Education Choices and Returns on the Labour and Marriage Markets: Subjective Expectations, Gender and Decision Making.

Orazio P. Attanasio* and Katja M. Kaufmann†

February 2011, first draft: June 2008

Abstract

The goal of this paper is to improve our understanding of educational decisions in two dimensions: First, we investigate what are important determinants of schooling decisions and whether they differ for male and female youths. In particular, we are interested in the role of expectations about monetary returns to schooling, perceived risks of earnings and unemployment for different schooling scenarios and returns to schooling in the marriage market. Second, we analyze the decision-making process within the household with specific focus on the role of the youth. We shed light on whose expectations matter in schooling decisions, the ones of the parents or the ones of the youths, and whether this depends on the age and gender of the youth. To address these questions we use a data set on Mexican junior and senior high school graduates that elicits their own and their parents' beliefs about future earnings for different scenarios of highest schooling degree. In addition we construct proxies for returns in the marriage market and have information about the actual schooling choice and an extensive set of controls.

We find that boys care more about monetary returns to schooling than girls –in particular in the decision to attend college–, while marriage market considerations seem important only for girls. Risk perceptions matter and they are particularly important from the perspective of the parents. In terms of the intra-household decision process, our results indicate that while mothers decide about the schooling of their daughters for both high school and college, they are not involved in the college enrolment decisions of boys.

JEL-Classification: I21, I22, O15
KEYWORDS: Schooling Decisions, Subjective Expectations, Perceived Earnings and Employment Risk, Marriage Market Returns, Intra-Household Allocation, Gender Differences, Mexico.

*University College London, Institute for Fiscal Studies, NBER, CEPR and BREAD. o.attanasio@ucl.ac.uk

†Department of Economics and IGIER, Università Bocconi, Address: Via Roentgen 1, 20136 Milano, Italy, e-mail: *katja.kaufmann@unibocconi.it*. We would like to thank Sandra Black, Janet Currie, Pascaline Dupas, Lena Edlund, Claudio Ferraz, Caroline Hoxby, Eliana La Ferrara, Lance Lochner, Shaun McRae, Matthias Messner, Karthik Muralidharan, Bernard Salanie, Aloysius Siow, Kevin Stange, Alessandro Tarozzi, Duncan Thomas, Wilbert van der Klaauw and Basit Zafar for valuable comments. We are thankful also to conference participants at the NBER Education meeting in Spring 2010, at the CEPR labour economics program on the economics of family policy and the CEPR development meeting in Spring 2010, at the NYU/Caltech La Pietra conference, at the BREAD summer workshop, at the SITE meeting on Women and the Economy and at the CESifo conference on the Economics of Education and to seminar participants at Boston Fed, New York Fed, Rochester and UVA. All remaining errors are of course our own.

1 Introduction

The process of human capital accumulation has long been identified as one of the most important factors in development. And yet, in many developing countries, even with relatively high levels of income, the stock of education has grown painfully slowly, despite high returns to acquiring education. This is definitely the case for many countries in Latin America and a cause of concern for policy makers.

The goal of this paper is to improve our understanding of educational decisions among poor Mexican households in two dimensions: first, we analyze the decision-making process within the household with specific focus on the role of the youth. We shed light on whose expectations matter in schooling decisions, the ones of the parents or the ones of the youths, and whether this depends on the gender and age of the youth. Second, we investigate what are important determinants of schooling decisions and whether they differ for male and female youths. In particular, we are interested in the role of expectations about monetary returns to schooling, perceived risks of earnings and unemployment for different schooling scenarios and returns to schooling in the marriage market. To address these questions we use a data set on Mexican junior and senior high school graduates that elicits their own and their parents' beliefs about future earnings for different scenarios of highest schooling degree. In addition, we construct proxies for returns in the marriage market and have information about the actual schooling choice and an extensive set of controls.

Whether and at which age children should be considered as economic agents in household decisions is an important empirical issue. Especially in the case of older children, it is quite likely that parents are not the only ones who have a say, for example in education investment decisions. Older children are likely to have better earning opportunities and are more likely to be autonomous from parents. We suggest a novel approach to analyze intra-household decision-making based on the use of subjective expectations of the different members within the household. In particular, we test whether parents' or youths' expectations matter for education decisions (to the extent that they are different). Understanding whether parents' or youths' expectations matter is important, because not taking into account that children might be playing a role in household decisions –when they actually do– could result in misleading explanations of investment into higher education. For example, one might conclude that (parents') expectations about returns do not matter, while those of the youths do. This issue may have important implications for the design of public policy: the program *Oportunidades,* for instance, is considering the possibility to pay part of their schooling grants directly to the youths. This might improve the effectiveness of the intervention if youths play an important role in the decision about school attendance and parents have only imperfect control over the actions of the child.

We know surprisingly little about the importance of gender differences in the determinants of schooling choices. Are there differences between boys and girls in terms of having a say in schooling decisions? To what extent are the decision maker's subjective expectations on the returns to education different for boys and girls? To what extent do expected returns in the labour and marriage markets (and other variables) play a different role in the schooling decision for boys

and girls? These questions are important for understanding gender differences in educational attainment as well as for the effective design of programs intended to increase schooling among poor households.

Much of the existing literature on schooling decisions uses earnings realizations to measure or proxy for expected returns to education. This approach relies on strong assumptions about people's information sets and about how people form expectations. Moreover, such an approach has to deal with the issue that individuals with different levels of education whose earnings realizations are observed are typically selected into that education level. We use data on subjective expectations of earnings under different scenarios, and by doing so we avoid these problems.

We address our questions in the context of two schooling decisions: that of youths who have just finished junior high school to enrol (or not) in senior high school, and that of youths who have completed senior high school to enrol (or not) in college. We construct measures of expected earnings, perceived earnings risk and perceived unemployment risk held by parents and youths for different schooling scenarios, to study the extent to which these expectations affect education choices of male and female youths.

We first show that there are important differences in parents' and youths' expectations about earnings of one and the same person, the youth herself. We thus provide evidence on important differences in information sets and/or ways of forming expectations between parents and youths. This evidence in itself constitutes a strong justification for using direct measures of expectations. Moreover, it provides us with the opportunity of testing whether parents' or youths' expectations are relevant for schooling decisions.

We therefore proceed to estimate a model of schooling decisions to measure the importance of expected returns to education for boys and girls. In addition, we provide some evidence on the role of returns in the marriage market in education choices. We use two different proxies for returns: (i) the ratio of unmarried men to women in the locality of residence to capture the availability of partners; and (ii) the ratio of unmarried men to women with a certain level of schooling, if the suitability of the partner depends on having a similar (or higher) education level.

The idea for using these proxies is simple: if the likelihood of finding a (suitable) partner in the locality of residence is low, leaving the locality to go to college or high school has the benefit of increasing the chances of finding a (suitable) partner and of improving the outside option in the case that no partner is found. Due to social norms in Mexico, the link between schooling and marriage market considerations is particularly strong for girls, because parents would be very reluctant to let their unmarried daughter move to a different city to work and live by herself, while more willing to let her leave for continuing school (where she can live with classmates). Anecdotal evidence strongly points to the importance of marriage market considerations in schooling decisions: in Mexico there even exists the acronym "mmc"s (*'mientras me caso'*) for girls going to college to find a husband.

Our results on the intra-household decision-making process indicate that while mothers decide about the schooling of their daughters for both high school and college enrolment, they appear not to be involved in the college enrolment decisions of boys. Furthermore, we find that boys care more about monetary returns to schooling than girls –in particular in the decision to attend college.

Marriage market considerations on the other hand seem important only for girls. Risk perceptions matter and they are particularly important from the perspective of the parents.

Related literature Our paper is related to a growing literature using data on people's subjective expectations to understand educational decisions. 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. [1] Three closely related papers investigating the link between subjective expectations of earnings and schooling choices are Jensen (2010), Kaufmann (2009) and Nguyen (2008).[2] Jensen (2010) investigates how perceived returns to education affect schooling investments using data from the Dominican Republic. He finds that the students in his sample of 8th graders significantly underestimate returns to schooling. Informing a random subset of the youths 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 students in Madagascar about high returns to schooling increases their attendance rates and their test scores. Kaufmann (2009) uses the same Mexican data set as this paper to analyze the causes and consequences of the large income gradient in college attendance in Mexico. In particular, data on subjective individual expectations enable her to analyze to what extent the gradient can be explained by the poor expecting lower returns to college than the rich, for example due to ability and/or information differences. She finds that while differences in expected returns only explain part of the gap, poor individuals require significantly higher returns to be induced to attend college. Kaufmann (2009) tests implications of a school choice model in the presence of credit constraints and finds that expected returns of students at the margin of attending are significantly higher than average returns of people who already attend college, consistent with credit constraints for poor high-expected-return students. Relative to these papers, we address several complementary questions making use of unique features of the Mexican data.

One novel aspect of this paper is our focus on gender differences in the determinants of schooling. The only paper we know of that uses data on subjective expectations (but not on risk perceptions) to understand gender differences is by Zafar (2009), who analyzes the college major choice of undergraduate students at Northwestern University. In a very recent paper, Wiswall and Zafar (2011) conduct an experiment informing a random subset of NYU students about average earnings in the population and analyze how the resulting changes in expected earnings affect their major choice. Another two recent papers also show the link between schooling choices and returns to schooling

[1] See also the survey paper on the use of subjective expectations data by Manski (2004). There is, of course, a large literature on the (ex-post) effects of schooling on earnings and other outcomes. For monetary returns, see for example Card (2001), for non-monetary returns see for example Currie and Moretti (2003) on intergenerational returns and the survey by Lochner (2011) on the effects of education on crime, health and civic duties.

[2] Another related paper is by Stinebrickner and Stinebrickner (2009) about how college students from low income families form expectations about academic ability. Their results show that learning about ability plays a very prominent role in the college drop-out decision. Arcidiacono, Hotz, and Kang (2011) use data on subjective expectations of students at Duke University to explain college major choices.

–with some interest in gender differences– by making use of (quasi) experimental changes in returns to schooling. Abramitzky and Lavy (2011) make use of a policy change in Israeli kibbutzim, which increased returns to schooling. They observe that students in early-reforming kibbutzim increased their investment in education in terms of matriculation rates and scores and male youths responded particularly strongly. Jensen (2011) examines whether educational investments in girls in India respond to changes in employment opportunities. He provides recruitment services to young women in randomly selected rural villages helping them get jobs in the business process outsourcing industry and finds that this intervention increases schooling of girls in treatment villages. While the latter two papers do not use data on subjective expectations, these data enable us to also analyze if perceived earnings and unemployment risk play a role in educational decisions (and if their role differs for boys and girls) and to shed some light on who participates in the schooling decisions, the parents and/or the youth.

Furthermore, we also aim to shed light on the role of marriage market returns for schooling decisions of male and female youths. One of the few papers addressing this issue is Lafortune (2010) who explores whether second generation Americans modify their human capital acquisition decision when faced with a shift in the sex ratio of their state-level marriage market induced by immigration of their own ethnicity.

Lastly, our paper is related to the large (theoretical and empirical) literature on intra-household decision-making: [3] we study how different members within the household decide about educational attainment of the children. We are particularly interested in the role of children/youths as decision-makers and how this depends on the gender of the youth, a question that has not been studied much. Instead children have been modeled as household public goods (see, e.g., Blundell, Chiappori, and Meghir (2005)) or incorporated through the "caring preferences" of their parents (see, e.g., Bourguignon (1999)). Two notable exceptions are the following: Dauphin, Lahga, Fortin, and Lacroix (2008) test restrictions of a theoretical model on consumption decisions and infer the minimum number of decision-makers in three-person households from parametric constraints. Using UK data, they provide evidence that children start playing a role in consumption decisions at age 16. Giustinelli (2010) analyzes high school curriculum choices of Italian youths modeling the decision based on a direct questions about who is the decision-maker. Berry (2011) uses a very different approach and analyzes whether the identity of the recipient of cash incentives –either the parent or the child– can influence the effectiveness of conditional cash transfer programs.[4]

While our data on subjective expectations allow us to address several additional interesting questions –which would be complicated to address otherwise or require strong additional assumptions–, there are certain concerns that are sometimes raised with respect to using data on subjective expectations, such as endogeneity or ex-post rationalization. We will discuss these concerns in detail

[3] For papers rejecting the unitary model, see for example Thomas (1990), Fortin and Lacroix (1997) and Duflo (2003) and for papers on the collective model, see for example Browning and Chiappori (1998)).

