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

Katja Maria Kaufmann*

Department of Economics and IGIER, Bocconi University

March 2009

Abstract

When examining reasons for low school attendance researchers face an important identification problem: On the one hand people might expect low returns to schooling 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 quantitative expectations of future returns to schooling, which can be shown to affect their schooling decisions. I use these data on Mexican high school graduates to analyze the causes and implications of the steep income gradient in college enrollment in Mexico. Data on people's expected returns and on their schooling decisions allow me to directly estimate and compare cost distributions of poor and rich individuals. I find that poor individuals require significantly higher expected returns to be induced to attend college, implying that they face higher costs than individuals with wealthy parents. I then test predictions of a simple model of college attendance choice in the presence of credit constraints, using parental income and wealth as a proxy for the household's (unobserved) interest rate. I find that poor individuals with high expected returns are particularly responsive to changes in direct costs such as tuition, which is consistent with credit constraints playing an important role. Evaluating potential welfare implications by applying the Local Instrumental Variables approach of Heckman and Vytlacil (2005) to my model, I find that a sizeable fraction of poor individuals would change their decision and attend in response to a reduction in the interest rate. Individuals at the margin have higher expected returns than the individuals already attending college, which suggests that policies such as governmental student loan programs could lead to large welfare gains.

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

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

---

*Address: Via Roentgen 1, 20136 Milano, Italy, e-mail: *katja.kaufmann@unibocconi.it*. I would like to thank Orazio Attanasio, Aprajit Mahajan, John Pencavel and Luigi Pistaferri for their advice and support, and Manuela Angelucci, David Card, Pedro Carneiro, Giacomo De Giorgi, Christina Gathmann, Caroline Hoxby, Seema Jayachandran, Michael Lovenheim, Thomas MaCurdy, Shaun McRae, Edward Miguel, Sriniketh Nagavarapu, Marta Rubio-Codina, Alejandrina Salcedo, Alessandro Tarozzi, Frank Wolak and Joanne Yoong for insightful discussions and comments. I am thankful also to conference participants at the Northeast Universities Development Consortium Conference (NEUDC) 2007 and at the CEPR conference on the Economics of Education 2009, and to seminar participants at Bocconi University, at the Collegio Carlo Alberto, at the Einaudi Institute for Economics and Finance (EIEF, Rome), at the Institute for International Economic Studies (IIES, Stockholm), at the Stanford Economic Applications Seminar and the Stanford Labor and Development Reading Groups, at University College London (UCL) and at Universitat Pompeu Fabra (UPF, Barcelona) for helpful comments. All remaining errors are of course my own. This project was supported by the Taube Scholarship Fund Fellowship (SIEPR) and the Sawyer Seminar Fellowship of the Center for the Study of Poverty and Inequality. Previous versions of this paper circulated under "Marginal Returns to Schooling, Credit Constraints, and Subjective Expectations of Earnings".

1 Introduction

In both developed and developing countries there is a strong association between children's college attendance rates and parental income. For example, in the U.S. the poorest 40% of the relevant age group (18 to 24 years old) represent around 20% of the student body, while the richest 20% constitute 45%. For Mexico, the country I will be studying in this paper, the poorest 40% represent only 8% of the student body. This is low even compared to other Latin American countries. The richest 20% on the other hand constitute 60% of the student body. In addition overall college enrollment is particularly low in Mexico.[1] These empirical facts might reflect an important welfare loss if returns to education are high, but people cannot take advantage of them because they are credit constrained. 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 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 quantitative expectations of future returns to college as well as on their college attendance choice.

A traditional explanation for the income gradient in college attendance is credit constraints. Suppose that credit markets are imperfect in that banks only lend to individuals with collateral. Since college attendance involves direct costs (such as tuition), individuals from poor families, who are unable to cover such costs with parental income or with borrowed funds due to lack of collateral, will choose not to attend college even in the presence of high expected returns.[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 expected returns from human capital investment. Several papers in the literature, such as Cameron and Heckman (1998), Cameron and Heckman (2001) and Carneiro and Heckman (2002), attribute differences in college attendance rates between poor and rich in the US to differences in "college readiness". As stated in Carneiro and Heckman (2002), "most of the family income gap in enrollment is due to long-run factors that produce abilities needed to benefit from participation in college." They disprove the importance of credit constraints in the U.S. by showing that once one controls for ability and parental background measures (which proxy for returns to college and

[1] A strong correlation between children's educational attainment and parental resources is well-documented for most 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) for the case of Latin America. In Appendix C, I compare several 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 (see table 9) 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. But because such contracts are illegal, banks may choose to lend only to individuals who offer collateral to be seized in case of default.

preferences), parental income at the time of college attendance ceases to have a significant effect on the attendance decision. I cannot show this in my data. Nevertheless, it would be premature to conclude that this proves the importance of credit constraints.

Consider the conventional model of educational choices under uncertainty. In such a model, the decision to attend college depends on expected returns and risk from investing in college education, preferences, and potentially credit constraints. All these determinants are at least partly unobserved by the econometrician, posing an important identification problem (see, e.g., Manski (2004) and Cunha and Heckman (2006)). I address this identification problem using a particularly suitable data set with information on individuals' subjective expectations of earnings and perceived earnings risk.[3]

The existing "credit constraints" literature derives measures of earnings expectations using earnings realizations. This approach has the following problems. First, one has to make assumptions about the individuals' information sets as well as the mechanisms behind how they form expectations. These assumptions include whether earnings shocks were anticipated at the time of the choice (which is particularly problematic if large and unpredictable earnings shocks are the norm, as they are in developing countries) and whether people have precise information about their own ability. Second, computing expected returns to college requires constructing expected earnings in a counterfactual state. Thus, researchers have to make assumptions about how individuals form these expectations, i.e. whether and how they solve the problem that the observed earnings are from individuals who have self-selected into schooling. Another potentially important determinant of college attendance is perceived earnings risk. Taking into account earnings risk is relevant for the credit constraints issue, as it might not be optimal for poor individuals to attend college, despite high expected returns, if they face particularly risky college earnings.[4] Most papers in the literature neglect the importance of risk as a determinant of educational choice and assume no uncertainty or certainty equivalence (see, e.g., Cameron and Taber (2004) and Carneiro, Heckman, and Vytlacil (2005)).

If there are differences in expected returns or perceived earnings risk, which are correlated with parental income, this could lead to a spurious positive correlation between parental income and college attendance. Having data on each individual's (subjective) distribution of future earnings (for both high school and college as the highest degree) enables me to address this concern directly. Since what matters for the college attendance decision is each individual's perception of her own

[3]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.

[4]Papers that take into account this determinant include Padula and Pistaferri (2001) and Belzil and Hansen (2002). Only the former paper employs subjective expectations but aggregates perceived employment risk for education groups to analyze whether the implicit return to education is underestimated when not taking into account effects of different schooling levels on later earnings and employment risk.

skills and how these skills affect her future earnings, these data ideally provide respondent's earnings expectations and perceptions of earnings risk conditional on their information sets at the time of the decision.

The first finding of this paper is that even though expected returns to college are important determinants of college attendance decisions, they are not sufficient to explain the poor's low college attendance rates.[5] Data on people's expected returns and on their schooling decisions allow me to directly estimate and compare cost distributions of poor and rich individuals. I find that poor individuals require significantly higher expected returns to be induced to attend college, implying that they face higher costs than individuals with wealthy parents.

To understand the role of different cost components, I test predictions of a simple model of college attendance choice in the presence of credit constraints, using parental income and wealth as proxies for the unobserved household's interest rate. I find that poor individuals with high expected returns are particularly responsive to changes in direct costs such as tuition, which is consistent with credit constraints playing an important role.

Evaluating potential welfare implications by applying the Local Instrumental Variables approach of Heckman and Vytlacil (2005) to my model, I find that a sizeable fraction of poor individuals would change their decision and attend in response to a reduction in the interest rate. Individuals at the margin have higher expected returns than the individuals already attending college, which suggests that they might be prevented from attending because they face high borrowing costs.

The findings of this paper suggest that credit constraints could be one of the driving forces of Mexico's large inequalities in access to higher education and low overall enrollment rates. Mexico's low government funding for student loans and fellowships for higher education, which is low even by Latin American standards, is consistent with this view. The results of my policy experiments suggest that the introduction of a governmental student loan program could lead to large welfare gains by removing obstacles to human capital accumulation and fostering Mexico's development

[5] The following three related papers, by Attanasio and Kaufmann (2007), Nguyen (2008) and Jensen (2008), also find that "perceived" returns to schooling matter for people's schooling decisions: Jensen (2008) finds that the students in his sample of 8th graders 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 among these students. Nguyen (2008) finds that informing a random subset of a sample of students in Madagascar about high returns to schooling increases their attendance rates and their test scores. Attanasio and Kaufmann (2007) address two additional questions concerning the link between schooling choice and expectations (using the same Mexican survey data as this paper). In addition to expected returns they also take into account a potential role of perceived earnings and employment risk for different schooling degrees. Second, they have data on mothers' expectations about potential earnings of their children as well as adolescents' own expectations about their future earnings and can thus shed some light on whose expectations matter in the intra-household decision process for two important educational choices, that is high school and college attendance. They find that mothers' perceptions of risk are significant predictors of high school decisions. In terms of the college attendance decision on the other hand, only adolescents' expectations turn out significant and only expected returns seem to matter, while risk perceptions appear less relevant with respect to this decision.

and growth.

It is important to note that the evidence above could be consistent with other factors also driving the poor's low college attendance rates. One alternative explanation could be heterogeneity in time preferences. Even if none of the empirical patterns found in the data were driven by credit constraints, high expected returns of a sizable fraction of non-attenders could still justify government policies such as student loan programs, if there are externalities from college attendance and social returns are correlated with private returns or if people have time-inconsistent preferences, e.g. they become more patient when getting older.

2 Model of College Attendance Choice

In the Mexican case parental income and wealth remain significant in a reduced-form regression of the decision to attend college, even after controlling for an extensive set of other parental background measures and individual characteristics such as cognitive skills (see section 4). Thus the questions remains if income has a causal effect on college attendance suggesting that credit constraints are an obstacle for the poor to attend, or if the significance of income is due to an omitted variable bias. One potential determinant that has been neglected in this analysis of credit constraints are people's subjective expectations about future returns to schooling.

I use a simple model to illustrate that people's (subjective) expectations about future earnings are likely to be an important determinant in the college attendance decision and to show how using data that elicits people's subjective expectations directly can relax strong assumptions of conventional approaches about people's information sets. The model enables me to derive testable implications of credit constraints and to perform counterfactual policy experiments, for example to evaluate the welfare implications of a governmental student loan program.

