ability and the economic situation of the family at the moment of investments in human capital of the offspring is very likely to be influenced by parental ability and education. In a very interesting article, Behrman and Rosenzweig (2002) show that intergenerational correlations between mother's and children's education might be biased when not considering such system aspects. For instance, they mention that such estimates are upward biased, when not controlling for the ability channel or for assortative mating. Considering social mobility as a system and using the related estimation techniques allows not only to avoid such biases but also to estimate the relative importance of channels and their interactions. Additionally it shows clearly that variables oftentimes considered to be exogenous, such as parental education or the economic situation, are actually endogenous and not taking this link into account might downward bias of their determinants, as the indirect effects would be neglected.

3.2 Econometric model

The econometric model is directly based on the conceptual framework presented in section 3.1, focusing on the transmission of parental circumstances on children's schooling. Mother's and father's circumstances are taken individually, since they might differ substantially. The model to estimate can be illustrated with a path diagram shown in figure 2.

The core ideas of this diagram are directly drawn from the conceptual framework and only converted in a more empirical language. Only very few variables are considered to be exogenous, while most of the parental variables are clearly identified to be influenced by others. The variable *economic situation* is considered to be a latent variable we can only observe through a couple of manifest variables. As I will show later, this variable is split

Figure 2: Path diagram of the Structural Equation Model
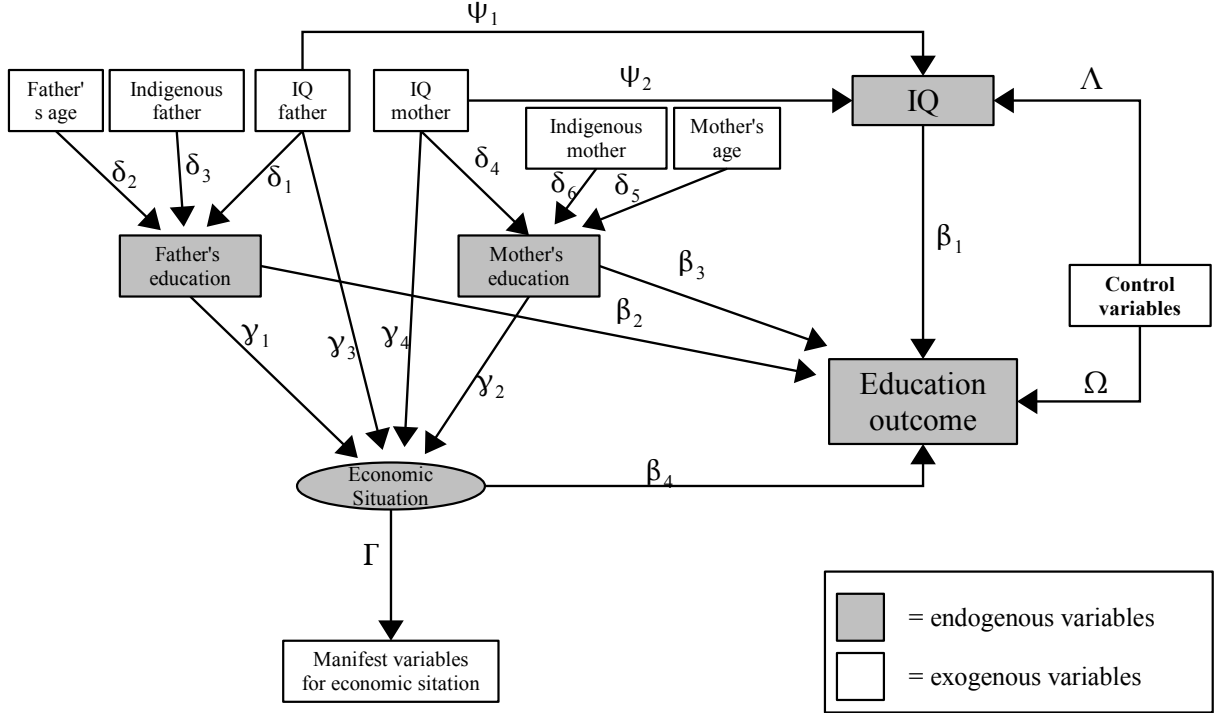
into long run and short run effects in the empirical analysis. For simplicity, these two effects are regrouped into one in figure 2. Taking deviations from the mean in order to avoid constant terms, we can write the model in equation form as follows:

