Understanding the Income Gradient in College Attendance in Mexico: The Role of Heterogeneity in Expected Returns

Abstract

Differences in college enrollment between poor and rich are striking in Latin America. Explanations such as differences in "college preparedness" and "credit constraints" have been advanced. An alternative explanation could be differences in information sets between poor and rich, for example about career opportunities, translating into different expected returns to college. Poor people might expect low returns and thus decide not to attend. Or they might face high (unobserved) costs that prevent them from attending despite high expected returns. I use data on people's subjective expectations of returns to address this identification problem. I find that poor individuals require higher expected returns to be induced to attend college than individuals from rich families. Testing predictions of a model of college attendance shows that poor individuals are particularly responsive to changes in direct costs, which is consistent with them being credit constrained. Performing counterfactual policy experiments, I find that a sizeable fraction of poor individuals would change their decision in response to a reduction in direct costs and that these individuals at the margin have expected returns that are as high or higher than the individuals already attending college.

JEL-Classification: I21, I22, I38, O15, O16

Keywords: Schooling Choice, Credit Constraints, Subjective Expectations, Marginal Returns to Schooling, Local Instrumental Variables Approach, Mexico.

1 Introduction

Differences in college enrollment rates between poor and rich individuals are a prevalent phenomenon, but particularly striking in Latin America. In Mexico, the country which is the focus of this paper, the richest 20% represent around 60% of the student body (compared to 45% in the US), while the poorest 40% constitute only 8% (compared to 20% in the US). In addition overall college enrollment is low in Mexico.[1] These empirical facts might reflect an important welfare loss if returns to education are high, but poor people cannot take advantage of them, for example because they are credit constrained.

A traditional explanation for the income gradient in college attendance is credit constraints. Credit market imperfections are a likely scenario in the case of human capital investments given the lack of collateral (since human capital is embodied in the person) and moral hazard problems (for example in terms of work effort to repay the loan). Suppose that credit markets are indeed imperfect in that banks only lend to individuals with collateral. Since college attendance involves direct costs (such as tuition and costs of living), individuals from poor families, who are unable to cover such costs with parental income, might choose not to attend college even in the presence of high expected returns, since they are unable to borrow (or can only borrow under less favorable conditions than the rich).[2] An alternative explanation for the gradient is that it may be optimal for poor individuals not to attend college –even if they could borrow to finance higher education– because of low returns from human capital investment.

One explanation that has been neglected in this analysis consists of differences in information sets between the poor and the rich, for example about career opportunities, translating into different perceptions of individual returns to college. Conditional on their information sets poor people might expect low returns and thus decide not to attend. Or they might face high (unobserved) costs that prevent them from attending despite high expected returns. This constitutes an important identification problem, because expected returns are not directly observable.

There are two ways in which this identification problem can be addressed. The first

[1] A strong correlation between children's educational attainment and parental resources has been documented for many countries, see e.g. the cross-country overview of Blossfeldt and Shavit (1993). The correlation is particularly strong for developing countries, see e.g. Behrman, Gaviria, and Szekely (2002) on Latin America. In Online Appendix B, I compare Latin American countries and the US and OECD in terms of attendance rates, inequality in access to higher education, and availability of fellowship and student loan programs and I give detailed background information on costs and financing of college attendance in Mexico.

[2] Conventionally, an individual is defined as credit constrained if she would be willing to write a contract in which she could credibly commit to paying back the loan ("enslave herself in the case of default") taking into account the riskiness of future income streams and of default. Since such contracts are illegal, banks may choose to lend only (or to much more favorable conditions) to individuals who offer collateral.

option is to reconstruct expectations from the observable (ex-post) outcomes. This approach departs from an assumption that is a corner stone of most economic models: individuals' ex-ante expectations about their ex-post outcomes are (in an expectation sense) correct (given their ex-ante information). In this sense their expectations are rational. This implies that for the alternative that the individual chose, one can reconstruct the ex-ante expectations from the observed ex-post outcomes (and other pieces of evidence about the individual's information set). This aspect constitutes a very important advantage of the approach. The main complication associated with this approach is that the expectations regarding the consequences of the alternative(s) that has (have) not been chosen need to be constructed in some other way, since for those alternatives no ex-post outcomes are observed.

There is a large number of papers that tackle the two tasks (reconstructing the expectations for the chosen and the not chosen alternatives) under more or less restrictive assumptions. For a method that requires only very weak assumptions to resolve the two tasks, see the recent paper by Cunha, Heckman, and Navarro (2005).

The alternative approach consists of eliciting expectations directly and using them in the empirical analysis. The main advantage is that expectations data can be obtained both for the chosen and the not-chosen alternatives (i.e. for counterfactual outcome(s)). Moreover, if the goal is to understand what determines the decisions of individuals, this approach does not require any assumption with respect to the questions a) Are the expectations (in some sense) rational/correct? and b) How do individuals form expectations?

Reaping the potentially important advantages of working with expectations data is possible only to the extent to which the elicitation of expectations is feasible. While the literature has reached the consensus that it is possible to obtain meaningful measures of expectations through survey methods (see Manski (2004), Attanasio (2009) and Delavande, Giné, and McKenzie (2011) for surveys of the literature, the latter two on developing countries), it is also clear that there are limits to the amount and complexity of information that can be elicited in a survey.[3]

I make use of data on subjective expectations of returns of a sample of Mexican high school graduates to analyze the importance of information differences in explaining the income gradient in college attendance. Since what matters for the college attendance decision is an individual's perception of her own skills and how these skills (and other characteristics) affect her future earnings, these data ideally provide each individual's earnings expectations conditional on her information set at the time of the decision.

The first finding of this paper is that the expected return to college is an important

[3] I discuss the constraints that are relevant in my context in detail in Section 2.

determinant of the college attendance decision. At the same time, differences in expected returns are not sufficient to explain the differences in attendance rates between poor and rich Mexicans. Instead, data on subjective expectations allow me to show that poor individuals require significantly higher expected returns to be induced to attend college.

I then test predictions of a model of college attendance choice in the presence of credit constraints, using parental income and wealth as proxies for the household's unobserved interest rate. I find that poor individuals are particularly responsive to changes in direct costs such as tuition. This finding is consistent with the poor facing a higher interest rate. To address the concern that differences in time preferences might be driving the results, I present suggestive evidence that in my sample there are no systematic differences in discount rates between the poor and the rich.

Lastly, I evaluate potential welfare implications of policies such as governmental fellowship programs, by applying the Local Instrumental Variables approach of Heckman and Vytlacil (2005) to my model of college attendance (see also Carneiro, Heckman, and Vytlacil (2010) and Carneiro, Heckman, and Vytlacil (2011)). I find that a sizeable fraction of poor individuals would change their decision and attend in response to a reduction in direct costs. Individuals at the margin have expected returns that are as high or higher than the ones of individuals already attending college, suggesting that such policies could lead to large welfare gains.

The goal of this paper is to contribute to a growing literature investigating the role of information in schooling decisions and to the literature on credit constraints in higher education decisions. In the context of the first literature, the following three papers analyze the link between "perceived" returns to schooling and people's schooling decisions.[4]

Jensen (2010) finds that children in 8th grade in the Dominican Republic significantly underestimate returns to schooling. Informing a random subset of them about higher measured returns leads to a significant increase in perceived returns and in attained years of schooling. Nguyen (2008) finds that informing a random subset of children in Madagascar about high returns to schooling increases their attendance rates and their test scores. Attanasio and Kaufmann (2009) address several complementary issues concerning the link between schooling choice and expectations (using the same data as this paper). In addition to using expected returns -as the first two papers- they also take into account perceived earnings and employment risk. Second, they have data on mothers' expectations about

[4]The seminal paper eliciting subjective expectations of earnings for different schooling degrees is by Dominitz and Manski (1996). They illustrate for a small sample of Wisconsin high school and college students that people are willing and able to answer subjective expectations questions in a meaningful way, but do not analyze the link between earnings expectations and investment into schooling.

earnings of their children as well as adolescents' own expectations and can thus shed light on whose expectations matter for educational choices. Third, they show that schooling decisions are more sensitive to changes in expected returns for rich than for poor students, which is consistent with the existence of credit constraints, as those could break the link between expected returns (or risk perceptions) and schooling decisions. A new version of this paper (Attanasio and Kaufmann (2010)) focuses on the intra-household decision process, where data on subjective expectations are used to analyze whose expectations matter and thus who participates in the decision, the parents and/or the youth.[5]

Also the following papers investigate the role of information in schooling decisions: Bettinger, Long, Oreopoulos, and Sanbonmatsu (2009) conduct an experiment for low- and moderate-income families in the US, in which they provide aid eligibility information, while a second treatment combines the information treatment with assistance in the federal application for financial aid. Dinkelman and Martinez (2011) conduct a field experiment in Chile to investigate whether children in 8th grade from poor backgrounds increase their effort in school upon learning about financial aid options for post-secondary schooling. Stinebrickner and Stinebrickner (2012) analyze how college students from low income families in the US form expectations about their own academic ability. Their results show that learning about ability plays a very prominent role in the college drop-out decision.

In this paper it would –in principle– have been interesting to ask people not only about expected benefits to college, but also about their knowledge about costs and financial aid possibilities and about their perceptions about their academic ability. In this context it is important to stress that at the time of the survey in 2005, financial aid opportunities for post-secondary education were rare in Mexico (see Online Appendix B). While it would be interesting to have data on students' perceptions about their own ability, I make use of detailed information on past school performance to proxy for students' perceptions about own future performance. Results suggest that –while learning about ability appears to be an important determinant for the decision to drop out of school– expectations about returns to schooling are important for enrollment decisions.

This paper is also closely linked to the literature on credit constraints in educational choices. Several papers in the literature investigate the importance of credit constraints in the US, such as Cameron and Heckman (1998), Cameron and Heckman (2001), Cameron and Taber (2004) and Carneiro and Heckman (2002), and attribute differences in college attendance rates between poor and rich in the US to differences in "college readiness".

[5] Three papers that use data on subjective expectations to explain college major choices are Arcidiacono, Hotz, and Kang (2011), Stinebrickner and Stinebrickner (2013) and Zafar (2009).

Cunha (2007) finds that credit constraints at the time of deciding about college enrollment are not very important in the US (compared to college readiness), but that the inability to borrow against future income is important earlier in life, thereby affecting college readiness later on. According to Navarro (2011) ability, preferences and uncertainty all play important roles. He finds that eliminating borrowing constraints (at the same time as uncertainty), college attendance increases by roughly 8%, and that, in particular when credit constraints are defined in terms of consumption smoothing, they play a stronger role than previously found.

Most of the existing literature on credit constraints uses earnings realizations to infer expectations about earnings. The important advantage of data on subjective expectations is that (earnings) expectations can be elicited directly for all possible schooling scenarios, that is including counterfactual states. This paper shows how these data can be used in the estimation of a simple school choice model. In a different context, Mahajan and Tarozzi (2011) and Mahajan, Tarozzi, Yoong, and Blackburn (2011) study identification and estimation of key preference parameters in a model of technology adoption, when data on subjective expectations about technology's impact are available.

The following papers use alternative approaches for investigating the importance of credit constraints in higher education: Stinebrickner and Stinebrickner (2008) analyze college drop-out decisions in the US. They show that drop-out rates would remain high even if credit constraints were removed entirely, that is when excluding students who state in the survey that they would like to borrow to smooth consumption during studying but cannot. Brown, Scholz, and Seshadri (2011) base their analysis on the assumption that only children of non-altruistic parents could potentially be borrowing constrained (while assuming that parents are not constrained). The authors then exploit the fact that the amount of subsidized loans that children can receive increases in the number of siblings who are currently eligible for loans. The authors find that children who are spaced more closely together complete more years of education, but only among the subsample of non-altruistic parents, thus providing evidence of borrowing constraints for this type of families. A very different methodological approach is taken by Lochner and Monje-Naranjo (2011), who develop a human capital model with borrowing constraints explicitly derived from government student loan programs and private lending under limited commitment. Using the calibrated model, they are able to predict the observed rise in students borrowing from private lenders, as well as the persistent strong positive correlation between ability and schooling and the rising importance of family income in the US in the 80s and 90s. Lovenheim (2011) uses short-run housing wealth changes to identify the effect of housing wealth on college attendance.

This paper aims to contribute to both literatures on credit constraints and on the role of information in educational decisions by analyzing the importance of heterogeneity in expected returns to education and of credit constraints in explaining the income gradient in college attendance in Mexico. The findings of this paper suggest that credit constraints are an important driving force of Mexico's large inequalities in access to higher education and low overall enrollment rates. Mexico's low government funding for college student loans and fellowships (low even compared to other Latin American countries) around the time of my survey (2005) is consistent with this view. The results of my counterfactual policy experiments point to the possibility of large welfare gains from introducing a governmental fellowship program by removing obstacles to human capital accumulation and fostering Mexico's development and growth.

2 Model of College Attendance Choice

Studies such as Carneiro and Heckman (2002) on the US have shown that the observed correlation between parental income and children's college attendance is driven by differences in cognitive skills and parental education between the poor and the rich. I do not find this in the Mexican context. In particular, parental income and wealth remain strong predictors of children's likelihood to attend college even after controlling for an extensive list of individual and family background characteristics (including cognitive ability and parental education). Nevertheless it would be premature to conclude that this is evidence of credit constraints. Instead parental income might still capture differences in information sets between poor and rich students that could translate into differences in expected returns and thereby affect the decision to attend college. For example, a student from a poor background might think (and rationally so) that even with a college degree she will not be hired for certain jobs that someone from a richer background with parental "connections" will be hired for (even at the same level of skills). While variables such as "quality of parental network" are usually not included in the information set of the researcher, they might be contained in the individual's information set, affecting her expectations and thereby also her college attendance decision. Neglecting these factors can lead to wrong conclusions about what is driving college attendance decisions. Data on people's subjective expectations of returns to college allow me to address this concern directly.

I show formally how direct information on people's subjective expectations can be used in a simple model of college attendance. In this model I abstract from a consumption smoothing motif and simplify the college enrollment problem to one of maximizing the expected present value of earnings given an individual-specific interest (or borrowing) rate. In this context, income differences (and interest rates) matter because if an individual is rich and expects high returns to college, he/she can pay for the investment (e.g. by foregoing the interest on savings). A poor individual with high expected returns on the other hand does not have the resources to cover the direct college costs, while not being able to borrow or only at an interest rate that is too high to make the investment worthwhile. This model enables me to derive testable implications of credit constraints and to perform counterfactual policy experiments, such as evaluating the welfare implications of a governmental fellowship program.

While a dynamic educational choice model a la Attanasio, Meghir, and Santiago (2011), Todd and Wolpin (2006) and others could be interesting and insightful, data limitations –in particular having data on expected returns to schooling only in the context of a single cross-section of data on two cohorts of individuals– do not allow me to estimate a full dynamic model on all educational decisions over the whole schooling history of an individual. In this context, a simple model on college enrollment allows me to illustrate in a straightforward and transparent way how data on subjective expectations can be used to help understand education decisions and to identify the importance of credit constraints. Furthermore, the model I am using allows me to provide evidence on the importance of credit constraints *at the margin of college enrollment*, which is a relevant margin for the following reason: One relatively simple and frequently discussed policy to raise college enrollment among the poor, is to provide fellowships or student loans in order to affect individuals' decision to enroll in college. This paper's goal is to provide some evidence on whether such a policy could be effective.[6]

I model the college attendance decision of a high school graduate at age 18 as follows (compare Carneiro, Heckman, and Vytlacil (2005)): The high school graduate decides to enroll in college ($S = 1$), if the expected present value of earnings when enrolling in college (conditional on the information she has at age 18, $EPV_{18}(S = 1)$) minus the expected present value of high school earnings (again conditional on the information she has at age 18, $EPV_{18}(S = 0)$) is larger than the costs of attending college (direct costs C_i, such as tuition, transportation, room and board –if necessary– and monetized psychological costs or benefits):

$$S = 1 \Longleftrightarrow S_i^* = EPV_{18}(S = 1) - EPV_{18}(S = 0) - C_i > 0$$

If the individual decides to enroll in college, she will complete college with probability p_i^C and receive the expected present value of college earnings, $EPV_{18}(Y_i^1)$. If she drops out

[6]At the same time, the focus on the decision of college enrollment should not be read as an indication that there are no credit constraints that are relevant for individuals' decisions earlier in their schooling history.

(D), she receives $EPV_{18}(Y_i^D)$, which I assume to be equal to the expected present value of high school earnings $EPV_{18}(Y_i^0)$.

$$
\begin{aligned}
S_i^* &= p_i^C EPV_{18}(Y_i^1) + (1-p_i^C) EPV_{18}(Y_i^0) - EPV_{18}(Y_i^0) - C_i \\
&= p_i^C \sum_{a=22}^{A} \frac{E_{18}(Y_{ia}^1)}{(1+r_i)^{a-18}} - p_i^C \sum_{a=18}^{A} \frac{E_{18}(Y_{ia}^0)}{(1+r_i)^{a-18}} - C_i \geq 0,
\end{aligned}
\tag{1}
$$

where i denotes the individual, a the age of the individual, A the age at retirement. $E_{18}(Y_{ia}^1)$ represents expected earnings with a college degree, $E_{18}(Y_{ia}^0)$ expected high school earnings and r_i the interest rate that individual i faces. It is important to stress that the expectations should be taken conditional on the information that the individual has at the time of making the decision.

Before discussing in detail the assumptions of this model, I first show formally how data on subjective expectations can be used in such a model of school choice and how this compares to conventional approaches using earnings realizations.

Assume that the economic model generating the data for the two potential outcomes, that is for earnings with a high school degree ($j = 0$) and for earnings with a college degree ($j = 1$), is of the following form ("Generalized Roy Model"):

$$
\begin{aligned}
\ln Y_{ia}^j &= \alpha_j + \beta_j' X_i + \gamma_j E_{ij} + U_{ia}^j \\
&= \alpha_j + \beta_j' X_i + \gamma_j E_{ij} + \theta_j' f_i + \epsilon_{ia}^j,
\end{aligned}
\tag{2}
$$

over the whole life-cycle, $a = 18, ..., A$. In terms of observable variables a labels age, A age at retirement, E_{ij} labor market experience, and X_i denotes other observable time-invariant variables.

U^j represents the unobservables in the potential outcome equation, which are unobserved from the perspective of the researcher. They are composed of a part that is anticipated by the individual at the time of the college attendance decision, $\theta_j' f_i$, and an unanticipated part ϵ_{ia}^j, where $E(\epsilon_{ia}^j) = 0$ for $j = 0, 1$. f_i is the individual's skill vector which captures cognitive and social skills (and any other characteristics of the individual and family that affect future earnings), and θ_j is a vector of skill prices, which can vary across individuals. Both f_i and θ_j are in the information set of the individual, while they are –at least in part– unobservable for the researcher.[7] In the conventional approach using earnings realizations $\theta_j' f_i$ is unobserved,

[7] Kaufmann and Pistaferri (2009) address the issue of superior information of the individual compared to

while $\theta_j' f_i$ is implicitly 'observed' in the approach using data on subjective expectations of earnings. For each individual I have data on her expectations of earnings for age a for both potential schooling degrees, that is on the left-hand sides of the following equations:

$$
\begin{aligned}
E_{18}(\ln Y_{ia}^0) &= \alpha_0 + \beta_0' X_i + \gamma_0(a-18) + \theta_0' f_i \\
E_{18}(\ln Y_{ia}^1) &= \alpha_1 + \beta_1' X_i + \gamma_1(a-22) + \theta_1' f_i,
\end{aligned}
\tag{3}
$$

where the expected labor market experience is the number of years in the labor market, $a - s^j - 6$ (where $s^0 = 12$ and $s^1 = 16$, since high school implies 12 years of schooling and college 16 years). Beliefs about future skill prices, θ_0, θ_1, can be allowed to differ across individuals. Individuals' perceptions about their own skills enter via f_i.

Thus in my model I can allow for self-selection into schooling on unobservables, which arises from the anticipated part of the earnings, $\theta_j' f_i$, while the unanticipated ϵ_{ia}^j can obviously not be acted upon.[8] In the 'conventional' Generalized Roy model there is self-selection on U_0 and U_1 (see equation (2)) and no distinction between anticipated and unanticipated idiosyncratic returns. For example, Carneiro, Heckman, and Vytlacil (2005) analyze *ex-post* returns in a framework without uncertainty as is common in the literature. I analyze school choice under uncertainty and *ex-ante* expected returns. Subjective expectations allow me to take into account the part of the idiosyncratic returns that is anticipated and (potentially) acted upon at the time of the schooling decision.

In this framework, the individual ex-post (gross) return to college, which can obviously never be observed due to unobserved counterfactual, can be written as:

$$
\begin{aligned}
\widetilde{\rho}_{ia} &= \ln Y_{ia}^1 - \ln Y_{ia}^0 \\
&= \alpha + (\beta_1 - \beta_0)' X_i + \gamma_1 E_{i1} - \gamma_0 E_{i0} + (\theta_1 - \theta_0)' f_i + (\epsilon_{ia}^1 - \epsilon_{ia}^0),
\end{aligned}
$$

where $\alpha = (\alpha_1 - \alpha_0)$.

the researcher in the context of intertemporal consumption choices. They analyze the empirical puzzle of excess smoothness of consumption, i.e. the fact that people respond less to permanent shocks than predicted by the permanent income hypothesis. Data on people's subjective expectations of earnings allow them to disentangle two competing explanations, insurance of even very persistent shocks versus superior information of the individual compared to the researcher. They show that people respond less to permanent shocks than predicted because they anticipate part of what the researcher labels as "shocks", while the role of insurance of very persistent shocks is only minor.