I model the college attendance decision of a high school graduate at age 18 as follows: An individual decides to attend college, $S = 1$, if the expected present value of college earnings ($EPV(Y_i^1)$) minus the expected present value of high school earnings ($EPV(Y_i^0)$) is larger than the costs of attending college (C_i) (direct costs, i.e. tuition, transportation, room and board if necessary, etc. and monetized psychological costs or benefits):

$$\begin{aligned} S_i^* &= EPV(Y_i^1) - EPV(Y_i^0) - C_i \\ &= \sum_{a=22}^{\infty} \frac{E(Y_{ia}^1)}{(1+r_i)^{a-18}} - \sum_{a=18}^{\infty} \frac{E(Y_{ia}^0)}{(1+r_i)^{a-18}} - C_i \geq 0, \end{aligned} \tag{1}$$

where i denotes the individual, a age of the individual, $E(Y_{ia}^1)$ represents expected college earnings, $E(Y_{ia}^0)$ expected high school earnings, C_i direct costs and r_i the interest rate that individual i faces. According to this model we 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), so 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,* $X_{i(a+1)} = X_{ia} + 1$*, until retirement. Returns to experience are the same in both schooling states and for each individual.*

The assumption of log earnings being additively separable in education and experience is commonly used in the literature (compare, e.g., Mincer (1974)). Assuming a deterministic relationship for experience is equivalent to using potential labor market experience as a proxy for actual experience in a Mincer earnings regression. I abstract from work during studying, and thus 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. In a similar framework, Carneiro, Heckman, and Vytlacil (2005) also make the assumption about returns to experience being the same in both schooling states and for all individuals.

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 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. As shown in section 4, perceived earnings and unemployment risk are not significant in a regression of college attendance choice (while they are for the decision to attend high school, see Attanasio and Kaufmann (2007)). For this reason and because taking into account risk would significantly complicate the model, I do not take into account risk considerations here.

Assumption 4 *Individuals have a common discount factor.*

The literature on credit constraints in general faces the problem of how to distinguish heterogeneity in borrowing rates from heterogeneity in time preferences. For example, Cameron and Taber (2004) assume one common discount factor for every individual and normalize the interest rate of the unconstrained individuals to be equal to this discount factor. If high-return individuals do not attend college because of a high discount rate, a policy intervention would have to be justified by high social returns to college that are correlated with private returns or with time-inconsistent preferences, e.g. people becoming more patient when getting older.

Assumption 5 *The problem is infinite horizon.*

In the following I will discuss how I use data on subjective expectations of earnings in this model and how this compares to conventional approaches that use earnings realizations instead. Assume that the economic model generating the data for the two potential outcomes ($j = 0, 1$) over the whole life-cycle ($a = 18, ..., A$) is of the form of the so called "Generalized Roy Model":

$$\begin{aligned} \ln Y_{ia}^{j} &= \alpha_j + \beta_j' X_i + \gamma(a - s^j - 6) + U_{ia}^{j} \\ &= \alpha_j + \beta_j' X_i + \gamma(a - s^j - 6) + \theta_j' f_i + \epsilon_{ia}^{j}, \end{aligned} \tag{2}$$

where $j = 0$ denotes high school degree (12 years of schooling, $s^0 = 12$), and $j = 1$ college degree (16 years of schooling, $s^1 = 16$). In terms of observable variables a labels age, A age at retirement and $(a - s^j - 6)$ represents potential labor market experience, while X denotes other observable time-invariant variables. I assume that log earnings profiles are parallel in experience across schooling levels. Thus the coefficient on experience is the same in both schooling states, $\gamma_1 = \gamma_0 = \gamma$ (compare Mincer (1974) and Carneiro, Heckman, and Vytlacil (2005)).

U^j represents the unobservables in the potential outcome equation, which are composed of a part that is anticipated 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 $j = 0, 1$.[6] f_i is the individual's skill vector which captures for example cognitive and social skills, and θ_j is a vector of (beliefs over future) skill prices. Both f_i and θ_j are in the information set of the individual, while they are –at least in part– unobservable

[6] 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* returns. Therefore I distinguish between a part of the idiosyncratic returns that is anticipated and (potentially) acted upon at the time of the schooling decision and a part that is not anticipated and can thus not be acted upon (compare Cunha, Heckman, and Navarro (2005) whose goal is to understand, which part of idiosyncratic returns is anticipated). Subjective expectations incorporate this information directly, as they only include the part that is anticipated.

for the researcher.[7] In this model self-selection into schooling on unobservables arises from the anticipated part of the returns, $\theta'_j f_i$, while the unanticipated ϵ^j_{ia} can obviously not be acted upon. $\theta'_j f_i$ is unobserved in the conventional approach using earnings realizations, while $\theta'_j f_i$ is implicitly 'observed' in the approach using information on subjective expectations of earnings (see equation (4)).

Thus the two potential outcomes relevant for the college attendance decision are:

$$\begin{aligned} \ln Y^0_{ia} &= \tilde{\alpha_0} + \beta'_0 X_i + \gamma a + \theta'_0 f_i + \epsilon^0_{ia} \\ \ln Y^1_{ia} &= \tilde{\alpha_1} + \beta'_1 X_i + \gamma a + \theta'_1 f_i + \epsilon^1_{ia}, \end{aligned} \tag{3}$$

with $\tilde{\alpha_j} = \big(\alpha_j - \gamma(s^j + 6)\big)$ for $j = 0, 1$. The individual (gross) return to college in this framework can be written as:

$$\begin{aligned} \widetilde{\rho}_i &= \ln Y^1_{ia} - \ln Y^0_{ia} \\ &= \widetilde{\alpha} + (\beta_1 - \beta_0)' X_i + (\theta_1 - \theta_0)' f_i + (\epsilon^1_{ia} - \epsilon^0_{ia}), \end{aligned}$$

where $\widetilde{\alpha} = (\tilde{\alpha_1} - \tilde{\alpha_0})$. The individual's ex-post return to college can never be observed, as only one of the two potential outcomes is observable.

From the individual's answers on her expectations of earnings for age 25 ($a = 25$), one can derive the following information on expected earnings:

$$\begin{aligned} E(\ln Y^0_{ia}) &= \tilde{\alpha_0} + \beta'_0 X_i + \gamma a + \theta'_0 f_i \\ E(\ln Y^1_{ia}) &= \tilde{\alpha_1} + \beta'_1 X_i + \gamma a + \theta'_1 f_i, \end{aligned} \tag{4}$$

Data on subjective expectations allow me to relax the assumption of rational expectations. 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. Nevertheless, I do need assumptions about the way returns to experience enter and I can not allow for heterogeneity in returns to experience, because the questions about earnings expectations have only been asked for one point of the life-cycle. Using the information given in equation (4), I can derive an expression for the expected gross return of individual i:

$$\begin{aligned} \rho_i &= E(\ln Y^1_{ia} - \ln Y^0_{ia}) \\ &= \widetilde{\alpha} + (\beta_1 - \beta_0)' X_i + (\theta_1 - \theta_0)' f_i. \end{aligned} \tag{5}$$

[7] Kaufmann and Pistaferri (forthcoming) address the question of superior information 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. Using data on people's subjective expectations of earnings allows 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.

To estimate the model of college attendance choice (see equation (1)), I make use of the data on subjective earnings expectation applying the following approximation $E(Y_{ia}) \equiv E(e^{\ln Y_{ia}}) \cong e^{E(\ln Y_{ia})+0.5Var(\ln Y_{ia})}$. Given the assumptions about experience, I can rewrite the participation equation (1) in terms of expected gross returns to college ρ (see Appendix B for the derivation):

$$\begin{aligned} S_i^* &= f\left(r_i, \rho_i, C_i, E(\ln Y_{i25}^0), \sigma_i^0, \sigma_i^1\right) \\ S_i &= 1 \text{ if } S_i^* \geq 0 \\ S_i &= 0 \text{ otherwise,} \end{aligned} \tag{6}$$

where S_i is a binary variable indicating the treatment status. The decision to attend college depends upon the (unobserved) interest rate r, expected return ρ, direct costs of attendance C, opportunity costs $E(\ln Y_{i25}^0)$ and the (subjective) standard deviations of future earnings σ_i^0, σ_i^1 (due to the approximation).

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

3 Data Description

In this section I describe the survey data that I am using in the following analysis. In particular I will discuss in detail the module eliciting subjective expectations of earnings and present 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 (2007)). 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.[8] As I analyze the college attendance decision in this paper, I restrict the sample to 18/19 year old high school graduates, who either start to work (or look for work) or decide to attend college.

[8]Due to the last eligibility criteria the sample only comprises the poorest third of the high school graduate population. Thus even the "high" income individuals are not rich. 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.

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/remittances to and from the household. The youth questionnaire contains a section on individuals' subjective expectations of earnings as discussed in the next section.

One important remark about the timing of the survey and the college attendance decision: 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. 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 this short period or have changed, but left expectations unchanged. To provide supporting evidence for this assumption I use the following as the counterfactual of the expectations of the high school graduates *before* they decided about attendance: I use the cross-section of earnings expectations of a cohort that is one year younger (just starting grade 12). The distributions of expected earnings of the two cohorts (for high school and college) are not significantly different suggesting that expectations have not changed significantly in these three months (see Attanasio and Kaufmann (2007)). 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 (and/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, which I do not find to be the case.

3.2 Elicitation of the Subjective Distribution of Future Earnings and Calculation of Expected Earnings and Expected Returns to College

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 gross returns to college (compare Guiso, Jappelli, and Pistaferri (2002) and Attanasio and Kaufmann (2007)). 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, $y^S_{mid} = (y^S_{min} + y^S_{max})/2$, 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 (see figure 1), which is more plausible than a stepwise uniform distribution as it puts less weight on extreme values.[9]

Thus I can express expected earnings $E(Y^S)$ and perceived earnings risk $Var(Y^S)$ for schooling degrees $S = 0, 1$ for *each* individual as follows:

$$\begin{aligned} E(Y^S) &= \int_{y^S_{min}}^{y^S_{max}} y f_{Y^S}(y) dy \\ Var(Y^S) &= \int_{y^S_{min}}^{y^S_{max}} \left(y - E(Y^S)\right)^2 f_{Y^S}(y) dy. \end{aligned}$$

I will perform the following analysis in terms of log earnings, so that I compute, for example, 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:

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

[9]The first moment of the individual distribution is extremely robust with respect to the underlying distributional assumption (see Attanasio and Kaufmann (2007) for more details on the triangular distribution, alternative distributional assumptions and robustness checks).

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

In this section I discuss some descriptive evidence of the validity of the data on subjective expectations of future earnings and returns. The validity of these data is analyzed in more depth in Attanasio and Kaufmann (2007), who conclude that people have a good understanding of the questions on subjective expectations. In particular their findings suggest that people have decent knowledge about skill prices and about local earnings for different schooling degrees. Investigating how well people are informed has important policy implications, as lack of information leading to an underestimate of returns to schooling or an overestimate of earnings risk could be one explanation for low enrollment rates. Furthermore, Attanasio and Kaufmann (2007) find that the measured subjective expectations capture –at least in part– the beliefs that people base their decisions on. This aspect determines if data on subjective expectations can improve our understanding of people's schooling decisions. In the following I will discuss a few of these results.

Attanasio and Kaufmann (2007) 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. This is informative, but not a test of whether people have "correct" expectations, because the expectations are about future earnings which will only be realized around 2012. Expected high school earnings are 1940 pesos 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.

Attanasio and Kaufmann (2007) show that earnings expectations vary with individual and family background characteristics in a similar way like observed earnings in Mincer earnings regressions. Even after controlling for these characteristics, expectations are strongly correlated with local average earnings for the relevant schooling level and gender (again using Census 2000 data).