[4] Berry (2011) conducts a field experiment in India giving incentives to achieve a specific reading goal. While this approach has the advantage to directly shed light on preferences of children versus parents, a potential caveat of this type of analysis could arise from the fact that the act of giving conditional cash to children might change the dynamics in the household. For example, before the introduction of the policy children might not have participated in educational decisions, but once they receive conditional cash they start playing a role.

and argue why we believe that these issues are not driving our results.

Outline The rest of the paper is organized as follows: Section 2 presents a basic model of education choices. Section 3 describes in detail the data, in particular the module on subjective expectations, and presents some summary statistics on attendance rates and a variety of background variables. Section 4 provides evidence that people are able to give meaningful answers to the subjective expectation questions and investigates how informed parents and male/female youths are about potential returns to schooling as well as if there are differences in expectations between parents and youth about future earnings of the youth. Section 5 discusses the results (including a robustness section on concerns related to the use of subjective expectation data). Section 6 concludes.

2 Schooling Decisions and Returns to Schooling.

The goals of this paper are two: First, we aim to improve our understanding of the main determinants of education choices and analyze whether they differ between boys and girls. For this purpose, we show how schooling decisions of young poor Mexicans relate to their expectations on the return and risk to that investment as well as to returns in the marriage market. Second, we intend to shed some light on the intra-household decision-process of schooling decisions in terms of whose expectations play a role and investigate again if there are gender differences.

2.1 Determinants of Enrolment.

One possible approach to analyze the relationship between enrolment and expected returns would be the construction of a full dynamic optimization model where individuals choose current activities taking into account current and future benefits and costs of the alternative choices. This type of models has been proposed, for instance, by Keane and Wolpin (1997) and used in a variety of contexts (see, for instance, Todd and Wolpin (2006) and Attanasio, Meghir, and Santiago (2005)). Kaufmann (2009) shows how data on people's subjective expectations can be used in a simple model of college enrolment choice. In this paper, rather than following this route, we present probit regressions that relate the probability of enrolment to several control variables, and to subjective expectations of earnings and proxies for returns in the marriage market. In a fully specified model, schooling decisions are determined by the entire probability distribution of future earnings under alternative scenarios. Here we assume that the effect of this distribution can be summarized by a few moments of the distribution of earnings at age 25. In addition, we control for current labour market conditions through state dummies and for family background and ability through several variables we observe in our data set.

Specifically, to model the decision to enrol in senior high school, having completed junior high, we use a latent index model we estimate on the sample of junior high school graduates. Denoting with S the enrolment decision ($S = 1$ if the individual decides to attend and $S = 0$ otherwise) we have:

$$S = 1 \Leftrightarrow S^* = \alpha + \sum_{z=2}^{3} \beta_z * \rho_z + \sum_{z=1}^{3} \gamma_z * Var(\log Y)_z + \sum_{z=1}^{3} \delta_z * P_z^W + R^M \mu + X'\theta + U > 0. \qquad (1)$$

where $z = 1, 2, 3$ denotes junior high school, senior high school and college, respectively. The vector X contains a number of control variables that are likely to affect the schooling decision, ranging from measures of individual ability to parental background and state fixed effects and θ denotes the vector of corresponding parameters. In terms of subjective expectation measures ρ_z denotes the expected (gross) returns to senior high school ($z = 2$) and college ($z = 3$) defined as $\rho_z = E(\log Y)_z - E(\log Y)_{z-1}$, where $E(\log Y)_z$ is the expected value of the distribution of (log) earnings at age 25 for the scenario that degree z ($z = 1, 2, 3$) is the highest completed by the youth. $Var(\log Y)_z$ denotes the variances of future earnings under different schooling scenarios and P_z^W the subjective probability of employment under different scenarios, as discussed further below. R^M is a proxy for marriage market returns.

We obviously expect the return to senior high school relative to junior high school, as perceived by the individual, to have a positive effect on the decision to attend senior high school. In equation (1), we also include the expected return to college to capture the fact that attending (and completing) high school provides the option to attend college later on.

In addition to the expected return, we also want our empirical model to take into account the possibility that the riskiness of a given investment might affect schooling decisions. For this reason, we enter, as determinants of the schooling decision, the variances of the future earnings under different schooling scenarios $Var(\log Y)_z$. Moreover, as the questions on future expected earnings are conditional on working, we enter the subjective probability of employment under different scenarios, P_z^W.

One would expect a high perceived earnings risk with a junior high school degree to have a positive effect on the probability of continuing to senior high school, and a high variance of log earnings with a senior high school degree to have a negative effect. On the other hand, for the decision to continue to senior high school, a high variance of log earnings after *college* increases the option value of continuing to senior high school. By enroling in senior high school, one can wait for additional information while still having the option to go to college. The specification in equation (1) is flexible enough to be able to capture these aspects.

Usually, proxies for ability, such as GPA and parental education, are supposed to capture differences in psychological costs of attending college as well as in the ability to benefit from high school or college through higher expected returns.[5] However, such variables are only imperfect proxies of the returns that an individual can obtain from her education. Skills are likely to be multi-dimensional and can hardly be captured even with good data on test scores. For this rea-

[5]For example, Cameron and Heckman (1998) and Cameron and Heckman (2001) address the question of credit constraints in college attendance decisions in the US by controlling for ability measures, such as AFQT score and parental education, that are supposed to capture differences in how much people can benefit from attending college. They show that as a result parental income loses significance, which they interpret as evidence against credit constraints in higher education in the US.

son, as individuals have idiosyncratic knowledge about these skills, data on individual subjective expectations can be very useful. More importantly, what matters for the individual's decision is her perception of her skills and her beliefs about how they affect future earnings, conditional on her information set at the time of the schooling decision. This provides a strong rationale for using "perceived" returns and "perceived" risk. We nevertheless also control for GPA and parental education both of which turn out to be very important determinants of the two schooling decisions we model. These variables might be proxying for the probability of completing senior high school or college and for preferences for education.

To investigate whether marriage market considerations are important for education choices, we include in the regressions proxies for returns in the marriage market, R^M. Using census data, we compute the ratio of unmarried men to women with at least a junior high school degree in the locality of residence. If there are few suitable partners available in the locality of residence, this is a push-factor to leave the locality.[6] In Mexico social norms are such that Mexican parents are generally reluctant to let their unmarried daughter move to a bigger city to work and live by herself, while they are more willing to let her leave home to attend school and live with classmates. This creates a link between schooling decisions and marriage market considerations, which is likely to be less strong for boys, as they could migrate to a bigger city to work and find a spouse. But even for boys in the presence of strong patterns of assortative mating it might be less costly to find an educated partner at school or college. For this reason we interpret "suitable partner" in terms of a partner who has a similar education level (or higher) and thus use the ratio of unmarried men to women with an education level as high or higher than the youth deciding about schooling.

Our proxy for the marriage market return could, potentially, hide different effects. One possibility is that such a variable captures 'peer' effects. For instance, if in a locality there are many girls that go to college, this fact might affect the probability that a given girl goes to college, all being equal. Notice, however, that our variable is the ratio of girls to boys with a certain level of education. To assuage these worries, we also consider an alternative proxy for the marriage market return: the ratio of unmarried men to women without constraining the measure to a specific education level. This measure might be more appropriate if the youth (for example, the male youth) cares less about the education level of the spouse.

We construct our proxies for marriage market returns using locality level data because we believe that this is the appropriate reference area (marriage "market") for most youths. This proxy is more likely to capture people's actual knowledge about availability of partners than when using a larger reference area such as a municipality.

To model the decision to enrol in college –taken by youths who have just completed senior high school–, an equation analogous to (1) can be used. The only modification we need to make is that we consider only the distributions of earnings under the two relevant scenarios for a youth who has just graduated from senior high school: to stay with the degree that she received already or to get a college degree ($z = 2, 3$). We therefore get:

[6] Edlund (2005) uses a similar argument to explain migration patterns that lead to young women outnumbering young men in urban areas. To proxy for marriage market considerations she uses the ratio of men to women in the municipality of residence and finds supportive evidence for this explanation using Swedish municipality data.

$$S = 1 \Leftrightarrow S^* = \alpha + \beta_3 * \rho_3 + \sum_{z=2}^{3} \gamma_z * Var(\log Y)_z + \sum_{z=2}^{3} \delta_z * P_z^W + R^M \mu + X'\theta + V > 0. \quad (2)$$

As in the case of the high school enrolment decision, we include measures of expected returns and perceived risks, while controlling for individual and family background characteristics and for state fixed effects. To capture marriage market considerations in college choices, we include the ratio of unmarried men to women –in general and with at least a senior high school degree– in the locality of residence.

To examine the question of gender differences in the determinants of schooling, we completely interact each regressor with a male and a female dummy to test for differential effects.

2.2 Whose Expectations Matter for Schooling Decisions, Parents' or Youths'?

To learn more about the decision-making process within the household, we ask whose expectations matter for education choices. Obviously, the answer is likely to depend on the age and potentially also on the gender of the child/youth. One is therefore likely to obtain different results when modeling the decision to attend primary, secondary and higher education. Especially for youths who have finished high school and are deciding whether to enrol in college, the assumption that all decisions are taken by parents might be too strong.

Of course, if youths' and parents' expectations were fully rational and based on the same information, they would coincide. However, if either the information set or the way it is processed differ, subjective expectations of the different actors might differ. One goal of this paper is to test whether expectations differ systematically between parents and youths and whose expectations are relevant for the decisions, while we allow for gender differences.

If one had data on both parents' and youths' expectations one could address this question from an empirical point of view. In particular, neglecting the variance terms, the model we would want to estimate for both school attendance decisions is as follows:

$$S = 1 \Leftrightarrow S^* = \alpha + X'\beta + R^M \mu + \gamma_P * Parents\prime Expectations + \gamma_Y * Youths\prime Expectations + W > 0. \quad (3)$$

Clearly the parameters γ_P and γ_Y are separately identified only if the two sets of expectations differ. In what follows we will show that this is the case in our data. As we discuss below, our data does contain information on the subjective expectations held by mothers as well as by the youth, but unfortunately this information is not as complete as we would like. Therefore the estimation of equation (3) involves the solution of a number of econometric problems we discuss below.

3 Data

The conditional cash transfer program *Oportunidades,* previously known as PROGRESA, has been associated since its inception with attempts to evaluate its impacts. In this spirit, when the program introduced in 2002/3 a new component known as *Jóvenes con Oportunidades,* an evaluation aimed at measuring its impact was started. The data we use was collected in 2005 as part of that evaluation. As we discuss below, in addition to standard variables, the survey contained a detailed subjective expectation module which we use extensively. In this section, we describe the data and its structure. We also describe in some detail the module used to elicit information on subjective expectations and report some evidence on the quality of these data.

3.1 The Survey

The evaluation survey of *Jóvenes con Oportunidades* was conducted in the fall 2005. *Jóvenes con Oportunidades* provides an additional grant to youths in the last three years of high school (preparatoria). This grant is deposited into a bank and can be accessed only upon graduation, if the recipient engages in one of several activities (such as going to college or starting a micro enterprise). Alternatively, the recipient has to wait for a year during which time the amount, about US$300, accumulates at the market interest rate.