$$\mathbf{L} = \mathbf{\Gamma}\text{econSR} + \varepsilon_1 \tag{1}$$

$$\mathbf{S} = \mathbf{\Gamma}\text{econLR} + \varepsilon_2 \tag{2}$$

$$\text{econLR} = \gamma_1\text{feduc} + \gamma_2\text{meduc} + \gamma_3\text{fiq} + \gamma_4\text{miq} + \varepsilon_3 \tag{3}$$

$$\text{econSR} = \gamma_5\text{feduc} + \gamma_6\text{meduc} + \gamma_7\text{fiq} + \gamma_8\text{miq} + \varepsilon_4 \tag{4}$$

$$\text{feduc} = \delta_1\text{fiq} + \delta_2\text{fage} + \delta_3\text{findi} + \varepsilon_5 \tag{5}$$

$$\text{meduc} = \delta_4\text{miq} + \delta_5\text{mage} + \delta_6\text{mindi} + \varepsilon_6 \tag{6}$$

$$IQ = \psi_1\text{fiq} + \psi_2\text{miq} + \mathbf{Z}\Lambda + \varepsilon_7 \tag{7}$$

$$schooling = \beta_1 IQ + \beta_2\text{feduc} + \beta_3\text{meduc} + \beta_4\text{econ} + \mathbf{Z}\Omega + \varepsilon_8 \tag{8}$$

where *econSR* and *econLR* are latent variables of the family short- and long run economic situation estimated from a set of indicator variables in the matrices **S** and **L**. *feduc* and *meduc* is the education of the father and the mother respectively, while *fiq* and *miq* are their respective ability measures. The variable *IQ* is the ability measure of the child and the *schooling* is an education outcome variable of the child. Finally the matrix **Z** is a matrix of exogenous control variables describing some characteristics of the child.

The use of short- and long run indicators for the economic situation is directly motivated by the findings in the literature, suggesting that short-term credit constraints oftentimes are less important than long run economic environment (Carneiro and Heckman, 2002).

The coefficients of main interest are the β's and to a lesser extend the ψ's. The former estimate the direct impact on the educational attainment of family background variables and child's ability. The latter permits to estimate

the causal relationship between parental child's ability, i.e. estimating the biological transmission. Through the φ's and β_1 the total effect of the biological transmission on the educational outcome can be estimated. This setting allows to estimate the relative importance of the different channels in the educational transmission and to distinguish channels of inequality of opportunity from biological transmissions that one cannot consider to be inequality of opportunity.

3.3 Identification

To simplify the notation, the endogenous left-hand side variables of the equations (3) to (8) can be combined in the matrix $\mathbf{Y}$ and all exogenous variables in matrix $\mathbf{X}$, which includes parental ability measures and control variables from matrix $\mathbf{Z}$. This allows rewriting the model in the standard *structural equation model* (SEM) notation:

$$\mathbf{W} = \mathbf{\Gamma}\text{econ} + \varepsilon_1 \tag{9}$$

$$\mathbf{Y} = B\mathbf{Y} + A\mathbf{X} + \xi \tag{10}$$

where ξ is a vector containing the error terms of the equations.