These results suggest that people understand the questions on subjective expectations well and are –at least on average– relatively well informed about skill prices and about how individual characteristics affect earnings.[10] At the same time there is still a considerable amount of heterogeneity in expected earnings. It is likely that the individual has superior information, for example about her skills, i.e. heterogeneity in information sets is only partially captured by variables that are observable to the researcher. This would suggest an important value of expectations, which could reflect (unobserved) heterogeneity in ability and in information about skill prices (compare

[10] Attanasio and Kaufmann (2007) test if results depend on how exactly the questions on subjective expectations were asked. In particular they compare people's answers when point expectations are elicited with answers to questions on the individual distribution of earnings (see the previous section). Furthermore, they can compare expectations of mothers about their children's earnings with adolescents' expectations about their own earnings to test if differences in information sets between parents and children are important. In another exercise they control for family-specific effects using data on expectations of different siblings and find that individual characteristics still have predictive power.

Kaufmann and Pistaferri (forthcoming)).

The most important question concerning the value of data on subjective expectations is whether elicited expectations capture the beliefs that people base their decisions on. Attanasio and Kaufmann (2007) provide evidence that this is indeed the case: After controlling for an extensive list of individual and family background characteristics in a reduced-form regression of schooling choices, they find that people's expectations remain important in explaining these decisions.[11] Furthermore, their data enables them to improve our understanding of the following aspect of the process of intra-household decision making: They can address the question whose expectations matter and find that older adolescents play an important role in this process in that their expectations matter for investment decisions into higher education. Attanasio and Kaufmann (2007) make use of the fact that for part of their sample they have data on adolescents' expectations as well as on point estimates of mothers' expectations about their children's future earnings. They find that for the high school attendance decision, mothers' expectations are important (the effect of the adolescents' expectations are similar in magnitude but not significant), while for the college attendance decision only the adolescents' expectations matter. Furthermore, they find that for the high school decision, mothers' risk perceptions (about unemployment and earnings risk) matter. For the college attendance decision on the other hand only adolescents' expectations about returns to college are significant.

For this reason I only use data on expectations of the adolescents and focus on expected returns in the following analysis. Unfortunately, the survey was not randomized upon who answered the questions on the subjective distribution of earnings (while the questions on point expectations were asked to the whole sample of mothers): In cases where the adolescent was not present mothers answered also the youth questionnaire –including the questions on the subjective distribution of earnings (see Appendix C for further details and summary statistics in table 10). I address the concern of sample selection bias as follows: In the reduced-form regressions I correct for sample selection by estimating jointly a latent index model for college attendance and a sample selection equation. As an exclusion restriction I use information on the date and time of the interview, which are strongly significant determinants of whether the respondent is the adolescent. Results suggest that sample selection on unobservables is not an important problem (the correlation between the error terms of the two equations is not significantly different from zero.)

3.4 Data on Educational Costs

The model in section 2 illustrates that –apart from expected earnings– college attendance decisions should also be afffected by direct costs of attending college. In Mexico these costs pocket a large

[11] This suggests that at least part of the unexplained heterogeneity of subjective expectations is driven by heterogeneity in information sets, such as in ability and information about skill prices and thus addresses the concern that the unexplained heterogeneity in expectations could be entirely driven by measurement error.

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)). For example, if the adolescent lives far away from the closest university, she will have to move to a different city and pay room and board. This will be an important additional cost factor compared to someone who can live with his 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.[12] About half of the adolescents live within a distance of 20 kilometers to the closest university, where a daily commute with public transportation might be possible. Twenty-five percent live within 20 to 40 kilometers distance, while the other quarter lives more than 40 kilometers away.

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. Forty percent of adolescents face tuition costs of at least 750 pesos. This is equivalent to 15% of median per capita parental income, while it only represents a fraction of total college attendance costs. Thus college attendance would imply a substantial financial burden for poor families.

To address the question 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, durables and remittances. I create the following two measures: per capita parental income and an index of parental income and wealth.[13] Median yearly per capita income is 5370 pesos (approximately 537 US$).

As I do not expect a linear effect of income and wealth on the interest rate that families face, I

[12] 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 our sample and in all neighboring states– from a web page provided by INGI (National Institute of Statistics, Geography and Information). My special thanks to Shaun McRae who helped extracting these data.

[13] 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.

use the following per capita parental income thresholds: twice the minimum monthly salary (44% of the sample fall into this first category of income below 5,000 pesos) and four times the minimum monthly salary (34% have per capita income between 5,000 and 10,000 pesos). The reason for using these income thresholds is their approximate correspondence with eligibility requirements for receiving a fellowship (see Appendix C). Therefore I can analyze whether eligibility for fellowships has an effect on college attendance. Nevertheless one should keep in mind that fellowships are quantitatively not very important: only 5% of the undergraduate student population received a fellowship in 2004.

4 Testable Implications of the Model of College Attendance Choice and Empirical Results

Before testing the implications of credit constraints of the model of college attendance choice, I show that even after controlling people's beliefs about returns to schooling –in addition to conventional measures used in the existing literature such as parental background and skills–, parental income remains a significant predictor of the college attendance choice (see table 1).

Thus with data on subjective expectations I can exclude the possibility that parental income is significant, *only* because it picks up differences in earnings expectations and perceived risk between poor and rich individuals.

To perform a more rigorous analysis of what is causing the income gradient in college attendance, I test implications of the model of college attendance in the presence of credit constraints.

4.1 The Distribution of Costs of College Attendance for Rich and Poor Individuals

Data on people's expected returns to college as well as on their attendance decision allows me to directly estimate the distribution of college attendance costs. Thus I can evaluate if poorer individuals face higher costs of attending college than rich individuals or if –on the other hand– the lower attendance of the poor is entirely driven by lower expected returns. The latter could be due to differences in "college preparedness" (see, e.g., Carneiro and Heckman (2002)) or to differences in information about skill prices.

To illustrate how data on expected returns enables me to estimate the distribution of costs, consider the stylized model of schooling investments by Becker (1967). In this model direct schooling costs are zero and credit constraints are modeled as heterogeneity in individuals' interest rates. People decide to attend college if expected returns are larger than the interest rate they face:

$$S = 1 \Leftrightarrow \rho \geq r.$$

Thus with data on schooling decisions ($S = 0, 1$), and on expected returns ρ, and the assumption that ρ and r are orthogonal, it is possible to derive the cumulative distribution function of the interest rate r:[14]

$$Pr(S=1|\rho=\tilde{\rho}) = Pr(r \le \tilde{\rho}|\rho=\tilde{\rho}) = F_{r|\rho=\tilde{\rho}}(\tilde{\rho}) = F_r(\tilde{\rho}). \tag{7}$$

Intuitively, the fraction of people who decide to attend college given that they expect return $\tilde{\rho}$ is equivalent to the fraction of people who face an interest rate r smaller than expected return $\tilde{\rho}$.

In the case of my more general model (see section 2), which allows for nonzero direct costs of attendance, I show that it is possible to write the participation equation in additively separable form between expected return ρ and total college attendance costs K (including direct costs and financing costs) (for the derivation see Appendix B).[15] Then total costs K take the place of the interest rate r in the equations of this section, and I can perform the analysis estimating the distribution of total costs.

I estimate the cost distribution $F_r(\tilde{\rho}) = Pr(S=1|\rho=\tilde{\rho})$ by performing Fan's (1992) locally weighted linear regression of college attendance S on the expected return ρ.[16] To compare the distribution function of costs for different income classes, I perform this analysis for "low", "middle" and "high" income individuals, i.e. yearly per capita income less than 5,000 pesos, between 5,000 and 10,000 pesos and more than 10,000 pesos.

Figure 2 shows that that poor individuals face higher costs than the rich, as the c.d.f. of costs for poorer individuals is shifted to the right. Take for example an interest rate of $r = 0.6$. More than 75% of the poor face an interest rate that is higher than this interest rate, $r = 0.6$, while only 55% of the rich individuals face costs $r > 0.6$. To put it differently, among individuals with expected returns around $\rho = 0.6$, 45% of rich individuals attend, but only 25% of the poor. Poor individuals require higher expected returns to be induced to attend college.

To analyze whether the cost distributions of the poor and the rich differ significantly, I calculate point-wise confidence intervals applying a bootstrap procedure. Figure 3 plots the c.d.f. of poor and rich individuals with 95%-confidence intervals and illustrates that the c.d.f. of costs of the poor is significantly shifted to the right compared to the one the "rich" for a wide range of the support, $r \in [0.25, 1.1]$. This is also true comparing the middle and the rich income group as well as middle and low income group for part of the support (results from the author upon request).

[14]The orthogonality assumption has the following caveat: It will be violated in a framework with search costs if people with higher expected returns exert more effort into the search for a lower interest rate.

[15]This is an approximation to the participation equation as derived from the model, as it neglects higher order terms of ρ, i.e. ρ^2, ρ^3 etc (see Appendix B).

[16]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 (see equation (7)). I drop large outliers of expected returns in the plotted graphs (upper 5%), thus being left with a support of [0.05, 1.2]. Including them would cause the nonparametric regression line to start sloping down for returns larger than 1.1, which seems to be driven by a few outliers where people state very high returns but do not attend.

4.2 Excess Responsiveness of the Poor to Changes in Direct Costs

In the last section I have shown that poorer individuals face significantly higher costs of college attendance and thus require higher expected returns to be induced to attend college. To understand the role of the different cost components and whether credit constraints play an important role in the low enrollment rate of poor Mexicans, I derive a testable prediction of the presence of credit constraints from the model of college attendance choice. As discussed in section 2, in this model credit constraints are captured by heterogeneity in the interest rate that people face.

My model of college attendance choice implies that individuals who face a high interest rate r react more strongly to changes in direct costs C (see equation (22) in Appendix B):

$$\left|\frac{\partial S^*}{\partial C}\right| \quad \text{is increasing in} \quad r. \tag{8}$$

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 should thus be 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.[17]

Compared to conventional approaches, data on subjective expectations has the following two advantages: First, I can control directly for people's expectations about their potential returns to college and thereby avoid a bias that could arise from omitting this determinant.[18] 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

[17] 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 U.S.A. using the NLSY 1979. This is consistent with increased availability of fellowships and loans in the U.S.A. over the relevant time period.

[18] Omitting expected returns in the participation equation could cause a problem as expected returns might be correlated with the interaction term of parental income and direct college costs. This would hamper the interpretation of the results of this test. Making use of data on subjective expectations directly can avoid this endogeneity concern.

poor individuals with high expected returns are potentially prevented from attending college due to high financing costs, while poor low-return individuals would decide not to 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 data section 3.4). As shown in table 1 living further away from the closest university has significantly negative effects on the probability to attend college. Table 2 illustrates that the negative effect of a larger distance is only significant for poor individuals: living 20 to 40 kilometers away from college instead of less than 20 kilometers decreases the probability of attending by about 6 percentage points for the poorest income category. In this case being able to control for earnings expectations does not change the results.