The primary sampling units of the evaluation survey are individuals who have just graduated from senior high school or from junior high school and who are eligible for *Jóvenes con Oportunidades*. There are three eligibility criteria: being in the last year of junior high school (9th grade) or attending senior high school (10 to 12th grade), being younger than 22 years of age, and being from a family that benefits from Oportunidades.

The survey consists of a family questionnaire and a youth questionnaire administered to each household member aged 15 to 25. As a consequence, the youths for whom we have data are not only the primary sampling units but also their siblings, provided they are aged 15 to 25. In total we have about information on about 23,000 youths. The questions in the youth questionnaire were addressed directly to the youth. However, if a specific youth was not present during the interview, the module was answered by the main respondent, who is generally the mother.

It is important to keep in mind that the overall sample includes all youths aged 15 to 25, regardless of their schooling status. In our analysis we concentrate only on students who just finished grade 9 (junior high school graduates) and decide about enroling in senior high school, and students who just finished grade 12 (senior high school graduates) and decide about enroling in college or not. There are about 2500 observations for each of the two groups of interest. We use the data on the siblings sample for a number of robustness exercises on the expectations data.

The survey provides 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 next.

3.2 The Expectations Module

The subjective expectations module was designed to elicit information on the individual distribution of future earnings and the probability of working for different scenarios about the highest completed school degree. The module starts with a simple explanation of probabilities. In particular, individuals are shown a ruler, graded from zero to one hundred, which is then used to express the probability of future events. The example that is used to illustrate the concept of probability is the event of rain the following day.

After explaining the use of the ruler to express probabilities and having practised that with the rain example, the interviewer moves on to discuss future earnings and the probability of working under different schooling scenarios. The scenarios differ for students graduating from junior high school and those graduating from senior high school. For the former, the interviewer asks to consider three different possibilities: that the student stops after junior high, that the student goes on to senior high, completes it and stops and that the student goes on to college and completes it. For the latter, only two scenarios are considered: that the students stops at senior high school and that the student goes on to college and completes it.

For each of the relevant scenarios, the youth is then asked questions about the probability of working at the age of 25 and about future earnings at age 25. For example, in the case of the last scenario for a senior high school student, the questions are:

1. *Assume that you complete 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. *Assume that you complete 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?*

where 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. It is important to remember that these questions were included in the youth questionnaire and therefore are available for every youth in the household aged 15 to 25. Moreover, when a specific youth was not present, these questions were answered by the mother. This implies that, for some households, we have mother's expectations about returns to education. Notice, however, that we do not have father's expectations and that, in any case, the structure of the data poses a number of selection issues that we discuss at length below.

This type of subjective expectations questions has been used extensively in a variety of contexts. In a companion paper (Attanasio and Kaufmann (2008)), we discuss the internal and external validity of the answers to these questions in our survey. In that paper, we show that respondents seem to have understood the questions reasonably well and that the data pass a number of internal and external validity tests. Below, we report briefly some of those results and refer the interested reader to our paper for further details.

In what follows, we relate educational decisions to subjective expectations. This is possible because of the timing of the survey. The *Jóvenes* survey was conducted in October/November 2005, that is two or three months after high school and college had started and enrolment decisions had been made.

While the timing of the survey (which occurred just after the school enrolment decision had been taken) allows to use the expectations data to estimate an enrolment model, it might also raise a potential concern that individuals might try to rationalize their choices. For example, individuals who decided to enrol in college rationalize their choice by stating higher expected returns to college (that is higher college earnings and/or lower expected high school earnings), and those, who decided not to enrol, state lower expected returns to college. This would lead to a more dispersed cross-section of returns after the decision compared to before the enrolment decision.[7]

To check whether ex-post rationalization introduces biases in expectations, we look at the expectations data of the siblings of our main subjects. Let us consider, for instance, the college enrolment decision of senior high school graduates. We can compare the cross sectional distribution of the expectations of our senior high school graduates to the expectations of the cohort of youths who are one year younger (just starting grade 12 at the time of the survey in October/November). If the older youths are ex-post rationalizing their choices, we would expect a cross-sectional distribution of expected earnings and returns that is more spread out for the cohort of senior high school graduates who have decided already whether to enrol in college or not, compared to the distribution of the cohort that is one year younger and just starting grade 12. The same reasoning should hold for perceptions of the probability of working. In Section 5.3 we discuss this issue at length and present evidence that our main results cannot be explained by ex-post rationalization.

3.3 Calculation of Expected Earnings, Perceived Earnings Risk, and Expected Gross Returns to Schooling

We use the answers to the three survey questions (2(a)-(c)) described above, together with some additional assumptions, to compute moments of the individual earnings distributions and expected gross returns to high school and college. We are interested in the individual subjective distribution of future earnings $f(Y^z)$ under the three possible scenarios of final education attainment we are considering: junior high school ($z = 1$), senior high school ($z = 2$) and college ($z = 3$). The survey provides, for each individual, information on the support of the distribution $[y^z_{min}, y^z_{max}]$

[7] This is true unless people switch positions in the distribution in such a way that the resulting cross-section looks exactly the same as before. This can only be the case if people with *low* expected returns to college decide to enrol in college and now state high returns to college and vice versa.

and on the probability mass to the right of the midpoint, $y^z_{mid} = (y^z_{min} + y^z_{max})/2$, of the support, $p = Pr\left(Y^z > y^z_{mid}\right)$. Thus we need an additional distributional assumption, $f(\cdot)$, in order to be able to calculate moments of these individual earnings distributions, using the three pieces of information on y^z_{min}, y^z_{max} and p. Given such an assumption on $f(.)$, we can compute all the moments of any function of future earnings under different scenearios we are interested in. For example, the expected value of log future earnings for *each* individual and each scenario ($z = 1, 2, 3$) will be given by $E(\ln Y^z) = \int_{y^z_{min}}^{y^z_{max}} \ln y f_{Y^z}(y) dy$.

For many of the specifications we estimate below, we consider the (gross) returns to college and senior high school. We compute these as the difference between expected log college (senior high) earnings and expected log senior high school (junior high school) earnings, $\rho^z = E(\ln(Y^z)) - E(\ln(Y^{z-1}))$ for $z = 2, 3$.[8]

In Attanasio and Kaufmann (2008), we use three different distributional assumptions, step-wise uniform, bi-triangular and triangular (compare Guiso, Jappelli, and Pistaferri (2002)). The last two assumptions give more weight to the middle of the support and less to the extremes. The first, instead, implies a relatively large value for the total variance. We rule out the possibility that the density function is U-shaped, giving more weight to the extremes. In our companion paper, we show that the first moment of the individual distribution is extremely robust with respect to the underlying distributional assumption, while the second moment is obviously larger for the step-wise uniform distribution that puts more weight on extreme values. In this paper we present results based on the triangular distribution, but we perform robustness checks using the other two distributional assumptions and point out differences if they occur.

3.4 Mother and Youth Expectations.

To determine whose expectations matter for enrolment decisions, one would ideally want to have data on both parents' and youth' expectations. Unfortunately, we have to deal with two important issues. First, we do not have information on fathers' expectations. We discuss and interpret our results in the light of this drawback and present evidence for the around 20% of households in which there is no father present, because the mother is single, separated, divorced or widowed. Second, the questions on the subjective distributions of earnings were not asked to both mother and children at the same time.

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 youths between age 15 and 25 using the "Jovenes" (youth) questionnaire. If a youth was not present, the household head or spouse answered the Jovenes questionnaire as well. As a result, for almost half the sample, the questions on the subjective distribution of future earnings were not answered by the youth herself. Instead mothers stated their expectations about future earnings of her child(ren) that are not present during the interviewer's visit.

[8] Notice that, as we have no information on the covariance of earnings under different scenarios, we cannot compute the variance of the return. For this reason, to capture the risk aspect of education choices we consider the variance of (log) earnings.

While it is unfortunate that we do not have information on the expectations of both mothers and children, the fact that for half the sample the earnings expectations questions were answered by youths, while for the other half the questions were answered by the mother allows to ask whether expectations held by mothers and youths about the youths' future earnings are different. If subjective expectations of mothers and youths were objective and rational expectations based on the same information, it would not matter who would answer (and the issue of whose expectations matter would be a moot one). It is therefore interesting to establish, whether the expectations of future earnings are systematically different depending on who answered the question.

In Section 4, we compare the expectations of mothers and youths. A straight comparison shows that these expectations are systematically different. Given the structure of the data, these differences can arise either because the questions answered by the mother and by the youth are measuring two different and distinct objects (the subjective probability held by the mother and the subjective probability held by the youth) or because the sample of youths absent from the interview (and for whom the question is answered by the mother) is systematically different from those present during the visit.

To take into account the possibility that the observed differences are due to sample selection and test for the presence of actual differences in expectations, we use a standard Heckman two-step approach (see Heckman (1979)). To achieve non-parametric identification of such a selection model, we need one or more variables that determine whether the question is answered by the youth rather than the mother and that, plausibly, do not affect the expectations directly. For such a purpose, we use information on the timing of the interview (the time of day of the interview, whether it took place on a weekday or weekend and whether it took place during weeks of holiday or not) to estimate the following equation:

$$R = 1 \Leftrightarrow R^* = \delta + X'\kappa + Z'\lambda + \epsilon \geq 0, \tag{4}$$

where $R = 1$ indicates a youth respondent and the vector Z includes the timing of the interview variables. We will show that the timing of the interview is a strong predictor of whether the youth is at home to answer the expectation questions. In what follows, we will be using this selection model both to test for differences between youths' and mothers' expectations and to estimate the school decision model, as we discuss below.

3.5 Marriage Markets Returns.

To proxy for the returns that education might have in the marriage market, we measure the ratio of unmarried men to women in the locality of residence of the youth in the sample. For the young cohort we construct the ratio of unmarried men to women who are in a similar age range as the youth, that is age 15 to 30 in the locality of residence, for the old cohort we use the ratio of unmarried men to women in the age range of 18 to 35.

To construct those proxies, we make use of Census data of the year 2000,[9] for which information

[9]The 2000 Census, "XII Censo General de Poblacion y Vivienda", can be found on the website of the Mexican

is available at the locality level. Having information at the locality level is important, because we believe that our "marriage market return" proxy works best in terms of capturing people's knowledge about availability of partners in the case of a small reference area (marriage "market") and we believe that the locality of residence is the correct reference area for most of the youths.

Unfortunately, Census information is only available for part of the localities in which the youths of our sample live. For the sample of junior high school graduates we lose around 30% of our observations. For the sample of senior high school graduates, who live in significantly smaller localities than the junior high school graduates (see descriptive statistics in the next section), we lose close to 60% of observations. For this reason we only use this significantly smaller sample for our analysis of marriage market returns, while we use the full sample for the first set of results on intra-household decision-making and labor market returns.

4 Subjective Expectations.

Before using the data on subjective expectations to model schooling choices, we describe the general patterns and provide some evidence on their quality. Further details on internal and external validity checks are contained in the companion paper Attanasio and Kaufmann (2008). We also discuss some of the econometric issues involved with the use of these data.

4.1 Descriptive Statistics: Characteristics of the Youth.

Tables 1 and 2 report summary statistics of individual and family background characteristics for the "young" cohort (junior high school graduates) and the "old" cohort (senior high school graduates), respectively. We present results separately for male and female youths and for the two samples of mother and youth respondents to investigate whether there are potential sample selection problems when performing the analysis separately for each type of respondent.

Around 79% of boys with a junior high school degree decide to enrol in senior high school compared to 83% of girls. In terms of college enrolment 29% of male high school graduates enrol in college compared to 32% of girls. Mexico has reached the stage where boys' and girls' educational attainment is very similar (or even higher in the case of high school enrolment), while for example in the US girls have overtaken boys in terms of both high school and college attainment (see Goldin, Katz, and Kuziemko (2006)). Comparing the samples of mother and youth respondents, enrolment rates are significantly lower for the youth sample in the case of the old cohort. For example, enrolment rates for girls are 24% for the youth sample compared to 40% of the mother sample (24% versus 33% for boys), which can be explained by the fact that youths who enrol in college are less likely to be at home when the interviewer arrives to conduct the survey (and they are thus less likely to be in the youth sample). In the case of the young cohort, there are no significant differences in enrolment rates between the samples of mother and youth respondents.