The covariance matrix of $\mathbf{Y}$ and $\mathbf{X}$ is given by:

$$\begin{bmatrix} \Sigma_{yy} & \Sigma_{yx} \\ \Sigma_{xy} & \Sigma_{xx} \end{bmatrix} \tag{11}$$

Since we can rewrite the matrix $\mathbf{B}$ as a lower triangular matrix with zero diagonal, the coefficients in $\mathbf{B}$ are identified (Heuchenne, 1997). Furthermore the identification of the coefficients in $\mathbf{A}$ is satisfied since Σ_{xx} is invertible and we can estimate the coefficients as follows:

$$\mathbf{A} = (\mathbf{I} - \mathbf{B})\Sigma_{yx}\Sigma_{xx}^{-1} \tag{12}$$

A stronger sufficient condition of identification is the *equivalent structure condition* saying that the model is identified, whenever any equivalent structure yields exactly to the same model. The condition states that the unique solution for the matrix $\mathbf{M}$ in the system of equations:

$$\mathbf{MA} = \mathbf{A} \tag{13}$$

$$\mathbf{MB} = \mathbf{B} \tag{14}$$

is the identity matrix, being $\mathbf{M} = \mathbf{I}$. The model presented in equation (10) satisfies this condition and is therefore non-parametrically identified.

From a more economic point of view, the variable helping to identify the model is the age of the father and the mother in their education equations. The age is here a proxy for the cohort average of education and, due to the overall increase of education over the past decades, older parents tend to have less education than younger do. On the other hand, the age does not play a direct role in the school attainment of the children. In order to avoid problems of biological transmission when parents are very young or very old, two dummy variables in the IQ equation are included, in order to capture such effects. The inclusion of indigenous dummy in parents' education production function can be justified by discrimination of indigenous people in the access to schooling. The discrimination from the time when the parents went to school can be assumed to be exogenous to the current education production of their children.

4 Data

The data used in this study come from the *Mexican Family Life Survey (MXFLS)*[2], a very rich and award winning panel data project from Mexico. At the moment when this study was elaborated, 2 waves from 2002 and 2005 were available. A particularly interesting feature of the MXFLS is the inclusion of a cognitive ability test for children and adults, based on Raven's progressive matrices[3]. Different tests were applied to children (5 to 12 years) and adults (13 to 65 years). In order to have comparable scores across age groups, the values were normalized to a distribution with mean 100 and standard deviation 15 for each age group. Given the panel structure of the survey, two test scores per person are available, allowing to compute the average score of the person in order to reduce measurement errors. Moreover, observations where the two scores had a difference of more than 25 points were dropped from the analysis. For a series of explanatory variables that should not change over time, incoherent answers were identified and corrected whenever possible . Observations where it was impossible to find the correct value were dropped from the analysis as well[4].

Parental education in years was obtained by computing the average time spent in school to achieve the reported education. Repeated years are therefore not considered as schooling years, as one can argue that they do not provide additional human capital.

The family log-consumption per capita was obtained out of a series of information consumption and normalized to the consumption per equivalent adult following the methodology proposed by Rojas (2007), who provides estimates for Mexico based on the subjective well-being approach.

The remaining variables included in the study were constructed in a straightforward way according to the standards in the literature and are reported with a short description in table 2.

A major challenge is the dependent variable, as years of schooling would not make much sense for young people still going to school and considering only people that are no longer in the schooling age would yield to two concerns. First, the schooling period of older people lies far in the past and therefore the economic situation for that time would be hard to proxy and secondly, the sample selection would no longer be random because people still studying, for instance a university degree, would not be considered.

For this reason, the dependent variable used in this study is a constructed variable based on the average schooling by age. In a first step, the average schooling in years per age group was computed and then the actual years of education of each individual was divided by the average value of its age group. This gives an indicator of how many percent of average education a child has, where the value of 1 would indicate that the child is just on average. The use of this variable permits to use a larger sample, as children from different ages are comparable. After deleting observations with missing or implausible values, the final sample contains 4855 people between 7 and 25 years. This relatively small sample size compared to the sample size of the survey might be surprising at a first glance. However, in order to estimate the simultaneous effect of both parents, the sample must be reduced to those children were both parents are observed in the data.