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, which is equivalent to 15% of median yearly per capita income and thus represents an important financial burden for poor individuals. Table 3 would suggest that tuition costs do not have any effect on attendance. But once we take into account that what matters is being poor *and* having high expected returns, results change (see table 4): Poor individuals with high expected returns are excess responsive with respect to a change in tuition costs, which is consistent with the predictions of a model with credit constraints.

5 Counterfactual Policy Experiments

In the previous section I have shown that poor people face significantly higher costs of college attendance than rich people and that poor high-expected-return individuals are most sensitive to changes in direct costs. These results provide (suggestive) evidence that credit constraints affect college attendance decisions of poor Mexicans with high expected returns. Nevertheless I cannot exclude the possibility that other factors are also driving the low college attendance rates among poor.[19] Even if the empirical fact mostly reflects heterogeneity in time preferences, for example, government policies such as student loan programs might still be recommendable. This would be the case, if there are externalities from college attendance (correlated with private returns), or if people have time-inconsistent preferences, e.g. they become more patient when getting older.

As credit constraints would create scope for policy interventions, I perform counterfactual policy experiments by applying the Local Instrumental Variables methodology of Heckman and Vytlacil

[19]The literature on credit constraints faces three partially unobserved determinants of schooling decisions that are hard to disentangle: expected returns (capturing unobserved skills and information about skill prices), credit constraints (heterogeneity in borrowing rates) and heterogeneity in preferences (e.g. discount rate). Data on subjective expectations help to address part of the identification problem, while the problem of disentangling heterogeneity in interest rate and time preferences remains. For example Cameron and Taber (2004) assume a common discount factor for all individuals. To address this additional identification problem one could add survey questions not only on expectations but also on time preferences.

(2005) to my model of college attendance making use of data on subjective expectations of earnings. In particular I evaluate potential welfare implications of the introduction of a means-tested student loan program. I estimate the fraction of people changing their decisions in response to a reduction in the interest rate, 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.

This idea follows Card's interpretation of the finding that in many studies devoted to estimating the "causal" effect of schooling, instrumental variable (IV) estimates of the return to schooling exceed ordinary least squares (OLS) estimates (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 of this finding 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 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.[20] 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 $E(\beta|S=1)$.[21]

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

[20] 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.

[21] $E(\ln Y_1|S=1) - E(\ln Y_0|S=0) = E(\beta|S=1) + (E(\ln Y_0|S=1) - E(\ln Y_0|S=0))$, 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.

situation with earnings realizations subjective expectations are asked for both possible states of highest potential schooling degree, i.e. 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 idea.

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 enables me to relax the rational expectations assumption with strong requirements on coinciding information sets of individuals and the researcher.

To test the validity of the instrument used here, I plot expected returns and distance to college and perform a locally weighted linear regression of expected returns on distance to college (see figure 9 see Appendix D). No pattern is apparent. I also regress expected returns on polynomials of distance to college (in addition to observable characteristics of the individual and her family background, such as GPA of junior high school, father's education, per capita parental income). Table 11 (see Appendix D) shows that the coefficients on the polynomials of distance to college are not significantly different from zero. Note that the table presents results for distance and squared distance, but adding further polynomials does not change the result. These results are consistent with the validity of the instrument that I use.

5.1 Implications of Credit Constraints for Marginal Returns to College

From the latent index model (see equation (6)), 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,

$$S_i^* = f\left(r_i, \rho_i^M, C_i, E(\ln Y_{i25}^0), \sigma_i^0, \sigma_i^1\right) = 0 \tag{9}$$

The presence of credit constraints has the following implication for marginal returns: implicit differentiation of equation (9) 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 (ceteris paribus) than those individuals on the margin who are not credit constrained:

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

In the next subsections I illustrate how the marginal return to college can be derived, and how it can be used to perform policy experiments.

5.2 Derivation of the Marginal Return to College

For the purpose of the test that Card proposed and to perform counterfactual policy experiments, I derive the "Marginal Treatment Effect" (MTE) in this section.[22]

One important first step in the derivation and estimation of the marginal return to college is the estimation of the propensity score $P(Z) \equiv P(S = 1|Z = z)$. $P(Z)$ represents the probability of attending college conditional on observables Z. To estimate the participation equation as derived from the school choice model in section 2, I perform the following monotonic transformation of $S^* = \nu(Z) - V$:

$$S^* \geq 0 \Leftrightarrow \nu(Z) \geq V \Leftrightarrow F_V(\nu(Z)) \geq F_V(V),$$

and define $\mu(Z) \equiv F_V(\nu(Z))$ and $U_S \equiv F_V(V)$. In this case U_S is distributed uniformly, $U_S \sim$ Unif$[0, 1]$.[23] Therefore, the participation equation can be written as follows:

$$S^* \geq 0 \Leftrightarrow P(Z) = \mu(Z) \geq U_S.$$

An individual indifferent between attending college or not is characterized by $U_S = \mu(Z) = P(Z)$. It is thus possible to estimate U_S, i.e. the (marginal) costs which are equal to r in my model, for the indifferent individual by estimating the propensity score $P(Z)$.

This will allow me to derive the marginal return to college or Marginal Treatment Effect (MTE), which is defined as:

$$\Delta^{MTE}(u_S) = E(\ln Y_1 - \ln Y_0|U_S = u_S) = E(\rho|U_S = u_S). \tag{10}$$

It represents the average gross gain to college for individuals who are indifferent between attending college or not at the level of unobservable costs $U_S = u_S$.

One important drawback of the LIV methodology is that the analysis 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 (2005) and Heckman, Vytlacil, and Urzua (2006)).

---

[22] This section follows Carneiro, Heckman, and Vytlacil (2005) and Heckman and Vytlacil (2005) in their derivation of the "Marginal Treatment Effect" (MTE). Their goal is to get estimates of summary measures of the return-to-schooling distribution purged from selection bias and the MTE is just a tool in this encounter, while I am interested in the marginal return to college for its own sake.

[23] U_S is distributed uniformly, because $Pr(U_S \leq \mu(Z)) = Pr(V \leq F_V^{-1}(\mu(Z))) = F_V(F_V^{-1}(\mu(Z))) = \mu(Z)$. Thus the propensity score is equal to $P(Z) \equiv Pr(S = 1|Z = z) = Pr(S^* \geq 0|Z) = Pr(U_S \leq \mu(Z)) = \mu(Z)$.

In a model with heterogeneity in interest rates I can only write the participation equation in additively separable form with data on subjective expectations of earnings: The participation equation as derived from the model can be expressed as a fourth-order polynomial in the unobservable interest rate, $1+r$ (see Appendix B for the derivation):

$$S_i^* \geq 0 \Leftrightarrow (1+r_i)^4 - A(Z_i;\theta)(1+r_i)^3 - B(Z_i;\theta) \leq 0, \tag{11}$$

where $A(Z_i;\theta), B(Z_i;\theta) > 0$ are functions of the observables $Z_i = \left(\rho_i, C_i, E(\ln Y^0), \sigma_i^0, \sigma_i^1\right)$ 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:

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

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

$$\begin{aligned} S_i^* &= -(1+\bar{r}) + g\left(Z_i;\theta\right) - V_i \\ S_i &= 1 \text{ if } S_i^* \geq 0 \\ S_i &= 0 \text{ otherwise.} \end{aligned} \tag{12}$$

I assume $V_i \sim N(0,1)$ and estimate the propensity score $P(Z)$ using a Maximum Likelihood procedure.

With the help of the predicted values of the propensity score, $\widehat{P(z)}$, I can define the values $u_S = F_V(V)$ over which the marginal return to college (MTE) can be identified: The MTE is defined for values of $\widehat{P(z)}$, for which one obtains positive frequencies for both subsamples $S=0$ and $S=1$. The observations for which $\widehat{P(z)}$ is outside of the support are dropped.[24]

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

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

The integral can be rewritten as (see Appendix B):

$$\begin{aligned} \int_0^p E\left(lnY_{it}^1 - lnY_{it}^0 | U_S = p\right) dU_S &= pE\left(lnY_{it}^1 - lnY_{it}^0 | U_S \leq p\right) \\ &= pE\left(lnY_{it}^1 - lnY_{it}^0 | P(Z) = p, S = 1\right). \end{aligned} \tag{13}$$

[24]The derivation of the policy relevant treatment effect is not affected by support problems, as the MTE only has positive weight over its' support. Even after trimming, the sparseness of data in the tails results in a large amount of variability in the estimation of the MTE for values of p closer to the corners of the support.

With subjective expectations of earnings one has data on each individual's expectation of earnings in both schooling states, and thus also for those individuals who decide to attend college, $E\left(lnY_{it}^1 - lnY_{it}^0|S=1\right)$. I estimate $P(Z)$ in a first step and therefore have a value $\widehat{P(z)} = p$ for each individual.

Finally I fit a nonparametric regression of

$$m(p) = pE[\ln Y_{it}^1 - \ln Y_{it}^0|P(Z) = p, S = 1]$$

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 $p = u_S$:

$$\Delta^{MTE}(u_S) = \frac{\partial m(p)}{\partial p} = \frac{\partial\{pE[\ln Y_{it}^1 - \ln Y_{it}^0|P(Z) = p, S = 1]\}}{\partial p}$$

I calculate standard errors by applying a bootstrap over the whole procedure described in this section.

To perform policy experiments, I introduce the following notation: the "Policy Relevant Treatment Effect" ($PRTE$) is a weighted average of the marginal returns to college ($\Delta^{MTE}(u_S)$), where the weights depend on who changes participation in response to the policy of interest (compare Heckman and Vytlacil (2001)). One important assumption underlying this analysis is that the participation equation continues to hold under hypothetical interventions. The $PRTE$ can be written as:

$$PRTE = \int_0^1 MTE(u)\omega(u)du, \quad \text{where} \quad \omega(u) = \frac{F_P(u) - F_{P^*}(u)}{E(P^*) - E(P)}. \tag{14}$$

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 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.

The following is a special case of a $PRTE$: 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. Suppose that $S^* = Z'\gamma + V$, and that γ_k (the kth element of γ) is nonzero. The resulting $PRTE$ is:

$$PRTE_\varepsilon = E(\rho_i|Z'\gamma \leq V \leq Z'\gamma + \varepsilon\gamma_k), \tag{15}$$

i.e., $PRTE_\varepsilon$ is the average return among individuals who are induced into university by the incremental subsidy.

I will use the $PRTE$ to evaluate different policies by deriving the average marginal expected return of individuals induced to change their schooling status as a response to these policies, and compare the results to the average return of those attending.

5.3 Estimation of the Marginal Return to College

This section describes how the estimation of the marginal return to college is implemented, 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 participation equation (12) using a Maximum Likelihood procedure. In order to empirically implement the notion of costs, C, I use the following auxiliary regression containing distance to the closest university ("Univ Dist"), distance squared ("Univ Dist Sq"), and state fixed effects to capture differences in direct costs. To proxy for preferences and capture monetized psychological costs, I include mother's education and GPA of junior high school:[25]

$$C = \delta_0 + \delta_1 \text{Univ Dist} + \delta_2 \text{Univ Dist Sq} + \delta_3 \text{GPA} + \delta_4 \text{Mother's Schooling} + \delta_5 \text{State Dummies}. \quad (16)$$

The results of the Maximum Likelihood Estimation of the propensity score are displayed in table 5. All cost variables (see equation (16)) as well as expected returns are highly significant and with the expected sign.