[8] Compare Cunha, Heckman, and Navarro (2005) who analyze which part of idiosyncratic returns is anticipated. Subjective expectations incorporate this information, as they only include the part that is anticipated. Thus the two approaches could complement each other in learning about individuals' information sets.

Using the information given in Equation (3), I can derive an expression for the expected, i.e. ex-ante anticipated, gross return of individual i, which I can observe for each individual given my subjective expectation data:

$$\begin{aligned}\rho_{ia} &= E_{18}(\ln Y_{ia}^1 - \ln Y_{ia}^0) \\ &= \alpha + (\beta_1 - \beta_0)' X_i + \gamma_1(a-22) - \gamma_0(a-18) + (\theta_1 - \theta_0)' f_i. \end{aligned} \tag{4}$$

According to my model of college attendance (see Equation (1)), one would ideally want data on expected future earnings over the whole life-cycle of each individual. Unfortunately, I only have data on expected earnings for age 25 (see Section 3). Thus I need to make an assumption about how earnings (expectations) evolve over the life-cycle.

I model the college attendance decision based on the following assumptions:

Assumption 1 *Log earnings are additively separable in education and years of post-schooling experience. Individuals enter the labor market with zero experience and experience is increasing deterministically until retirement.*

The assumption of log earnings being additively separable in education and experience is commonly used in the literature (compare, e.g., Mincer (1974)). I assume that individuals enter the labor market –either at age $a = 18$ or at age $a = 22$ depending on the college attendance decision– with zero experience and experience is increasing deterministically until retirement. In Section 4, I discuss why the assumption that individuals do not work while studying cannot be driving my results, but would –if anything– lead to an underestimation of the role of credit constraints.

Assumption 2 *Credit constraints are modeled as unobserved heterogeneity in interest rates, r_i.*

One special case would be two different interests rates, one for the group of credit constrained individuals, r_{CC}, and one for the group of individuals that is not constrained, r_{NC}, with $r_{CC} > r_{NC}$. In the literature, heterogeneity of credit access has often been modeled as a person-specific rate of interest (see, e.g., Becker (1967), Willis and Rosen (1979) and Card (1995)). This approach has the unattractive feature that a high lifetime r implies high returns to savings after labor market entry. The testable prediction that I derive from this model (see Section 4) –that is excess responsiveness of credit-constrained individuals with respect to changes in direct costs– is robust with respect to this assumption: It can also be derived, for example, from the model of Cameron and Taber (2004), who use a similar framework, but assume that constrained individuals face higher borrowing rates than unconstrained individuals during school, while both groups face the same (lower) borrowing rate once they graduate.

Assumption 3 *Individuals are risk-neutral.*

In a framework with uncertainty this assumption implies that the decision problem of college attendance simplifies to maximizing the expected present value of earnings net of direct costs (see Carneiro, Heckman, and Vytlacil (2005)). Of course this is a strong assumption and we might be worried that the poor are more risk averse than the rich, which could explain part of the income gradient if college is risky. Interestingly, I find that individuals perceive unemployment and earnings risk to be lower with a college degree than with a high school degree (see Table 2), i.e. they believe that college insures against labor market risk. In this respect the poor should be even more likely to enroll in college than the rich, if they are more risk averse. As I will show in Section 4, perceived earnings and unemployment risk are not significant in a regression of college attendance choice (while they are significant in the decision to attend high school, see Attanasio and Kaufmann (2009), suggesting that the risk measures I use are not simply too noisy). This suggests that risk considerations might not be of first-order importance in this context, and for this reason I do not take them into account in this simple model. On the other hand, college might be more risky for the poor in other respects, for example they might be facing a higher risk of dropping out of college.

Since I do not have data on individuals' perceived risk of dropping out of college, I use performance in high school and parental education as proxies for the dropout risk. The idea to use high school performance is based on the findings of Stinebrickner and Stinebrickner (2012) who show that academic performance in college is a crucial determinant of college dropout. Since my goal is to explain the college enrollment decision, the preceding academic experience that could determine an individual's perceived dropout risk is given by the performance in high school. The educational background of the parents can be taken as measure of a student's prior for his own ability (which is updated upon observing own performance).

Assumption 4 *Individuals have a common discount factor.*

This assumption is stronger than necessary in this context, but helps to keep the model simple. The assumption needed is that the discount factor is not correlated with people's income/wealth or with the interest rate they face. Thus, in a first step, I exclude –by assumption– the possibility that the income gradient in college attendance is due to systematic differences in time preferences and use data on subjective expectations to disentangle the role of expected returns versus heterogeneity in interest rates in explaining the income gradient. In a second step, I provide empirical evidence in Section 4.4 that there are no systematic differences in time preferences between income groups.

Assumption 5 *The problem is infinite horizon.*

To estimate the model of college attendance choice (see equation (1)), I make use of the data on subjective earnings expectation using the following relationship $E(Y_{ia}) \equiv E(e^{\ln Y_{ia}}) = e^{E(\ln Y_{ia})+0.5Var(\ln Y_{ia})}$ (which holds with equality in the case of log-normally distributed earnings, which is the traditional parameterization, otherwise it is an approximation). Given the assumptions about returns to experience, I can rewrite the participation equation (1) in terms of expected gross returns to college ρ_i (see the Appendix for the derivation):

$$\begin{aligned} S_i^* &= f\left(r_i, \rho_i, C_i, E_{18}(\ln Y_{i25}^0), p_i^C, p_i^{W1}, p_i^{W0}, \sigma_i^0, \sigma_i^1\right) \\ S_i &= 1 \text{ if } S_i^* \geq 0 \\ S_i &= 0 \text{ otherwise,} \end{aligned} \tag{5}$$

where S_i is a binary variable indicating the treatment status. The decision to attend college depends upon the (unobserved) interest rate r_i, expected return ρ_i, direct costs of attendance C_i, opportunity costs $E(\ln Y_{i25}^0)$, the probability of completing college p_i^C, the probability of being employed with college and high school degree, p_i^{W1} and p_i^{W0}, and the (subjective) standard deviations of future earnings σ_i^0, σ_i^1.

Before deriving and testing implications of this model to analyze the role of credit constraints in college attendance decisions, I describe the data that I will be using.

3 Data Description

In this section I describe the data and discuss in detail the module eliciting subjective expectations of earnings and several validity checks of these data.

3.1 Survey Data

The survey "Jovenes con Oportunidades" was conducted in fall 2005 on a sample of about 23,000 15 to 25 year old adolescents in urban Mexico (compare Attanasio and Kaufmann (2009)). The sample was collected to evaluate the program "Jovenes con Oportunidades", which was introduced in 2002/03 and which gives cash incentives to individuals to attend high school and get a high school degree.

Primary sampling units are individuals, who are eligible for this program. There are three eligibility criteria: being in the last year of junior high school (9th grade) or attending high school (10 to 12th grade), being younger than 22 years of age, and being from a family that receives Oportunidades transfers.[9] Due to the last eligibility criteria the sample only comprises the poorest third of the high school graduate population. Thus even the individuals that I denote as "high" income individuals are not rich.[10] Since I analyze the college attendance decision in this paper, I restrict the sample to high school graduates, who decide to either attend college or start to work (or look for work).

The survey consists of a family questionnaire and a questionnaire for each 15 to 25 year old adolescent in the household. The data comprises detailed information on demographic characteristics of the young adults, their schooling levels and histories, their junior high school GPA, and detailed information on their parental background and the household they live in, such as parental education, earnings and income of each household member, assets of the household and transfers including remittances to and from the household. The youth questionnaire contains a section on individuals' subjective expectations of earnings as discussed in the next section.

The following important remark about the timing of the survey and the college attendance decision is necessary: One might be surprised about the fact that the following analysis – which requires knowledge of earnings expectations as well as of the actual college attendance decision– is possible with just one single cross-section. In principle I would want to have data on people's expectations at the time when they are deciding about attending college, that is some time before college starts in August or September 2005. Instead the Jovenes survey was conducted in October/November 2005 and thus two or three months after college had started.

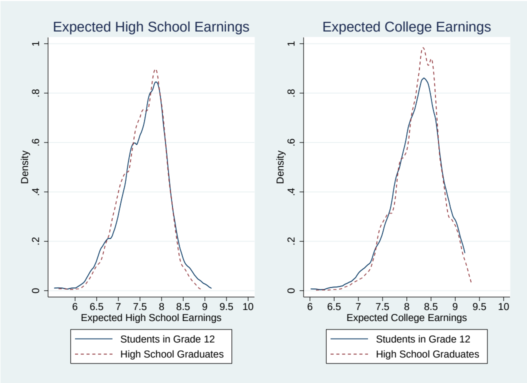
To use this survey for the following analysis I have to make the assumption that individuals' information sets have not changed during these two or three months or have changed, but left expectations about future earnings at age 25 –i.e. earnings seven years later– unchanged.

[9] Individuals can be above 18 while in high school, since some individuals start school late or have to repeat grades. Furthermore, the age of the individuals of the sample varies between 15 and 25, because the sample also includes the siblings of the primary sampling units. Lastly, individuals in grades 9 and 12 (i.e. last grade of junior and senior high school) were oversampled to study their enrollment decisions.

[10] One might wonder whether the participation in the program might have affected expectations in a way that drives my results. Since *all* individuals in my sample have participated in the program, participation is unlikely to lead to *differences* in the behavior of the very poor and richer individuals. Second, even if earnings expectations were affected in some way, this would only be relevant for the external validity of my results but not for the internal validity since my analysis conditions on earnings expectations. Of course, the external validity would depend on many other variables as well and not only on Progresa/Oportunidades' impact on expectations (e.g. on how well developed is the financial sector of a country, other institutional variables etc.).

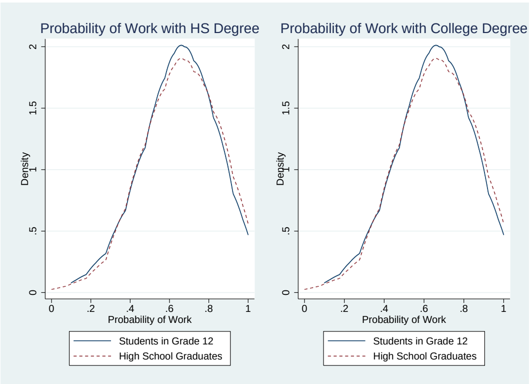
As I do not observe expectations of high school graduates *before* college starts, I perform the following consistency check of this assumption: I use the cross-section of earnings expectations of a cohort that is one grade below (just starting grade 12 in the survey months October/November) and compare it to the cross-section of expectations of my sample of high school graduates. The distributions of expected earnings (for high school and college as highest degree) do not differ significantly between the two cohorts nor do the distributions of the perceived probability of work, suggesting that expectations have not changed significantly in these three months (see Figures 1 and 2, also see Table 9 in Online Appendix B).

Figure 1: Comparing Expectations of High School Graduates with a One-Year Younger Cohort: Expected Earnings



Notes: The difference between the two cross-sectional distributions is not significant in either case (also see Appendix B, Table 9).

Figure 2: Comparing Expectations of High School Graduates with a One-Year Younger Cohort: Probability of Work



Notes: The difference between the two cross-sectional distributions is not significant in either case (also see Appendix B, Table 9).

These results can also address the following potential concern: individuals might try to rationalize their choice two or three months later, i.e. individuals, who decided to attend college, rationalize their choice by stating higher expected college earnings (or lower expected high school earnings), and those, who decided not to attend, state lower expected college and higher high school earnings. This would lead to a more dispersed cross-section of earnings after the decision.[11] A similar argument applies for the perceived probability of working. I do not find any evidence of ex-post rationalization in my data, thus lending support to my assumption.

3.2 The Subjective Distribution of Future Earnings

The subjective expectations module was designed to elicit information on the individual distribution of future earnings and the probability of working for different scenarios of highest completed schooling degree. After showing the respondent a scale from zero to one hundred to explain the concept of probabilities and going over a simple example, the following four questions on earnings expectations and employment probabilities were asked:

1. Each high school graduate was asked about the probability of working conditional on two different scenarios of highest schooling degree:
   Assume that you finish High School (College), and that this is your highest schooling degree. From zero to one hundred, how certain are you that you will be working at the age of 25?

2. The questions on subjective expectations of earnings are:
   Assume that you finish High School (College), and that this is your highest schooling degree. Assume that you have a job at age 25.

   (a) *What do you think is the maximum amount you can earn per month at that age?*

   (b) *What do you think is the minimum amount you can earn per month at that age?*

   (c) *From zero to one hundred, what is the probability that your earnings at that age will be at least x?*

   x is the midpoint between maximum and minimum amount elicited from questions (a) and (b) and was calculated by the interviewer and read to the respondent.

In the following paragraph I briefly describe how the answers to the three survey questions (2(a)-(c)) are used to compute moments of the individual earnings distributions and expected

[11] This is true unless people switch positions in the distribution in such a way that the resulting cross-section looks exactly the same as before. This can only be the case if the people who decide to enroll in college are the ones with particularly *low* expected returns, and they later report high returns to college to justify their decision. And similarly, the people who decide not to enroll in college are the ones with particularly high returns and they later state low expected returns.

gross returns to college (compare Guiso, Jappelli, and Pistaferri (2002) and Attanasio and Kaufmann (2009)). As a first step, I am interested in the individual distribution of future earnings $f(Y^S)$ for both scenarios of college attendance choice, where $S = 0$ ($S = 1$) denotes having a high school degree (college degree) as the highest degree. The survey provides information for each individual on the support of the distribution $[y^S_{min}, y^S_{max}]$ and on the probability mass to the right of the midpoint of the support, $Pr\left(Y^S > (y^S_{min} + y^S_{max})/2\right) = p$. Thus I need to make a distributional assumption, $f(\cdot)$, in order to be able to calculate moments of these individual earnings distributions. I assume a triangular distribution, which is more plausible than a stepwise uniform distribution as it puts less weight on extreme values.[12]

Thus I can calculate expected earnings $E(Y^S)$ and perceived earnings risk $Var(Y^S)$ for schooling degrees $S = 0$ and $S = 1$ for *each* individual. I will perform the following analysis in terms of log earnings, so that I compute expected log earnings as $E(\ln(Y^S)) = \int_{y^S_{min}}^{y^S_{max}} \ln(y) f_{Y^S}(y)dy$ and I can thus calculate expected (gross) returns to college as:

$$\rho \equiv E(\text{return to college}) = E(\ln(Y^1)) - E(\ln(Y^0)).$$

The module on expectations was supposed to be answered by the youths. In cases where the adolescent was not present, mothers answered also the youth questionnaire –including the questions on the subjective distribution of earnings– in addition to the household questionnaire. Attanasio and Kaufmann (2009) make use of the fact that the data contains information on parents' expectations for part of the sample and information on youths' own expectations for the rest of the sample. They analyze whose expectations are relevant for schooling decisions, the ones of the adolescent or the ones of the parents. They find that for the high school attendance decision, only mothers' expectations are important, while for the college attendance decision adolescents' expectations matter.

For this reason I use the subsample for which the adolescents answer themselves and address the concern of sample selection bias as follows (for summary statistics of the two samples, see Online Appendix B, Table 10): I correct for sample selection using a Heckman selection correction (see Heckman (1979)) applied to a non-linear context, i.e. by estimating jointly a latent index model for college attendance and a sample selection equation. As an exclusion restriction I use information on the exact date and time of the interview, which is a strongly significant determinant of whether the respondent is the adolescent. For example, adolescents are significantly more likely to be at home –and thus able to respond themselves– on weekends and during holidays (see Table 1). Results suggest that sample

[12]The first moment of the individual distribution is extremely robust with respect to the underlying distributional assumption (see Attanasio and Kaufmann (2009) for more details on the triangular distribution, alternative distributional assumptions and robustness checks).

selection on unobservables is not an important concern, as I find that the correlation between the error terms of the two equations is never significantly different from zero, once I control for individual and family background characteristics (see Section 4.3). Also, the results are similar and lead to the same conclusions when using the full sample, i.e. including the adolescents for whom the mother answers using mothers' expectations (results from the author upon request).

3.3 Validity Checks of the Data on Expected Earnings and Returns to College

In this section I compare the data on subjective expectations of future earnings to data on actual earnings and provide evidence of their value-added (for summary statistics of the variables used in the following analysis, see Table 2).

It is important to stress that the possibility that individuals might be misinformed or might have convictions that are 'distorted/biased' is not an argument *against* the use of expectations data, but instead is one of the main arguments *in their favor.* It is exactly the observation that individuals make different experiences and that they are exposed to different pieces of information, that has led economists to the conclusion that it is unlikely that they all hold the same belief. But if the beliefs of two individuals can differ substantially, then it immediately follows that it is important to control for these differences in beliefs if we want to correctly understand the differences in the observed behavior of these two individuals.[13] Therefore it is important to be able to obtain an (at least noisy) measure of the beliefs that people base their decision on. For that reason the goal of this section is to convince the reader that the high school graduates in my sample were able to understand the questions on expectations and to give meaningful answers.

First, I compare the level of earnings expectations of Mexican high school graduates to the level of contemporaneous earnings realizations using Census data of the year 2000. In particular, I compare observed high school earnings to expected high school earnings for those individuals who decided to stop school after high school. I thereby take into account that realized high school earnings are only observable for this subgroup of people (analogously for college earnings). This exercise is informative, but not a test of whether people have "correct" expectations, because the expectations are about future earnings which are only realized in the year 2012. Expected monthly high school earnings are 1940 pesos (and thus

[13]Even if poor individuals' earnings expectations were downward biased this would still not invalidate my results or conclusions since they are obtained through a comparison of poor and rich individuals who hold the *same* earnings expectations. Suppose that poor students are indeed more likely to underestimate their true potential than rich students. Then my findings tell us that *on top of this latter problem (downward biased income expectations)* poor students also face credit constraints: they are less likely to attend college than the rich even when holding the same expectations.

Table 1: Selection Equation: Probit Model for Who Responds to the Expectation Questions

Dependent Variable	**Adolescents**	**Responds**
	(1)	(2)
	Marg Eff	Marg Eff
	(SE)	(SE)
Interview Sunday	0.110*	0.092
	(0.059)	(0.061)
Interview Thursday	-0.087**	-0.089**
	(0.037)	(0.038)
Interview Thursday*Aftern.	0.079*	0.067
	(0.042)	(0.043)
Interview Saturday*Aftern.	0.106**	0.114**
	(0.052)	(0.053)
Interview Saturday*Even.	0.285***	0.336***
	(0.083)	(0.074)
Interview Week 40	0.149**	0.144**
	(0.060)	(0.061)
Interview Week 41	0.133***	0.160***
	(0.032)	(0.032)
Interview Week 42	0.112***	0.117***
	(0.028)	(0.029)
Interview Week 45	-0.053**	-0.070**
	(0.026)	(0.027)
Interview Week 46	-0.047	-0.077**
	(0.037)	(0.038)
Female		0.102***
		(0.018)
GPA - top tercile		-0.089***
		(0.021)
Father's Educ - Jr HS		-0.036
		(0.029)
Father's Educ - Sr HS		-0.005
		(0.056)
Father's Educ - Univ		-0.154
		(0.102)
Mother's Educ - Univ		0.285**
		(0.143)
Per cap Income 5 - 10k		-0.012
		(0.022)
Per cap Income more than 10k		0.017
		(0.025)
Dist to Univ 20 - 40km		0.022
		(0.022)
Dist to Univ above 40km		-0.016
		(0.026)
Tuition Above 750 Pesos		-0.011
		(0.030)
State FE	Yes	Yes
Observations	3342	3342
Log Likelihood	-2264.413	-2172.746
P-value	0.000	0.000

Notes: Table displays marginal effects and standard errors in brackets. * p<0.1 ** p<0.05 *** p<0.01. Excl. categories: Interview on Monday, Interview in the morning, Interview in week 43 (week 40 to 42 are holidays), male, lowest GPA tercile, parents' education primary or less, per capita income less than 5000 pesos. GPA second tercile and mother's education lower than university are insignificant and not displayed due to space constraints.

Table 2: Summary Statistics.