---

[2]The original Spanish name is *Encuesta Nacional sobre Niveles de Vida de los Hogares (ENNVIH)*

[3]More details on Raven's progressive matrices and its implementation in the MXFLS can be found in Raven et al. (1986, 1983); Rubalcava and Teruel (2006)

[4]For instance if in one wave the father was younger than the child and in the other wave the difference was plausible, then only the plausible value was taken. However, if there was no plausible value, the observation was dropped

Table 2: Variables used in the study

Abbreviation	Description	Mean	Std. dev.
Dependent outcome variable			
Educ. index	Years of education divided by the average years of education of the age group	*1.006	0.301
Endogenous regressors:			
Educ. father	Father's years of education	6.407	4.261
Educ. mother	Mother's years of education	5.879	3.762
IQ	Child's ability measure	*100.756	13.183
Exogenous regressors:			
Age father	Age of the father	44.693	8.517
Age mother	Age of the mother	41.294	7.765
IQ father	Ability measure for the father	*100.318	13.580
IQ mother	Ability measure for the mother	*97.221	13.412
Indig. father	Dummy variable for indigenous father	0.170	.
Indig. mother	Dummy variable for indigenous mother	0.143	.
Female	Dummy for girls (=1)	0.490	.
Age	Age in years	14.767	4.878
Rural	Dummy variable for rural areas	0.510	.
Latent regressors:			
Econ situation short run	Latent variable for the short run economic situation	.	.
Econ situation long run	Latent variable for the long run economic situation	.	.
Manifest variables for short run economic situation:			
Log consumption 2002	Log consumption per equivalent adult in 2002	4.522	0.771
Log consumption 2005	Log consumption per equivalent adult in 2005	4.492	0.794
Manifest variables for long run economic situation:			
Electricity	Dummy for availability of electricity	0.986	.
Clean floor	Dummy for clean floor in the house	0.829	.
Good roof	Dummy for good quality roof	0.876	.
Number of bedrooms	Number of bedrooms in the house	2.204	0.983
Phone availability	Dummy for having a phone in the household	0.312	.
Kitchen	Dummy for having a kitchen	0.932	.
Drainage	Dummy for clean drainage of feces	0.830	.
Clean garbage evac.	Dummy for clean garbage evacuation	0.641	.
Clean cooking energy	Dummy for clean cooking energy	0.815	.

Note: Descriptive statistics based on the sample of 4855 observations

* = Normalization was made with the full sample in order to use a maximum of information. The mean can deviate slightly from the normalized value due to missing values in other variables.

5   Results

5.1   Descriptive statistics

While table 2 presented some univariate descriptive statistics of the sample, I present now some simple correlations among important variables. On the one hand, they should provide a good impression of the data and outline some potentially interesting phenomena and, on the other hand, to compare them better with data used in other studies.

In the sample, the correlation of the IQ with the father's IQ is 0.398. This value is very close to the 0.347 and 0.38 estimated by Björklund et al. (2010) and Black et al. (2009) respectively. The same correlation with respect to the mother was found to be 0.348, which is slightly lower than the father-son correlation. Considering only the two oldest siblings in a family yields to a siblings IQ-correlation of 0.433, which is again relatively close to the values found by Björklund et al. (2010) who find estimates between 0.473 and 0.510.

Interestingly, and giving a first evidence for assortative mating, the spousal IQ-correlation is 0.408. The spousal education-correlation based on the years of schooling is with 0.665 even higher and supports the idea of an important role of non-random spousal selection.

Regarding the simple educational attainment correlation between parents and children, a very interesting pattern can be found when splitting the sample into age groups. Table 3 presents the correlation between the education index used in this study to proxy the educational attainment of children and their parents' actual schooling attainment.

Table 3: Intergenerational correlation of education

Age group	Sample size	Correlation with	
		Father	Mother
All (7 to 25 y.)	4855	0.2641	0.2752
7 to 10 years	1141	0.1048	0.1061
11 to 15 years	1623	0.2341	0.2556
16 to 20 years	1385	0.3709	0.4089
21 to 25 years	706	0.4894	0.4931

Generally, the correlation of the mother and the father with the children's educational index are relatively similar and follow the same increasing pattern. While for the youngest children the correlation is quite small (about 0.1) it attains almost a correlation of 0.5, which is already substantial and approaches the correlation fo 0.55 estimated by de Hoyos et al. (2010) for children born between 1972 and 1981 in Mexico.