Tables 6 and 7 illustrate the magnitude of the effects of these variables on attendance. In table 6 I illustrate the effect of "moving" an individual closer to the closest university, of increasing mother's schooling and of increasing the GPA. The baseline case evaluates all explanatory variables at their medians, that is living about 18km from the closest university, having a mother with five years of schooling and a GPA of 82 out of 100. This leads to a predicted probability of attending college of 22%. Living 5km closer increases the probability of attending by 1.3 percentage points, i.e. by about 6%. Increasing mothers' schooling by one year increases this probability by 1.8 percentage points, while the probability is increased by 0.7 percentage points in case of a one percentage point higher GPA.

Table 7 displays the effect of an increase in expected returns to college for different baseline cases. A 10 percentage point increase in expected returns (equivalent to 16%), increases the probability of attending college by about 1 percentage point (or 5%). The effect of an increase in return doubles if the individual faces lower direct costs, i.e.lives 5km closer to the closest university. This

[25] I use mother's schooling as the father is not present in some of the households and would thus lead to more missing values. The results are robust with respect to using father's schooling instead. In contrast to previous specifications, I add mother's years of schooling and distance directly instead of using dummies, as it is hard to achieve convergence otherwise.

is consistent with the presence of credit constraints as individuals, who face lower costs are less likely to be credit constrained, and can thus act upon higher expected returns more easily.

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. Figure 4 shows that the probability of attending college is generally relatively low for adolescents of the Jovenes sample, but that there is a right-shift in the density for high school graduates, who decided to attend college. Their mean (median) probability is about 36% (34%), while the mean (median) probability of attending for those who stopped is around 29% (27%). Figure 4 illustrates that there is little mass outside of the interval 0.1 and 0.8. Therefore I estimate the marginal return to college over the support p in the interval $[0.1, 0.8]$.

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 $u_s = P(Z)$. Figure 5 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 0.05 leads to a wiggly line that is clearly undersmoothed, a large bandwidth of 0.2 seems to lead to an oversmoothed graph. Reassuringly the marginal return to college is upward sloping independent of the choice of the bandwidth. Individuals facing higher (unobservable) borrowing rates, who have to be compensated by a higher $P(Z)$ to be made indifferent, have higher expected returns on the margin.

Lastly, I calculate standard errors by performing a bootstrap over the whole procedure discussed above. Figure 6 displays the marginal return to college with 95% confidence intervals using a bandwidth of 0.15. Unfortunately error bands are wide in particularly for large values of $P(Z)$ for which there are few data points.[26]

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 (means-tested) student loan program. Therefore I analyze the effect of a change in interest rate for poor (or poor and able) individuals and the effect of a change in direct costs (distance to college). 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

[26] Carneiro, Heckman, and Vytlacil (2005) have the same problem of very 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.

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 test whether the average "marginal" expected return is significantly larger than the average expected return of individuals attending college. Thus with subjective expectations I can improve on the test suggested by Card (1995) (compare analysis in section 4.1). 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, for example by building new universities in places that previously did not have higher education institutions. In section 4.2 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 affects individuals at the margin. I perform the analysis by decreasing the distance to college by 10 kilometers (for different target groups). This counterfactual policy leads to an increase in college attendance of about 10% (2.3 percentage points), and to an average marginal expected return of 1.10 (see table 8). Decreasing the distance only for very poor individuals (less than 5,000 pesos per capita income), leads to a change in attendance of 3%, while those individuals who change college attendance have an average marginal expected return of 1.08. For poor and able individuals (per capita income less than 5,000 pesos and a GPA higher than the median), this policy would lead to a change in attendance of 2%, and an average marginal expected return of 1.11. 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 despite high expected returns. For a full cost-and-benefit analysis of this policy the costs of building new universities would obviously have to be taken into account.

A more efficient policy could consist of the introduction of a governmental student loan program. Therefore, as a second policy experiment, I consider the effect of a decrease in the interest rate of poor (and able) individuals. A 10% change in the interest rate for very poor individuals leads to an average marginal return of 1.11 (the college attendance rate increases by about 50%, that is about 11 percentage points), while this change for poor and able individuals leads to an average marginal return of 1.14 (see table 8). In both cases, the average marginal return of individuals induced to change their college attendance decision is significantly higher than the average return of those individuals already attending college.

Again, for a full cost-and-benefit analysis one would have to take into account the costs of providing student loans, that is bureaucratic costs for giving out and recovering loans in addition to costs of interest.[27] If a large-scale policy is put in place, one would additionally have to take into account general equilibrium effects, in particular in terms of skill prices.

[27] Note that if poor credit-constrained individuals are extremely risk averse in terms of taking a loan, a government loan program might have no or very little effect, while a fellowship program could have much larger effects.

6 Conclusion

The goal of this paper has been to improve our understanding of both the causes and the welfare implications of the steep income gradient in college attendance in Mexico. In this context, researchers face an important identification problem. On the one hand people might expect low returns to schooling 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 used data on people's subjective quantitative expectations of future returns to schooling. 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 at the time of the decision.

Data on expected returns allowed me to directly estimate and compare cost distributions of poor and rich individuals. I found that poor individuals require significantly higher expected returns to be induced to attend college, implying that they face higher costs than individuals with wealthy parents.

I then tested predictions of a simple model of college attendance choice in the presence of credit constraints, using parental income and wealth as a proxy for the household's (unobserved) interest rate. I found that poor individuals with high expected returns are particularly responsive to changes in direct costs such as tuition, which is consistent with credit constraints playing an important role.

Evaluating potential welfare implications by applying the Local Instrumental Variables 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 the interest rate. Individuals at the margin have higher expected returns than the individuals already attending college, which again is consistent with credit constraints playing an important role.

These results suggest that credit constraints could be one of the driving forces of Mexico's large inequalities in access to higher education and low overall enrollment rates. This is consistent with Mexico's low government funding for student loans and fellowships for higher education. In this case the results of my policy experiments suggest that the introduction of a governmental student loan program could lead to large welfare gains by removing obstacles to human capital accumulation and fostering Mexico's development and growth.

It is important to note that the evidence above is consistent with other factors also driving the poor's low college attendance rates. Even if the steep income gradient mostly reflects heterogeneity in time preferences, for example, government policies such as student loan programs might still be recommendable. This could be the case if there are externalities from college attendance and social returns are correlated with private returns, or if people have time-inconsistent preferences, e.g. they become more patient when getting older.

References

Attanasio, O. P., and K. M. Kaufmann (2007): "Educational Choices, Subjective Expectations and Credit Constraints," *mimeo*.

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

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.

Belzil, C., and J. Hansen (2002): "Earnings Dispersion, Risk Aversion, and Education," *IZA Discussion Paper No. 513*.

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

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.

——— (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.

——— (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): "Understanding What IV Estimate: Estimating Marginal and Average Returns to Education," *mimeo*.

Cunha, F., and J. J. Heckman (2006): "A New Framework for the Analysis of Inequality," *NBER working paper 12505*.

Cunha, F., J. J. Heckman, and S. Navarro (2005): "Separating Uncertainty from Heterogeneity in Life Cycle Earnings," *Oxford Economic Papers*, 57, 191–261.

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.

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., and E. Vytlacil (2001): "Policy-Relevant Treatment Effects," *American Economic Review, Papers and Proceedings*, 91 (2), 107–111.

——— (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. (2008): "The (Perceived) Returns to Education and the Demand for Schooling," *mimeo*.

Kaufmann, K. M., and L. Pistaferri (forthcoming): "Disentangling Insurance and Information in Intertemporal Consumption Choices," *American Economic Review, Papers and Proceedings*.

Kling, J. R. (2001): "Interpreting Instrumental Variables Estimates of the Returns to Schooling," *Journal of Business and Economic Statistics*, 19 (July), 358–364.

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

Mincer, J. (1974): "Schooling, Experience and Earnings," *NBER, New York*.

Nguyen, T. (2008): "Information, Role Models and Perceived Returns to Education: Experimental Evidence from Madagascar," *mimeo*.

Padula, M., and L. Pistaferri (2001): "Education, Employment and Wage Risk," *Working Paper No. 67, Centre for the Studies in Economics and Finance*.

Willis, R. J., and S. Rosen (1979): "Education and Self-Selection," *Journal of Political Economy*, 87, S 7–36.

7 Appendix A

Figure 1: The triangular distribution of earnings

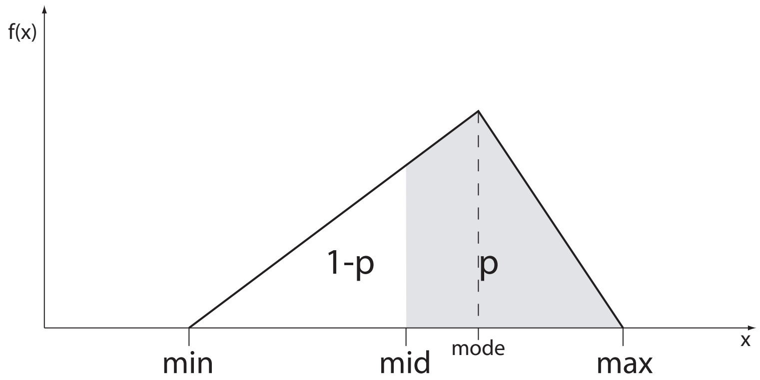

Figure 2: The Cumulative Distribution Function of Costs for Different Income Classes.

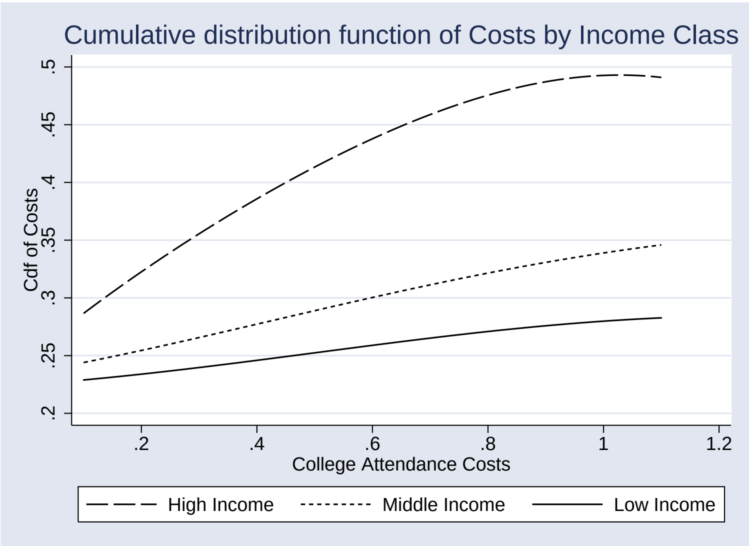

Figure 3: The Cumulative Distribution Function of Costs of Poor versus Rich Individuals with 95% Confidence Intervals.