Variable	Obs	Mean	Std. Dev.	Median
Expected Return	1612	0.6670	0.3820	0.6047
Expected Log High School Earnings	1612	7.5778	0.5004	7.6432
Var of Log High School Earnings	1612	0.0054	0.0079	0.0028
Var of Log College Earnings	1612	0.0039	0.0061	0.0019
Prob of Work High School	1612	0.6657	0.1817	0.7
Prob of Work College	1612	0.8250	0.1601	0.9
College Attendance Rate	1612	0.2308	0.4215	0
Female	1612	0.5813	0.4935	1
GPA (Scale 0 to 100)	1612	82.19	7.16	82
GPA Second Tercile	1612	0.2804	0.4493	0
GPA Top Tercile	1612	0.2773	0.4478	0
Father's Yrs of Schooling	951	5.33	2.96	6
Father's Educ Jr High School	1612	0.1067	0.3088	0
Sr High School	1612	0.0292	0.1683	0
College	1612	0.0050	0.0703	0
Mother's Yrs of Schooling	1140	5.03	2.77	5
Mother's Educ Jr High School	1612	0.1234	0.3291	0
Sr High School	1612	0.0174	0.1307	0
College	1612	0.0037	0.0609	0
Per Capital Parental Income (Pesos)	1187	7519.54	8010.08	5200
Per Capita Par Inc $<$ 5000 Pesos	1612	0.5906	0.4692	1
5000 to 10000 Pesos	1612	0.2407	0.4276	0
$>$ 10000 Pesos	1612	0.1687	0.3746	0
Distance to University (km)	1612	24.2312	22.8159	18.26
Distance to Univ $<$ 20 km	1612	0.5298	0.4993	1
20 to 40 km	1612	0.2599	0.4387	0
$>$ 40 km	1612	0.2103	0.4076	0
Tuition Costs (Pesos)	1171	608.8104	634.5729	750
Tuition Costs above 750 Pesos	1612	0.4187	0.4935	0

approximately 200 US$) compared to mean observed high school earnings of 1880 pesos. Expected college earnings are larger than college earnings observed in the year 2000 (3800 versus 3300 pesos). These results are consistent with people expecting a continuation of previous trends, that is stagnating high school earnings and increasing college earnings. The implied returns –defined as the difference between log college earnings and log high school earnings– are thus around 0.65 and very similar to other studies on Mexico (see, e.g. Binelli (2008) who finds a difference of 0.64 in log hourly wages between higher and intermediate education in 2002 using ENIGH data and compare Carneiro, Heckman, and Vytlacil (2005) who find a log difference of 0.4 for the US).[14]

While the Mexican high school graduates in my sample appear to have a decent knowledge about skill prices (at least on average), there is a large amount of heterogeneity in expected earnings. Part of this heterogeneity can be explained by individual and family characteristics. Interestingly, earnings expectations vary with individual and family background characteristics in a similar way as observed earnings in Mincer earnings regressions. For example, female youths expect significantly lower earnings, while the gender gap is smaller for college than for high school earnings (as observed in the case of realized earnings), and expected college earnings are positively correlated with the GPA of the youth (see Online Appendix B).[15]

Still a considerable amount of heterogeneity in expected earnings remains, which could reflect measurement error in subjective expectations or could be due to superior information of the individual compared to the researcher, for example about her own cognitive and non-cognitive skills, about how well her parents are "connected" and help her find a job etc (compare Kaufmann and Pistaferri (2009) for evidence on superior information of people in the labor force about future income, which helps in explaining the puzzle of excess smoothness of consumption). The following result suggests that at least part of the heterogeneity in subjective expectations can be explained by heterogeneity in people's information sets: People's expectations remain an important determinant of schooling decisions even after controlling for an extensive set of individual and family background characteristics, which reflect the information set of the researcher in conventional approaches (see Section 4).

The results of this section suggest that the data on subjective expectations are an (at least

[14]Studies differ in their findings about how well informed their subjects are. For example, Jensen (2010) finds that children in grade eight in the Dominican Republic significantly underestimate returns to schooling, while I find that the earnings expectations of Mexican high school graduates are –on average– relatively close to observed earnings. In this context it is important to keep in mind that the surveyed youths in this paper have completed at least 11 years of schooling and are thus more likely to understand the probabilistic questions well than individuals with lower education levels, as in many other studies in developing countries.

[15]Also I test for behavioral biases and provide evidence that those who decide to attend college do not exert more mental effort in responding to the questions, than those who decide not to go to college (results from the author upon request). As I have shown in Section 3.1 there is no evidence that people justify their schooling decisions ex-post.

noisy) measure of the beliefs that people base their decisions on and thereby help to bridge the usual differences in information sets between the researcher and the individuals that are studied. This points towards an important value-added of data on subjective expectations for our understanding of people's schooling decisions.

3.4 Data on Educational Costs

According to the model of college attendance choice (see Section 2) direct costs of attending college should be an important determinant of college attendance decisions in addition to expected earnings. In Mexico these costs pocket a large fraction of parental income for relatively poor families, as will be shown below. Thus they might play an important role in explaining low college attendance rates of the poor.

I collected data on the two most important cost factors, enrollment and tuition costs and costs of living. As costs of living during college depend heavily on the accessibility of universities, I use distance to college as a proxy (compare, e.g., Card (1995) and Cameron and Taber (2004) who use a dummy for whether there is a college in the same country). In my sample the majority of people who decide to go to college are indeed enrolled in the college closest to them (85% go to the college in their own municipality, 95% in their own state). Thus distance appears like a good measure of direct costs in my context.

For example, if an adolescent lives far away from the closest university, she will have to move to a different city and pay room and board. She thus has to incur important additional costs compared to someone who can live with her family during college. I collected information on the location of higher education institutions offering four-year undergraduate degrees and computed the actual distance between these institutions and the adolescent's locality of residence.[16] About half of the adolescents live within a distance of 20 kilometers to the closest university, which might permit a daily commute with public transportation. Twenty-five percent live within 20 to 40 kilometers distance, while the other quarter lives more than 40 kilometers away (see summary statistics in Table 2).

In terms of (yearly) tuition and enrollment fees I use administrative data from the National Association of Universities and Institutes of Higher Education (ANUIES). I determine the locality with universities that is closest to the adolescent's locality of residence and use the lowest tuition fee of all the universities in this locality as my cost measure. The median

[16]I use information on the location of public and private universities and technical institutes offering undergraduate degrees from the Department of Public Education (SEP, Secretaria de Educacion Publica - Subsecretaria Educacion Superior). I extracted geo-code information of all adolescents' localities of residence (around 1300) and of all localities with at least one university –in the states of my sample and in all neighboring states– from a web page provided by INEGI (National Institute of Statistics, Geography and Information). My special thanks to Shaun McRae who helped extracting these data.

tuition fee is 750 pesos (see Table 2).[17] This is equivalent to 15% of median per capita parental income in my sample, while it only represents a fraction of total college attendance costs. Thus college attendance would imply a substantial financial burden for poor families.

To analyze if the ability to finance college costs plays a major role in explaining the income gradient in college attendance, I need proxies for unobserved financing costs (reflected by the interest rate in my model, see Section 2). Financing costs depend mainly on parental income and wealth, which determine the availability of resources, the ability to collateralize and receive loans, and at what interest rate to receive loans or forego savings.

The survey provides detailed information on income of each household member, savings if existent, durable goods and remittances. I create the following two measures: per capita parental income and an index of parental income and wealth.[18] Median yearly per capita income is 5200 pesos (approximately 520 US$). I use these two measures, per capita parental income and an index of parental income and wealth, as proxies for the (unobserved) interest rate that the household faces when testing implications of borrowing constraints in my model of educational choices.

I use per capita parental income as a measure of the resources available to the youth, since in the standard framework, siblings compete for limited resources within the household, so that an increase in the number of children decreases average child investment (see, e.g., Becker and Lewis (1973)). On the other hand, in particular in developing countries, it is not uncommon that older siblings contribute to household resources that are used to invest in the education of their younger siblings. Therefore I show that using measures of total family income (and wealth) leads to very similar results (see Online Appendix B).

As the relationship between income/wealth and the interest rate that families face might not be linear, I use dummies for different categories of per capita parental income with the following income thresholds: twice and four times the minimum monthly salary (equivalent to around 5,000 and 10,000 pesos). These thresholds correspond to official thresholds that determine eligibility to government programs, such as fellowship programs, where families with per capital income below twice the minimum wage are classified as most in need and given priority, while families are still eligible with income up to four times the minimum wage. Again it is important to point out that fellowships and student loans played a very limited role for higher education in Mexico around the time of my survey: only 5% of the undergraduate student population received a fellowship in 2004, while about 2% benefited

[17]Unfortunately, the measure of tuition costs is missing for nearly a third of the sample. I include these missing observations in the excluded category of the dummy of tuition costs to avoid small sample sizes.

[18]Per capita parental income includes parents' labor earnings, other income sources such as rent, profits from a business, pension income etc. and remittances, divided by family size. The index of parental income and wealth is created by a principle component analysis of per capita income, value of durable goods and savings. Only a very selective and richer group of households saves or borrows: 4% of households have savings, while 5% borrow.

from a student loan (for further details on the system of higher education in Mexico, see Online Appendix B). With this classification 59% of the sample fall into this first category of income below 5,000 pesos and 24% have per capita income between 5,000 and 10,000 pesos, as shown in Table 2. For robustness I also use an index of parental income and wealth as a second proxy for the interest rate that a family faces and include this measure using quartiles, since the index does not have a natural unit of measurement.

4 What Explains the Income Gradient in College Enrollment?

In this section I analyze what explains the large differences in college enrollment rates between poor and rich Mexicans. In particular I am interested in distinguishing between the following two explanations: Data on individuals' expectations allow me to analyze if differences in expected (monetary) returns (or perceived risks) between the poor and the rich explain the gap in college enrollment. In that case I need to investigate further if poor Mexicans rationally expect lower returns than the rich (e.g. due to lower quality primary and secondary education, the family being less well "connected" etc) or if they *under*estimate their potential returns to college education or *over*estimate risks (e.g. they are not informed about certain career opportunities with a college degree).

If on the other hand, the poor expect similar returns as the rich, but require higher expected returns to be induced to attend, then they have to be facing higher direct costs of schooling (where costs are defined broadly as including, for example, tuition costs and psychological costs or benefits from college) or higher borrowing costs. To understand the role of different cost components, it is important to model the decision to enroll in college in dependence of all those potential determinants.

4.1 The Income Gradient and Expected Returns

The first exercise is to analyze if parental income is correlated with college attendance, *only* because it picks up differences in how much individuals can benefit from going to college. To address this issue, conventional approaches control for "long-run factors" such as parental education and individual ability to proxy for these benefits. I then add controls for individuals' expectations about their potential returns to college (and their perceptions about unemployment and earnings risk) to control in a more direct way for monetary returns to college and to allow for differences in information sets between the poor and the rich.

Table 3 shows that individuals' expected returns are an important predictor for the decision to enroll in college, even after controlling for an extensive set of individual and family background characteristics. The perceived probabilities of working and perceived earnings risk on the other hand are not significant (while Attanasio and Kaufmann (2010) find that these measures are relevant for the decision to enroll in senior high school). As higher-ability youths expect higher returns to college, the coefficient on the expected return to college becomes slightly smaller after controlling for youths' GPA and parents' education. In as far as higher ability affects college attendance via higher expected returns, one should not control for ability separately. The reasons for controlling for GPA and parental education are to control for the perceived probability of the youth to complete college, and second, to control for differences in tastes for education (psychological costs/benefits).

The main conclusion of Table 3 is that enrollment gaps between the poor and the rich remain even after controlling not only for conventional "long-run factors" such as ability and parental education, but also for the return to college that the individual expects (and thus for potential information differences between income groups).

Before I give some back-of-the-envelope calculations to quantify the potential importance of credit constraints, I first discuss three important reasons for why it is well possible that I underestimate their role. The first and most important reason is related to the sample on which my analysis is based. In particular, my sample comprises roughly the poorest third of the Mexican population, since all households are recipients of Progresa/Oportunidades. Thus also among the "richest" income group in my sample (i.e. the somewhat less poor), there might be individuals who are credit constrained. As my analysis is based on a comparison of the enrollment rates between income groups, it can only give an idea of how many *more* individuals are constrained among the poor compared to the somewhat less poor.

A second and third reason for why my analysis leads to a lower bound on the role of credit constraints are related to the meaning of the elicited expectations.

In Mexico, like in other countries, universities vary in their quality and consequently also in their tuition fees. Unfortunately I do not know which quality people had in mind when answering the expectation questions. If individuals already take into account budget constraints when stating their expectations, then the poor and the richer might state expectations related to different types of colleges. The poor may report lower expected returns than an equally smart richer individual, because the latter has in mind a more expensive higher quality college. Even though the poor individual is constrained since she does not consider the more expensive high quality university, in my analysis the individual would be considered as "not constrained", that is low expected returns would explain the low enrollment rate of these individuals.

A third reason is related to the fact that students might consider working while studying. Working while studying would imply c.p. that the individual either takes longer to complete his studies (and thus receives college earnings one year later), which he would be unlikely

Table 3: Probit Model of the College Attendance Decision.

Dependent Variable	College Attendance		
	(1)	(2)	(3)
	Marg Eff	Marg Eff	Marg Eff
	(SE)	(SE)	(SE)
Expected Return to College	0.092***	0.078**	0.077**
	(0.033)	(0.034)	(0.034)
Prob of Work - Sr HS	0.032	0.013	0.005
	(0.087)	(0.085)	(0.081)
Prob of Work - College	-0.008	-0.001	0.032
	(0.101)	(0.099)	(0.092)
Var of Log Earn - Sr HS	-2.625	-3.016	-2.959
	(1.919)	(2.008)	(1.958)
Var of Log Earn - College	-0.310	0.036	0.196
	(2.351)	(2.291)	(2.164)
Female	-0.055*	-0.059*	-0.046
	(0.029)	(0.033)	(0.032)
GPA - Second Tercile		0.055*	0.055*
		(0.031)	(0.031)
GPA - Top Tercile		0.187***	0.174***
		(0.038)	(0.045)
Father's Educ - Jr HS		0.099**	0.073*
		(0.042)	(0.042)
Father's Educ - Sr HS		0.151*	0.100
		(0.078)	(0.075)
Father's Educ - Univ		0.547***	0.574***
		(0.120)	(0.131)
Mother's Educ - Jr HS		0.100**	0.074*
		(0.040)	(0.039)
Mother's Educ - Sr HS		0.203**	0.173*
		(0.099)	(0.101)
Per Cap Income 5 - 10k			0.051*
			(0.031)
Per Cap Income $\geq$ 10k			0.119***
			(0.037)
Dist to Univ 20 - 40km			-0.076***
			(0.029)
Dist to Univ $\geq$ 40km			-0.106***
			(0.031)
Tuition $\geq$ 750 Pesos			-0.082**
			(0.039)
State FE	Yes	Yes	Yes
Observations	3342	3342	3342
Uncensored Obs	1612	1612	1612
Log Likelihood	-3041.971	-2990.349	-2972.964
Sample Sel: Corr betw Err	-0.487	-0.282	-0.131
Sample Sel: P-Val	0.055	0.314	0.654

Notes: Table displays marginal effects and standard errors in brackets. * $p<0.1$ ** $p<0.05$ *** $p<0.01$. Excl. categories: male, lowest GPA tercile, father's and mother's education primary or less (mother's education university not displayed, as not significant due to small number of obs), per capita income less than 5000 pesos, distance to university less than 20 km and tuition less than 750 pesos.

to do unless credit constrained, or the individual would have less time to study per course and therefore perform less well, in which case he is likely to graduate with worse grades and should therefore expect lower earnings.[19] This implies that an individual who anticipates that he has to work to fund himself will either expect lower college earnings (because of worse grades and/or studying longer, which sends a bad signal) and this would explain his low likelihood of college attendance, i.e. I would classify the individual as not credit constrained, while he should be classified as constrained. On the other hand, the individual might state expectations for the ideal case of going to college without having to work, but then does not go to college since he cannot borrow, in which case I would correctly classify him as constrained.

Keeping in mind these three observations, I give a back-of-the-envelope estimate of the importance of credit constraints. I follow the analysis of Carneiro and Heckman (2002), who regress the college attendance decision on ability (AFQT test score) and other "long-run" factors and use the coefficients on parental income quartiles to estimate what fraction is credit constrained. In particular, they compute a (weighted) average of the gaps in enrollment between highest and lower income quartiles. Of course, in this exercise, the fraction that is defined as credit constrained crucially depends on the enrollment rates of the highest income group, since enrollment gaps are determined by a comparison with the latter. Therefore, my estimate can only give an idea of how many *more* individuals are constrained among the poorer compared to the slightly less poor.

When conducting this exercise based on the coefficients of the parental income categories in Column 3 of Table 3, I find that –in my sample– among the poorer income groups 8% more individuals are credit constrained than among the highest (or least poor) income group.[20] To put this figure into perspective, the enrollment of the "highest" income group in my sample is about 33%. When taking into account the full Mexican population, people in the highest income quartile display college enrollment rates of 67% (see Online Appendix B, Figure 7).

4.2 Differences in Expected Returns between Poor and Rich

Having shown in the previous section, that family resources still matter for the likelihood to enroll in college –even after controlling for expected returns– and that an important fraction of individuals might be credit constrained in their college attendance choice, I show in this

[19]One might argue that working while studying leads individuals to enter the labor market with job experience, which could be rewarded in terms of higher earnings. At the same time, individuals who have to work to support themselves while studying usually work in lower-quality jobs, such as working at McDonalds, where the job experience is unlikely to be rewarded in terms of higher college earnings.

[20]This figure is based on the difference between lowest and highest income group, which is 11.9% where the low income group makes up 59% of my sample, and on the difference between middle and highest income group which is 5.1% with a weight of 24%.

section that poor individuals require significantly higher expected (monetary) returns that the rich to be induced to attend college. Data on people's subjective earnings expectations allow me to conduct this exercise without any further assumptions, since I have information on the expected return of every individual, while otherwise returns are unobservable at the individual level.

I estimate the probability of college enrollment conditional on expected returns $Pr(S = 1|\rho = \tilde{\rho})$ by performing Fan's (1992) locally weighted linear regression of college attendance S on the expected return ρ.[21] I perform this analysis for different income categories, that is for "low", "middle" and "high" income individuals (yearly per capita income less than 5,000 pesos, between 5,000 and 10,000 pesos and more than 10,000 pesos, where the thresholds correspond to twice and four times the minimum wage, see Section 3.4). I calculate point-wise confidence intervals applying a bootstrap procedure.

Figure 3: The Cumulative Distribution Function of Costs for Different Income Classes.

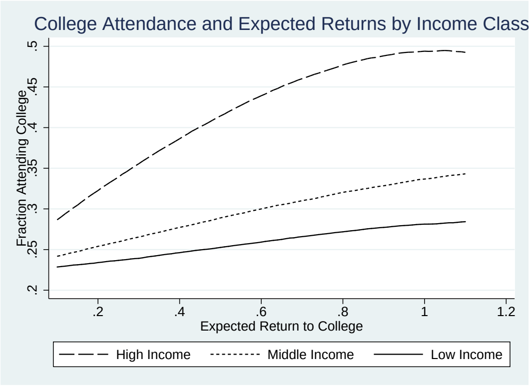


Figure 3 shows that poor individuals require significantly higher expected returns to be induced to attend college than the rich, as the c.d.f. of costs is shifted to the right for poorer individuals. Among individuals with expected returns of around $\rho = 0.6$ (which is equal to the median gross return defined as the difference between expected log college and high school earnings, see Section 3.3), 45% of rich individuals attend, but only 25% of the poor. Poor individuals thus require higher expected returns to be induced to attend college. These differences are significant (see Online Appendix B, Figure 6). In this context it is important to keep in mind that the individuals I call "rich" in my sample are still relatively poor (below the median income in society), as my sample only comprises families that are Oportunidades beneficiaries. Thus we would expect even larger differences when comparing the poor to truly rich individuals.

[21] I use a Gaussian kernel and a bandwidth of 0.3. A smaller bandwidth will lead to a more wiggly line, while the result of a significant right shift in the c.d.f. of costs for poorer individuals remains unchanged. Note that the c.d.f. of costs can only be estimated over the support of the expected return.

4.3 Testable Implications of a Model with Credit Constraints and Empirical Results

I have shown that differences in expected returns alone cannot explain the income gradient in college enrollment in Mexico. Instead poorer individuals require higher expected returns to be induced to attend college, which implies that they have to be facing higher costs of college attendance (where costs are broadly defined as including direct costs, such as tuition and psychological costs/benefits, and borrowing costs). For this reason, I make use of the model of college attendance choice introduced in Section 2, which allows for a potential role of credit constraints, while being able to take into account people's expectations about returns and controlling for differences in direct costs. As discussed, credit constraints are captured by heterogeneity in the interest rate that people face.

To understand whether credit constraints play an important role in driving low enrollment rates of poor Mexicans, I derive the following testable implications of credit constraints from my model of college attendance choice. The model implies that individuals who face a high interest rate r react more strongly to changes in direct costs C (see Equation (19) in the Appendix):

$$\left|\frac{\partial P(S=1)}{\partial C}\right| \quad \text{is increasing in} \quad r. \tag{6}$$

Intuitively, an increase in costs has to be financed through a loan (or foregone savings) with interest rate r. The negative impact of a cost increase is thus larger for people who face a large interest rate.

I test this prediction using dummies for groups that are likely to face different interest rates if credit constraints are important, that is I use dummies of parental income (and wealth). Thus I test for excess responsiveness of poor individuals with respect to changes in direct costs, such as tuition costs and distance to college.

The prediction of excess responsiveness of credit constrained groups to changes in direct costs is not specific to my model. This prediction can be derived from a more general class of school choice models, such as for example from the model of Cameron and Taber (2004). They have more general assumptions concerning heterogeneity in interest rate (see Section 2), i.e. they allow for r to be different between credit constrained and unconstrained individuals during school while r is the same for both groups after school. Cameron and Taber (2004), Card (1995) and Kling (2001) use a similar test interacting variables such as parental income and race with a dummy for the presence of a college in the residential county.[22]

[22] Card (1995) and Kling (2001) find evidence of important credit constraints for an older cohort of the National Longitudinal Survey (NLS Young Men), while Cameron and Taber (2004) do not find evidence of credit constraints for the US using the NLSY 1979. This is consistent with increased availability of fellowships and loans in the US over the relevant time period.