In what follows, we make use of data on GPA (grade point average between 0 and 100) of junior high school as a proxy for academic achievement. Tables 1 and 2 show that girls have a significantly

statistical institute, INEGI.

higher GPA than boys for both junior and senior high school graduates. This is consistent with empirical evidence for many countries that girls outperform boys at school (see Goldin, Katz, and Kuziemko (2006)). Comparing Tables 1 and 2, we notice that individuals of the older cohort have a slightly higher GPA than those of the young cohort. This most likely reflects the fact that the senior high school graduates are a self-selected sample compared to the junior high school graduates who might or might not attend and finish senior high school. In the case of the young cohort, there are significant (on 10%) but small differences between the sample of mother and youth respondents for girls, with higher GPA for the mother sample (82.8 versus 82.1). For the old cohort, there are significant but small differences between mother and youth sample, this time only for boys, again with higher GPA for the mother sample (81.8 versus 81).

To control for parental education as one of the most important determinants of children's schooling choices, we use information on parents' years of completed schooling in the form of four education dummies for both mother and father (unless the household is single headed): for primary education, junior and senior high school and university. For the young cohort, about 70% of mothers and fathers have only some primary education, while around 25% have attended junior high school. Fathers' education is slightly higher than that of mothers: about 6% have attended senior high school (3% for mothers) and 1% have some university education (0.7% for mothers). For the old cohort, parents are slightly less educated (75-80% primary and 17-18% secondary education). There are few and only small differences between mother and youth sample: for the young cohort, mothers in the mother sample are slightly less educated in the case of boys. For the old cohort, mothers in the mother sample are slightly less educated in the case of girls.

We create three per capita income categories, where the thresholds are equal to twice and four times the minimum wage.[10] For the young cohort, about 40% of the sample is in the lowest income category (that is yearly per capita income is below 5000 pesos) and thus relatively poor, reflecting the fact that our sample only consists of Oportunidades families. 35% are in the second highest category (5000 to 10000 pesos) and the remaining 25% in the top income category. The old cohort is slightly poorer with about 55% in the lowest income category. Differences between mother and youth sample are only significant in one instance for the young cohort, that is for girls the mother sample is poorer –consistent with lower parental education (see above).

The same pattern for young and old cohort can be found in terms of father's occupation: For the young cohort, 30% of fathers are unskilled workers, another 50% employees, around 18% are self-employed, 1.5% are family workers and less than 0.9% are employers. For the old cohort the numbers are 43%, 28%, 24%, 3.5% and 0.8%, respectively. Thus fathers in the youth sample are significantly less likely to be unskilled workers or self-employed and more likely to be employees. Comparing mother and youth samples, fathers in the youth sample are significantly more likely to be employees (and less likely to be self-employed or unskilled worker) than in the mother sample.

[10]Per capita parental income is constructed using parental income –such as parents' labour earnings, other income sources such as rent, profits from a business, pension income etc and remittances– divided by family size. Median yearly per capita income is 6066 pesos (approximately 606 US$). We add the income measures in the form of dummies to allow –in a flexible way– for nonlinear effects of income. The reason for the chosen income thresholds is their approximate correspondence with eligibility requirements for receiving fellowships (even though they are quantitatively not very important, see Kaufmann (2009)).

Part of the analysis that follows is conducted on the subsample of households where the father is not present because the mother is single, separated, divorced or widowed. In our sample about 22% (15%) of students of the young (old) cohort live in a household without father. There are no significant differences between boys and girls. Differences between mother and youth sample are only significant for the old cohort in the case of boys, where boys in the youth sample are more likely to live without a father. This can be explained by boys being more likely to stay at home, if there is no father in the household.

Another potentially important factor in schooling decisions is the number of siblings (in addition to the resource effect which we aim to capture by controlling for parental income divided by family size). For the young cohort, male and female youths have around 1.4 brothers and 1.4 sisters. In the case of the old cohort, male youths have around 1.44 brothers, female youths have significantly less brothers (1.36), while male youths have less sisters than female youths (1.26 versus 1.46). This pattern is similar for the young cohort but not significant. Differences between the youth sample and mother sample are only significant for the old cohort: Boys and girls in the mother sample have significantly more brothers (on 10%).

Turning to our proxies for marriage market returns, for the young cohort the ratio of unmarried men to women is 1.07 with a standard error of 0.18 for boys and 0.16 for girls, while the ratio of unmarried men to women with at least a junior high school degree in the locality of residence is 0.99 (standard deviation 0.27 for boys and 0.24 for girls). For the old cohort the ratio of unmarried men to women is 1.11 for boys and 1.12 for girls (the standard deviation is around 0.22 for boys and 0.26 for girls), while the ratio of unmarried men to women with at least a senior high school degree is 1.05 for boys and 1.08 for girls (standard deviation 0.56 and 0.66 respectively). There are no significant differences between the youth and mother sample nor between boys and girls.

As marriage market prospects also depend on the size of the marriage market, we control for locality size. Around 50% of junior high school graduates live in a locality of residence with below 15000 inhabitants, while around 20% live in localities between 15000 and 50000 people and the remaining 30% come from localities of more than 50000 people. When we compare the youth sample to the mother sample, adolescents from the youth sample generally come from significantly larger localities. This can be explained by the fact that youths in larger cities are more likely to have a school close to home and thus live at home, so that the interviewer is more likely to find them when conducting the survey. For the old cohort the majority comes from a locality of residence with below 15000 inhabitants (85% of the sample), while about 6% live in localities between 15000 and 50000 people and another 9% live in localities above 50000 people. Once again adolescents from the youth sample are more likely to come from larger cities.

Public schools in Mexico do not charge tuition fees, while public universities charge for enrolment and tuition. In addition, in the case of university education, youths are much less likely to have a university close to their locality of residence compared to senior high schools. Therefore direct costs will be particularly important for the decision to attend college. We proxy for costs of living using the distance from the locality of residence of the youth to the closest university and for tuition costs in nearby universities (for the data sources and the exact definition of the variables, see Kaufmann

(2009)). Around 50% of the youths live less than 20 km from the closest university (a distance which might allow daily commute), while around 25% live either between 20 and 40 km or more than 40 km from the closest university. There are no significant differences between boys and girls or mother and youth sample. Concerning tuition costs, about 40% of youths face tuition costs of more than 750 pesos, which corresponds to more than 15% of per capita median income and thus implies a substantial burden in particular for poor families. Tuition costs are significantly higher in the case of the youth sample compared to the mother sample, because youths in the youth sample live in larger cities with universities that charge higher fees.

We can summarize the evidence from Tables 1 and 2 by saying that the features of the sample reflect that we are working with families that are beneficiaries of a welfare program targeted to the poorest sector of Mexican society. The differences between girls and boys we observe are conform to our expectations and confirm findings of other empirical studies, for example the fact that girls outperform boys at school. There are some (usually small but) significant differences in individual characteristics and family background variables between the subsample where the earning expectations questions were answered by the youth and the subsample where they were answered by the mother. The selection process is very similar for boys and girls: For the young cohort, the youth sample is from slightly more educated and richer families who live in larger localities. For the old cohort, the youth sample is again more likely to be from larger localities, to have fathers who are employees and to have less siblings. Finding differences between mother and youth sample points towards a potential sample selection in our analysis of these questions, which we address, as discussed below. We now turn to analyze the differences in subjective expectations.

4.2 Youth vs Mother Respondents: A Selection Equation.

As mentioned above, to test whether observed differences in expectations of mothers and youths reflect actual differences in expectations or differences in the composition of the two samples, we estimate the selection equation (4). In particular, we model the probability that the youth answers the questionnaire (rather than her mother) as a function of all the variables in Tables 1 and 2 and of a set of variables capturing the timing of the interview. For brevity, we discuss the main findings here, while we report the complete set of results in Tables 16 and 17 in Appendix B.

For the young cohort, (male and female) youths who answered the subjective expectations questions themselves (because they were present at the moment of the interview) are more likely to live in larger localities, to have fathers who are employees instead of unskilled workers and to have a slightly lower GPA. For the old cohort, youths who answer themselves are more likely to have a father who is an employee, have less siblings and are more likely to be from a family without father present. Reassuringly, the selection process is very similar for boys and girls.

When considering the timing variables that identify the possible differences between the two sets of expectations (and below the schooling equations) we experimented with various set of variables. We find that, in our sample, the time of day of the interview does not affect much the probability that the youth is present (and therefore answers the expectations questions). However, the day of the week dummies and the indicator for holiday weeks are strongly significant determinants of who

is the respondent with F-statistics of 32 and 18 for young and old cohort respectively.

4.3 Descriptive Statistics: Subjective Expectations

In Tables 3 and 4, we present summary statistics for the variables derived from the subjective expectations questions for the young and old cohort. The top two panels of the Table report expected log earnings and expected (gross) returns. The bottom two panels, instead, focus on perceived earnings and employment risk for three different schooling degrees, junior and senior high school and college (or the latter two in the case of the old cohort). We summarize these measures separately for male and female youths, and separately for mother and youth respondents. In columns 3 and 6 of these Tables, we test the hypothesis that the youth's and mother's expectations are the same. We use the estimates from the selection equation (4) we have just discussed to correct for the possibility that the observed means differ because youths who answer the questionnaire might be systematically different, in some non observable dimension, from youths whose mother answers the questionnaire.

Not surprisingly, but reassuringly, expected log earnings increase in schooling level. Gross returns to schooling –measured as the difference between expected log earnings of two consecutive schooling degrees– are large and larger for college than high school. Comparing male and female youths when youths respond themselves, males expect slightly higher earnings and lower returns consistent with a gender earnings gap that decreases in schooling level as observed in actual earnings data from Mexico.[11] Mothers on the other hand expect surprisingly similar earnings and returns for boys and girls.

Standard deviations of log earnings are one possible measure of (perceived) earnings risk and commonly used for measuring risk in the case of observed earnings. It is important to point out that having information on the individual earnings distributions (e.g. in the form of minimum, maximum and probability above the midpoint) allows us to derive a measure of "true" risk, as perceived by the individual, while using the variability of observed earnings data will confound risk with unobserved heterogeneity. Moreover, observed variability will not distinguish between predicted and unpredicted changes.

Perceived earnings risk, as measured by the standard deviation of logs, decreases with education and is lowest for the expectations of earnings conditional on having a college degree. At the same time the probability of work increases with education. Thus lower income risk could be another important motivation, in addition to higher expected earnings, for achieving higher schooling. Male and female youth respondents perceive a very similar level of earnings risk, while males perceive a slightly higher probability of working, where the difference decreases again in schooling level.[12]

The third and sixth column of Tables 3 and 4 present results of the test of the differences between youths' and mothers' expectations, for each of the subjective expectation measures, i.e.

[11] Attanasio and Kaufmann (2008) compare the data on subjective expectations we are using to earnings data from the Census, see Table 18 in Appendix B.

[12] In addition to second moments, one can consider higher moments of the distribution. Looking at the skewness, one finds that, on average, individual earnings distributions are left-skewed (i.e. the probability to have earnings above the midpoint is larger than 0.5) and increasingly left-skewed with increasing schooling level.

for expected earnings, expected returns, perceived probability of working, and earnings risk. Table 3 shows results for the young cohort. Mothers' expectations about future earnings of her children are higher than the expectations for the youth sample. For girls, mothers expect 63% higher junior high school earnings, 44% higher senior high school earnings and 19% higher college earnings, where the first two differences are significant. For boys, mothers' expectations of junior high school earnings are around 20% higher than those of the male youths, though the difference is not significant.

In terms of expected returns, mothers expect lower high school and college returns than youth (difference is not significant for college for boys).