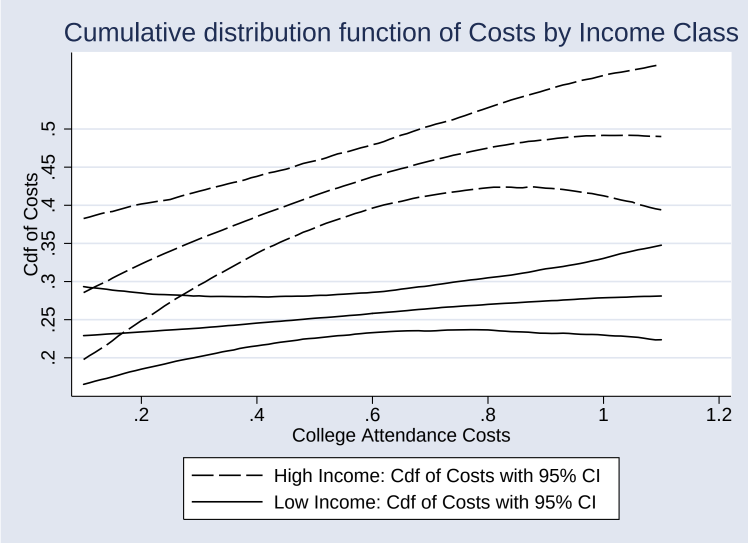

Figure 4: Predicted probability of attending college for high school graduates, who decided to attend college, and those, who stopped school after high school.

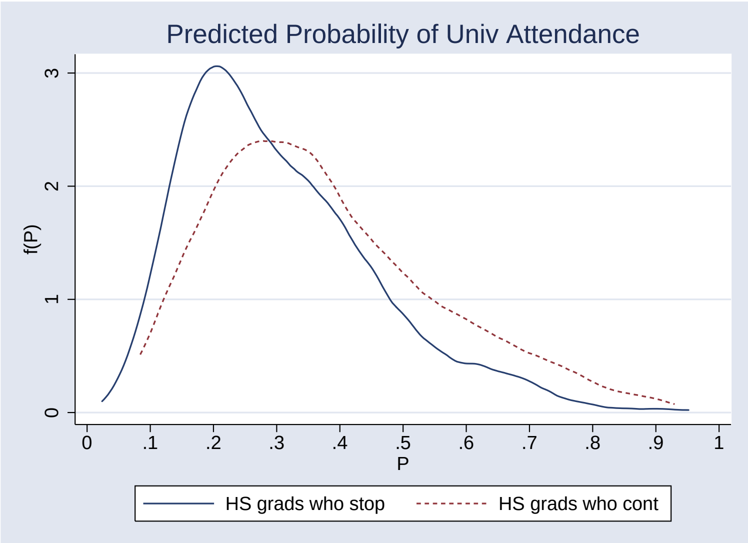

Figure 5: The Marginal Return to College ("MTE") conditional on different levels of unobserved costs (for different bandwidths).

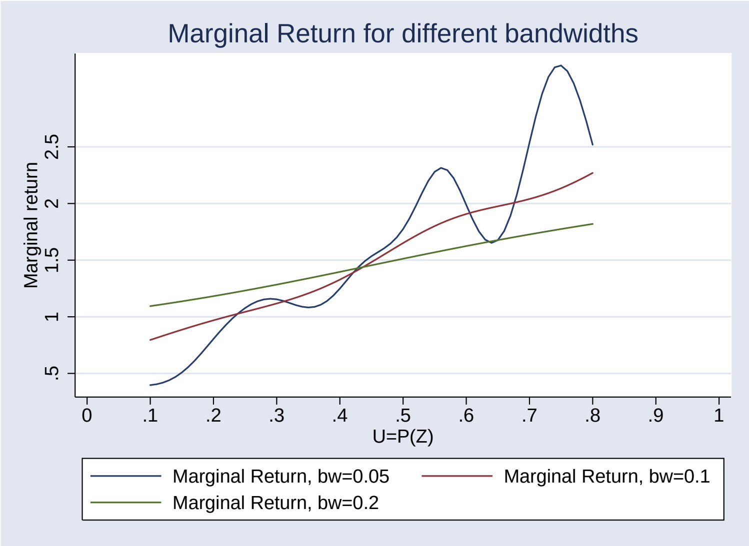

Figure 6: The Marginal Return to College ("MTE") conditional on different levels of unobserved costs with 95% Confidence Interval bands (for a bandwidth of 0.15).

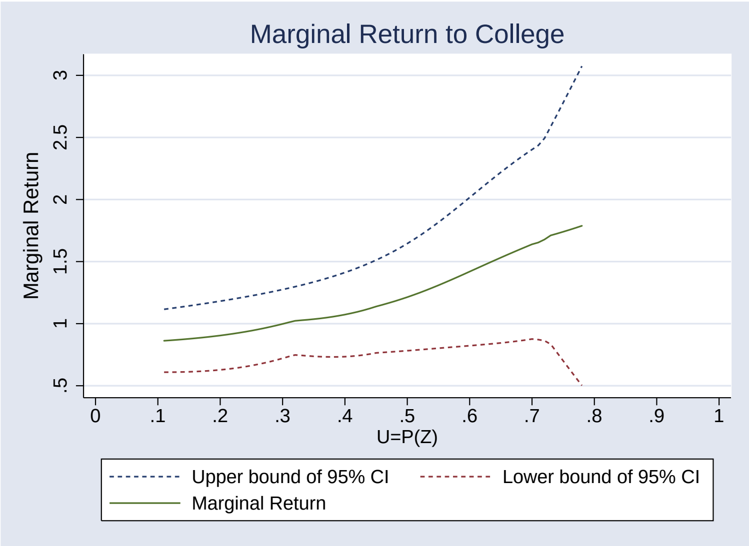

Table 1: Probit model for the college attendance decision.

Dep. Var.: Attend College	Model 1 Marg. Eff. (S.E.)	Model 2 Marg. Eff. (S.E.)	Model 3 Marg. Eff. (S.E.)
Expected Return to College	0.081**	0.077**	0.074**
	(0.032)	(0.033)	(0.032)
Prob of Work - HS	0.029	0.010	-0.009
	(0.083)	(0.084)	(0.073)
Prob of Work - College	0.043	0.030	0.032
	(0.096)	(0.097)	(0.084)
Var of Log Earn - HS	-2.264	-2.917	-2.750
	(1.803)	(1.927)	(1.784)
Var of Log Earn - College	-0.015	0.157	0.452
	(2.260)	(2.286)	(1.995)
GPA - second tercile		0.044	0.034
		(0.031)	(0.028)
GPA - top tercile		0.170***	0.149***
		(0.031)	(0.040)
Father's Educ - junior HS		0.100**	0.068*
		(0.040)	(0.039)
Father's Educ - HS		0.191***	0.131*
		(0.074)	(0.073)
Father's Educ - Univ		0.524***	0.575***
		(0.125)	(0.147)
Per cap Income - 5 to 10k			0.040
			(0.029)
Per cap Income - more than 10k			0.105***
			(0.040)
Dist to Univ 20 to 40km			-0.066**
			(0.027)
Dist to Univ more than 40km			-0.096***
			(0.031)
Tuition more than 750 pesos			-0.097**
			(0.041)
State, Gender, Marital Status Dummies	Yes	Yes	Yes
Observations	3680	3680	3680
Censored Obs	2009	2009	2009
Log Likelihood	-3290.394	-3261.734	-3229.325
Sample Sel.: Corr. betw. Errors	-0.534	-0.351	0.002
P-Value of LR test of Indep Eqns	0.057	0.299	0.995

Notes: Table displays marginal effects and standard errors in brackets. * $p<0.1$ ** $p<0.05$ *** $p<0.01$. Excl. categories: male, single, lowest GPA tercile, father's education primary or less, per capita income less than 5000 pesos, distance to university less than 20km and tuition less than 750 pesos.

Table 2: Probit model for the college attendance decision: Differential effect of direct costs (distance to college) for different per capita income categories.

Dep. Var.: Attend College	Model 4 Marg. Eff. (S.E.)	Model 5 Marg. Eff. (S.E.)
Univ 20 - 40km * Pcap Income $<$ 5k	-0.063**	-0.063**
	(0.030)	(0.029)
Univ 20 - 40km * Pcap Income 5 - 10k	-0.041	-0.043
	(0.036)	(0.035)
Univ 20 - 40km * Pcap Income $>$ 10k	0.017	0.011
	(0.049)	(0.046)
Univ $>$ 40km * Pcap Income $<$ 5k	-0.043	-0.044
	(0.029)	(0.028)
Univ $>$ 40km * Pcap Income 5 - 10k	-0.095***	-0.096***
	(0.037)	(0.035)
Univ $>$ 40km * Pcap Income $>$ 10k	-0.033	-0.033
	(0.048)	(0.045)
Expected Return to College		0.053*
		(0.028)
Per cap Income - 5 to 10k	0.061*	0.061*
	(0.032)	(0.031)
Per cap Income - more than 10k	0.095**	0.092**
	(0.041)	(0.041)
Controls for Exp Log Earn, Prob of Work and Var of Log Earn	Yes	Yes
Control's for Ability, Father's Educ, State, Gender, Marital Status Dummies	Yes	Yes
Observations	3680	3680
Censored Obs	2009	2009
Log Likelihood	-3253.562	-3247.898
Sample Sel.: Corr. betw. Errors	0.282	0.328
P-Value of LR test of Indep Eqns	0.308	0.243

Notes: Table displays marginal effects and standard errors in brackets. * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Excl. categories: male, single, lowest GPA tercile, father's education primary or less, per capita income less than 5000 pesos.

Table 3: Probit model for the college attendance decision: Differential effect of direct costs (tuition costs) for different per capita income categories.

Dep. Var.: Attend College	Model 6 Marg. Eff. (S.E.)	Model 7 Marg. Eff. (S.E.)
Tuition > 750 pesos * Pcap Income < 5k	-0.012 (0.028)	-0.016 (0.027)
Tuition > 750 pesos * Pcap Income 5 - 10k	-0.019 (0.036)	-0.022 (0.035)
Tuition > 750 pesos * Pcap Income > top	0.065 (0.054)	0.061 (0.052)
Univ 20 - 40km * Pcap Income < 5k	-0.060** (0.030)	-0.059** (0.029)
Univ 20 - 40km * Pcap Income 5 - 10k	-0.044 (0.036)	-0.046 (0.035)
Univ 20 - 40km * Pcap Income > 10k	0.024 (0.050)	0.018 (0.047)
Univ > 40km * Pcap Income < 5k	-0.042 (0.029)	-0.042 (0.028)
Univ > 40km * Pcap Income 5 - 10k	-0.094** (0.037)	-0.096*** (0.036)
Univ > 40km * Pcap Income > 10k	-0.032 (0.048)	-0.032 (0.046)
Expected Return to College		0.052* (0.028)
Per cap Income - 5 to 10k	0.066* (0.039)	0.068* (0.039)
Per cap Income - more than 10k	0.064 (0.043)	0.063 (0.042)
Controls for Exp Log Earn, Prob of Work and Var of Log Earn	Yes	Yes
Control's for Ability, Father's Educ, State, Gender, Marital Status Dummies	Yes	Yes
Observations	3680	3680
Censored Obs	2009	2009
Log Likelihood	-3252.215	-3246.406
Sample Sel.: Corr. betw. Errors	0.289	0.331
P-Value of LR test of Indep Eqns	0.303	0.247

Notes: Table displays marginal effects and standard errors in brackets. * p<0.1, ** p<0.05, *** p<0.01. Excl. categories: male, single, lowest GPA tercile, father's education primary or less, per capita income less than 5000 pesos.