Compared to conventional approaches, data on subjective expectations provide the following two advantages: First, I can control directly for people's expectations about their potential returns to college and thereby avoid biased estimates that could arise from omitting this determinant. This makes my test more robust and enables me to analyze the validity of the test used without controlling for people's expectations. Second, being poor does not necessarily imply being credit constrained: only poor individuals with high expected returns are potentially prevented from attending college due to high financing costs, as they are the ones likely to be close to the margin of indifference ($S^* = 0$). Poor low-return individuals on the other hand would not attend college anyways. Thus with information on expected returns I can refine the test and test for excess responsiveness of poor *high-expected-return* individuals to changes in direct costs.

The first cost measure that I use is distance of the adolescent's home to the closest university (see Section 3.4). As shown in the previous section, living further away from the closest university has a significantly negative effect on the probability to attend college. Table 4 illustrates that the negative effect of a larger distance is particularly strong for poor individuals as predicted by the model in the presence of credit constraints. Living 20 to 40 kilometers away from college instead of less than 20 kilometers decreases the probability of attending by about 9 percentage points for the poorest income category and this negative effect is significantly larger for the poor than for the rich (p-value 0.07). Increasing the distance to more than 40 kilometers has a large effect for the middle income category, but the coefficients for the different income categories are not significantly different from each other. In this context, it is important to keep in mind that credit constraints are identified by comparing the poorest individuals to the richer individuals in my sample, who are themselves relatively poor. This could explain why –in the case of a high cost shock– all income groups are similarly responsive.[23]

The conclusions remain unchanged when I use different proxies for being credit constrained, that is quartiles of an indicator of parental income and wealth and measures of total family income/wealth (see Online Appendix B).

In terms of the second cost measure I use yearly tuition and enrollment fees. In particular I use a dummy for tuition costs above 750 pesos (the median), which is equivalent to 15% of median yearly per capita income and thus represents an important financial burden for poor individuals. The first two columns of Table 5 would suggest that tuition costs do not have an effect on attendance, that is the coefficient for the poor is negative and positive for the rich, but neither of the coefficients is significant. At the same time, the difference

[23] A comparison between the first and second column of Table 4 shows that including measures of expectations does not change the results (with the exception that the coefficients on the dummies for distance become slightly more negative for poor and middle income families).

Table 4: Excess Responsiveness of the Poor to Changes in Direct Costs (Distance to College).

Dependent Variable	**College Attendance**		
	(1)	(2)	(3)
	Marg Eff	Marg Eff	Marg Eff
	(SE)	(SE)	(SE)
Univ 20 - 40km * Par Income < 5k	-0.089**	-0.092**	-0.076
	(0.044)	(0.044)	(0.059)
Univ 20 - 40km * Par Income < 5k * High Exp Ret			-0.067
			(0.083)
Univ 20 - 40km * Par Income 5 - 10k	-0.044	-0.049	-0.041
	(0.054)	(0.054)	(0.078)
Univ 20 - 40km * Par Income 5 - 10k * High Exp Ret			-0.022
			(0.109)
Univ 20 - 40km * Par Income > 10k	0.053	0.048	0.062
	(0.071)	(0.070)	(0.099)
Univ 20 - 40km * Par Income > 10k * High Exp Ret			-0.026
			(0.119)
Univ > 40km * Par Income < 5k	-0.048	-0.051	-0.041
	(0.043)	(0.043)	(0.058)
Univ > 40km * Par Income < 5k * High Exp Ret			-0.064
			(0.081)
Univ > 40km * Par Income 5 - 10k	-0.136***	-0.145***	-0.160**
	(0.051)	(0.050)	(0.069)
Univ > 40km * Par Income 5 - 10k * High Exp Ret			0.046
			(0.152)
Univ > 40km * Par Income > 10k	-0.045	-0.047	-0.152**
	(0.071)	(0.072)	(0.075)
Univ > 40km * Par Income > 10k * High Exp Ret			0.292
			(0.200)
Par Income < 5k * High Exp Ret			0.088
			(0.059)
Par Income 5 - 10k * High Exp Ret			0.184**
			(0.084)
Par Income > 10k * High Exp Ret			0.132
			(0.093)
Controls for Expected Return, Exp Log Earn, Prob of Work and Var of Log Earn	No	Yes	Yes
Controls: GPA, Par Income and Educ, Sex, State FE	Yes	Yes	Yes
Observations	3342	3342	3342
Uncensored Obs	1612	1612	1612
Log Likelihood	-2984.591	-2971.787	-2965.898
Sample Sel: Corr betw Err	-0.167	-0.172	-0.112
Sample Sel: P-Val	0.569	0.556	0.709

Notes: Table displays marginal effects and standard errors in brackets. * p<0.1, ** p<0.05, *** p<0.01. Excl. categories: male, lowest GPA tercile, parents' education primary or less, per capita income less than 5000 pesos, interactions of distance to university of less than 20km with parental income and low expected return interacted with parental (per capita) income.

between the coefficients of poor and rich is significant, i.e. they are differentially responsive to a cost increase. Once I take into account that what matters is being poor *and* having high expected returns, results become even more pronounced: Poor individuals with high expected returns, defined as returns above the median, are excess responsive with respect to a change in tuition costs. An increase in tuition to more than 750 pesos reduced the likelihood to attend by 12 percentage points for poor high-return individuals. The negative effect of an increase in costs is significantly larger for the poor than for the rich (p-value 0.09). The same picture arises using quartiles of the parental income and wealth indicator or when using total family income/wealth (see Online Appendix B). For individuals in the lowest income/wealth quartile with high expected returns an increase in tuition costs reduces their likelihood to attend by about 15 percentage points (significantly larger in absolute value than for the top quartile, with a p-value of 0.08).

While I found that an increase in costs in terms of distance has the largest effect on the poor –as predicted by my model–, I also want to investigate if the negative effect of an increase in distance to college is larger for poor *high-return* individuals. The results point in a similar direction but are less clear-cut when including the triple interaction (which slices the already small sample even further): Using income to proxy for the interest rate, the negative effect of an increase in distance is larger for high-return poor than for the average poor (the coefficient doubles), but the difference is not significant (see Table 4). Results are similar for parental income and wealth or total family income/wealth (see Online Appendix B).

To sum up, results of this section are consistent with the predictions of a model with credit constraints. At the same time one might still be worried that the results might be driven by the poor having a higher discount rate than the rich. I will investigate this issue in the next section.

4.4 Differences in Time Preferences between Poor and Rich

The goal of this section is to address the concern that results are driven by the poor having a higher discount rate than the rich (instead of facing a higher interest rate). For this purpose, I make use of survey questions on health-related variables associated with making tradeoffs between the present and future. In particular, I use questions on smoking and drinking alcohol. The literature on time preferences suggests that there is an important correlation between time preferences and health-related variables. For example, a study by Chabris, Laibson, Morris, Schuldt, and Taubinsky (2008) finds that the discount rate is significantly correlated with health-related variables such as body-mass index, exercise and smoking, and that it can explain 15 to 20 percent of the variation (across people) in each of these

Table 5: Excess Responsiveness of the Poor to Changes in Direct Costs (Tuition Costs).

Dependent Variable	College Attendance		
	(1)	(2)	(3)
	Marg Eff	Marg Eff	Marg Eff
	(SE)	(SE)	(SE)
Tuition > 750 * Par Income < 5k	-0.043	-0.052	-0.010
	(0.040)	(0.040)	(0.058)
Tuition > 750 * Par Income < 5k * High Exp Ret			-0.124*
			(0.064)
Tuition > 750 * Par Income 5 - 10k	-0.013	-0.021	-0.053
	(0.055)	(0.055)	(0.075)
Tuition > 750 * Par Income 5 - 10k * High Exp Ret			0.039
			(0.108)
Tuition > 750 * Par Income > 10k	0.073	0.069	0.042
	(0.069)	(0.070)	(0.102)
Tuition > 750 * Par Income > 10k * High Exp Ret			0.021
			(0.127)
Par Income < 5k * High Exp Ret			0.099
			(0.062)
Par Income 5 - 10k * High Exp Ret			0.149*
			(0.086)
Par Income > 10k * High Exp Ret			0.131
			(0.093)
Controls for Expected Return, Exp Log Earn, Prob of Work and Var of Log Earn	No	Yes	Yes
Controls: GPA, Par Income and Educ, Sex, State FE	Yes	Yes	Yes
Observations	3342	3342	3342
Uncensored Obs	1612	1612	1612
Log Likelihood	-2987.347	-2975.075	-2969.499
Sample Sel: Corr betw Err	-0.268	-0.305	-0.280
Sample Sel: P-Val	0.358	0.297	0.310

Notes: Table displays marginal effects and standard errors in brackets. * p<0.1, ** p<0.05, *** p<0.01. Excl. categories: male, lowest GPA tercile, parents' education primary or less, per capita income less than 5000 pesos, interactions of tuition costs less than 750 pesos with parental income and low expected return interacted with parental (per capita) income.

measures, while no other variable explains as much of the variation as the discount rate. For other studies that rely on smoking as proxies for time preferences, see the survey article by Grossman (2000) (also see Khwaja, Silverman, and Sloan (2007)).

To provide suggestive evidence on whether differences in time preferences between the poor and the rich might be driving my results, I make use of the following survey questions on smoking and drinking alcohol: "Do you currently smoke?",[24] "Do you drink (even if occasionally)?" and "On average, how many beers, cooler, via real, glasses of wine, brandy, mezcal etc. do you drink in a normal week?" (the last question was asked to those who answered "yes" to the previous question). In addition I use a question on how the youths would make use of 3000 pesos (around 300$), if they had this amount available in that moment, that is whether they would use it for immediate consumption or to save/invest (e.g. in education).

In Table 6, I compare the answers to the three survey questions on health-related variables and to the survey question on the usage of 3000 pesos for youths of different income groups. I find that 3% of the individuals smoke, irrespective of the income category they belong to. In terms of drinking alcohol, 12% of the poorest and the middle-income group state "yes" versus 17% of the richer income group (the difference between the poor and the rich is significant on 5%), i.e. rich youths are more likely to drink. To exclude those that occasionally have a drink, I also create a dummy for whether an individual has more than 2 drinks per week (in a normal week) and find that 4% of the poor and the rich have –on average– more than two drinks per week, compared to 3% of the middle-income category (differences are not significant). Answering to the question "If you had 3000 pesos now, what would you do with the money?", 17% of the poor state that they would use the money for immediate consumption instead of saving/investing the money. Among the middle income group 19% would use the money for immediate consumption and among the rich 22% would use the money for consumption. The differences between the poor and the rich are significant on 10%. These findings are hard to reconcile with the poor being more impatient than the rich.[25]

To sum up, the results in this section suggest that the poor are not more impatient than richer individuals in my sample. Thus differences in discount rates between poor and rich

[24] Of those who currently smoke, 94% had started smoking before age 18.

[25] The income groups used are equivalent to the ones I have used in the previous analysis and the thresholds correspond to two and four times the minimum monthly salary (see Section 3.4). If –instead of comparing income categories– I compare different income/wealth quartiles, I again find very similar results, none of which suggests that the poor are more impatient than the rich. Also, if I regress the variables "smoking", "drinking any alcohol", "having more than two drinks" and "using 3000 pesos for consumption" on measures of parental income/wealth while controlling for individual characteristics such as age, gender and state of residence, I do not find any significant correlation (see Online Appendix B).

Table 6: Time Preference of Different Per Capita Income Categories.

	Per Capita Income Category				
	$\leq$ 5000	5-10000	$\geq$ 10000	Compare	
	(1)	(2)	(3)	(1)-(2)	(1)-(3)
	Mean	Mean	Mean	Diff	Diff
	(SD)	(SD)	(SD)	(P-Val)	(P-Val)
Intertemp Behavior: Health					
Smoke	0.03	0.03	0.03	-0.00	0.00
	(0.17)	(0.18)	(0.16)	(0.693)	(0.749)
Drink Alcohol					
Yes	0.12	0.12	0.17	0.00	-0.05
	(0.32)	(0.32)	(0.37)	(0.889)	(0.043)
$\geq$ 2/week	0.04	0.03	0.04	0.01	0.00
	(0.20)	(0.18)	(0.19)	(0.420)	(0.647)
How Use 3000 Pesos?					
Immediate Consumption	0.17	0.19	0.22	-0.02	-0.05
(Alternative: Save/Invest)	(0.38)	(0.40)	(0.41)	(0.362)	(0.093)
Observations	952	388	272		

Notes: Columns 1 to 3 display means and standard deviations in brackets. Columns 4 and 5 display the difference of (1)-(2) and (1)-(3), respectively, and the p-value of the difference in brackets.

do not seem to explain the income gradient.[26]

5 Counterfactual Policy Experiments

In the previous section I have shown that poor people require significantly higher returns to be induced to enroll in college. Furthermore, I have shown that poor (high-expected-return) individuals are most sensitive to changes in direct costs, which is consistent with credit constraints affecting college attendance decisions of poor Mexicans with high expected returns. I have provided suggestive evidence that these results are not driven by differences in time preferences between the poor and the rich in my sample. Thus my results point towards the importance of credit constraints in college attendance decisions of poor Mexicans.

For this reason, I evaluate potential welfare implications of the introduction of a fellowship program that can be means-tested or performance-based. I perform counterfactual policy experiments by applying the Local Instrumental Variables methodology of Heckman and Vytlacil (2005) to my model of college attendance making use of data on subjective

[26]Including the variables “smoking”, “drinking” and “money usage” in my regressions analysis, I find that both “smoking” and “drinking” have a (large) negative coefficient but are not significant (smoking is close to significant in some specifications). A dummy for “using the money for immediate consumption” has a strong and significant negative effect on the decision to enroll in college. Including those variables in my analysis does not change the qualitative results and quantitatively the results become stronger (results available from the author upon request).

expectations of earnings. I estimate the fraction of people changing their decisions in response to a reduction in direct costs, and derive the expected returns of those individuals ("marginal" expected returns).

The comparison between "marginal" expected returns (of individuals who switch participation in response to a policy) and average expected returns of individuals attending college is interesting not only from a policy-evaluation point of view. If "marginal" expected returns are higher than expected returns of individuals who attend college, then individuals at the margin have to be facing particularly high unobserved costs, as they would otherwise also be attending college given their high expected returns.

One word of caution is necessary before describing the counterfactual policy experiments. As argued in this paper, data on people's subjective expectations can be very useful for understanding people's behavior, as the data appear to capture the beliefs that people base their decisions on (compare Section 3.3). For the welfare analysis on the other hand, one would like to know people's actual returns, which are never observed. Given that people seem to have a good understanding of their potential earnings (see Section 3.3) and most likely have a better knowledge of their own skills, people's expectations might be relatively realistic. Nevertheless it is very hard to evaluate the rationality of expectations. Thus the policy-experiments should be taken with caution in terms of quantitative evaluation of the welfare benefits, and seen more as an additional piece of evidence concerning the importance of borrowing constraints, as explained below.

The idea of the third test of credit constraints comparing marginal returns to returns of those attending school is directly linked to Card's interpretation of the finding that in many studies instrumental variable (IV) estimates of the return to schooling exceed ordinary least squares (OLS) estimates (see Card (2001)). Since IV can be interpreted as estimating the return for individuals induced to change their schooling status by the selected instrument, finding higher returns for "switchers" suggests that these individuals face higher marginal costs of schooling. In other words, Card's interpretation is that "marginal returns to education among the low-education subgroups typically affected by supply-side innovations tend to be relatively high, reflecting their high marginal costs of schooling, rather than low ability that limits their return to education."

This argument has two problems in terms of how the idea was implemented (compare Carneiro and Heckman (2002)) and one more fundamental problem in terms of assumptions about people's information sets. I will argue how these problems can be addressed using data on subjective expectations. In terms of the implementation, the validity of many of the instruments used in this literature has been questioned, thus challenging the IV results.[27]

[27] Carneiro and Heckman (2002) show for several commonly used instruments using the NLSY that they are either correlated with observed ability measures, such as AFQT, or uncorrelated with schooling.

Second, even granting the validity of the instruments, the IV-OLS evidence is consistent with models of self-selection or comparative advantage in the labor market even in the absence of credit constraints. The problem is that ordinary least squares does not necessarily estimate the average return of those individuals who attend college, $E(\beta|S=1) \equiv E(\ln Y_1 - \ln Y_0|S=1)$, which would be the correct comparison group to test for credit constraints. Rather OLS identifies $E(\ln Y_1|S=1) - E(\ln Y_0|S=0)$, which could be larger or smaller than $E(\beta|S=1)$.[28]

Data on subjective expectations allow me to directly test the validity of the "instrument" that I will be using to compute marginal returns and perform policy experiments: In contrast to the situation with earnings realizations, subjective expectations are asked for both possible states of highest potential schooling degree, which implies that I also have data on "counterfactual earnings". Therefore I can compute expected returns for each individual and test if returns are orthogonal to distance to college, which is the instrument that I will be using. With data on each individual's expected return I can also directly address the second problem of implementation: I can directly compute the average (expected) return of the adolescents who attend college and I do not have to rely on OLS. Therefore I can compare marginal returns with returns of the individuals who chose to attend in the spirit of Card's interpretation of the IV-OLS comparison.

Even if this test could be implemented with data on earnings realizations alone, the following fundamental problem concerning people's information sets would remain: People at the margin might have –ex-post– higher returns than those who attend. But these people might have decided not to attend because they expected low returns *ex-ante*. As argued before data on people's subjective expectations permit to relax the rational expectations assumption with strong requirements on coinciding information sets of individuals and the researcher.

I can test the validity of the instrument used here, by regressing expected returns on polynomials of distance to college and tuition costs in the first column (in addition to observable characteristics of the individual and her family background) and on the dummies I use for distance and tuition costs, and find that neither the coefficients on distance to college nor on tuition costs are significantly different from zero (adding further polynomials does not change the results) (see Online Appendix B).

[28] $E(\ln Y_1|S=1) - E(\ln Y_0|S=0) = E(\beta|S=1) + (E(\ln Y_0|S=1) - E(\ln Y_0|S=0))$, where the last bracket could be larger or smaller than zero. In particular, in the case of comparative advantage, the OLS estimate will be smaller than the average return of those attending. This could lead to a case in which IV estimates are larger than OLS estimates, but smaller than the average return of those attending, from which one would wrongly conclude that credit constraints are important.

5.1 Implications of Credit Constraints for Marginal Returns to College

From the latent index model (see Equation (5)), I can derive the return at which an individual is exactly indifferent between attending college or not, in which case $S^* = 0$:

An individual is indifferent between attending college or not at the following -implicitly defined- "marginal" return, ρ^M,

$$S_i^* = f\left(r_i, \rho_i^M, C_i, E(\ln Y_{i25}^0), p_i^C, p_i^{W1}, p_i^{W0}, \sigma_i^0, \sigma_i^1\right) = 0 \tag{7}$$

The presence of credit constraints has the following implication for marginal returns: implicit differentiation of equation (7) leads to:

$$\frac{d\rho_i^M}{dr_i} = -\frac{\partial f/\partial r_i}{\partial f/\partial \rho_i^M} > 0,$$

and thus credit constrained individuals, who face higher borrowing costs, $r_{CC} > r_{NC}$, have higher marginal returns than those individuals on the margin who are not credit constrained:

$$\rho^M(r_{CC}) > \rho^M(r_{NC}).$$

5.2 Derivation of the Marginal Return to College

To provide further evidence on the importance of credit constraints (by comparing expected returns of people at the margin of attending –"marginal returns to college"– to the return of those already attending) and to introduce a framework to perform counterfactual policy experiments, I show how the "Local Instrumental Variable" (LIV) methodology by Heckman and Vytlacil (2005) can be applied to my model of college attendance and data on subjective expectations of earnings (see also Carneiro, Heckman, and Vytlacil (2011) and Carneiro, Heckman, and Vytlacil (2010)).[29]

First, I show how I can derive a selection equation from my school choice model (see Section 2), which is characterized by heterogeneity in the unobserved interest rate r. The propensity score can then be estimated from this selection equation. Second, I show how the predicted value of the propensity score is used in the estimation of the "marginal returns to college" (or "marginal treatment effect", MTE).

The LIV methodology relies critically on the assumption that the selection equation has a representation in additively separable form, $S^* = \mu(Z) - U_S$ (see, e.g., Heckman and Vytlacil

[29] In Online Appendix B, I give a brief introduction to the derivation of the "marginal treatment effect" (MTE) and of the "policy-relevant treatment effect" (PRTE).

(2005) and Heckman, Vytlacil, and Urzua (2006)). In general, this is not the case in a school choice model with credit constraints. In my case, data on subjective expectations allow me to write the selection equation in additively separable form despite unobserved heterogeneity in interest rates, as I will show below. The key assumption is that all unobserved heterogeneity stems from the interest rate, while parental education, youths' ability, distance to college and tuition costs are sufficient to control for direct costs.