Mothers perceive a significantly lower earnings risk than boys and girls with respect to junior high school earnings and in addition with respect to senior high school earnings for boys. The perceived probability of working is similar for mother and youth respondents, except for boys perceiving a higher probability of working with a senior high school.

According to Table 4 differences in expectations between mothers and youths are small for the old cohort, also after correcting for potential sample selection. At the same time we have seen that the composition of the two samples is slightly different (see Tables 1 and 2). Therefore we will also control for compositional differences in the next section. We will see that there are important differences in expectations also for the old cohort.

Our results suggest that there are genuine differences in mothers' and youths' expectations. Informational differences between mother and youth respondents are particularly important for the young cohort. For example, girls' expectations of earnings are consistently lower than mothers' expectations for all schooling scenarios and differences are large.

These results underline the importance of understanding whose expectations are relevant for school choices, the ones of the mothers or the ones of the youths themselves. We pursue this issue in Section 5, after presenting some results to support the validity of the expectation data that we use in the rest of this paper.

4.4 Data Validation: Expected Earnings and Individual Characteristics

An additional interesting exercise consists in relating the expected value of future earnings to a number of observable variables, such as those listed in Tables 1 and 2. Such an exercise serves two purposes. First, as a validation exercise, we can check whether expected earnings vary with observables in a way similar to actual earnings. We expect people to draw inferences about their own potential earnings from what they observe from others. Thus finding that expectations vary with observable characteristics in a way similar to observed earnings lends support to their validity. We refer the interested reader to Attanasio and Kaufmann (2008) for a more detailed discussion.

Second, using again the selection equation discussed above, we can test the hypothesis that, after controlling for these observable variables, whose averages are slightly different in the two samples of youth and mother respondents, expected earnings are still different. That is, we can test formally the hypothesis that the differences in expectations are not induced by the different composition of the two samples of different respondents, but reflect genuine differences in expectations. We report the estimates of the coefficients obtained relating expected earnings to these variables in Tables

5 and 6. The coefficient on the youth respondent dummy, in the first row, should be interpreted as the test of differences in expectations between mothers and children, conditional on observables and corrected for potential sample selection.

Table 5 presents results for the young cohort. For boys, having a father who is an employer significantly increases expected senior high school and college earnings. Expectations of junior and senior high school earnings are decreasing in the number of brothers a boy has, possibly because this decreases the likelihood that he can take over the (small-scale) business of the father. For girls on the other hand, earnings expectations with senior high school and college are significantly lower when they have a father who is an employer, while expected earnings are larger when the girl comes from a richer family. Maybe surprisingly, expected junior high school earnings are significantly lower for girls whose mother is slightly more educated (on 10%), while expected college earnings are higher when the father is more educated. For the young cohort a higher GPA is not associated with higher expected earnings in contrast to what we find for the older cohort.

For the older cohort Table 6 shows that a higher GPA is associated with significantly higher earnings, in particular for college as the highest degree. For boys expected earnings increase in the number of sisters, which is somewhat harder to interpret, and expected high school earnings are larger if the boy is from a richer family. Expectations about girls' earnings on the other hand are lower if the mother is single, separated or divorced and there is no father in the household and lower if the father is a family worker. Again girls expected earnings are lower if the mother has a junior high school degree instead of less education and higher if parents have a college degree (only significant for fathers).

Testing for differences between mother's expectations about her child's earnings and the youth's own expectations, we find that youths (in particular girls) expect much lower earnings than mothers, consistent with previous results. This is true also for the old cohort (though only significant for junior high school earnings of girls).

Tables 3 to 6 show that there is a considerable amount of heterogeneity in expectations. This still holds after controlling for the extensive set of individual and family background characteristics reflecting the importance of unobserved heterogeneity in cognitive and social skills and differences in information sets e.g. about skill prices.[13]

To conclude, the evidence of this section suggests that our sample of Mexican youths and their mothers understand the expectation questions and give meaningful answers. At the same time there are important differences in information sets between mothers and youths, as they differ in their expectations about earnings of one and the same person –the youth herself–, suggesting that conventional approaches using earnings realizations and strong assumptions on rationality and information sets could be problematic. In particular it is impossible with such approaches to address the question whose expectations matter for schooling choices. Data on parents' and youths' subjective expectations on the other hand allow to gain insights into the intra-household

[13] An alternative explanation is that the remaining "heterogeneity" reflects noise. But we show that subjective expectations are able to predict schooling choices even after controlling for an extensive set of individual and family background characteristics. This suggests that at least part of the heterogeneity captures factors unobserved to the researcher, such as skills and information about skill prices, which influence earnings expectations.

allocation process of human capital investments. Furthermore, we perform this analysis separately for boys and girls and for the high school and college decision to understand if and how these intra-household decision processes differ for children of different gender and age and also to compare the role of different determinants of schooling, such as expected earnings and perceived risk, for male and female youths.

5 Schooling Decisions and Returns to Schooling.

In this section, we report the results we obtain modeling the schooling decisions of the two cohorts we consider: youths who have completed junior high school and decide whether to enrol in senior high school, and youths who decide whether to enrol in college, having completed senior high school. For each of these two cohorts, we estimate the schooling decisions jointly for boys and girls. However, to study gender differences, we interact all regressors with gender dummies. In the tables, we present the estimation results in two separate columns.

Before focussing our attention on gender differences and more generally on the determinants of education choices, we start our analysis investigating who makes the decision to enrol in senior high school or college: the youth or the mother.

5.1 Mother vs Youth Decision.

Having shown that mothers' expectations are significantly different from those of their children, we can sensibly address the question discussed in Section 2.2 about whose expectations are relevant for education choices. With data on parents' and youths' expectations, one could model schooling choices as potentially depending on both variables and check which, empirically, affect the enrolment decision.

Unfortunately this simple strategy is not feasible because, while we do have information on both mothers' and youths' expectations, we do not have them for the same individuals. As we discussed above, we have two different samples, one with mother and one with youth respondents. This feature of the data poses two problems. First, we cannot consider both expectations in the same regression. Second, if we run separately two regressions, one with the youth expectations and one with the mother expectations on the two different samples, we have to take into account the fact that the two subsamples were not randomly selected (compare Section 4.3).

Dealing with the second problem is relatively straightforward. We can use the same sample selection model that we use to correct the test of the differences between mother and youth expectations. When estimating a regression on one of the two subsamples, we can use once again the exclusion restriction that the timing of the interview does not affect education choices while it determines significantly which of the two samples an individual observation belongs to, and correct our estimates using a Heckman-type selection equation (adapted to a non-linear context).

The first problem, instead, is more difficult to deal with and, effectively, implies that we can consider only two extreme alternatives: one where only the mother's expectations matter and one

where only the youths' expectations matter. A model where both expectations matter is, effectively, not identifiable with our data.

Suppose that only mothers' expectations matter for the education decision. In this case, the results from the sample where mothers expectations are available would provide a consistent estimate of the effect of these expectations on school enrolment. Instead, the results from the sample where only youth expectations are available would provide inconsistent estimates of the schooling choice equation. The coefficient on youth expectations might be zero or, to the extent that youth expectations are somewhat correlated with mother expectations, could attract a significant coefficient whose size would probably be smaller in absolute value than its' true value because of attenuation bias.[14] On the other extreme, if only the youth's expectations matter, we would obtain consistent estimates only from the sample with youth respondents and inconsistent estimates from the sample with mothers respondents.

5.1.1 College Attendance Choice

In Table 2, we have seen that 29% of male high school graduates enrol in college compared to 31% of girls. We model the college attendance decision using equation (2), which relates the probability of enrolment to a set of control variables and expected (gross) returns to college, perceived unemployment and earnings risk under the two different schooling scenarios.

In Table 7, we present estimates for boys and girls in separate columns, even though they are estimated together in one model where all regressors are interacted with gender dummies to test for gender differences in coefficients. We report our estimates separately for youth respondents (Columns 1 and 2) and mother respondents (Columns 3 and 4) taking into account the selected nature of the sample. In the following tables, we report coefficient estimates, while we present the complete set of tables with marginal effects in Appendix B, where regressors are evaluated at their median values.

In terms of individual characteristics, academic performance, as measured by the GPA, is an important determinant of the decision to attend senior high school for both boys and girls. Past academic performance is both a measure of the psychological costs or benefits of getting further education and also captures the likelihood of being able to complete senior high school (compare Stinebrickner and Stinebrickner (2009)). As expected, parental education is an important determinant of the probability of going to senior high school.

Interestingly, higher parental income appears to be more important for girls than boys in the decision to attend college. Cost variables, such as distance from university and the level of tuition fees in nearby universities are also important determinants of the decision, even though the level of tuition fees in a nearby university seems to affect only girls.

The key results of Table 7 are two: First, boys' own expectations about returns to college are a strongly significant predictor for their decision to enrol in college. The coefficient is more than twice as large as for girls, for whom the coefficient is not significantly different from zero. This

[14] A rigorous statement about attenuation bias is difficult to make in non-linear models such as the one we are using.

result is not driven by a larger standard error for the girls' coefficient, i.e. it is not the case that the impact of returns in girls choices is estimated less precisely.[15]

Second, mothers' expectations are significant predictors in girls' decisions to enrol in college (while girls' own expectations are insignificant). If the mother perceives a high likelihood of her daughter finding a job with a college degree, this significantly increases the probability of her daughter to attend college. For boys on the other hand, only their own expectations seem to matter. These results suggest that mothers take an important part in the decision of their daughter to attend college, but not in their son's decision.

Perceptions about earnings risk are never significant in the college enrolment decision, neither for youth respondents nor for mother respondents, while they are significant for mother respondents in the high school enrolment decision. As this suggests that risk perceptions of earnings are not important for the college decision, we do not include them in the regressions, though including them would not change the results.[16]

As we mentioned in Section 3.2, our data does not contain information on fathers' expectations. However, it is possible that fathers' expectations could be different from mothers and play a role in the schooling decisions. To investigate whether this is likely to be a problem, we let the coefficients on the expectations variables to be different for the subsample of households headed by single, separated, divorced or widowed women. For this subsample there is no need to have data on fathers' expectations. We report these coefficients (and the coefficients on the rest of the variables considered in Table 7) in Table 8.

In families without father present, boys' own expectations are strongly significant predictors of their college attendance decision (significant on 1%), while mothers' expectations are insignificant. For girls on the other hand, only mothers' expectations matter. Furthermore, for these types of families the coefficient on expected returns is significantly larger for boys than for girls (on 2%). Monetary returns seem significantly more important for boys than girls (at least in families without father).

As we will discuss in detail in Section 5.3, it is unlikely that these results are driven by people rationalizing their choices ex-post. Although we control for potential self-selection problems because of non-randomness of who responds to the expectation questions, it is worthwhile to point out that self-selection does not appear to be a problem in the case of the old cohort (we never reject that the correlation between the error terms of selection equation and main equation is zero).

To summarize, our evidence shows that in the context of college enrolment decisions, there are important gender differences in the determinants of school choices as well as in the intra-household decision process. In particular, results suggest that mothers are involved in the college enrolment decision of their daughters, but not in their sons' decision. Boys in households without father on the other hand seem to be the sole decision-maker with respect to their decision to attend college.

[15]The difference between the two coefficients is not significant on conventional levels for the whole sample, but the difference is significant on 2% for the subsample of households without father, where the mother is single, separated or divorced (see Table 8). Also the difference between the coefficient on expected returns for boys and girls is significant when we include proxies for marriage market returns (see Section 5.2).

[16]If it was just a matter of the risk perceptions of earnings being measured in a too noisy way, then we would expect them to also be insignificant in the high school choice.

5.1.2 High School Attendance Choice

We now turn our attention to the decision to enrol or not in high school for youths who have just finished junior high school. As shown in Table 1, around 78% of boys and 81% of girls graduating from junior high school decide to enrol in senior high school. In Table 9, we report our estimates separately for youth respondents (Columns 1 and 2) and mother respondents (Columns 3 and 4), taking into account the selected nature of the sample. Again we present estimates for boys and girls in separate columns, although they are estimated simultaneously.