Table 4: Probit model for college attendance decision: Excess responsiveness of poor high-expected-return individuals to changes in direct costs.

Dep. Var.: Attend College	Model 8 Marg. Eff. (S.E.)	Model 9 Marg. Eff. (S.E.)
Tuition > 750 pesos * Pcap Income < 5k	-0.030	0.037
	(0.031)	(0.045)
Tuition > 750 pesos * Pcap Income < 5k * Exp Return high		-0.110**
		(0.044)
Tuition > 750 pesos * Pcap Income 5 - 10k	-0.022	-0.052
	(0.043)	(0.051)
Tuition > 750 pesos * Pcap Income 5 - 10k * Exp Return high		0.054
		(0.075)
Tuition > 750 pesos * Pcap Income > top	0.068	0.031
	(0.060)	(0.076)
Tuition > 750 pesos * Pcap Income > 10k * Exp Return high		0.050
		(0.090)
Expected Return to College		0.075**
		(0.035)
Per cap Income - 5 to 10k	0.056	0.055
	(0.035)	(0.035)
Per cap Income - more than 10k	0.096**	0.093**
	(0.042)	(0.041)
Controls for Exp Log Earn, Prob of Work and Var of Log Earn - HS	Yes	Yes
Control's for Ability, Father's Educ, State, Gender, Marital Status Dummies	Yes	Yes
Observations	3680	3680
Censored Obs	2009	2009
Log Likelihood	-3247.959	-3238.761
Sample Sel.: Corr. betw. Errors	0.092	0.131
P-Value of LR test of Indep Eqns	0.751	0.649

Notes: Table displays marginal effects and standard errors in brackets. * p<0.1, ** p<0.05, *** p<0.01. Excl. categories: male, single, lowest GPA tercile, father's education primary or less, per capita income less than 5000 pesos.

Table 5: Maximum Likelihood Estimation of the participation equation of college attendance.

Participation Equation		
Dep. Var.: Attend College	Coefficients	Std. Err.
Costs		
University Distance	-.0096	.0035***
University Distance Squared	.0001	.0000**
GPA of Junior High School	.0181	.0042***
Mother's Schooling	.0479	.0141***
Benefits		
Exp Return to College	.2518	.1084**
Difference in Var of College and HS Earnings	2.4630	1.2998*
(Exp Return + Var Difference) Squared	-.1146	.06083*
Constant		
$-(1+r)$	-10.2934	.4874***
Log-Likelihood	-550.8655	
Wald Chi Square (8)	26.98***	
N of observations	1057	

Notes: Significance levels: *** 1%, ** 5%, * 10%.

Table 6: Effect of changes in variables compared to the baseline case.

	Dist to Univ	Mother's schooling	GPA of Junior HS	Exp Gross Return to College	Prob of Attending	Diff
1	18.24	5	82	0.6213	0.2213	0.013
	13.24	5	82	0.61	0.2341	
2	18.24	5	82	0.6213	0.2213	0.018
	18.24	6	82	0.6213	0.2397	
3	18.24	5	82	0.6213	0.2213	0.007
	18.24	5	83	0.6213	0.2280	

Notes: For the baseline case all variables are evaluated at their median. One variable at a time is changed.

Table 7: Effect of changes in expected returns to college at different baseline cases.

	Dist to Univ	Mother's schooling	GPA of Junior HS	Exp Gross Return to College	Prob of Attending	Diff
4	18.24	5	82	0.6213	0.2213	0.010
	18.24	5	82	0.7213	0.2318	
5	13.24	5	82	0.6213	0.2213	0.024
	13.24	5	82	0.7213	0.2449	
6	18.24	6	82	0.6213	0.2396	0.011
	18.24	6	82	0.7213	0.2506	

Notes: For the first baseline case all variables are evaluated at their median. Then the effect of a change in expected returns is evaluated at different baselines.

Table 8: Results of Policy Experiments: Change in the College Attendance Rate, and the Average Marginal Expected Return compared to the Average Expected Return of the Individuals Attending College.

Policy Change	Individuals changing college attendance decision: Change in Overall Attendance Rate (in %)	Individuals changing college attendance decision: **Marginal** Expected Return	Individuals attending college: **Average** Expected Return	P-Value of One-Sided Test
Decrease Interest Rate by 10%	80%	0.88	0.72	0.25
for very poor	50%	1.11	0.72	0.08
for very poor and able	32%	1.14	0.72	0.04
Decrease Dist by 10km	10%	1.10	0.72	0.11
for very poor	3%	1.08	0.72	0.15
for very poor and able	2%	1.11	0.72	0.08

8 Appendix B

8.1 Derivation of the Participation Equation from the Model of College Attendance Choice

In order to use the potential outcome equations (3) and the subjective expectation information (4), and rewrite the participation equation in terms of expected returns to college, I use the following approximation

$$E(Y_{ia}) \equiv E(e^{\ln Y_{ia}}) \cong e^{E(\ln Y_{ia})+0.5Var(\ln Y_{ia})} \tag{17}$$

and assume that $Var(\ln Y_{ia}^S) = (\sigma_i^S)^2$ for all a and $S = 0,1$. Thus I can rewrite the expected present value of university earnings (analogously for high school earnings) as

$$\begin{aligned}
EPV(Y_i^1) &= \sum_{a=22}^{\infty} \frac{\exp(E(lnY_{ia}^1) + 0.5Var(lnY_i^1))}{(1+r)^{a-18}} \\
&= \sum_{a=22}^{\infty} \frac{\exp\left(\tilde{\alpha_1} + \beta_1' X_i + \gamma a + \theta_1' f_i + 0.5(\sigma_i^1)^2\right)}{(1+r)^{a-18}} \\
&= \frac{\exp\left(\tilde{\alpha_1} + \beta_1' X_i + \theta_1' f_i + 0.5(\sigma_i^1)^2\right)}{(1+r)^4} \cdot \left(\sum_{a=22}^{\infty} \frac{\exp(\gamma a)}{(1+r)^{a-22}}\right) \\
&= \frac{\exp\left(\tilde{\alpha_1} + \beta_1' X_i + \theta_1' f_i + 0.5(\sigma_i^1)^2\right)}{(1+r)^4} \cdot \exp(\gamma 22) \left(\frac{1}{1-\left(\frac{\exp(\gamma)}{(1+r)}\right)}\right),
\end{aligned} \tag{18}$$

where I assume that $\exp(\gamma) < 1 + r$ to apply the rule for a geometric series. Analogously, I can derive the following expression for $EPV(Y_i^0)$

$$EPV(Y_i^0) = \exp\left(\tilde{\alpha_0} + \beta_0' X_i + \theta_0' f_i + 0.5(\sigma_i^0)^2\right) \cdot \exp(\gamma 18) \cdot \left(\frac{1+r}{1+r-\exp(\gamma)}\right). \tag{19}$$

Using expression (18) and (19), I can write the decision rule in the following way:

An individual decides to attend college if $EPV(Y_i^1) - EPV(Y_i^0) \geq C$, and thus if

$$\begin{aligned}
&\exp\left(\tilde{\alpha_1} + \beta_1' X_i + \theta_1' f_i + 0.5(\sigma_i^1)^2\right) \cdot \left(\frac{\exp(\gamma 22)}{(1+r)^4}\right) \cdot \left(\frac{1+r}{1+r-\exp(\gamma)}\right) \\
&- \left(\exp(\tilde{\alpha_0} + \beta_0' X_i + \theta_0' f_i + 0.5(\sigma_i^0)^2)\right) \cdot \exp(\gamma 18) \cdot \left(\frac{1+r}{1+r-\exp(\gamma)}\right) \geq C,
\end{aligned}$$

which I can rewrite in the following way

$$\begin{aligned}
\exp\left(\tilde{\alpha_1} + \beta_1' X_i + \gamma \cdot 25 + \theta_1' f_i + 0.5(\sigma_i^1)^2\right) - \left(\exp(\tilde{\alpha_0} + \beta_0' X_i + \gamma \cdot 25 + \theta_0' f_i + 0.5(\sigma_i^0)^2)\right) \cdot \\
\exp(-\gamma 4) \cdot (1+r)^4 \geq C \exp(\gamma 3)(1+r)^3(1+r-\exp(\gamma)).
\end{aligned}$$

Making use of the 'subjective' expectation information, this is equivalent to

$$\exp\left(E(\ln Y_{i25}^1) + 0.5(\sigma_i^1)^2\right) - \left(\exp(E(\ln Y_{i25}^0) + 0.5(\sigma_i^0)^2)\right) \cdot \exp(-\gamma 4) \cdot (1+r)^4$$
$$\geq C \exp(\gamma 3)(1+r)^3(1+r-\exp(\gamma)). \tag{20}$$

In order to express the decision rule (20) in terms of expected gross returns to university and use the information on expected returns from 'subjective' expectations of earnings (see expression (5)), 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

$$\begin{aligned} B - A &= \left(E(\ln Y_{i25}^1) + 0.5(\sigma_i^1)^2\right) - \left(E(\ln Y_{i25}^0) + 0.5(\sigma_i^0)^2\right) \\ &= \rho_i + 0.5\left((\sigma_i^1)^2 - (\sigma_i^0)^2\right), \end{aligned}$$

I can write the decision rule as

$$\exp\left(E(\ln Y_{i25}^0) + 0.5(\sigma_i^0)^2\right) \cdot \left(\sum_{j=0}^{\infty}\frac{\left(\rho_i + 0.5\left((\sigma_i^1)^2 - (\sigma_i^0)^2\right)\right)^j}{j!}\right)$$
$$-\left(\exp(E(\ln Y_{i25}^0) + 0.5(\sigma_i^0)^2)\right) \cdot \exp(-\gamma 4) \cdot (1+r)^4 \geq C\exp(\gamma 3)(1+r)^3(1+r-\exp(\gamma))$$

and rearranging will lead to

$$\left(\sum_{j=0}^{\infty}\frac{\left(\rho_i + 0.5\left((\sigma_i^1)^2 - (\sigma_i^0)^2\right)\right)^j}{j!}\right) - \exp(-\gamma 4)\cdot(1+r)^4 \geq \frac{C\exp(\gamma 3)(1+r)^3(1+r-\exp(\gamma))}{\exp\left(E(\ln Y_{i25}^0) + 0.5(\sigma_i^0)^2\right)}.$$

Thus using the 'subjective' expectation information, the latent variable model for attending university can be written as

$$\begin{aligned} S^* &= \left(\sum_{j=0}^{\infty}\frac{\left(\rho_i + 0.5\left((\sigma_i^1)^2 - (\sigma_i^0)^2\right)\right)^j}{j!}\right) \\ &\quad -(1+r)^4\left(\exp(-\gamma 4) + \frac{C\exp(\gamma 3)}{\exp\left(E(\ln Y_{i25}^0) + 0.5(\sigma_i^0)^2\right)}\right) \\ &\quad +(1+r)^3\left(\frac{C\exp(\gamma 4)}{\exp\left(E(\ln Y_{i25}^0) + 0.5(\sigma_i^0)^2\right)}\right) \qquad (21) \\ S &= 1 \text{ if } S^* \geq 0 \\ S &= 0 \text{ otherwise,} \end{aligned}$$

where S is a binary variable indicating the treatment status.