Under this assumption, the selection equation as derived from my model can be expressed as a fourth-order polynomial in the unobservable interest rate, $1 + r$ (see the Appendix for the derivation):

$$S_i^* \geq 0 \Leftrightarrow (1 + r_i)^4 - A(Z_i;\theta)(1 + r_i)^3 - B(Z_i;\theta) \leq 0, \tag{8}$$

where $A(Z_i;\theta), B(Z_i;\theta) > 0$ are functions of $Z_i = \left(\rho_i, C_i, E(\ln Y^0), p_i^{W1}, p_i^{W0}, p_i^C, \sigma_i^0, \sigma_i^1\right)$ including the expected return ρ_i from the data on subjective expectations, and a coefficient vector θ. One can show that this fourth-order polynomial equation has exactly one positive root with $1 + r_i \geq 0$, which can be analytically computed, so that the following holds:[30]

$$g\left(Z_i;\theta\right) \geq 1 + r_i \Rightarrow (1 + r_i)^4 - A(Z_i;\theta)(1 + r_i)^3 - B(Z_i;\theta) \leq 0.$$

The selection equation can thus be rewritten in the following additively separable form (defining V_i as deviations from the mean interest rate $r_i = \bar{r} + V_i$):

$$\begin{aligned} S_i^* &= -(1 + \bar{r}) + g\left(Z_i;\theta\right) - V_i \\ S_i &= 1 \text{ if } S_i^* \geq 0 \\ S_i &= 0 \text{ otherwise.} \end{aligned} \tag{9}$$

I estimate the propensity score $P(Z)$ using a Maximum Likelihood procedure. I can then define the values u_S over which the marginal return to college (MTE) can be identified with the help of the predicted values of the propensity score: The MTE is defined for values of $\widehat{P(z)}$, for which one obtains positive frequencies for both subsamples $S = 0$ and $S = 1$ (i.e. observations outside of the support are dropped).

As a second step in the derivation of the marginal return to college one can show that

[30]The intuition is as follows: We are interested in whether the function $f(x) = x^4 - ax^3 - b$ has exactly one root on the positive real line (which is the relevant range for the interest rate), i.e. for $x \geq 0$. For values of x smaller than or equal to a the function is negative, as $f(x) = x^3(x - a) - b < 0$ if $x \leq a$. For values of x larger than a, the function is always increasing ($f'(a) = a^3$ and $f'(x) = 4x^3 - 3ax^2 > 0$ for $x \geq a$) and the slope is bounded below by a^3 ($f''(x) = 6x(2x - a) > 0$ for $x \geq a$), so there is exactly one positive root.

the MTE can be written as follows:

$$\Delta^{MTE}(u_S) \equiv E\left(lnY_{it}^1 - lnY_{it}^0 | U_S = p\right) = \left.\frac{\partial\left\{\int_0^p E\left(lnY_{it}^1 - lnY_{it}^0 | U_S = p\right) dU_S\right\}}{\partial p}\right|_{p=u_S} = \left.\frac{\partial m(p)}{\partial p}\right|_{p=u_S}$$

where the integral in the numerator can be rewritten as (see the Appendix):

$$\begin{aligned} m(p) \equiv \int_0^p E\left(lnY_{it}^1 - lnY_{it}^0 | U_S = p\right) dU_S &= pE\left(lnY_{it}^1 - lnY_{it}^0 | U_S \leq p\right) \\ &= pE\left(lnY_{it}^1 - lnY_{it}^0 | P(Z) = p, S = 1\right). \end{aligned} \tag{10}$$

With subjective expectations of earnings one has data on each individual's expectation of earnings in both schooling states ($E(lnY_{it}^1 - lnY_{it}^0)$). In addition I can use the predicted value of the propensity score, $\widehat{P(z)} = p$, for each individual, which I calculated in the first step after estimating $P(Z)$, and I have data on the actual school choice S. Thus I can compute $m(p)$.

Finally I estimate the $\Delta^{MTE}(u_S) = \frac{\partial m(p)}{\partial p}$ for different values of $p = u_S$, by fitting a nonparametric regression of $m(p)$ on the propensity score using a locally weighted regression approach (Fan (1992)). The predicted value of this regression at p is then the estimated value of the regression function at the grid point, i.e., $\hat{m}(p) = \hat{\beta}_0(p) + \hat{\beta}_1(p)p$. $\hat{\beta}_1(p)$ is a natural estimator of the slope of the regression function at p and thus estimates the MTE for different values of $p = u_S$. I calculate standard errors by applying a bootstrap over the whole procedure described in this section (including estimation and prediction of $P(Z)$).

In a third step, I make use of the estimated MTE to conduct policy experiments –such as evaluating the introduction of fellowships– by estimating the "policy-relevant treatment effect" or $PRTE$ (again I calculate standard errors of the $PRTE$ by applying a bootstrap around the procedure described above, including the computation of the $PRTE$).

5.3 Estimation of the Marginal Return to College

This section describes the estimation of the marginal return to college, and discusses the empirical results of this estimation, while the next section discusses the results of the policy experiments.

First I estimate the propensity score from a reduced-form version of the participation equation (9) using a Maximum Likelihood procedure (compare Carneiro, Heckman, and Vytlacil (2011)). In order to empirically implement the notion of costs, C, I use the following auxiliary regression containing dummies for the distance to the closest university, a dummy for tuition costs above 750 pesos, and state fixed effects to capture differences in direct costs. To proxy for preferences (i.e. psychological costs or benefits from college) and for the probability of completing college, I include parents' education and past school performance (GPA of junior high school). The results of the Maximum Likelihood Estimation of the propensity score are displayed in Tables 4 and 5 and discussed in Section 4.3.

Second, I determine the relevant support for the MTE by estimating the density of the predicted probability of attending college. I compare the density for high school graduates, who decided to attend college ($S = 1$), with the one of those, who stopped school after high school ($S = 0$), using smoothed sample histograms. The probability of attending college is generally relatively low for adolescents of the Jovenes sample, but there is a right-shift in the density for high school graduates, who decided to attend college. Their mean (median) probability is about 34% (32%), while the mean (median) probability of attending for those who stopped is around 26% (24%) (also see Figure 7 in Online Appendix B). Since there is little mass outside of the interval $[0.08, 0.67]$, I estimate the marginal return to college over the support of $p \in [0.08, 0.67]$.

Third, I estimate the MTE. I estimate a series of locally weighted regressions on each point on the grid of $u_S = P(Z)$ using a step size of 0.01 over the support of $P(Z)$. The estimators of the slope of these regressions for the different points on the grid are the marginal returns for different levels of unobservables $u_s = P(Z)$.[31]

Lastly, I calculate standard errors by performing a bootstrap over the whole procedure discussed above. Unfortunately error bands are wide in particularly for large values of $P(Z)$ for which there are few data points (see Figure 9 in Online Appendix B displaying the marginal return to college with 95% confidence intervals using a bandwidth of 0.15).[32] In the next section I will use these estimation results to perform policy experiments.

5.4 Results of the Policy Experiments

The goal of this section is twofold: First, I evaluate potential welfare implications of government policies, such as the introduction of a governmental fellowship program or tuition subsidies. Therefore I analyze the effect of a change in direct costs on the likelihood to attend college. To simulate the effect of a means-tested and a merit-based policy, I perform this analysis separately for poor and for poor and able individuals. Means-tested policies should help to target the policy to those individuals most likely to be constrained, since resources are limited. Eligibility based on merit –determined for example in terms of previous school

[31]Figure 8 in Online Appendix B displays the marginal return to college for three different bandwidths using a Gaussian kernel. One can see that the choice of bandwidth controls the trade-off between bias and variance: while a relatively small bandwidth of 0.1 leads to a wiggly line that is clearly undersmoothed, a large bandwidth of 0.2 seems to lead to an oversmoothed graph.

[32]Carneiro, Heckman, and Vytlacil (2011) and Carneiro, Heckman, and Vytlacil (2010) have the same problem of wide confidence bands using the NLSY. The fact that my sample only contains relatively poor individuals all of which have a low probability of attending college is likely to aggravate the problem.

performance–, has the advantage that the policy supports individuals who are more likely to actually complete college instead of dropping out.

In this analysis I compute the fraction of people changing their decisions as a result of the policy and derive the average "marginal" expected returns of these individuals. I estimate the "Policy Relevant Treatment Effect" ($PRTE$) for the policies of interest, which will be a weighted average of the marginal returns to college (MTE), as determined in the previous section. For the evaluation of policies it is crucial to derive the "marginal" return instead of the "average" return of a randomly selected individual, because only the people "at the margin" are the ones who will respond to policies.

Second, I compare the average "marginal" expected return to the average expected return of individuals attending college. Thus with subjective expectations I can improve on the test suggested by Card (2001). Larger "marginal" returns indicate that individuals at the margin face higher unobserved costs.

The first policy I evaluate is a decrease in the distance to the closest university. This could be seen as a literal decrease in the distance by building new universities in places that previously did not have higher education institutions or as a reduction in direct costs via fellowships for costs of living. Of course the implied costs of the two policies are likely to be very different and difficult to determine. In addition the analysis in this section does not take into account general equilibrium effects of such policies. Thus the goal of this section is not a complete cost-benefit-analysis, but to test for credit constraints by comparing expected returns of people at the margin to the ones of those already attending and to give an idea of potential welfare benefits of government policies such as fellowship programs in Mexico.

In Section 4.3 I have shown that a change in distance to college affects poor high-return individuals most. In addition I take into account in this section, that a change in costs can only affect individuals at the margin. I perform the analysis by decreasing the distance to college by 20 kilometers (for different target groups). This counterfactual policy leads to an increase in college attendance of about 4% (1 percentage point), and to an average marginal expected return of 0.89 (see Table 7).[33] Decreasing the distance only for very poor individuals (per capita income less than 5,000 pesos), leads to a change in attendance of 2%, while those individuals who change college attendance have an average marginal expected return of 0.88. For very poor and very able individuals (per capita income less than 5,000 pesos and GPA in the top tercile), this policy would lead to a change in attendance of about 1%, and an average marginal expected return of 0.90.

We cannot reject that the average marginal return (between 0.88 and 0.90 for the three groups) is as high or higher than the average expected return of those already attending

[33]The (expected) return is defined as the log difference between expected college and high school earnings. As discussed in Section 3.3, the average expected return is close to estimates of realized returns in Mexico.

college (0.71). In the case of the last group (very poor and high-performing), their expected return is even close to being significantly larger than the average expected return of those attending college (p-value 0.14).

As a second policy experiment, I consider the effect of a 10% decrease in tuition costs, for example via tuition subsidies. A 10% reduction in tuition costs leads to an average marginal return of 0.83, 0.79 for the poor and 0.81 for the poor and able, which is as high as the average expected return of those individuals attending (see Table 7). Unfortunately, tuition costs are very noisily measured, so standard errors for the fraction of "switchers" and for the marginal returns are large.

To conclude, these results imply that individuals at the margin have to be facing high unobserved costs to explain the fact that they did not attend college in the absence of the policy, despite having expected returns as high or higher than of those (rich) individuals already attending college.

6 Conclusion

The goal of this paper has been to improve our understanding of the huge differences in college enrollment rates between poor and rich individuals in Mexico and to show how data on people's subjective expectations of earnings can help in this endeavor.

When examining reasons for low school attendance among the poor, researchers face the following identification problem: On the one hand poor people might expect particularly low returns to schooling –due for example to lower cognitive skills or perceptions of limited career opportunities even with a college degree–, and thus decide not to attend. On the other hand they might face high attendance costs that prevent them from attending despite high expected returns.

To address this identification problem, I use data on people's subjective expectations of their idiosyncratic returns to college. Since what matters for people's decisions is the perception of their own cognitive and social skills and their beliefs about future skill prices, these data ideally provide people's expectations conditional on their information sets.

Using data on subjective expectations, I can show that poor individuals require significantly higher expected returns to be induced to attend college than individuals with wealthy parents. I found that poor individuals are particularly responsive to changes in direct costs, which is consistent with the predictions of a model with credit constraints. Furthermore, I have provided suggestive evidence that there are no systematic differences in time preferences between people of different income categories, so that my results are unlikely to be driven by the poor being more impatient than the rich.

Evaluating potential welfare implications by applying the Local Instrumental Variables

Table 7: Counterfactual Policy Experiments.

Policy Change	Individuals Changing College Attendance Decision		Individuals Attending College	
	Change in Overall Attendance Rate in pp (in %) (p-value)	**Marginal** Expected Return (MTE)	**Average** Expected Return (TTE)	**Diff MTE-TTE** (p-value)
Decrease Dist by 20km				
for all	1pp (4%) (p-val 0.02)	0.89	0.71	0.18 (0.16)
for very poor	0.4pp (2%) (p-val 0.07)	0.88	0.71	0.17 (0.26)
for very poor and very able	0.2pp (1%) (p-val 0.07)	0.90	0.71	0.19 (0.14)
Decrease Tuition by 10%				
for all	0.4pp (2%) (p-val 0.39)	0.83	0.71	0.12 (0.29)
for very poor	0.3pp (1.5%) (p-val 0.28)	0.79	0.71	0.08 (0.37)
for very poor and very able	0.3pp (1.5%) (p-val 0.28)	0.81	0.71	0.10 (0.36)

approach of Heckman and Vytlacil (2005) to my model, I found that a sizeable fraction of poor individuals would change their decision and attend in response to a reduction in direct costs. Individuals at the margin have expected returns that are as high or higher than the ones of individuals already attending college, which is consistent with credit constraints playing an important role.

My results suggest that credit constraints are one of the driving forces of Mexico's large inequalities in access to higher education and low overall enrollment rates and point to large welfare gains of introducing a governmental fellowship program by removing obstacles to human capital accumulation and fostering Mexico's development and growth.

Appendix

Derivation of the Participation Equation

The goal of this section is to use the potential outcome equations (2) and the subjective expectation information (3) in my model of college attendance according to which an individual decides to attend college if

$$EPV_{18}(Y_i^1) - EPV_{18}(Y_i^0) - \frac{C_i}{p_i^C} \geq 0. \tag{11}$$

To express the expected present value of earnings for both schooling scenarios in terms of subjective expectations of earnings, I need to take into account that the questions on subjective expectations of earnings were asked conditional on working ($E_{18}(Y_{ia}^S|W^S = 1)$ for $S = 0, 1$) in addition to asking about the perceived probability of working in the two different schooling scenarios (p_i^{WS} for $S = 0, 1$):[34]

$$EPV_{18}(Y_i^1) = \sum_{a=22}^{A} \frac{p_i^{W1} E_{18}(Y_{ia}^1|W^1 = 1)}{(1 + r_i)^{a-18}} \tag{12}$$

To then use the potential outcome equations (2) and the subjective expectation information (3), and rewrite the participation equation in terms of expected returns to college, I use

[34] I can also allow the increase in experience to differ across people depending on their perceived probability of being employed with a high school and college degree (p_i^{W0} and p_i^{W1}), which should capture the fraction of the year that they expect to be employed (in principle, one would like to use the perceived probability of working for each year over the whole life-cycle, but in my data questions on subjective expectations were only asked for age 25, so that I would have to assume $p_{ia}^{Wj} = p_{i25}^{Wj} = p_i^{Wj}$ for all a and for $j = 0, 1$). In that case $E_{18}(\ln Y_{ia}^0) = \alpha_0 + \beta_0' X_i + \gamma_0 p_{ia}^{W0}(a - 18) + \theta_0' f_i$ and analogously for college earnings. The following derivation goes through with the adjustment that γ_S would have to be substituted by $\gamma_S p_i^{WS}$ for $S = 0, 1$ in all following equations (results from the author upon request).

the following relationship

$$E(Y_{ia}) \equiv E(e^{\ln Y_{ia}}) = e^{E(\ln Y_{ia})+0.5Var(\ln Y_{ia})}, \tag{13}$$

which holds exactly in the case of earnings that are distributed log-normally, which is the traditional parameterization, (otherwise approximately).

Thus I can rewrite the expected present value of college earnings (analogously for high school earnings) as

$$\begin{aligned} EPV_{18}(Y_i^1) &= \sum_{a=22}^{\infty} \frac{p_i^{W1}\exp(E_{18}(lnY_{ia}^1|W^1=1)+0.5Var_{18}(lnY_{ia}^1|W^1=1))}{(1+r_i)^{a-18}} \\ &= \sum_{a=22}^{\infty} \frac{p_i^{W1}\exp\left(\alpha_1+\beta_1'X_i+\gamma_1(a-22)+\theta_1'f_i+0.5(\sigma_i^1)^2\right)}{(1+r_i)^{a-18}} \\ &= \frac{p_i^{W1}\exp\left(\alpha_1+\beta_1'X_i+\theta_1'f_i+0.5(\sigma_i^1)^2\right)}{(1+r_i)^4} \cdot \left(\sum_{a=22}^{\infty} \frac{\exp\left(\gamma_1(a-22)\right)}{(1+r_i)^{a-22}}\right) \\ &= \frac{p_i^{W1}\exp\left(\alpha_1+\beta_1'X_i+\theta_1'f_i+0.5(\sigma_i^1)^2\right)}{(1+r_i)^4} \left(\frac{1+r_i}{1+r_i-\exp(\gamma_1)}\right), \end{aligned} \tag{14}$$

where I assume that $\exp(\gamma_j) < 1+r_i$ for $j=0,1$ to apply the rule for a geometric series,[35] that $Var(\ln Y_{ia}^S|W^1=1) = (\sigma_i^S)^2$ for all a and $S=0,1$ and $A \to \infty$ as an approximation.

Data on subjective expectations of earnings for age $a=25$ thus allow me to rewrite the expected present value of college earnings as follows (see equation (3)):

$$EPV_{18}(Y_i^1) = \frac{p_i^{W1}\exp\left(E_{18}(lnY_{i25}^1|W^1=1)+0.5(\sigma_i^1)^2-3\gamma_1\right)}{(1+r_i)^3} \cdot \left(\frac{1}{1+r_i-\exp(\gamma_1)}\right),$$

Analogously, I can derive the following expression for $EPV_{18}(Y_i^0)$

$$EPV_{18}(Y_i^0) = p_i^{W0}\exp\left(\alpha_0+\beta_0'X_i+\theta_0'f_i+0.5(\sigma_i^0)^2\right) \cdot \left(\frac{1+r_i}{1+r_i-\exp(\gamma_0)}\right). \tag{15}$$

Substituting the expressions for the expected present value of college and high school earnings into equation (11), an individual decides to attend college if

[35] Some back-of-the-envelope calculations suggest that this assumption is reasonable in the given context: Papers such as Connolly and Gottschalk (2006) and Heckman, Lochner, and Taber (1998) find returns to experience well below 0.05 for the US, while interest rates in Mexico are clearly significantly higher than 0.05 in the relevant period (see for example McKenzie (2006)).

$$\left[\frac{p_i^{W1}\exp\left(E_{18}(lnY_{i25}^1|W^1=1)+0.5(\sigma_i^1)^2-3\gamma_1\right)}{(1+r_i)^3}\cdot\left(\frac{1}{1+r_i-\exp(\gamma_1)}\right)\right]$$

$$-\left[p_i^{W0}\exp\left(E_{18}(lnY_{i25}^0|W^0=1)+0.5(\sigma_i^0)^2-7\gamma_0\right)\cdot\left(\frac{1+r_i}{1+r_i-\exp(\gamma_0)}\right)\right]-\frac{C_i}{p_i^C}\geq 0,$$

which I can rewrite in the following way

$$\begin{aligned}\exp\left(E_{18}(lnY_{i25}^1|W^1=1)+0.5(\sigma_i^1)^2\right)-\exp\left(E_{18}(lnY_{i25}^0|W^0=1)+0.5(\sigma_i^0)^2\right)\\ \cdot\left[(1+r_i)^4\frac{p_i^{W0}}{p_i^{W1}}\cdot\frac{\exp(3\gamma_1)}{\exp(7\gamma_0)}\left(\frac{1+r_i-\exp(\gamma_1)}{1+r_i-\exp(\gamma_0)}\right)\right]\\ \geq(1+r_i)^3\frac{C_i}{p_i^C p_i^{W1}}(1+r_i-\exp(\gamma_1))\end{aligned}$$

In the following, I assume: $\left(\frac{1+r_i-\exp(\gamma_1)}{1+r_i-\exp(\gamma_0)}\right)\approx 1$, which is approximately satisfied given estimates of returns to experience of around 0.03 for college and 0.02 for high school and an interest rate of around 10% (see, for example, as mentioned above Connolly and Gottschalk (2006) using SIPP data for the US or Heckman, Lochner, and Taber (1998) who show that differences in returns to experience between high school and college educated are small).