All these correlations show that the data is comparable to previous studies and the use of the educational index might proxy reasonably well the schooling attainment generally used for older people. Moreover, the fact of having comparable values makes it also possible to consider the results found in this study to be pertinent outside the Mexican context.

5.2   Econometric results

The possibility of estimating several channels simultaneously permits not only to avoid some biases due to omitted variables, but also to quantify these biases by running regressions with some excluded variables on the same data. This idea influenced the estimation strategy and made it straightforward to estimate besides the complete model also some simplified models. In table 4, I present a first series of estimation results considering the analysis of potential biases mentioned before. Standardized coefficients are reported and should be interpreted as changes in standard deviations of the dependent variable upon a one standard deviation change in the regressors.

Model 1 is the complete model including the IQ, father's and mother's education, and the latent variable measuring the economic situation of the family as regressors of the schooling attainment. These main regressors were accompanied by control variables such as the gender and a rural area dummy, which are not reported in table 4[5]. The models 2 to 6 are based on the same estimation structure where all but the indicated coefficient were set to be zero. These auxiliary estimations have the purpose to analyze the potential bias when the information on some important variables is not available.

Considering model 1, the estimation is quite precise and all coefficients are significant at a 1%-level. The coefficient related to the child's IQ measure attains by far the highest value with 0.192. The remaining effects

---

[5]The full estimation results of model 1 can be found in table 6 in the appendix

Table 4: Estimation results of equation (8)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Short run econ. situation	0.063***	0.094***				0.067***
	(0.017)	(0.016)				(0.017)
Long run econ. situation	0.117***	0.280***				0.159***
	(0.025)	(0.019)				(0.024)
IQ	0.192***		0.275***			
	(0.014)		(0.014)			
Father's education	0.062***			0.245***		0.069***
	(0.019)			(0.014)		(0.020)
Mother's education	0.095***				0.255***	0.108***
	(0.019)				(0.014)	(0.019)
Control variables	yes	yes	yes	yes	yes	yes
R squared	0.138	0.110	0.079	0.063	0.068	0.108
N	4855	4855	4855	4855	4855	4855

Notes: Standardized coefficients. Standard deviations in parentheses.

Significance level: *** = 1%, ** = 5% and * = 10%.

attain broadly spoken between one third and half of the IQ-coefficient's size. Even when controlling for parental education, the economic situation plays an important and significant role in the schooling outcome of the children. The long-term economic situation appears to matter almost twice as much as the short run indicator, based on consumption. This finding is coherent with findings in the literature previously mentioned. When considering short- and long run economic situation, the effect becomes almost as important as the biological transmission. Mother's education matters slightly more than father's education when taking into account the direct effect.

Direct versus indirect effects

An interesting feature of simultaneous equation models is that they enable us to compute direct and indirect effects. It is clear that parental education does not only affect the schooling outcome through the direct effect discussed before, but also through the economic situation of the family. Figure 3 shows the direct and indirect effects based on the results of model 1, fully reported in table 6. As in the discussion before, one can easily see that the IQ of the child and the economic situation[6] have the most important impacts. Much more interesting is the finding regarding parental education, where, besides the direct effect discussed earlier, a relatively large indirect effect through the economic situation matters for the schooling outcome of the children. Parental IQ plays by definition only a role through indirect channels, nevertheless the total effect composed of three different indirect effects should not be neglected. This becomes even clearer when we recall the spousal IQ-correlation of about 0.4, meaning that children with more able mothers normally also have more able fathers.

Biases when neglecting channels

Models 2 to 5 only include one of the four possible channels assuming the others to have no impact. The last model includes the oftentimes available data on the education of parents and the economic situation by

---

[6]For this graphical representation the short and the long run effects were summed up.