8.2 Testable Predictions about Excess Responsiveness to Changes in Direct Costs

Making use of the participation equation for college attendance (21), the following results show that individuals who face a higher interest rate are more responsive to changes in direct costs.

$$\frac{\partial S^*}{\partial C} = \frac{-(1+r)^4 \left(\exp(\gamma)\right)^3 + (1+r)^3 \left(\exp(\gamma)\right)^4}{\exp\left(E(\ln Y^0_{i25}) + 0.5(\sigma^0_i)^2\right)} < 0$$

as $\exp(\gamma) < 1+r$ by assumption to apply the rule for geometric series (see previous section), and

$$\frac{\partial^2 S^*}{\partial C \partial r} = \frac{-4(1+r)^3 \left(\exp(\gamma)\right)^3 + 3(1+r)^2 \left(\exp(\gamma)\right)^4}{\exp\left(E(\ln Y^0_{i25}) + 0.5(\sigma^0_i)^2\right)} < 0 \tag{22}$$

as $4(1+r) > 3\exp(\gamma)$.

Thus $\left|\frac{\partial S^*}{\partial C}\right|$ is increasing in r, that is individuals who face a higher interest rate are more responsive to changes in direct costs.

8.3 Derivation of the Marginal Return to College

Proof for deriving equation (13):

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

9 Appendix C

9.1 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 9). 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 9). Figure 7 displays college attendance rates of 18 to 24 year old high school graduates for different parental income quartiles.[28] High school graduates are already a selective group, as only about 54% of the relevant age group attain a high school degree. The attendance rate of individuals 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 8, 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 (see section 3.4). Forty 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 a university is close to the location of residence. In the latter case she can commute while taking advantage of the economies of scale of living with her family. I therefore construct a measure of distance to the closest university for each individual (see section 3.4).

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

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 9). 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% 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 (8.0) 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.

Figure 7: College enrollment rates of 18 to 24 year old high school completers by parental income quartile (Mexican Family Life Survey, 2003).

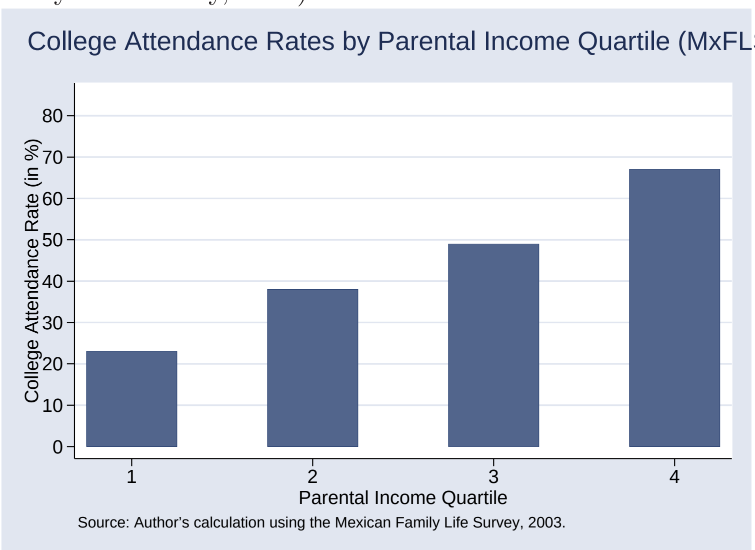

Figure 8: College enrollment rates of 18 to 24 year old high school completers by parental income quartile (Jovenes con Oportunidades Survey, 2005).

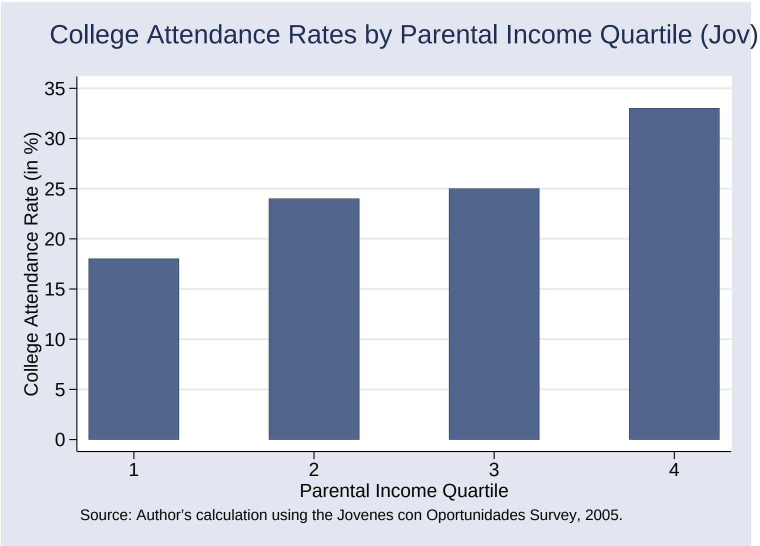

Table 9: Comparison of enrollment rates, fraction of poorest 40% in percent of the student population, fraction of GDP spend on higher education, fraction of expenditures on higher education on fellowships and student loans: Mexico, other Latin American countries, OECD and USA.

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 ".".

9.2 Potential Sample Selection Problem

The interviewer visited the primary sampling units and their families in October and November 2005 and interviewed the household head or spouse using the family questionnaire and adolescents between age 15 and 25 using the "Jovenes" (youth) questionnaire. If the adolescent was not present, the household head or spouse answered the Jovenes questionnaire as well. As a result the questions on expected earnings were not answered by the adolescent herself for about half the sample, i.e. mothers state their expectations about future earnings of her child(ren) that are not present during the interviewer's visit.

Table 10 compares summary statistics of important variables for the two groups of respondents. College attendance rates are significantly lower in the case that the adolescent responds, which raises concerns about sample selection in the case of using only adolescent respondents. Individuals who attend college –in particular if they live far from the closest university– are less likely to be at home at the time of the interview. Sample selection can –at least partially– be explained by observable variables: adolescent respondents live significantly closer to the closest university, are significantly more likely to be female (as many families do not want their female children to live on their own away from home) and have lower per capita household expenditures. On the other hand, variables such as expected returns to college as well as ability, father's years of schooling and per capita parental income do not differ significantly between the two groups.

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

Respondent	Mother Mean (SE)	Adolescent Mean (SE)	Diff. Mean
Attend College	35.8%	23.1%	-12.7%***
	(0.48)	(0.42)	
Female	50%	58%	8%***
	(0.50)	(0.49)	
Ability (GPA)	82.3	82.2	-0.1
	(10.34)	(7.16)	
Father's Yrs of Schooling	5.3	5.4	0.1
	(3.00)	(2.99)	
Per Cap Parental Income	7493.25	7472.02	-21.23
	(7635.84)	(7909.00)	
Expected Gross Return	0.65	0.66	0.01
$(E(\ln Y_{\text{Col}}) - E(\ln Y_{\text{HS}}))$	(0.36)	(0.38)	
Distance to Univ	27.40	24.16	-3.24***
	(23.35)	(22.72)	
Observations	2010	1670	

Notes: Significance of difference in means is displayed as follows, * p<0.1 ** p<0.05 *** p<0.01.

10 Appendix D: Robustness Checks

Figure 9: Scatter Plot and Nonparametric Regression of Expected Returns on Distance to College.

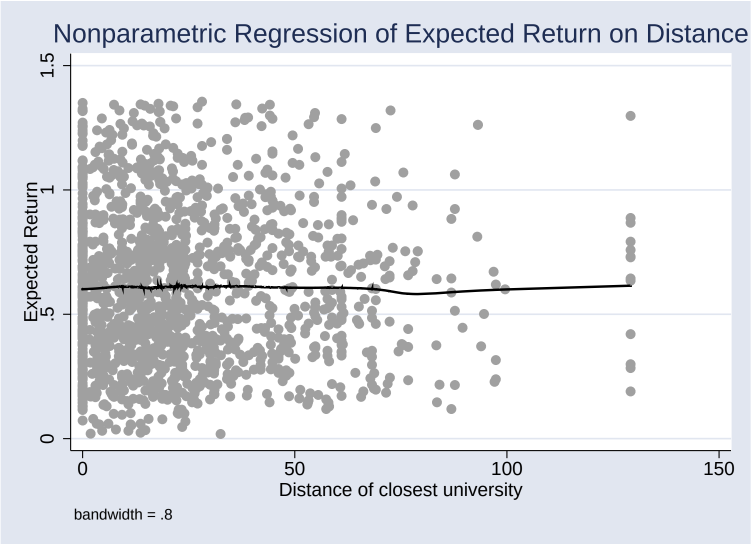

Table 11: Correlation between Expected Returns and Distance to College.

Dep. Var.: Expected Return	Coeff. (S.E.)
Distance to Univ	0.001
	(0.001)
Distance Squared	-0.000
	(0.000)
GPA - second tercile	0.044**
	(0.020)
GPA - top tercile	0.069***
	(0.019)
Father's Educ - junior HS	0.025
	(0.025)
Father's Educ - senior HS	0.084*
	(0.045)
Father's Educ - Univ	-0.126
	(0.115)
Per cap Income - 5 to 10k	-0.004
	(0.019)
Per cap Income - more than 10k	0.005
	(0.021)
Observations	3493
Censored Obs	1916
Lambda	-0.028
S.E. of Lambda	0.054

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

Table 12: First stage results of the sample selection correction in the probit model for college attendance

Dep. Var.: Adolescent Responds	Model 1	Model 2
Interview Sunday	0.117**	0.095
	(0.058)	(0.059)
Interview Thursday	-0.070**	-0.078**
	(0.035)	(0.036)
Interview Sunday*Evening	-0.202*	-0.194*
	(0.106)	(0.109)
Interview Saturday*Evening	0.313***	0.344***
	(0.082)	(0.078)
Interview Week 40	0.162***	0.162***
	(0.058)	(0.059)
Interview Week 41	0.133***	0.159***
	(0.031)	(0.031)
Interview Week 42	0.098***	0.105***
	(0.027)	(0.028)
Sex		0.092***
		(0.017)
Married		0.344***
		(0.068)
GPA - second tercile		0.066***
		(0.021)
GPA - top tercile		-0.023
		(0.020)
Father's Educ - junior HS		-0.028
		(0.027)
Father's Educ - HS		0.015
		(0.054)
Father's Educ - Univ		-0.141
		(0.096)
Per cap Income - 5 to 10k		0.009
		(0.021)
Per cap Income - more than 10k		0.040
		(0.024)
Dist to Univ 20 to 40km		0.019
		(0.021)
Dist to Univ more than 40km		-0.037
		(0.025)
Tuition more than 750 pesos		-0.009
		(0.028)
Observations	3680	3680
Log Likelihood	-2483.362	-2376.194
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, male, single, lowest GPA tercile, father's education primary or less, per capita expenditures and income less than 5000 pesos.