In order to express the decision rule in terms of expected gross returns to college and use the information on expected returns from subjective expectations of earnings (see expression (4)), I use a Taylor series approximation of $\exp(B)$ around A, $\exp(B)=\exp(A)\sum_{j=0}^{\infty}\frac{(B-A)^j}{j!}$, to rewrite the decision rule, which has the form $\exp(B)-\exp(A)\cdot L\geq K$. Noting that in this context

$$\begin{aligned}B-A &= \left(E_{18}(\ln Y_{i25}^1|W^1=1)+0.5(\sigma_i^1)^2\right)-\left(E_{18}(\ln Y_{i25}^0|W^0=1)+0.5(\sigma_i^0)^2\right)\\ &= \rho_{i25}+0.5\left((\sigma_i^1)^2-(\sigma_i^0)^2\right),\end{aligned}$$

I can write the decision rule as

$$\begin{aligned}&\exp\left(E_{18}(\ln Y_{i25}^0|W^0=1)+0.5(\sigma_i^0)^2\right)\cdot\left(\sum_{j=0}^{\infty}\frac{\left(\rho_{i25}+0.5\left((\sigma_i^1)^2-(\sigma_i^0)^2\right)\right)^j}{j!}\right)\\ &-\left(\exp(E_{18}(\ln Y_{i25}^0|W^0=1)+0.5(\sigma_i^0)^2)\right)\cdot(1+r_i)^4\frac{p_i^{W0}}{p_i^{W1}}\cdot\frac{\exp(3\gamma_1)}{\exp(7\gamma_0)}\\ &-(1+r_i)^3\frac{C_i}{p_i^C p_i^{W1}}(1+r_i-\exp(\gamma_1))\geq 0.\end{aligned}$$

Rearranging will lead to

$$\sum_{j=0}^{\infty}\frac{\left(\rho_{i25}+0.5\left((\sigma_i^1)^2-(\sigma_i^0)^2\right)\right)^j}{j!}$$
$$-(1+r_i)^4\left[\frac{p_i^{W0}}{p_i^{W1}}\cdot\frac{\exp(3\gamma_1)}{\exp(7\gamma_0)}+\frac{C_i}{p_i^C p_i^{W1}\exp\left(E(\ln Y_{i25}^0|W^0=1)+0.5(\sigma_i^0)^2\right)}\right]$$
$$+(1+r_i)^3\frac{C_i}{p_i^C p_i^{W1}\exp\left(E(\ln Y_{i25}^0|W^0=1)+0.5(\sigma_i^0)^2\right)}\exp(\gamma_1)\geq 0.$$

Thus using the data on subjective expectations, the latent variable model for attending university can be written as

$$S = \begin{cases} 1 & \text{if } S^*\geq U \\ 0 & \text{otherwise.}\end{cases}$$

where, (16)

$$S^* = \sum_{j=0}^{\infty}\frac{\left(\rho_{i25}+0.5\left((\sigma_i^1)^2-(\sigma_i^0)^2\right)\right)^j}{j!}$$
$$-(1+r_i)^4\left[\frac{p_i^{W0}}{p_i^{W1}}\cdot\frac{\exp(3\gamma_1)}{\exp(7\gamma_0)}+\frac{C_i}{p_i^C p_i^{W1}\exp\left(E(\ln Y_{i25}^0|W^0=1)+0.5(\sigma_i^0)^2\right)}\right] \quad (17)$$
$$+(1+r_i)^3\frac{C_i}{p_i^C p_i^{W1}\exp\left(E(\ln Y_{i25}^0|W^0=1)+0.5(\sigma_i^0)^2\right)}\exp(\gamma_1),$$

(18)

Derivation of the Testable Prediction of Excess Responsiveness

Making use of the participation equation for college attendance (see equation (17)), the following results show that individuals who face a higher interest rate are more responsive to changes in direct costs.

$$\frac{\partial S^*}{\partial C}=\frac{-(1+r_i)^4+(1+r_i)^3\exp(\gamma_1)}{p_i^C p_i^{W1}\exp\left(E(\ln Y_{i25}^0|W^0=1)+0.5(\sigma_i^0)^2\right)}<0$$

as $\exp(\gamma_1)<1+r_i$ (see the previous section). Furthermore

$$\frac{\partial^2 S^*}{\partial C\partial r}=\frac{-4(1+r_i)^3+3(1+r_i)^2\exp(\gamma_1)}{p_i^C p_i^{W1}\exp\left(E(\ln Y_{i25}^0|W^0=1)+0.5(\sigma_i^0)^2\right)}<0, \quad (19)$$

as $4(1+r_i)>3\exp(\gamma_1)$.

Thus $\left|\frac{\partial S^*}{\partial C}\right|$ is increasing in r_i. As $P(S=1)=\Phi(S^*)$, which is a monotonic transformation of S^*, also $\left|\frac{\partial P(S=1)}{\partial C}\right|$ is increasing in r and thus individuals who face a higher interest rate are more responsive to changes in direct costs.

Participation Equation as a Fourth-Order Polynomial in the Interest Rate

The participation equation (17) can be expressed as polynomial in the interest rate under the assumption that all unobserved heterogeneity stems from the unobserved interest rate r:

$$(1+r)^4 - (1+r)^3 \frac{C \exp(\gamma_1)}{\left(p_i^C p_i^{W1} \exp\left(E(\ln Y_{i25}^0) + 0.5(\sigma_i^0)^2\right)\right) \cdot \frac{p_i^{W0}}{p_i^{W1}} \cdot \frac{\exp(\gamma_1 3)}{\exp(\gamma_0 7)} + C}$$
$$- \frac{p_i^C p_i^{W1} \exp\left(E(\ln Y_{i25}^0 | W^0 = 1) + 0.5(\sigma_i^0)^2\right)}{\left(p_i^C p_i^{W1} \exp\left(E(\ln Y_{i25}^0 | W^0 = 1) + 0.5(\sigma_i^0)^2\right)\right) \cdot \frac{p_i^{W0}}{p_i^{W1}} \cdot \frac{\exp(\gamma_1 3)}{\exp(\gamma_0 7)} + C} \sum_{j=0}^{\infty} \frac{\left(\rho_i + 0.5\left((\sigma_i^1)^2 - (\sigma_i^0)^2\right)\right)^j}{j!}.$$

Derivation of the Marginal Return to College

Derivation of equation (10):

$$\begin{aligned}
E(U_1 - U_0 | U_S \leq p) &= \int_{-\infty}^{\infty} (U_1 - U_0)\, f(U_1 - U_0 | U_S \leq p) d(u_1 - u_0) \\
&= \int_{-\infty}^{\infty} (U_1 - U_0) \frac{\int_0^p f\left(U_1 - U_0, U_S\right) du_S}{Pr(U_S \leq p)} d(u_1 - u_0) \\
&= \int_{-\infty}^{\infty} (U_1 - U_0) \frac{\int_0^p f\left(U_1 - U_0 | U_S\right) f(u_S) du_S}{p} d(u_1 - u_0) \\
&= \frac{1}{p} \int_0^p E\left(U_1 - U_0 | U_S = u_S\right) du_S.
\end{aligned}$$

References

Arcidiacono, P., V. J. Hotz, and S. Kang (2011). Modelling college major choices using elicited measures of expectations and counterfactuals. *Journal of Econometrics*.

Attanasio, O. and K. M. Kaufmann (2010). Educational choices and subjective expectations of returns: Evidence on intra-household decision-making and gender differences. *mimeo*.

Attanasio, O. P. (2009). Expectations and perceptions in developing countries: Their measurement and their use. *The American Economic Review 99(2)*, 87–92.

Attanasio, O. P. and K. M. Kaufmann (2009). Educational choices, subjective expectations and credit constraints. *NBER Working Paper 15087 July*.

Attanasio, O. P., C. Meghir, and A. Santiago (2011). Education choices in mexico: Using a structural model and a randomized experiment to evaluate progresa. *Review of Economic Studies*.

Becker, G. S. (1967). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. Chicago: University of Chicago Press.

Becker, G. S. and H. G. Lewis (1973). On the interaction between the quantity and quality of children. *The Journal of Political Economy, 81(2)*, S279–S288.

Behrman, J. R., A. Gaviria, and M. Szekely (2002). *Social Exclusion in Latin America: Introduction and Overview. Social Exclusion in Latin America and the Caribbean.* Washington, D.C.: Inter-American Development Bank.

Bettinger, E. P., B. T. Long, P. Oreopoulos, and L. Sanbonmatsu (2009). The role of simplification and information in college decisions: Results from the h&r block fafsa experiment. *NBER Working Paper 15361*.

Binelli, C. (2008). Returns to education and increasing wage inequality in latin america. *mimeo*.

Blossfeldt, H. and Y. Shavit (1993). *Persistent Inequality: Changing Educational Attainment in Thirteen Countries*. Boulder, Co: Westview Press.

Brown, M., J. K. Scholz, and A. Seshadri (2011). A new test of borrowing constraints for education. *Review of Economic Studies*.

Cameron, S. V. and J. J. Heckman (1998). Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of american males. *Journal of Political Economy 106 (2)*, 262–333.

Cameron, S. V. and J. J. Heckman (2001). The dynamics of educational attainment for black, hispanic and white males. *Journal of Political Economy 109 (3)*, 455–499.

Cameron, S. V. and C. Taber (2004). Borrowing constraints and the returns to schooling. *Journal of Political Economy 112 (1)*, 132–182.

Card, D. (1995). *Using Geographic Variation in College Proximity to Estimate the Return to Schooling, In: Aspects of Labour Market Behaviour: Essays in Honour of John Vanderkamp*. Toronto: Univ. Toronto Press.

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica 69 (5)*, 1127–60.

Carneiro, P. and J. J. Heckman (2002). The evidence on credit constraints in post-secondary schooling. *The Economic Journal 112 (October)*, 705–734.

Carneiro, P., J. J. Heckman, and E. Vytlacil (2005). Estimating marginal returns to education. *Working Paper*.

Carneiro, P., J. J. Heckman, and E. Vytlacil (2010). Evaluating marginal policy changes and the average effect of treatment for individuals at the margin. *Econometrica 78(1)*, 377–394.

Carneiro, P., J. J. Heckman, and E. Vytlacil (2011). Estimating marginal returns to education. *The American Economic Review 101(6)*.

Chabris, C., D. Laibson, C. Morris, J. Schuldt, and D. Taubinsky (2008). Individual laboratory-measured discount rates predict field behavior. *Journal of Risk and Uncertainty 37(2)*, 237–269.

Connolly, H. and P. Gottschalk (2006). Differences in wage growth by education level - do less-educated workers gain less from work experience? *mimeo*.

Cunha, F. (2007). A time to plant and a time to reap. *mimeo*.

Cunha, F., J. J. Heckman, and S. Navarro (2005). Separating uncertainty from heterogeneity in life cycle earnings. *Oxford Economic Papers 57*, 191–261.

Delavande, A., X. Giné, and D. McKenzie (2011). Measuring subjective expectations in developing countries: A critical review and new evidence. *Journal of Development Economics 94(2)*, 151–163.

Dinkelman, T. and C. Martinez (2011). Investing in schooling in chile: The role of information about financial aid for higher education. *mimeo*.

Dominitz, J. and C. F. Manski (1996). Eliciting student expectations of the returns to schooling. *Journal of Human Resources 31 (1)*, 1–26.

Fan, J. (1992). Design-adaptive nonparametric regression. *Journal of the American Statistical Association 87*, 998–1004.

Grossman, M. (2000). *The Human Capital Model. In: Handbook of Health Economics.* Elsevier, Amsterdam, Vol 1, pp 347408.

Guiso, L., T. Jappelli, and L. Pistaferri (2002). An empirical analysis of earnings and employment risk. *Journal of Business and Economic Statistics 20 (2)*, 241–253.

Heckman, J. J. (1979). Sample selection as a specification error. *Econometrica 47 (1)*, 153–161.

Heckman, J. J., L. Lochner, and C. Taber (1998). Explaining rising wage inequality: Explorations with a dynamic general equilibrium model of labor earnings with heterogeneous agents. *Review of Economic Dynamics 1*, 1–58.

Heckman, J. J. and E. Vytlacil (2001). Policy-relevant treatment effects. *The American Economic Review, Papers and Proceedings 91 (2)*, 107–111.

Heckman, J. J. and E. Vytlacil (2005). Structural equations, treatment effects, and econometric policy evaluation. *Econometrica 73 (3)*, 669–738.

Heckman, J. J., E. Vytlacil, and S. S. Urzua (2006). Understanding instrumental variables in models with essential heterogeneity. *mimeo*.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *Quarterly Journal of Economics 125(2)*, 515–548.

Kaufmann, K. M. (2008). Understanding the income gradient in college attendance in mexico: The role of heterogeneity in expected returns to college. *SIEPR Working Paper, January 2008 07-040*.

Kaufmann, K. M. and L. Pistaferri (2009). Disentangling insurance and information in intertemporal consumption choices. *The American Economic Review, Papers and Proceedings 99(2)*, 387–92.

Khwaja, A., D. Silverman, and F. Sloan (2007). Time preference, time discounting, and smoking decisions. *Journal of Health Economics Vol 26 (5)*, 927–949.

Kling, J. R. (2001). Interpreting instrumental variables estimates of the returns to schooling. *Journal of Business and Economic Statistics 19 (July)*, 358–364.

Lochner, L. J. and A. Monje-Naranjo (2011). The nature of credit constraints and human capital. *The American Economic Review*.

Lovenheim, M. F. (2011). The effect of liquid housing wealth on college enrollment. *Journal of Labor Economics*.

Mahajan, A. and A. Tarozzi (2011). Time inconsistency, expectations and technology adoption: The case of insecticide treated nets. *mimeo, Stanford University and Duke University.*.

Mahajan, A., A. Tarozzi, J. Yoong, and B. Blackburn (2011). Bednets, information and malaria in orissa. *mimeo, Stanford University, Duke University and RAND.*.

Manski, C. F. (2004). Measuring expectations. *Econometrica 72 (5)*, 1329–76.

McKenzie, D. (2006). The consumer response to the mexican peso crisis. *Economic Development and Cultural Change 55(1)*, 139–172.

Mincer, J. (1974). Schooling, experience and earnings. *NBER, New York*.

Navarro, S. (2011). Using observed choices to infer agent's information: Reconsidering the importance of borrowing constraints, uncertainty and preferences in college attendance. *CIBC Working Paper 2011-8*.

Nguyen, T. (2008). Information, role models and perceived returns to education: Experimental evidence from madagascar. *mimeo*.

Stinebrickner, R. and T. R. Stinebrickner (2008). The effect of credit constraints on the college drop-out decision: A direct approach using a new panel study. *The American Economic Review 98(5)*, 2163–84.

Stinebrickner, R. and T. R. Stinebrickner (2012). Learning about academic ability and the college drop-out decision. *Journal of Labor Economics 30(4)*, 707–748 (lead article).

Stinebrickner, T. R. and R. Stinebrickner (2013). A major in science? initial beliefs and final outcomes for college major and dropout. *The Review of Economic Studies*.

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

Willis, R. J. and S. Rosen (1979). Education and self-selection. *Journal of Political Economy 87*, S 7–36.

Zafar, B. (2009). College major choice and the gender gap. *Federal Reserve Bank of New York Staff Reports*.

Appendix B: NOT FOR PUBLICATION

Background Information on College Enrollment and on Costs and Financing of College Attendance in Mexico

In 2004 around 22% of adolescents of the relevant age group (18 to 24 years) were attending college in Mexico to receive an undergraduate degree ("licenciatura") (ANUIES, annual statistics 2004). This attendance rate is significantly lower than in many other Latin American countries (see Table 8). Mexico is characterized by large inequalities in access to college education for different income groups. In comparison to other Latin American countries, such as Colombia, Argentina and Chile, only Brazil has a smaller fraction of poor students attending college (see Table 8). Figure 4 displays college attendance rates of 18 to 24 year old high school graduates for different parental income quartiles.[36] High school graduates are already a selective group, for example for urban Mexico about 75% of the relevant age group attain a high school degree. The attendance rate of high school graduates in the lowest parental income quartile is around 22% compared to 67% for the highest parental income quartile. The "Jovenes con Oportunidades" sample (2005) used in this paper consists of high school graduates from Oportunidades families and is thus only representative of about the poorest third of the high school graduate population. The positive correlation between parental income and college attendance rate can also be found for this sample, but differences between poorest quartile (17%) and richest quartile (35%) are smaller, as every individual in the sample is relatively poor (see figure 5, Jovenes con Oportunidades 2005).

College attendance costs in Mexico pocket a large fraction of parental income for relatively poor families. Costs consist of enrollment and tuition fees, fees for (entrance) exams and other bureaucratic costs, costs for transport and/or room and board, health insurance (mandatory for some universities), costs for schooling materials such as books. Administrative data on tuition and enrollment fees per year from the National Association of Universities and Institutes of Higher Education (ANUIES) reveals a large degree of heterogeneity: Yearly tuition and enrollment costs vary between 50 pesos ("Universidad Autónoma de Guerrero", Guerrero) and 120,000 pesos ("Tecnológico de Monterrey", I.T.E.S.M. - Campus Puebla), which is equivalent to approximately 5 and 12,000 US$. The tuition cost measure that I use in my analysis is the minimum yearly tuition/enrollment fee of universities in the closest locality with at least one university. Fifty percent of the high school graduates face (minimum) tuition costs of over 750 pesos, which is equivalent to about 15% of median yearly per capita parental income. The other important cost factor depends on whether the adolescent has to move to a different city and pay room and board or whether she can live with her family during college. I therefore construct a measure of distance to the closest university for each individual.

In Mexico funding for higher-education fellowships and student loan programs is very limited and only about 5% of the undergraduate student population receive fellowships, while 2% receive student loans, which is low even compared to other Latin American countries (see Table 8). The national scholarship program PRONABES was created in 2001 with the goal of more equal access to higher education at the undergraduate level. In 2005 funding of PRONABES amounted to 850 million pesos (equal to 40 US$ per student per year) and 5%

[36]Parental income is measured in the last year before the college attendance decision.

of the undergraduate student population received a fellowship ("beca") in 2005 compared to 2% in 2001/02 (see Department of Public Education (SEP)), 2005). Eligibility for a fellowship is subject to three conditions: first, a maximum level of family income, where priority is given to families with less than two times the minimum monthly salary, while in special cases people are still eligible with less than four times the minimum monthly salary. Second, students need a minimum GPA (80) and third, they have to have been accepted at a public university or technical institute. After each year, the student has to prove that economic eligibility criteria are still met and that she is in good academic standing. In 2004/05 the fellowship consisted of a monthly stipend of 750 pesos –slightly more than half the minimum wage per month– in the first year of studies, and increased to 1000 pesos in the fourth year of studies. Student loan programs are also of minor importance in Mexico. Only about 2% of the national student population benefit from a student loan, which is low even compared to poorer Latin American countries, such as Colombia (9%) and Brazil (6%). In Mexico there are four different programs that offer student loans. The largest program, SOFES, offers loans to 1.5% of students and was implemented by a collaboration of private universities. It is need-and-merit based, but students with collateral are preferred. The other three are very small state programs, ICEES in Sonora state, ICEET in Tamaulipas, and Educafin in Guanajuato, which are not part of my sample.

Figure 4: College enrollment rates of 18 to 24 year old high school completers by parental income quartile (Mexican Family Life Survey, 2003).

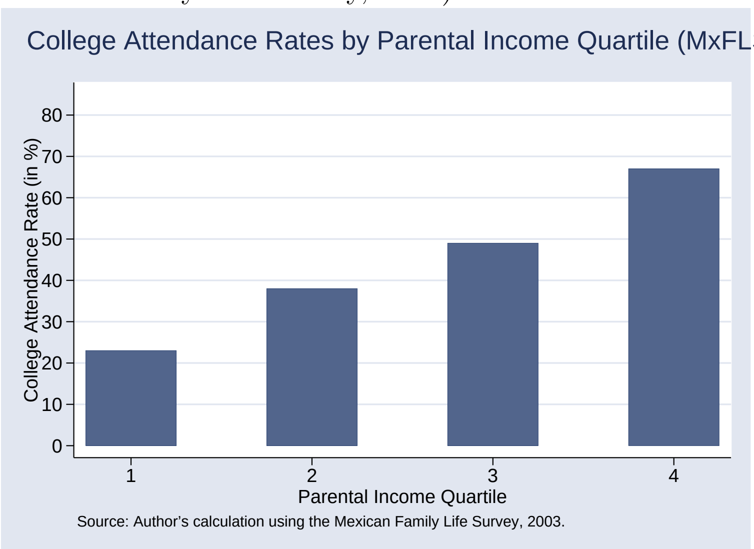


Figure 5: College enrollment rates of 18 to 24 year old high school completers by parental income quartile (Jovenes con Oportunidades Survey, 2005).

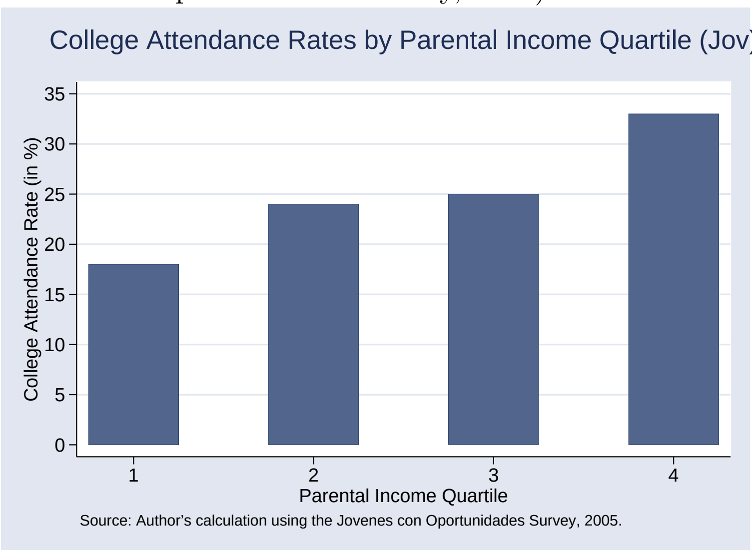


Table 8: Comparison of enrollment rates, fraction of poorest 40% in percent of the student population, fraction of GDP spend on higher education, fraction of expenditures on higher education on fellowships and student loans: Mexico, other Latin American countries, OECD and USA.