Figure 3: Direct and indirect effects on child's schooling attainment

Direct and indirect effects on school attainment
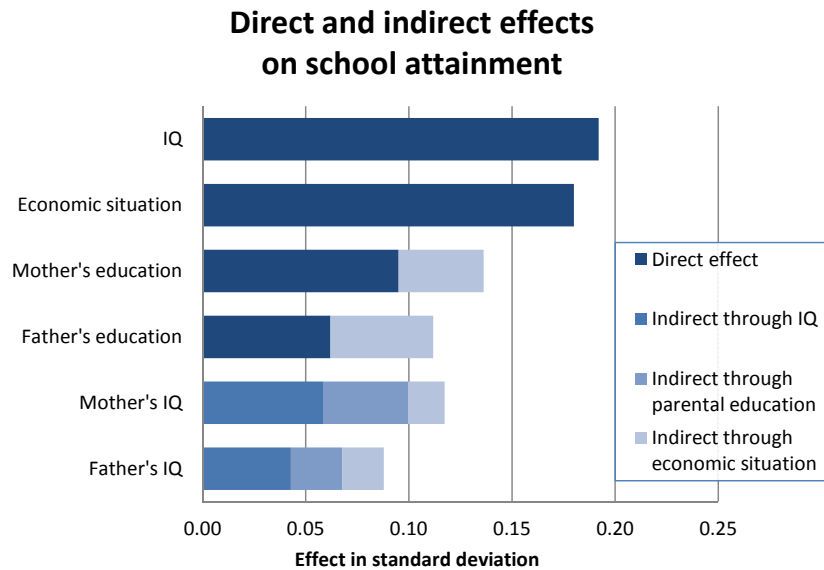
excluding the ability measure. One can easily notice that the one-covariate models always give strongly upward biased estimators of the coefficients. Not surprisingly, the bias for the IQ measure is the lowest in relative terms, but the coefficient is still 1.4 times higher than in model 1. In comparison, the coefficients attain 1.5 times the value of model 1 for the short run and 2.4 for the long run economic situation, almost 4 times for father's and 2.7 times for mother's education. These findings suggest very important biases when regressing schooling outcome on one single channel. However, the biases become much smaller, when all but the ability measure are included. Due to missing information of ability measures in most of the surveys, this setting corresponds to the best one can due oftentimes. The bias represents about 35% of the original estimation for the long run and is with 6% to 15% substantially lower for the other regressors. These estimations of the bias might give an idea about biases when data on ability is not available, although they are not of primary importance to the present study.

Regression by age groups

A second series of estimation results is presented in table 5 where model 1 is estimated for different age groups - an estimation strategy directly motivated by the descriptive findings presented in table 3.
As for the simple correlation, I find very different results according to the age groups[7]. While the model generally explains very little variation in the schooling outcome for the youngest children, the power of prediction increases with every age group. The coefficient of the IQ is significant at the 1%-level in all samples, but its size is almost three times higher for the two oldest cohorts with respect to the youngest. For the youngest cohort only the IQ has a significant effect, while the remaining variables do not explain any variance in the schooling attainment. The older the children, the more important these variables become. The short and long run economic situation attains values close to 0.15 for the oldest cohort. The importance of father's education increases for children

---

[7]The analysis based on even smaller age groups of 2 year intervals showed essentially the same pattern

Table 5: Estimation results of equation (8) for different age groups

Age group:	7 to 10	11 to 15	16 to 20	21 to 25
Short run econ. situation	-0.021	0.068**	0.106***	0.133***
	(0.038)	(0.029)	(0.027)	(0.037)
Short run econ. situation	0.098	0.140***	0.122***	0.146**
	(0.06)	(0.043)	(0.043)	(0.058)
IQ	0.099***	0.204***	0.262***	0.254***
	(0.031)	(0.026)	(0.025)	(0.033)
Father's education	0.030	0.025	0.071**	0.138***
	(0.044)	(0.033)	(0.034)	(0.048)
Mother's education	0.032	0.076**	0.178***	0.181***
	(0.043)	(0.033)	(0.032)	(0.048)
R squared	0.043	0.138	0.264	0.354
N	1141	1623	1385	706