Individual and family background variables have the expected signs: the probability of senior high school attendance is increasing in GPA and in parental education. A larger number of siblings decreases the likelihood to continue schooling, in particular the presence of brothers. Surprisingly, the negative effect is stronger on boys than girls.[17] Being from a family in the lowest per capita parental income category decreases the likelihood to enrol in high school (significant only for girls).

In terms of the role of subjective expectations, Table 9 shows that youths' own expectations are not significant in the decision of high school attendance. For girls, mothers' expectations are a predictor of girls' decision to enrol in high school. If the mother perceives a higher probability that her daughter will be employed with a junior high school degree, this lowers the probability of the daughter to enrol in senior high school. If the mother on the other hand perceives a high probability of employment with a college degree, this increases the likelihood of enrolment in senior high school, as one would expect. This result suggests that having the option to enrol in college with a high school degree can be important for the high school enrolment decision. For boys on the other hand, mothers' expectations are not significant.

To check whether the lack of information on fathers' expectations is a problem, again we allow the coefficients on the expectations to be different for the subsample of households without a father and report the results in Table 10. In families with a father present, the mother appears to be involved in both boys' and girls' decision to enrol in high school. For example, a higher perceived earnings risk with a junior high school degree increases the likelihood to enrol in senior high school (for boys), while a higher earnings risk with a senior high school degree decreases the likelihood (for girls). The youths' expectations on the other hand do not seem to matter. In families without father present, boys' own expectations are significant while mothers' are not. The likelihood to enrol in high school increases in boys' expected returns to college (significant on 5%), in the probability of working with a college degree and with a higher perceived variance of college earnings (as this increases the option value). For girls, mothers' expectations are once again significant (girls' own expectations are significant in one instance with a –surprisingly– negative coefficient on the probability of working with a college degree, which is significant on 10%).

Thus in the case of junior high school graduates, as for the older cohort, boys without father seem to be the main decision maker for their schooling decision. In families with father on the other hand, the expectations of the 15-year-old youths do not matter at all for the decision to enrol in high school.

[17] We controlled for the number of siblings also in the college enrolment decision, but these variables were never significant and thus do not appear in the tables.

5.2 Labour and Marriage Market Returns

In the last section, we found differences in the role of expected returns and perceived risks in schooling decisions of boys versus girls. In what follows, we want to explore further how determinants of schooling decisions differ for boys and girls, and consider another potentially very important determinant of schooling: returns in the marriage market.

As discussed in the introduction, there is surprisingly little evidence on the importance of this determinant of schooling, despite the fact that finding an appropriate partner is certainly one of the big challenges individuals face during their lifetime. Thus, one would expect that individuals assign a high value to being in an environment where it is more likely to find a high quality partner.

In this section, we provide suggestive evidence on the importance of marriage market returns in the decision to enrol in college and senior high school. As we discussed in the previous section, with our data, we can only estimate a model with one set of expectations (either the youths' or the mothers'). The evidence we have just discussed seems to indicate that, for boys, youth's expectations seem to matter (for college), while for girls mothers' expectations determine schooling decisions. For completeness, however, we report the full set of results we obtain introducing the returns in the marriage market as a determinant of education choice in both subsamples.

5.2.1 College Attendance Choice

Tables 11 and 12 report estimates for youth and mother respondents using the two different proxies for marriage market returns, that is the sex ratio conditional and not conditional on education, respectively. In Column 1 (and 2), we show results for boys (and girls) in the youth sample using the ratio of unmarried men to women (women to men) with at least a senior high school degree and in Columns 3 and 4 we show results for the mother sample (for boys and girls).

As discussed in Section 3.5, we lose almost 60% of the observations when constructing the proxy with locality level Census data, as the Census does not contain information for all localities of the youths in our sample. Given the small sample size (approximately 1200 overall, which means around 300 observations for each group of male and female youth respondents and for mother respondents responding for sons and daughters), we can only control for the most important variables to achieve convergence in our bivariate probit estimation.

Table 11 reports results based on the sex ratio conditional on education. The key results of this table are two: First, as discussed above, expectations about returns to college are a strongly significant predictor for boys' decision to enrol in college. Also in the case of this smaller sample, the coefficient on expected returns is again (significantly) larger for boys than girls and more than twice in magnitude. But now girls' expectations are marginally significant as well (at the 10% level), probably because of the more parsimonious specification we use on this smaller sample.[18]

Second, for girls in the youth sample, the coefficient on the ratio of unmarried women to men conditional on education is significant and has the expected sign: fewer available suitable partners, i.e. many women per man, in the locality of residence increase the likelihood of girls to attend

[18] It seems that expectations are more strongly significant in localities in which the Census was conducted.

college. For boys, the coefficient on the analogous ratio of unmarried men to women is negative and not significantly different from zero. The difference between boys and girls of the coefficient on the marriage market proxy is significant at the 5% level. For the mother sample on the other hand, the coefficient is positive for both girls and boys, but not significant.

Table 12 presents results for the proxy without conditioning on education. Again the coefficient for girls is positive and significant (at the 5% level), while the coefficient for boys is insignificant. The difference between boys and girls is significant at the 3% for the youth sample, while for the mother sample the coefficient is about three times larger for girls than boys but neither is significant.

The results in this section provide suggestive evidence that marriage market considerations might play an important role in the college enrolment decision of girls. At the same time there seem to be significant gender differences in the importance of this determinant.

5.2.2 High School Attendance Choice

Lastly, we want to investigate the importance of marriage market returns in the decision to enrol in senior high school. Table 13 shows the importance of labor and marriage market returns for the high school decision. As in the case of college, Column 1 (and 2) show the coefficient estimates for male (and female) youth respondents, Columns 3 and 4 the results for the mother sample.

The coefficient on the ratio of unmarried women to men with at least a junior high school degree is significant only for girls and significantly larger in magnitude than for boys (on 10% for the youth sample and on 3% for the mother sample). At the same time the other proxy –without conditioning on education– is not significant in the case of high school, neither for boys nor girls. To ensure that the significance of the proxy is not driven by a few outliers, we also ran regressions in which we dropped the top and bottom 1% of the variable "ratio of unmarried women to men with at least a junior high school degree". The results become even stronger (for the youth sample the coefficient on the proxy for girls is now significantly different from zero on 5%). The same holds true for the college enrolment decision.[19]

5.3 Robustness: Subjective Expectations and Endogeneity

In this section we discuss in detail concerns that are sometimes raised with respect to data on subjective expectations. In particular, we discuss potential problems of endogeneity due to omitted variables and due to reversed causality (ex-post rationalization).

Ex-Post Rationalization (or Reversed Causality)

The term "ex-post rationalization" describes the behavior of people who state beliefs to justify their choices (ex-post), that is the decision affects the beliefs instead of beliefs affecting the decision. For example, an individual might eat a whole cake and justify this decision by stating the belief that otherwise the cake would have gone bad. In the context of this paper, one might be worried that

[19]Results available from the authors upon request.

people decide to go to college for reasons other than expected monetary returns, and that they justify this decision by stating high expected returns.

We think that ex-post rationalization of college enrolment, which would imply that the main reason to attend college is its consumption value (or similar reason) rather than expected future returns, is unlikely for poor Mexican youths. One could think that such an explanation might ring true for rich or even middle class individuals, but not for individuals from extremely poor families for whom college enrolment is often extremely difficult from a financial point of view. However, to test the possibility that our respondents answer the expected return questions to justify ex-post their choices, as the survey on expectations was conducted two-three months after the schooling decision, we conduct the test that we discussed in Section 3.2. In particular, we compare the cross-section of expected earnings for our cohorts of interest (i.e. students who had just graduated from junior and senior high school before the survey) and cohorts that are one year younger (and thus just starting grade 9 or 12). Let us first concentrate on the college enrolment decision of senior high school graduates. If individuals rationalize their choices, we would expect the following pattern: Individuals who decided to enrol in college rationalize their choice by stating higher expected college earnings or lower high school earnings. Those individuals who decided not to enrol state lower college earnings (or higher high school earnings). This would lead to a cross-sectional distribution of expected earnings and returns that is more spread out for the cohort of senior high school graduates who have decided already, compared to the distribution of the cohort that is one year younger and just starting grade 12. The same reasoning holds for perceptions of the probability of working.

To test for differences between the distribution of expectations of those two adjacent cohorts, we use a Kolmogorov-Smirnov test, which is a nonparametric test for the equality of continuous, one-dimensional probability distributions. The null hypothesis is that the cross-sectional distribution of -for example- expected returns is the same for the sample of junior/senior high school graduates and the sample of a cohort that is one year younger and just starting grade 9/12.

In Tables 14 and 15 we report p-values of Kolmogorov-Smirnov tests for senior high school graduates and junior high school graduates, respectively. We conduct the test for youth and mother respondents separately and separately by gender for the different measures of subjective expectations, i.e. for expected earnings for different schooling scenarios, expected returns (which is the relevant variable we use in the main regression analysis instead of expected earnings) and perceived risks.

Table 14 for senior high school graduates illustrates that for youth respondents the distribution of expectations is the same for the senior high school graduates themselves compared to the cohort that is one year younger. We never reject equality of distributions apart from one instance, that is girls' perceptions about the probability of working with a senior high school degree. In that case the distribution of the older cohort is not more spread out, but some of the older girls seem to have updated their beliefs about probability of working upward compared to the one-year younger cohort (see Figure 1). At the same time, this is not driving any of our results because in any case girls' own expectations are never significant in the college attendance choice regression.

For mother respondents we reject equality of distributions of expected return for both boys and girls on 1%. Again the distribution is not more spread out, but some mothers seem to have updated their beliefs about expected returns upward (see Figure 2). In principle this could mean that some mothers rationalize their children's choice to go to college by stating higher expected returns to college. But at the same time, mothers' expectations about returns are never significant in our regressions of college attendance, neither for girls nor for boys.

Table 15 illustrates the results of Kolmogorov-Smirnov tests for the cohort of junior high school graduates. For youth respondents, we can reject equality of distributions in three instances: for boys in terms of expected returns to high school and college and for girls in terms of perceived probability of working with a senior high school degree. We only find one instance in which a measures of subjective expectations is significant in the main results where rationalization could potentially be a problem, that is expected college returns for boys when there is no father in the household (otherwise boys' expected returns are not significant, see main results Tables 9 to 10). Therefore we perform the Kolmogorov-Smirnov test again for this subgroup of boys for expected returns to high school and to college (see Figure 3). We cannot reject equality of distributions of junior high school graduates (230 observations) and the cohort that is one year younger (70 observations) with p-values of 0.374 for return to high school and 0.514 for return to college. Thus the result that expected returns to high school are significant for boys without father is not threatened by the concern of rationalization.

For mother respondents, equality of distributions can be rejected for expected high school returns for boys and expected college returns for girls (see Table 15). Again neither of the two measures of expected returns is significant in any of the regressions. Also the distributions of perceived probability of working with high school and college degree have shifted for both boys and girls to a very similar extent (not in terms of larger spread, but some mothers have revised their beliefs upward, see Figure 4). This cannot explain why mothers' beliefs about the probability of work with a college degree is significant in explaining high school choices of girls, but not for boys. We should also mention that if the results we have shown were driven by ex-post rationalization, one would have to explain why the relationship is significant for boys but not for girl. This would imply that boys engage in ex-post rationalization but not girl.

Classical Endogeneity

Another concern with using subjective expectations as determinants of schooling choices is another forms of endogeneity: the one that would arise from unobserved taste heterogeneity (or omitted variables). Individuals who have high expected returns and go to college more often might also be the ones who have high preferences for being in college or attach weight to other outcomes, which are not measured, to which college is conducive. While we have discussed above the concern that individuals go to college because of the consumption value of college, in the following we discuss the concern that individuals might go to college because they expect returns other than labor market returns. We have to ask ourselves what could be other outcomes that are (as or) more important than expected monetary returns and correlated with the latter. One possibility of course are returns

in the marriage market.