Countries Ranked by Per Cap GDP	Enrollment in Higher Education in % of 18-24 Year Old	Fraction of Poorest 40% of 18-24 Year Old as % of Student Body	Expenditures on Higher Education in % of GDP	Spending on Fellowships and Loans in % of Exp. on Higher Educ	Beneficiaries of Student Loans in % of students
Brazil	16%	4%	1.5%	11.2%	6%
Colombia	23%	14%	1.7%	.	9%
Peru	29%	.	.	.	.
Mexico	**20%**	**8%**	**1.1%**	**6.2%**	**2%**
Chile	39%	16%	2.2%	34.8%	.
Argentina	37%	16%	1.1%	.	.
OECD	56%	.	.	17.5%	.
USA	54%	20%	.	.	35%

Sources: World Bank (2005) for Enrollment and Fraction of Poorest 40%, OECD Indicators (2007) for Expenditures on Higher Education and on Spending on Fellowships and Loans. CIA World Factbook (2006) and IMF Country Ranking for Ranking of Per Capita GPD (PPP). For Beneficiaries of Student Loans: Ministery of Education, Brazil (2005); ICETEX, Colombia (2005); SOFES (2005), ICEES (2006), ICCET (2007) and Educafin (2007) in Mexico; US Office of Post-Secondary Education Website, 2006. Information not available indicated as ".".

Brief Introduction to the Local Instrumental Variable Methodology

To introduce the "Local Instrumental Variable (LIV)" methodology (see Heckman and Vytlacil (2005), Carneiro, Heckman, and Vytlacil (2011) and Carneiro, Heckman, and Vytlacil (2010)), the framework of the "Generalized Roy Model" is a useful starting point (compare Section 2):[37]

$$\begin{aligned} \ln Y_0 &= \alpha + U_0 \\ \ln Y_1 &= \alpha + \bar{\rho} + U_1 \\ S^* &= \mu(Z) - U_S \\ S = 1 &\Leftrightarrow S^* \geq 0. \end{aligned}$$

In the context of this framework, if U_S is independent of U_0, U_1, the average treatment effect can be calculated as the simple difference between the outcome of the "treated" ($\ln Y_1$) and the "untreated" ($\ln Y_0$). If on the other hand U_S is correlated with U_0, U_1, that is people self-select into treatment based on U_S which is correlated with the potential outcomes, then the simple difference will be a biased estimate of the average treatment effect. The problem is that one compares "treated" and "untreated" individuals who differ in their unobserved costs, U_S, and these unobserved costs are correlated with the potential outcomes.

The LIV methodology addresses this endogeneity problem as follows: Imagine U_S was observable and one could thus condition on U_S when computing the simple difference. In other words, one could use as counterfactual outcomes for people who were treated those individuals with the same U_S who were not treated. This approach would solve the usual endogeneity problem. This is exactly the key idea of the "Marginal Treatment Effect" (MTE), which is defined as follows:

$$\Delta^{MTE}(u_S) = E(\ln Y_1 - \ln Y_0 | U_S = u_S) = E(\rho | U_S = u_S). \tag{20}$$

The obvious question is how one can condition on U_S that is unobserved. Even though U_S is generally unobserved, it is known for individuals who are exactly indifferent between selecting into or out of treatment ("on the margin"), as can be seen from the selection equation: $S^* = 0 \Leftrightarrow \mu(Z) = U_S$. One can compute U_S for those individuals who are indifferent, by estimating the selection equation and calculating the propensity score $P(Z) \equiv P(S = 1|Z = z)$, which is the probability of selecting into treatment conditional on observable characteristics Z. The "Marginal Treatment Effect" can then be estimated for those individuals who are indifferent and characterized by $U_S = \mu(Z) = P(Z)$. For example in my context, the MTE represents the average gross gain to college for individuals who are indifferent between attending college or not and who have unobservable costs of $U_S = u_S$.

In a second step, policy experiments can be performed using the estimated MTE in the following way (see Heckman and Vytlacil (2001)): The "Policy Relevant Treatment Effect" ($PRTE$) is a weighted average of the marginal treatment effects (MTE), where the weights

[37] $\ln Y_1$ and $\ln Y_0$ denote log earnings with and without college ($S = 1, 0$), α denotes average earnings without college, $\overline{\rho}$ average returns to college and U_1 and U_0 the error terms in the earnings equations. For notational simplicity, I omit conditioning on observable characteristics X. The latent variable S^* depends on observable characteristics Z, which contain at least one element that is not in X and the error term U_S. Individuals choose college if and only if the value of the latent variable S^* is larger than zero.

depend on who changes participation in response to the policy of interest. One important assumption underlying this analysis is that the selection equation continues to hold under hypothetical interventions. The $PRTE$ can be written as:

$$PRTE = \int_0^1 MTE(u)\omega(u)du, \quad \text{where} \quad \omega(u) = \frac{F_P(u) - F_{P^*}(u)}{E(P^*) - E(P)}. \tag{21}$$

P is the baseline probability of $S = 1$ with cumulative distribution function F_P, while P^* is defined as the probability produced under an alternative policy regime with cumulative distribution function F_{P^*}. The intuition for the $PRTE$ is as follows: Given a certain level of unobservable costs, u, those individuals with $P(Z) > u$ will attend college, which is equivalent to a fraction $1 - F_P(u)$. A reduction, for example, in direct costs, Z, will lead to a new larger probability of attending, $P(Z^*)$. Thus for a given u, there are now more people deciding to attend college, $1 - F_{P^*}(u)$, and the change can be expressed as $F_P(u) - F_{P^*}(u)$. The weight is normalized by the change in the proportion of people induced into the program, $E(P^*) - E(P)$, to express the impact of the policy on a per-person basis.[38]

[38]The intuition is even more straightforward in the following special case: Suppose that $S^* = Z'\gamma + V$. Consider a policy that shifts Z_k (the kth element of Z) to $Z_k + \varepsilon$. For example, Z_k might be the tuition faced by an individual and the policy change might be to provide an incremental tuition subsidy of ε dollars. The resulting $PRTE_\varepsilon$ is the average return among individuals who are induced into university by the incremental subsidy, $PRTE_\varepsilon = E(\rho_i | Z'\gamma \leq V \leq Z'\gamma + \varepsilon\gamma_k)$.

Robustness Checks A

Figure 6: The Cumulative Distribution Function of Costs with 95% Confidence Intervals.

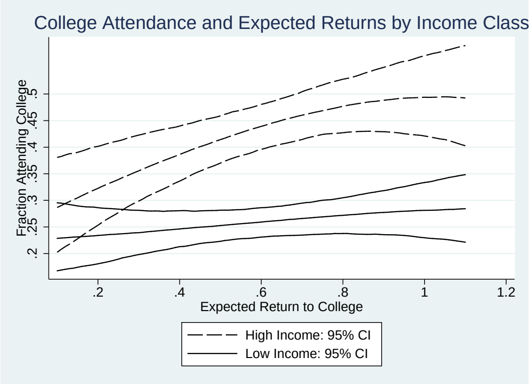


Figure 7: The Predicted Probability of Attending College.

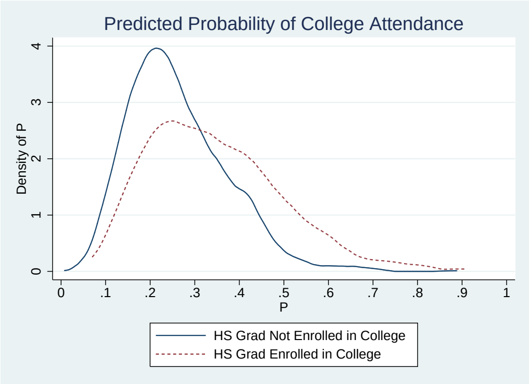


Figure 8: The Marginal Return to College for Different Levels of Unobserved Costs.

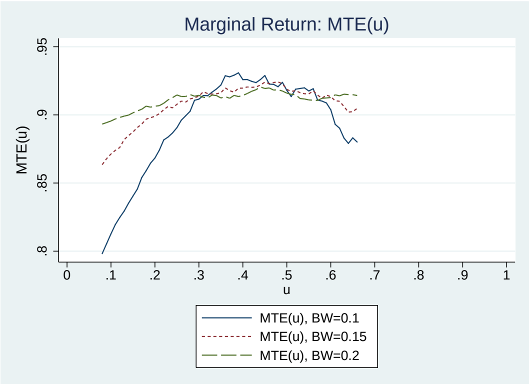


Figure 9: The Marginal Return to College with 95% Confidence Interval Bands.

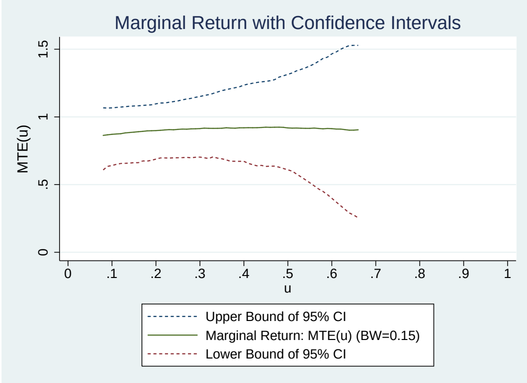


Table 9: Rationalization of Choices

	P-Val of KS-Test
Exp Log Earnings	
- Senior HS	0.417
- College	0.677
Exp Return	
- College	0.188
Prob of Work	
- Senior HS	0.236
- College	0.349
Observations	
(Sen HS Grads/Grade 12)	1612/469

Notes: Table displays the p-values of Kolmogorov-Smirnov tests of equality of distributions. The null hypothesis is that the cross-sectional distribution of -for example- expected returns is the same for the sample of senior high school graduates (whose schooling decision we are analyzing) and the sample of a cohort that is one year younger and just starting grade 12 (who have thus not decided yet about whether to enrol in college or not).

Table 10: Summary statistics of important variables of the two groups of respondents.

Respondent	Adolescent Mean (Std. Dev.)	Mother Mean (Std. Dev.)	P-Val of Diff
Expected Return	0.6670	0.6550	0.347
	(0.3820)	(0.3592)	
Expected Log High School Earnings	7.5778	7.6477	0.000
	(0.5004)	(0.4338)	
Var of Log High School Earnings	0.0054	0.0046	0.003
	(0.0079)	(0.0062)	
Var of Log College Earnings	0.0039	0.0034	0.022
	(0.0061)	(0.0054)	
Prob of Work High School	0.6657	0.6505	0.015
	(0.1817)	(0.1780)	
Prob of Work College	0.8250	0.8142	0.046
	(0.1601)	(0.1544)	
College Attendance Rate	0.2308	0.3636	0.000
	(0.4215)	(0.4812)	
Female	0.5813	0.4954	0.000
	(0.4935)	(0.5001)	
GPA (Scale 0 to 100)	82.19	82.27	0.783
	(7.16)	(10.34)	
Father's Yrs of Schooling	5.33	5.34	0.902
	(2.96)	(3.03)	
Mother's Yrs of Schooling	5.03	5.06	0.794
	(2.77)	(2.76)	
Per Capital Parental Income (Pesos)	7519.54	7925.42	0.371
	(8010.08)	(13638.29)	
Distance to University (km)	24.2312	26.4647	0.005
	(22.8159)	(22.8688)	
Tuition Costs (Pesos)	608.8104	503.4896	0.000
	(634.5729)	(338.1346)	

Table 11: Correlation between Earnings Expectations and Individual and Family-Background Characteristics.

Dependent Variable	**Expected Earnings**	
	High School	College
Female	-0.116***	-0.069***
	(0.026)	(0.026)
GPA of Junior HS (0-100)	0.001	0.004**
	(0.002)	(0.002)
Mother's Educ - Jr HS	-0.056	-0.046
	(0.036)	(0.035)
Mother's Educ - Sr HS	-0.021	0.013
	(0.089)	(0.087)
Mother's Educ - Univ	0.092	0.234
	(0.194)	(0.189)
Father's Educ - Jr HS	-0.023	0.004
	(0.039)	(0.038)
Father's Educ - Sr HS	0.060	0.114
	(0.071)	(0.069)
Father's Educ - Univ	0.164	0.121
	(0.167)	(0.163)
Per cap Income - 5 to 10k	0.015	0.015
	(0.029)	(0.028)
Per cap Income - more than 10k	0.050	0.044
	(0.033)	(0.032)
Observations	3342	3342
Cens. obs.	1730	1730
Chi-Square	211.983	157.746
Inverse Mills Ratio	0.096	0.046
S.E. of Inv Mills	0.076	0.075

Notes: Table displays coefficients and standard errors in brackets. * p<0.1 ** p<0.05 *** p<0.01. Excl. categories: not obese, father in household, father's and mother's education primary or less, lowest per capita parental income category, father's occupation unskilled worker, size of locality of residence less than 15k.

Table 12: Correlation between Expected Returns and Direct Costs of Schooling.

Dep Var: Expected Return	Coeff./(S.E.)	Coeff./(S.E.)
Mother's Educ - Jr HS	-0.009	0.011
	(0.034)	(0.030)
Mother's Educ - Sr HS	0.048	0.036
	(0.076)	(0.073)
Mother's Educ - Univ	0.168	0.115
	(0.192)	(0.158)
Father's Educ - Jr HS	0.001	0.027
	(0.035)	(0.032)
Father's Educ - Sr HS	0.066	0.054
	(0.061)	(0.058)
Father's Educ - Univ	-0.186	-0.054
	(0.144)	(0.136)
Per cap Income - 5 to 10k	0.022	-0.002
	(0.028)	(0.023)
Per cap Income - more than 10k	-0.007	-0.007
	(0.031)	(0.027)
GPA - second tercile	0.004	0.026
	(0.027)	(0.023)
GPA - top tercile	0.042	0.053**
	(0.028)	(0.024)
Distance to University	0.002	
	(0.002)	
Distance Squared	-0.000	
	(0.000)	
Tuition Costs	0.000	
	(0.000)	
Tuition Squared	0.000	
	(0.000)	
Tuition Above 750 Pesos		0.046
		(0.031)
Dist to Univ 20 to 40km		0.013
		(0.023)
Dist to Univ above 40km		0.043
		(0.028)
Observations	2327	3342
Censored Observations	1156	1730
Lambda	-0.086	-0.070
S.E. of Lambda	0.063	0.064

Notes: * p<0.1 ** p<0.05 *** p<0.01. Excl. categories: male, lowest GPA tercile, parents' education primary or less, per capita income less than 5000 pesos.

Robustness Checks B

Figure 10: The Cumulative Distribution Function of Costs for Different Income Classes.

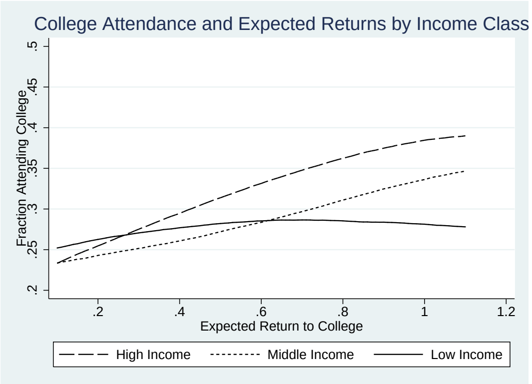


Figure 11: The Cumulative Distribution Function of Costs with 95% Confidence Intervals.

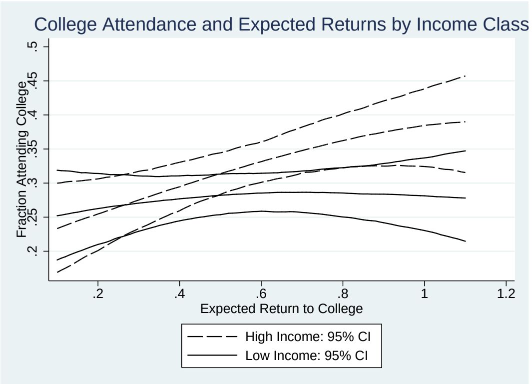


Table 13: Determinants of College Attendance: Total Household Income.

Dependent Variable	College Attendance		
	(1)	(2)	(3)
	Marg Eff	Marg Eff	Marg Eff
	(SE)	(SE)	(SE)
Expected Return to College	0.092***	0.078**	0.076**
	(0.033)	(0.034)	(0.034)
Prob of Work - Sr HS	0.032	0.013	0.012
	(0.087)	(0.085)	(0.077)
Prob of Work - College	-0.008	-0.001	0.023
	(0.101)	(0.099)	(0.089)
Var of Log Earn - Sr HS	-2.625	-3.016	-2.701
	(1.919)	(2.008)	(1.900)
Var of Log Earn - College	-0.310	0.036	0.029
	(2.351)	(2.291)	(2.092)
Female (d)	-0.055*	-0.059*	-0.044
	(0.029)	(0.033)	(0.032)
GPA - second tercile (d)		0.055*	0.057*
		(0.031)	(0.030)
GPA - top tercile (d)		0.187***	0.170***
		(0.038)	(0.047)
Father's Educ - Jr HS (d)		0.099**	0.078*
		(0.042)	(0.042)
Father's Educ - Sr HS (d)		0.151*	0.109
		(0.078)	(0.074)
Father's Educ - Univ (d)		0.547***	0.569***
		(0.120)	(0.142)
Mother's Educ - Jr HS (d)		0.100**	0.076*
		(0.040)	(0.039)
Mother's Educ - Sr HS (d)		0.203**	0.172*
		(0.099)	(0.101)
Mother's Educ - Univ (d)		0.196	0.234
		(0.209)	(0.208)
Total Fam Income - T2 (d)			0.025
			(0.028)
Total Fam Income -T3 (d)			0.060*
			(0.032)
Dist to Univ 20 to 40km (d)			-0.076***
			(0.028)
Dist to Univ above 40km (d)			-0.105***
			(0.030)
Tuition Above 750 Pesos (d)			-0.078**
			(0.038)
Observations	3342	3342	3342
Censored Obs	1730	1730	1730
Log Likelihood	-3041.971	-2990.349	-2975.200
Sample Sel: Corr betw Err	-0.487	-0.282	-0.061
Sample Sel: P-Val	0.055	0.314	0.835

Notes: Table displays marginal effects and standard errors in brackets. * p<0.1 ** p<0.05 *** p<0.01. Excl. categories: male, lowest GPA tercile, father's and mother's education primary or less, lowest family income tercile, distance to university less than 20 km and tuition less than 750 pesos.

Table 14: Excess Responsiveness of the Poor to Changes in Direct Costs (Distance to College): Per Capital Income and Wealth.

Dependent Variable	**College Attendance**		
	(1)	(2)	(3)
	Marg Eff (SE)	Marg Eff (SE)	Marg Eff (SE)
Univ 20 - 40km * Par Inc/Wealth Q1	-0.123**	-0.124**	-0.145*
	(0.054)	(0.053)	(0.075)
Univ 20 - 40km * Par Inc/Wealth Q1 * High Exp Ret			0.023
			(0.148)
Univ 20 - 40km * Par Inc/Wealth Q2	-0.009	-0.006	0.014
	(0.073)	(0.073)	(0.109)
Univ 20 - 40km * Par Inc/Wealth Q2 * High Exp Ret			-0.042
			(0.136)
Univ 20 - 40km * Par Inc/Wealth Q3	-0.078	-0.081	-0.064
	(0.062)	(0.060)	(0.095)
Univ 20 - 40km * Par Inc/Wealth Q3 * High Exp Ret			-0.018
			(0.141)
Univ 20 - 40km * Par Inc/Wealth Q4	0.074	0.071	0.116
	(0.073)	(0.072)	(0.109)
Univ 20 - 40km * Par Inc/Wealth Q4 * High Exp Ret			-0.065
			(0.115)
Univ > 40km * Par Inc/Wealth Q1	-0.064	-0.064	-0.020
	(0.053)	(0.052)	(0.078)
Univ > 40km * Par Inc/Wealth Q1 * High Exp Ret			-0.127
			(0.096)
Univ > 40km * Par Inc/Wealth Q2	-0.030	-0.030	-0.029
	(0.072)	(0.071)	(0.102)
Univ > 40km * Par Inc/Wealth Q2 * High Exp Ret			-0.006
			(0.147)
Univ > 40km * Par Inc/Wealth Q3	-0.178***	-0.177***	-0.214**
	(0.058)	(0.057)	(0.085)
Univ > 40km * Par Inc/Wealth Q3 * High Exp Ret			0.106
			(0.235)
Univ > 40km * Par Inc/Wealth Q4	-0.088	-0.087	-0.177**
	(0.064)	(0.063)	(0.076)
Univ > 40km * Par Inc/Wealth Q4 * High Exp Ret			0.266
			(0.188)
Interaction of Par Inc/Weath Quartiles and High Ret	Yes	Yes	Yes
Controls: Expected Return, Exp Log Earn, Prob of Work and Var of Log Earn	No	Yes	Yes
Controls: GPA, Par Inc/Wealth and Educ, Sex, State FE	Yes	Yes	Yes
Observations	3342	3342	3342
Uncensored Obs	1612	1612	1612
Log Likelihood	-2981.146	-2978.124	-2968.895
Sample Sel: Corr betw Err	-0.208	-0.177	-0.209
Sample Sel: P-Val	0.488	0.565	0.504

Notes: Table displays marginal effects and standard errors in brackets. * p<0.1, ** p<0.05, *** p<0.01. Excl. categories: male, lowest GPA tercile, parents' education primary or less, lowest parental income/wealth quartile, interactions of distance to university less than 20km with parental income/wealth and low expected return interacted with parental income/wealth quartiles.