Notes: Standardized coefficients. Significance level: *** = 1%, ** = 5% and * = 10%. Standard deviations in paretheses.

from 16 to 20 and again for the oldest cohort. The coefficients of mother's education increase sharply between the cohort of 11 to 15 year old children and the next older, while remaining at that level for the oldest cohort. The increasing importance of the regressors by age groups has several relevant impacts. Differences in the schooling of the youngest cohort are most likely due to late or no enrollment. The fact of not finding significant effects at this stage of the educational path suggests that inequality is not due to different access at the primary school level. Inequality generated by the family background variable is generated for the subsequent educational levels and becoming more important over time.

6 Conclusion

This study tries to contribute to the understanding of the mechanisms generating the high intergenerational education correlations observed all over the world and especially in Latin American Countries. A particularly important issue is to distinguish the different channels of transmission outlined by the literature over the past years. Using very rich data from Mexico a structural equation model of the educational transmission can be estimated, allowing to distinguish among the different channels, being the biological transmission of ability, the transmission through the economic situation and the education-to-education channel.

The results suggest that the economic situation of the family is the most important direct intergenerational channel. The IQ of the child has still a higher impact on the schooling outcome, but it is only partly explained by parental IQ, thus the total effect of parental ability on the schooling outcome is smaller than the effect of the economic situation. The parental education also plays an important role, however, a large part of the total effect is due to the indirect effect through the economic situation.

These results reject clearly the hypothesis that the intergenerational correlations of education are only due to the transmission of ability and not to inequality of opportunity. Given that the economic situation plays such

a crucial role, the current inequality in educational attainment is most likely not efficient, since not the most able but the economically most favored children tend to get the highest education.

Sound policy interventions targeting the link between the economic situation of people and the schooling outcome of their offspring might therefore increase educational mobility and equality of opportunity. However, the results also show clearly that there are other factors influencing educational attainment that are not necessarily socially unfair.

Finally the analysis demonstrates clearly that estimates ignoring important alternative channels of transmission tend to overestimate the effects of the analyzed variables.

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* (0), pp. 1255–1282.

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

Banerjee, Abhijit V. and Andrew F. Newman, "Poverty, Incentives, and Development," *The American Economic Review*, 1994, *84* (2), 211–215.

Becker, Gary S. and Nigel Tomes, "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," *Journal of Political Economy*, 1979, *87* (6).

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*, 2009.

_ , **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, "Recent Developments in Intergenerational Mobility," IZA Discussion Paper No. 4866. Prepared for the Handbook in Labor Economics 2010.

_ , **Paul J Devereux, 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.

Boudon, R., "L'inégalité des chances," 1973.

_ , "Education, Opportunity and Social Inequality," Wiley, New York 1974.

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

de Hoyos, Rafael, Juan M. Martínez de la Calle, and Miguel Székely, *Educación y movilidad social en México*, Centro de Estudios Espinosa Yglesias, 2010.

Hertz, Tome, 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.

Heuchenne, Christian, "A Sufficent Rule for Identification in Structural Equation Modeling Including the Null B and Recursive Rules as Extreme Cases," *Structural Equation Modeling*, 1997, *4* (3).

Loury, Glenn C., "Intergenerational transfers and the distribution of earnings," *Econmetrica*, 1981, *49* (4), 843–867.

López-Calva, Luis Felipe and Alejandra Macías, "¿Estudias o trabajas?," In: Serrano, Julio and Florencia Torche (eds): Movilidad social en México - Población, desarrollo y crecimiento. Centro de Estudios Espinosa Yglesias. Mexico City. 2010.

Merton, R., "Reference group theory and social mobility," in R. Bendix and S.M. Lipset, eds., *Class, Status and Power*, The Free Press, New York, 1953.