Can this phenomenon explain our main results? One result is that expected returns to college significantly increase the likelihood to go to college for boys, but not for girls. To explain a positive coefficient on expected labor market returns for boys with an "omitted variable bias" story, one would have to argue that there is a positive correlation between marriage market and labor market returns for boys and boys care mostly about returns in the marriage market instead of labor market returns. This seems hard to believe and runs contrary to the vast majority of papers on educational choices. Also including proxies for returns in the marriage market does not decrease the coefficient on expected labor market returns.

Even if one believes that the coefficient on labor market returns could pick up effects of marriage market returns for boys, we are less interested in the magnitude of the coefficient and more in the differences between boys and girls and the differences between (male/female) youth respondents and mother respondents. Thus, if it is true that labor and marriage market returns are positively correlated and therefore the coefficient on monetary returns picks up marriage market effects, then it is hard to explain, why the coefficient on expected monetary returns is positive for boys but zero for girls. One explanation could be that girls neither care about their marriage market outcomes nor about their labor market outcomes, but again it seems hard to believe that girls (in particular from poor families) would decide to go to college based on something entirely different than labor or marriage market returns.

Can the differential effect between mother and youth respondents be explained by endogeneity? In particular, we find a positive and significant effect of boys' expectations on their likelihood to go to college, but no effect of mothers' expectations for their sons. Also we find positive coefficients on mothers' expectations for girls' decision to go to college, but no effects for girls' own expectations. It seems hard to come up with a story about omitted variables that can explain these results.

To conclude, in this section we provided evidence and discussed in detail why we believe that endogeneity concerns do not affect our two main conclusions and cannot be driving these results: there are significant differences in determinants between boys and girls and differential treatment of boys and girls.

6 Conclusion

The results of this paper speak to several important questions about the determinants of investment in human capital among poor households in Mexico. In particular, we have analyzed how expectations of future returns affect schooling decisions of two cohorts of Mexican children, one who has just completed junior high school and one that has just completed senior high school. In doing so we consider not only the expected monetary returns (as expressed by future earnings under different schooling alternatives) but also the risks involved with such choices and the returns on the marriage market.

Our main result is that the schooling decision is quite different for boys and girls in many dimensions. Our results indicate that boys care more about monetary returns to schooling than girls

–in particular in the decision to attend college. Marriage market considerations on the other hand seem important only for girls. Risk perceptions are particularly important from the perspective of the parents. In terms of the intra-household decision process, mothers are always involved in the schooling decisions of their daughters, but not of their sons. Even in households where the father is absent, boys seem to be the sole decision-maker with respect to both high school and college attendance choice.

We provide evidence that not only expected (monetary) returns matter for educational decisions, but also risk perceptions and marriage market considerations. This result is important as these considerations have been neglected in the literature, partly for the lack of appropriate data. We show how data on subjective expectations allow us to address these issues directly, as well as help us to shed light on the intra-household decision-making process. Our data on subjective expectations show that there are significant and important differences in the expectations held by youths and their parents on the return to education for the same youth. These differences can be due either to different information sets or to differences in the way information is processed.

Direct data on people's beliefs enable us to be agnostic about differences in information sets and ways of forming expectations between parents and youths and boys and girls. Avoiding strong assumptions about how people form expectations and what information they possess is crucial for our analysis, as results on gender differences in determinants of schooling and on intra-household decision processes might be very sensitive to these assumptions.

Unfortunately, our data does not contain direct questions on the returns to education on the marriage market. For this reason, we are forced to use some indirect proxies for such a return. In this sense, our evidence on marriage market considerations might be interpreted as only suggestive and subject to a number of caveats. However, such evidence, at the very least, warrants further research on this issue, which is potentially very important and which has not received much attention in the literature so far.

Our results have important policy implications for the design of programs aiming at increasing schooling, such as conditional cash transfer programs, fellowship programs, information campaigns etc. For example, an improved understanding of intra-household decision processes for human capital investments is crucial to determine who should receive the conditional cash or fellowship. Furthermore, for the effective design of such programs, it is indispensable to understand, whether there are differences in the determinants of schooling choices for boys and girls, and differences in the intra-household decision process depending on the gender of the child.

Our paper adds to the literature on subjective expectations in illustrating that –also in developing countries, at least conditional on a certain level of education– people seem able and willing to respond meaningfully to questions about their perceptions of future earnings and employment and that these data can improve our understanding of important economic decisions, such as investment into human capital.

References

Abramitzky, R., and V. Lavy (2011): "How Responsive is Investment in Schooling to Changes in Redistribution Policies and in Returns?," *mimeo*.

Arcidiacono, P., V. J. Hotz, and S. Kang (2011): "Modelling College Major Choices using Elicited Measures of Expectations and Counterfactuals," *Journal of Econometrics*.

Attanasio, O. P., and K. M. Kaufmann (2008): "Subjective Returns to Schooling and Risk Perceptions of Future Earnings - Elicitation and Validation of Subjective Distributions of Future Earnings," *mimeo*.

Attanasio, O. P., C. Meghir, and A. Santiago (2005): "Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate Progresa," *IFS Working Papers, EWP05/01*.

Berry, J. (2011): "Child Control in Education Decisions: An Evaluation of Targeted Incentives to Learn in India," *mimeo, MIT*.

Blundell, R., P.-A. Chiappori, and C. Meghir (2005): "Collective Labor Supply With Children," *Journal of Political Economy*, 113(6), 1277–1306.

Bourguignon, F. (1999): "The Cost of Children: May the Collective Approach to Household Behavior Help?," *Journal of Population Economics*, 12(4), 503–522.

Browning, M., and P.-A. Chiappori (1998): "Efficient Intra-household Allocations: A General Characterization and Empirical Tests," *Econometrica*, 66(6), 1241–1278.

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.

Card, D. (2001): "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems," *Econometrica*, 69 (5), 1127–60.

Currie, J., and E. Moretti (2003): "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings," *Quarterly Journal of Economics*, 118(4), 1495–1532.

Dauphin, A., A.-R. E. Lahga, B. Fortin, and G. Lacroix (2008): "Are Children Decision-Makers Within the Household?," *IZA Discussion Paper No. 3728*.

Dominitz, J., and C. F. Manski (1996): "Eliciting Student Expectations of the Returns to Schooling," *Journal of Human Resources*, 31 (1), 1–26.

Duflo, E. (2003): "Grandmothers and Granddaughters: Old-Age Pensions and Intrahousehold Allocation," *World Bank Economic Review*, 17, 1–25.

Edlund, L. (2005): "Sex and the City," *Scandinavian Journal of Economics*, 107(1), 25–44.

Fortin, B., and G. Lacroix (1997): "A Test of the Unitary and Collective Models of Household Labor Supply," *Economic Journal*, 107(443), 933–955.

Giustinelli, P. (2010): "Understanding Choice of High School Curriculum: Subjective Expectations and Child-Parent Interactions," *Mimeo*.

Goldin, C., L. Katz, and I. Kuziemko (2006): "The Homecoming of American College Women: The Reversal of the College Gender Gap," *Journal of Economic Perspectives*, 20(4), 133–156.

Guiso, L., T. Jappelli, and L. Pistaferri (2002): "An Empirical Analysis of Earnings and Employment Risk," *Journal of Business and Economic Statistics*, 20 (2), 241–253.

Heckman, J. J. (1979): "Sample Selection as a Specification Error," *Econometrica*, 47 (1), 153–161.

Jensen, R. (2010): "The (Perceived) Returns to Education and the Demand for Schooling," *Quarterly Journal of Economics*, 125(2), 515–548.

——— (2011): "Economic Opportunities and Gender Differences in Human Capital: Experimental Evidence for India," *mimeo*.

Kaufmann, K. M. (2009): "Understanding the Income Gradient in College Attendance in Mexico: The Role of Heterogeneity in Expected Returns to College," *mimeo*.

Keane, M. P., and K. I. Wolpin (1997): "Career Decisions of Young Men," *Journal of Political Economy*, 105 (June), 473–522.

Lafortune, J. (2010): "Making Yourself Attractive: Pre-Marital Investments and the Returns to Education in the Marriage Market," *mimeo*.

Lochner, L. (2011): "Non-Production Benefits of Education: Crime, Health, and Good Citizenship," *NBER Working Paper 16722*.

Manski, C. F. (2004): "Measuring Expectations," *Econometrica*, 72 (5), 1329–1376.

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

Stinebrickner, R., and T. R. Stinebrickner (2009): "Learning about Academic Ability and the College Drop-Out Decision," *mimeo*.

Thomas, D. (1990): "Intra-household resource allocation: An inferential approach," *Journal of Human Resources*, 25(4), 635–64.

Todd, P. E., and K. I. Wolpin (2006): "Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility," *American Economic Review*, 96(5), 1384–1417.

Wiswall, M., and B. Zafar (2011): "Determinants of College Major Choice: Identification using an Information Experiment," *mimeo*.

Zafar, B. (2009): "College Major Choice and the Gender Gap," *Federal Reserve Bank of New York Staff Reports*.

Table 1: Individual and Family-Background Characteristics: Junior High School Graduates