Table 15: Excess Responsiveness of the Poor to Changes in Direct Costs (Distance to College): Total Household Income.

Dependent Variable	College Attendance		
	(1)	(2)	(3)
	Marg Eff	Marg Eff	Marg Eff
	(SE)	(SE)	(SE)
Univ 20 - 40km * Fam Income Tercile 1 (d)	-0.107**	-0.108**	-0.085
	(0.042)	(0.042)	(0.057)
Univ 20 - 40km * Fam Inc 1 * High Exp Ret (d)			-0.058
			(0.081)
Univ 20 - 40km * Fam Income Tercile 2 (d)	-0.019	-0.022	-0.034
	(0.051)	(0.050)	(0.075)
Univ 20 - 40km * Fam Inc 2 * High Exp Ret (d)			0.020
			(0.112)
Univ 20 - 40km * Fam Income Tercile 3 (d)	0.102	0.095	0.115
	(0.066)	(0.065)	(0.093)
Univ 20 - 40km * Fam Inc 3 * High Exp Ret (d)			-0.038
			(0.106)
Univ > 40km * Fam Income Tercile 1 (d)	-0.066*	-0.068*	-0.053
	(0.040)	(0.039)	(0.058)
Univ > 40km * Fam Inc 1 * High Exp Ret (d)			-0.039
			(0.082)
Univ > 40km * Fam Income Tercile 2 (d)	-0.108**	-0.115**	-0.160**
	(0.050)	(0.047)	(0.070)
Univ > 40km * Fam Inc 2 * High Exp Ret (d)			0.115
			(0.158)
Univ > 40km * Fam Income Tercile 3 (d)	0.002	-0.001	-0.127
	(0.072)	(0.071)	(0.083)
Univ > 40km * Fam Inc 3 * High Exp Ret (d)			0.323*
			(0.185)
Fam Inc 1 * High Exp Ret (d)			-0.111**
			(0.048)
Fam Inc 2 * High Exp Ret (d)			0.028
			(0.055)
Fam Inc 3 * High Exp Ret (d)			0.007
			(0.053)
Controls for Expected Return, Exp Log Earn, Prob of Work and Var of Log Earn	No	Yes	Yes
Controls: GPA, Fam Income and Educ, Sex, State FE	Yes	Yes	Yes
Observations	3342	3342	3342
Censored Obs	1730	1730	1730
Log Likelihood	-2985.843	-2981.618	-2960.931
Sample Sel: Corr betw Err	-0.133	-0.096	-0.144
Sample Sel: P-Val	0.648	0.748	0.623

Notes: Table displays marginal effects and standard errors in brackets. * p<0.1, ** p<0.05, *** p<0.01. Excl. categories: male, lowest GPA tercile, parents' education primary or less, lowest family income tercile, interactions of distance to university of less than 20km with family income and low expected return interacted with family income.

Table 16: Excess Responsiveness of the Poor to Changes in Direct Costs (Distance to College): Total Household Income and Wealth.

Dependent Variable	College Attendance		
	(1)	(2)	(3)
	Marg Eff (SE)	Marg Eff (SE)	Marg Eff (SE)
Univ 20 - 40km * Fam Inc/Wealth Q1 (d)	-0.098** (0.042)	-0.097** (0.041)	-0.093** (0.042)
Univ 20 - 40km * Fam Inc/Wealth Q1 * High Exp Ret (d)			-0.078 (0.069)
Univ 20 - 40km * Fam Inc/Wealth Q2 (d)	-0.071 (0.045)	-0.075* (0.044)	-0.056 (0.052)
Univ 20 - 40km * Fam Inc/Wealth Q2 * High Exp Ret (d)			0.010 (0.093)
Univ 20 - 40km * Fam Inc/Wealth Q3 (d)	-0.050 (0.047)	-0.049 (0.046)	-0.070 (0.049)
Univ 20 - 40km * Fam Inc/Wealth Q3 * High Exp Ret (d)			0.041 (0.100)
Univ 20 - 40km * Fam Inc/Wealth Q4 (d)	-0.066 (0.043)	-0.064 (0.043)	-0.081* (0.045)
Univ 20 - 40km * Fam Inc/Wealth Q4 * High Exp Ret (d)			0.171 (0.109)
Univ > 40km * Fam Inc/Wealth Q1 (d)	-0.112*** (0.042)	-0.114*** (0.040)	-0.110*** (0.042)
Univ > 40km * Fam Inc/Wealth Q1 * High Exp Ret (d)			-0.073 (0.072)
Univ > 40km * Fam Inc/Wealth Q2 (d)	-0.081* (0.044)	-0.081* (0.043)	-0.061 (0.050)
Univ > 40km * Fam Inc/Wealth Q2 * High Exp Ret (d)			0.005 (0.095)
Univ > 40km * Fam Inc/Wealth Q3 (d)	-0.085* (0.049)	-0.087* (0.048)	-0.100* (0.051)
Univ > 40km * Fam Inc/Wealth Q3 * High Exp Ret (d)			-0.047 (0.094)
Univ > 40km * Fam Inc/Wealth Q4 (d)	-0.060 (0.049)	-0.058 (0.048)	-0.061 (0.048)
Univ > 40km * Fam Inc/Wealth Q4 * High Exp Ret (d)			0.155 (0.121)
Interaction of Fam Inc/Weath Quartiles and High Ret	Yes	Yes	Yes
Controls: Expected Return, Exp Log Earn, Prob of Work and Var of Log Earn	No	Yes	Yes
Controls: GPA, Fam Inc/Wealth and Educ, Sex, State FE	Yes	Yes	Yes
Observations	3342	3342	3342
Censored Obs	1730	1730	1730
Log Likelihood	-2982.507	-2979.743	-2974.615
Sample Sel: Corr betw Err	-0.130	-0.097	-0.030
Sample Sel: P-Val	0.657	0.747	0.923

Notes: Table displays marginal effects and standard errors in brackets. * p<0.1, ** p<0.05, *** p<0.01. Excl. categories: male, lowest GPA tercile, parents' education primary or less, lowest family income/wealth quartile, interactions of distance to university less than 20km with family income/wealth and low expected return interacted with family income/wealth quartiles.

Table 17: Excess Responsiveness of the Poor to Changes in Direct Costs (Tuition Costs): Per Capital Income and Wealth.

Dependent Variable	College Attendance		
	(1)	(2)	(3)
	Marg Eff (SE)	Marg Eff (SE)	Marg Eff (SE)
Tuition > 750 * Par Inc/Wealth Q1	-0.064	-0.067	-0.000
	(0.048)	(0.047)	(0.071)
Tuition > 750 * Par Inc/Wealth Q1 * High Exp Ret			-0.148*
			(0.084)
Tuition > 750 * Par Inc/Wealth Q2	-0.037	-0.037	-0.006
	(0.065)	(0.064)	(0.095)
Tuition > 750 * Par Inc/Wealth Q2 * High Exp Ret			-0.055
			(0.118)
Tuition > 750 * Par Inc/Wealth Q3	-0.051	-0.055	-0.087
	(0.062)	(0.061)	(0.094)
Tuition > 750 * Par Inc/Wealth Q3 * High Exp Ret			0.038
			(0.137)
Tuition > 750 * Par Inc/Wealth Q4	0.069	0.066	0.117
	(0.070)	(0.070)	(0.104)
Tuition > 750 * Par Inc/Wealth Q4 * High Exp Ret			-0.106
			(0.101)
Interaction of Par Inc/Weath Quartiles and High Ret	Yes	Yes	Yes
Controls: Expected Return, Exp Log Earn, Prob of Work and Var of Log Earn	No	Yes	Yes
Controls: GPA, Par Inc/Wealth and Educ, Sex, State FE	Yes	Yes	Yes
Observations	3342	3342	3342
Uncensored Obs	1612	1612	1612
Log Likelihood	-2987.524	-2984.668	-2972.787
Sample Sel: Corr betw Err	-0.329	-0.309	-0.326
Sample Sel: P-Val	0.236	0.275	0.247

Notes: Table displays marginal effects and standard errors in brackets. * p<0.1, ** p<0.05, *** p<0.01. Excl. categories: male, lowest GPA tercile, parents' education primary or less, lowest parental income/wealth quartile, interactions of tuition costs less than 750 pesos with parental income/wealth and low expected return interacted with parental income/wealth quartiles.

Table 18: Excess Responsiveness of the Poor to Changes in Direct Costs (Tuition Costs): Total Household Income.

Dependent Variable	College Attendance		
	(1)	(2)	(3)
	Marg Eff	Marg Eff	Marg Eff
	(SE)	(SE)	(SE)
Tuition > 750 * Fam Income T1 (d)	-0.070	-0.078	-0.038
	(0.049)	(0.049)	(0.067)
Tuition > 750 * Fam Income T1 * High Exp Ret (d)			-0.084
			(0.076)
Tuition > 750 * Fam Income T2 (d)	-0.043	-0.045	-0.024
	(0.050)	(0.050)	(0.071)
Tuition > 750 * Fam Income T2 * High Exp Ret (d)			-0.048
			(0.086)
Tuition > 750 * Fam Income T3 (d)	-0.044	-0.050	-0.087
	(0.052)	(0.052)	(0.069)
Tuition > 750 * Fam Income T3 * High Exp Ret (d)			0.058
			(0.100)
Fam Inc 1 * High Exp Ret (d)			-0.090
			(0.061)
Fam Inc 2 * High Exp Ret (d)			0.060
			(0.069)
Fam Inc 3 * High Exp Ret (d)			0.003
			(0.062)
Controls for Expected Return, Exp Log Earn, Prob of Work and Var of Log Earn	No	Yes	Yes
Controls: GPA, Fam Income and Educ, Sex, State FE	Yes	Yes	Yes
Observations	3342	3342	3342
Censored Obs	1730	1730	1730
Log Likelihood	-2991.555	-2987.575	-2978.676
Sample Sel: Corr betw Err	-0.321	-0.318	-0.315
Sample Sel: P-Val	0.295	0.310	0.264

Notes: Table displays marginal effects and standard errors in brackets. * p<0.1, ** p<0.05, *** p<0.01. Excl. categories: male, lowest GPA tercile, parents' education primary or less, family income lowest tercile, interactions of tuition costs less than 750 pesos with family income and low expected return interacted with family income.

Table 19: Excess Responsiveness of the Poor to Changes in Direct Costs (Tuition Costs): Total Household Income and Wealth.

Dependent Variable	College Attendance		
	(1)	(2)	(3)
	Marg Eff	Marg Eff	Marg Eff
	(SE)	(SE)	(SE)
Tuition > 750 * Fam Inc/Wealth Q1 (d)	-0.060	-0.063	0.028
	(0.059)	(0.059)	(0.088)
Tuition > 750 * Fam Inc/Wealth Q1 * High Exp Ret (d)			-0.141*
			(0.080)
Tuition > 750 * Fam Inc/Wealth Q2 (d)	-0.093*	-0.097*	-0.066
	(0.053)	(0.053)	(0.073)
Tuition > 750 * Fam Inc/Wealth Q2 * High Exp Ret (d)			-0.062
			(0.093)
Tuition > 750 * Fam Inc/Wealth Q3 (d)	-0.017	-0.016	0.017
	(0.058)	(0.058)	(0.081)
Tuition > 750 * Fam Inc/Wealth Q3 * High Exp Ret (d)			-0.067
			(0.093)
Tuition > 750 * Fam Inc/Wealth Q4 (d)	-0.031	-0.035	-0.095
	(0.054)	(0.054)	(0.070)
Tuition > 750 * Fam Inc/Wealth Q4 * High Exp Ret (d)			0.123
			(0.108)
Fam Inc/Wealth Q1 * High Exp Ret (d)			0.008
			(0.081)
Fam Inc/Wealth Q2 * High Exp Ret (d)			-0.049
			(0.072)
Fam Inc/Wealth Q3 * High Exp Ret (d)			0.069
			(0.078)
Fam Inc/Wealth Q4 * High Exp Ret (d)			-0.032
			(0.060)
Controls for Expected Return, Exp Log Earn, Prob of Work and Var of Log Earn	No	Yes	Yes
Controls: GPA, Fam Income and Educ, Sex, State FE	Yes	Yes	Yes
Observations	3342	3342	3342
Censored Obs	1730	1730	1730
Log Likelihood	-2995.767	-2993.140	-2982.624
Sample Sel: Corr betw Err	-0.318	-0.313	-0.282
Sample Sel: P-Val	0.250	0.264	0.311

Notes: Table displays marginal effects and standard errors in brackets. * p<0.1, ** p<0.05, *** p<0.01. Excl. categories: male, lowest GPA tercile, parents' education primary or less, lowest family income/wealth quartile, interactions of tuition costs less than 750 pesos with family income/wealth and low expected return interacted with family income/wealth quartiles.

Table 20: Time Preference of Different Per Capita Income Categories: Total Household Income.

	Total Family Income Category				
	Tercile 1 (low)	Tercile 2	Tercile 3	Compare	
	(1)	(2)	(3)	(1)-(2)	(1)-(3)
	Mean	Mean	Mean	Diff	Diff
	(SD)	(SD)	(SD)	(P-Val)	(P-Val)
Intertemp Behavior: Health					
Smoke	0.03	0.02	0.03	0.01	0.00
	(0.18)	(0.16)	(0.18)	(0.450)	(0.946)
Drink Alcohol					
Yes	0.12	0.11	0.15	0.01	-0.03
	(0.33)	(0.31)	(0.36)	(0.511)	(0.111)
$\geq$ 2/week	0.04	0.04	0.05	-0.00	-0.01
	(0.19)	(0.19)	(0.21)	(0.837)	(0.416)
How Use 3000 Pesos?					
Immediate Consumption	0.16	0.19	0.21	-0.03	-0.05
(Alternative: Save/Invest)	(0.36)	(0.39)	(0.41)	(0.136)	(0.024)
Observations	585	523	504		

Notes: Columns 1 to 3 display means and standard deviations in brackets. Columns 4 and 5 display the difference of (1)-(2) and (1)-(3), respectively, and the p-value of the difference in brackets.

Table 21: Counterfactual Policy Experiments: Total Household Income.

Policy Change	Individuals Changing College Attendance Decision: Change in Overall Attendance Rate in pp (in %) (p-value)	Individuals Changing College Attendance Decision: **Marginal** Expected Return (MTE)	Individuals Attending College: **Average** Expected Return (TTE)	**Diff MTE-TTE** (p-value)
Decrease Dist by 20km				
for all	1pp (4%) (p-val 0.04)	0.89	0.71	0.18 (0.17)
for very poor	0.4pp (2%) (p-val 0.12)	0.88	0.71	0.16 (0.29)
for very poor and very able	0.2pp (1%) (p-val 0.09)	0.90	0.71	0.19 (0.18)
Decrease Tuition by 10%				
for all	0.3pp (1.5%) (p-val 0.49)	0.85	0.71	0.14 (0.27)
for very poor	0.3pp (1.5%) (p-val 0.28)	0.79	0.71	0.08 (0.40)
for very poor and very able	0.3pp (1.5%) (p-val 0.29)	0.81	0.71	0.10 (0.36)

Table 22: Determinants of College Attendance: Including Ability-Return Interactions.

Dependent Variable	College Attendance	
	(1)	(2)
	Marg Eff	Marg Eff
	(SE)	(SE)
Expected Return to College	0.118**	0.116**
	(0.050)	(0.050)
Exp Return * GPA 2	-0.059	-0.061
	(0.074)	(0.071)
Exp Return * GPA 3	-0.087	-0.081
	(0.074)	(0.071)
Prob of Work - Sr HS	0.004	0.010
	(0.081)	(0.078)
Prob of Work - College	0.034	0.025
	(0.093)	(0.090)
Var of Log Earn - Sr HS	-3.063	-2.794
	(1.973)	(1.916)
Var of Log Earn - College	0.212	0.033
	(2.180)	(2.106)
Female (d)	-0.046	-0.044
	(0.032)	(0.032)
GPA - second tercile (d)	0.097	0.102
	(0.063)	(0.063)
GPA - top tercile (d)	0.241***	0.233***
	(0.074)	(0.075)
Father's Educ - Jr HS (d)	0.074*	0.079*
	(0.042)	(0.042)
Father's Educ - Sr HS (d)	0.096	0.105
	(0.075)	(0.074)
Father's Educ - Univ (d)	0.570***	0.564***
	(0.132)	(0.144)
Mother's Educ - Jr HS (d)	0.073*	0.075*
	(0.039)	(0.039)
Mother's Educ - Sr HS (d)	0.176*	0.175*
	(0.101)	(0.101)
Mother's Educ - Univ (d)	0.215	0.225
	(0.206)	(0.207)
Dist to Univ 20 to 40km (d)	-0.076***	-0.076***
	(0.029)	(0.028)
Dist to Univ above 40km (d)	-0.106***	-0.105***
	(0.031)	(0.030)
Tuition Above 750 Pesos (d)	-0.083**	-0.080**
	(0.039)	(0.038)
Per cap Income - 5 to 10k (d)	0.054*	
	(0.031)	
Per cap Income - more than 10k (d)	0.120***	
	(0.037)	
Total Family Income T2 (d)		0.026
		(0.028)
Total Family Income T3 (d)		0.061*
		(0.033)
Observations	3342	3342
Censored Obs	1730	1730
Log Likelihood	-2972.170	-2974.427
Sample Sel: Corr betw Err	-0.140	-0.069
Sample Sel: P-Val	0.634	0.816

Notes: Table displays marginal effects and standard errors in brackets. * p<0.1 ** p<0.05 *** p<0.01. Excl. categories: male, lowest GPA tercile, father's and mother's education primary or less, lowest household income tercile, distance to university less than 20 km and tuition less than 750 pesos.

Table 23: Time Preferences: Comparison between Per Cap Income $\leq 5k$ and $\geq 5k$ and $\leq 10k$.

Dependent Variable	Smoke (Yes/No)	Alcohol (Yes/No)	Alcohol ($\geq$ 2/week)	Immediate Consumption
Per cap Income - less than 5k	-0.001	0.014	0.012	0.002
	(0.010)	(0.019)	(0.012)	(0.023)
Female	-0.050***	-0.061***	-0.027**	-0.007
	(0.009)	(0.017)	(0.011)	(0.021)
Age	0.013***	0.019**	0.009**	0.023**
	(0.004)	(0.008)	(0.005)	(0.009)
Chiapas	-0.020	-0.066***	-0.006	-0.170***
	(0.012)	(0.023)	(0.014)	(0.027)
Guanajuato	-0.021	-0.102	-0.017	0.166
	(0.065)	(0.119)	(0.075)	(0.144)
Guerrero	0.012	0.003	0.034*	-0.095***
	(0.016)	(0.029)	(0.018)	(0.035)
Michoacan	0.005	0.230***	0.058***	-0.013
	(0.017)	(0.032)	(0.020)	(0.038)
Veracruz	0.015	-0.040	0.013	-0.055
	(0.016)	(0.030)	(0.019)	(0.036)
Observations	1340	1340	1340	1340
R-squared	0.032	0.080	0.019	0.039

Notes: Table displays coefficients and standard errors in brackets. * p<0.1 ** p<0.05 *** p<0.01. Excl. category: male. All regressions contain state fixed effects.

Table 24: Time Preferences: Comparison between Per Cap Income $\leq 5k$ and $\leq 10k$.

Dependent Variable	Smoke (Yes/No)	Alcohol (Yes/No)	Alcohol ($\geq$ 2/week)	Immediate Consumption
Per cap Income - less than 5k	0.003	-0.035	0.012	-0.015
	(0.012)	(0.023)	(0.014)	(0.027)
Female	-0.044***	-0.084***	-0.039***	-0.005
	(0.010)	(0.019)	(0.012)	(0.022)
Age	0.012***	0.010	-0.000	0.007
	(0.004)	(0.009)	(0.005)	(0.010)
Chiapas	-0.002	-0.037	0.003	-0.188***
	(0.013)	(0.025)	(0.015)	(0.029)
Guanajuato	-0.006	0.097	-0.011	0.315*
	(0.075)	(0.146)	(0.089)	(0.171)
Guerrero	0.020	0.073**	0.073***	-0.094**
	(0.016)	(0.031)	(0.019)	(0.037)
Michoacan	0.022	0.233***	0.052**	-0.033
	(0.018)	(0.035)	(0.022)	(0.041)
Veracruz	0.009	-0.026	0.004	-0.091**
	(0.017)	(0.032)	(0.020)	(0.038)
Observations	1224	1224	1224	1224
R-squared	0.027	0.075	0.028	0.044

Notes: Table displays coefficients and standard errors in brackets. * p<0.1 ** p<0.05 *** p<0.01. Excl. category: male. All regressions contain state fixed effects.