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

Raven, J.C., J.H Court, and J. Raven, *Manual for Raven's Progressive Matrices and Vocabulary Scales (Section 3) - Coloured Progressive Matrices (1983 edition)*, Lewis, London., 1983.

_ , _ , **and** _ , *Manual for Raven's Progressive Matrices and Vocabulary Scales (Section 2) - Coloured Progressive Matrices (1986 edition with U.S. norms)*, Lewis, London., 1986.

Rojas, Mariano, "A Subjective Well-being Equivalence Scale for Mexico: Estimation and Poverty and Income-distribution Implications," *Oxford Development Studies*, 2007, *35* (3), 273–293.

Rubalcava, Luis and Graciela Teruel, "Guía del usuario para la Primera Encuesta Nacional sobre Niveles de Vida de los Hogares," Online available at: `www.ennvih-mxfls.org` 2006.

Solon, Gary, "A Model of Intergenerational Mobility Variation over Time and Place," in Miles Corak, ed., *Generational income mobility in North America and Europe*, Cambridge: Cambridge University Press., 2004.

Stiglitz, Joseph E., "Distribution of Income and Wealth among Individuals," *Econometrica*, July 1969, *37* (3), 382–397.

Stinebrickner, Todd R. and Ralph Stinebrickner, "The Effect of Credit Constraints on the College drop-out Decision: A direct Approach Using a New Panel Study," NBER Working Paper 13340 2007.

van Leeuwen, Marieke, Stéphanie van den Berg, and Dorret Boomsma, "A twin-family study of general IQ," *Learning and Individual Differences*, 2008, *18*, pp. 76–88.

Wendelspiess Chávez Juárez, Florian, "A conceptual framework of social mobility based on complex system theory," Working Paper, University of Geneva 2011.

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

A Full estimation results

Table 6: Full estimation results of model 1

Measuring short run econ. sit. (Eq.(1))			Equation 5 (Father's educ.)		
Log consumption 2002	0.096***	0.007	Father's IQ	0.403***	0.012
Log consumption 2005	0.942***	0.007	FAGE	-0.222***	0.012
			Indigenous father	-0.153***	0.03

Measuring long run econ. sit. (Eq.(2))			Equation 6 (Mother's educ.)		
Electricity	0.498***	0.059			
Clean floor	0.778***	0.016	Mother's IQ	0.43***	0.011
Good roof	0.584***	0.022	MAGE	-0.256***	0.011
Number of bedrooms	0.272***	0.016	Indigenous mother	-0.242***	0.032
Phone availability	0.713***	0.017			
Kitchen	0.364***	0.032	Equation 7 (IQ)		
Drainage	0.629***	0.019	Father's IQ	0.223***	0.014
Clean garbage evac.	0.667***	0.016	Mother's IQ	0.306***	0.014
Clean cooking energy	0.86***	0.013	Female	-0.068***	0.026
			Old parents	-0.325***	0.072
Equation 3 (Short run econ. sit.)			Young parents	0.029	0.033
Father's education	0.185***	0.018			
Mother's education	0.257***	0.018	Equation 8 (Schooling outcome)		
Father's IQ	0.026*	0.015	Econ 1	0.063***	0.017
Mother's IQ	0.118***	0.015	Econ 2	0.117***	0.025
			IQ	0.192***	0.014
Equation 4 (Long run econ. sit.)			Father's education	0.062***	0.019
Father's education	0.326***	0.02	Mother's education	0.095***	0.019
Mother's education	0.214***	0.02	Female	0.109***	0.027
Father's IQ	0.069***	0.017	Rural	0.11***	0.031
Mother's IQ	0.203***	0.017			

Notes: The dependent variable for the structural equation is reported in parenthesis in the title of each panel. The coefficients are reported in the first column and the standard errors in the second. All coefficient are standardized, meaning that for continuous regressors the coefficients measures how many standard deviations the dependent variable changes when the regressor changes one standard deviation and for dichotomous variables when the variables turns from zero to one. Significance levels: *** = 1%, ** = 5% and * = 10%