Junior HS Graduates:		**Total**			**Boys**			**Girls**	
Respondent:	**Boys**	**Girls**	**Diff** (P-Val)	**Youth**	**Mother**	**Diff** (P-Val)	**Youth**	**Mother**	**Diff** (P-Val)
	Mean/(SE)			Mean/(SE)			Mean/(SE)		
Enrollment Rate: Senior HS	0.793	0.835	(0.009)	0.809	0.771	(0.135)	0.841	0.826	(0.498)
	(0.405)	(0.371)		(0.393)	(0.420)		(0.366)	(0.379)	
GPA of Jr HS (Scale 0-100)	79.323	82.358	(0.000)	79.087	79.642	(0.270)	82.059	82.818	(0.099)
	(8.216)	(7.702)		(8.035)	(8.452)		(7.173)	(8.438)	
Obese	0.022	0.025	(0.072)	0.027	0.016	(0.264)	0.019	0.032	(0.181)
	(0.148)	(0.155)		(0.161)	(0.129)		(0.139)	(0.177)	
Mother's Educ - Primary	0.692	0.715	(0.281)	0.684	0.702	(0.561)	0.699	0.738	(0.182)
	(0.462)	(0.451)		(0.465)	(0.457)		(0.458)	(0.439)	
Mother's Educ - Jr HS	0.269	0.244	(0.217)	0.272	0.265	(0.814)	0.252	0.233	(0.503)
	(0.444)	(0.430)		(0.445)	(0.442)		(0.434)	(0.423)	
Mother's Educ - Sr HS	0.032	0.030	(0.798)	0.041	0.021	(0.089)	0.038	0.017	(0.047)
	(0.176)	(0.171)		(0.198)	(0.143)		(0.191)	(0.132)	
Mother's Educ - Univ	0.005	0.010	(0.291)	0.002	0.010	(0.132)	0.009	0.010	(0.975)
	(0.175)	(0.099)		(0.045)	(0.102)		(0.099)	(0.100)	
Father's Educ - Primary	0.686	0.695	(0.727)	0.686	0.685	(0.972)	0.695	0.693	(0.963)
	(0.464)	(0.461)		(0.464)	(0.465)		(0.461)	(0.461)	
Father's Educ - Jr HS	0.244	0.234	(0.648)	0.251	0.234	(0.609)	0.239	0.225	(0.664)
	(0.429)	(0.423)		(0.434)	(0.424)		(0.427)	(0.418)	
Father's Educ - Sr HS	0.062	0.057	(0.708)	0.054	0.073	(0.329)	0.056	0.059	(0.848)
	(0.242)	(0.233)		(0.227)	(0.261)		(0.231)	(0.237)	
Father's Educ - Univ	0.007	0.014	(0.230)	0.007	0.007	(0.943)	0.009	0.021	(0.211)
	(0.085)	(0.116)		(0.086)	(0.083)		(0.094)	(0.144)	
Per Cap Income - 5 to 10k	0.350	0.334	(0.421)	0.363	0.331	(0.272)	0.325	0.348	(0.406)
	(0.477)	(0.471)		(0.481)	(0.471)		(0.468)	(0.476)	
Per Cap Income - above 10k	0.222	0.25	(0.106)	0.219	0.226	(0.789)	0.267	0.224	(0.084)
	(0.416)	(0.433)		(0.414)	(0.418)		(0.442)	(0.417)	
Father's Occup - Unsk. Work	0.301	0.273	(0.227)	0.277	0.334	(0.086)	0.288	0.251	(0.237)
	(0.459)	(0.446)		(0.447)	(0.472)		(0.453)	(0.434)	
Father's Occup - Employee	0.489	0.520	(0.209)	0.542	0.416	(0.000)	0.508	0.539	(0.373)
	(0.500)	(0.499)		(0.498)	(0.493)		(0.500)	(0.499)	
Father's Occup - Employer	0.009	0.008	(0.890)	0.009	0.009	(0.978)	0.012	0.003	(0.121)
	(0.094)	(0.091)		(0.094)	(0.095)		(0.108)	(0.055)	
Father's Occup - Self-Empl.	0.184	0.181	(0.908)	0.161	0.215	(0.054)	0.182	0.182	(1.000)
	(0.387)	(0.385)		(0.367)	(0.412)		(0.386)	(0.386)	
Father's Occup - Fam. Work	0.016	0.015	(0.851)	0.011	0.024	(0.182)	0.009	0.024	(0.133)
	(0.128)	(0.123)		(0.105)	(0.154)		(0.099)	(0.154)	
No Father in Household (Mother Single/Sep/Div)	0.208	0.254	(0.008)	0.215	0.199	(0.514)	0.262	0.241	(0.396)
	(0.406)	(0.435)		(0.411)	(0.399)		(0.440)	(0.428)	
Number of Male Siblings	1.43	1.354	(0.116)	1.38	1.496	(0.103)	1.35	1.36	(0.884)
	(1.163)	(1.179)		(1.137)	(1.195)		(1.198)	(1.149)	
Number of Female Siblings	1.411	1.446	(0.477)	1.381	1.452	(0.334)	1.403	1.514	(0.118)
	(1.206)	(1.221)		(1.166)	(1.259)		(1.208)	(1.239)	
Ratio Unmarried Men/Women	1.073	1.069	(0.537)	1.072	1.073	(0.913)	1.064	1.076	(0.222)
	(0.175)	(0.159)		(0.145)	(0.207)		(0.121)	(0.204)	
Ratio Unmarried Men/Women (With Jun HS)	0.990	0.995	(0.638)	1.001	0.975	(0.109)	1.002	0.984	(0.222)
	(0.269)	(0.243)		(0.261)	(0.280)		(0.213)	(0.283)	
Locality Size - 15 to 50k	0.185	0.189	(0.817)	0.177	0.196	(0.411)	0.167	0.223	(0.017)
	(0.389)	(0.392)		(0.382)	(0.397)		(0.373)	(0.416)	
Locality Size - above 50k	0.318	0.339	(0.288)	0.349	0.276	(0.009)	0.375	0.283	(0.001)
	(0.466)	(0.473)		(0.477)	(0.447)		(0.484)	(0.451)	
Observations	1111	1253		638	473		759	494	

Table 2: Individual and Family-Background Characteristics: Senior High School Graduates

Senior HS Graduates:		**Total**			**Boys**			**Girls**	
Respondent:	**Boys**	**Girls**	**Diff**	**Youth**	**Mother**	**Diff**	**Youth**	**Mother**	**Diff**
	Mean/(SE)		(P-Val)	Mean/(SE)		(P-Val)	Mean/(SE)		(P-Val)
Enrollment Rate: Senior HS	0.289	0.316	(0.114)	0.243	0.325	(0.001)	0.237	0.401	(0.000)
	(0.453)	(0.645)		(0.429)	(0.468)		(0.426)	(0.490)	
GPA of Jr HS (Scale 0-100)	81.469	82.922	(0.000)	81.032	81.816	(0.063)	83.021	82.818	(0.686)
	(7.819)	(9.623)		(6.722)	(8.578)		(7.419)	(11.519)	
Obese	0.021	0.031	(0.096)	0.018	0.023	(0.595)	0.032	0.030	(0.819)
	(0.144)	(0.173)		(0.136)	(0.150)		(0.176)	(0.171)	
Mother's Educ - Primary	0.800	0.782	(0.286)	0.794	0.805	(0.674)	0.791	0.772	(0.455)
	(0.400)	(0.413)		(0.405)	(0.397)		(0.407)	(0.419)	
Mother's Educ - Jr HS	0.173	0.179	(0.724)	0.181	0.168	(0.590)	0.171	0.188	(0.442)
	(0.378)	(0.383)		(0.385)	(0.374)		(0.376)	(0.391)	
Mother's Educ - Sr HS	0.020	0.036	(0.022)	0.018	0.022	(0.651)	0.033	0.039	(0.594)
	(0.141)	(0.187)		(0.133)	(0.147)		(0.180)	(0.194)	
Mother's Educ - Univ	0.005	0.002	(0.24)	0.007	0.005	(0.729)	0.005	0.000	(0.083)
	(0.076)	(0.050)		(0.082)	(0.071)		(0.071)	(0.000)	
Father's Educ - Primary	0.760	0.754	(0.772)	0.747	0.769	(0.445)	0.763	0.746	(0.529)
	(0.427)	(0.43)		(0.435)	(0.421)		(0.425)	(0.436)	
Father's Educ - Jr HS	0.183	0.187	(0.826)	0.191	0.176	(0.574)	0.181	0.192	(0.674)
	(0.386)	(0.389)		(0.394)	(0.381)		(0.386)	(0.394)	
Father's Educ - Sr HS	0.045	0.047	(0.831)	0.053	0.039	(0.372)	0.051	0.043	(0.575)
	(0.207)	(0.212)		(0.224)	(0.195)		(0.220)	(0.204)	
Father's Educ - Univ	0.011	0.011	(0.955)	0.008	0.014	(0.433)	0.004	0.018	(0.033)
	(0.107)	(0.105)		(0.091)	(0.117)		(0.064)	(0.135)	
Per Cap Income - 5 to 10k	0.262	0.283	(0.216)	0.252	0.271	(0.429)	0.269	0.298	(0.223)
	(0.440)	(0.451)		(0.434)	(0.445)		(0.444)	(0.457)	
Per Cap Income - above 10k	0.184	0.183	(0.935)	0.187	0.181	(0.811)	0.179	0.187	(0.661)
	(0.388)	(0.386)		(0.390)	(0.385)		(0.383)	(0.390)	
Father's Occup - Unsk. Work	0.443	0.430	(0.559)	0.408	0.469	(0.057)	0.413	0.448	(0.236)
	(0.497)	(0.495)		(0.492)	(0.499)		(0.493)	(0.497)	
Father's Occup - Employee	0.278	0.284	(0.727)	0.312	0.252	(0.038)	0.312	0.256	(0.035)
	(0.448)	(0.451)		(0.464)	(0.434)		(0.464)	(0.437)	
Father's Occup - Employer	0.006	0.009	(0.349)	0.005	0.007	(0.634)	0.014	0.005	(0.150)
	(0.077)	(0.098)		(0.069)	(0.084)		(0.116)	(0.073)	
Father's Occup - Self-Empl.	0.234	0.241	(0.725)	0.229	0.238	(0.745)	0.226	0.256	(0.247)
	(0.423)	(0.427)		(0.421)	(0.426)		(0.419)	(0.437)	
Father's Occup - Fam. Work	0.038	0.034	(0.599)	0.045	0.034	(0.362)	0.034	0.034	(0.987)
	(0.192)	(0.182)		(0.208)	(0.181)		(0.182)	(0.182)	
No Father in Household (Mother Single/Sep/Div)	0.153	0.160	(0.592)	0.182	0.130	(0.011)	0.166	0.155	(0.553)
	(0.360)	(0.367)		(0.386)	(0.337)		(0.372)	(0.362)	
Number of Male Siblings	1.444	1.361	(0.07)	1.379	1.496	(0.095)	1.312	1.413	(0.096)
	(1.27)	(1.171)		(1.274)	(1.265)		(1.122)	(1.22)	
Number of Female Siblings	1.263	1.457	(0.000)	1.237	1.285	(0.457)	1.417	1.499	(0.228)
	(1.164)	(1.314)		(1.187)	(1.145)		(0.301)	(1.327)	
Ratio Unmarried Men/Women	1.111	1.117	(0.604)	1.113	1.108	(0.804)	1.121	1.113	(0.675)
	(0.222)	(0.256)		(0.232)	(0.214)		(0.270)	(0.234)	
Ratio Unmarried Men/Women (With Sen HS)	1.053	1.084	(0.372)	1.027	1.078	(0.286)	1.084	1.083	(0.994)
	(0.563)	(0.656)		(0.593)	(0.532)		(0.706)	(0.579)	
Locality Size - 15 to 50k	0.058	0.065	(0.465)	0.081	0.041	(0.003)	0.056	0.074	(0.174)
	(0.234)	(0.246)		(0.272)	(0.198)		(0.231)	(0.262)	
Locality Size - above 50k	0.09	0.084	(0.617)	0.103	0.08	(0.155)	0.105	0.063	(0.003)
	(0.286)	(0.278)		(0.304)	(0.271)		(0.307)	(0.243)	
Distance to Univ - 20 to 40km	0.243	0.263	(0.242)	0.25	0.239	(0.626)	0.267	0.258	(0.698)
	(0.429)	(0.44)		(0.433)	(0.426)		(0.443)	(0.438)	
Distance to Univ - above 40km	0.237	0.218	(0.244)	0.235	0.239	(0.871)	0.204	0.234	(0.166)
	(0.425)	(0.414)		(0.424)	(0.426)		(0.403)	(0.423)	
Tuition Costs - above 750 pesos	0.391	0.366	(0.169)	0.449	0.344	(0.001)	0.416	0.312	(0.000)
	(0.488)	(0.482)		(0.498)	(0.475)		(0.493)	(0.463)	
Observations	1320	1509		583	737		778	731	

Table 3: Subjective Expectations of Future Earnings: Junior High School Graduates

Junior HS Graduates		**Boys**			**Girls**	
Resp:	**Youth**	**Mother**	**Corr Diff (Y-M)**	**Youth**	**Mother**	**Corr Diff (Y-M)**
	Mean/(SE)		(P-Val)	Mean/(SE)		(P-Val)
Exp Log Earnings						
- Junior HS	7.059	7.232	-0.201	7.005	7.217	-0.632
	(0.608)	(0.489)	(0.214)	(0.589)	(0.451)	(0.000)
- Senior HS	7.614	7.764	-0.004	7.585	7.724	-0.441
	(0.536)	(0.425)	(0.976)	(0.550)	(0.421)	(0.005)
- College	8.243	8.361	-0.023	8.209	8.31	-0.185
	(0.502)	(0.416)	(0.865)	(0.521)	(0.438)	(0.208)
Exp Return						
- Senior HS	0.555	0.532	0.197	0.579	0.506	0.191
	(0.321)	(0.307)	(0.037)	(0.315)	(0.269)	(0.035)
- College	0.628	0.597	-0.019	0.625	0.586	0.256
	(0.368)	(0.304)	(0.850)	(0.355)	(0.326)	(0.017)
Std Dev of Log Earn						
- Junior HS	0.077	0.071	0.044	0.076	0.07	0.045
	(0.048)	(0.045)	(0.003)	(0.047)	(0.045)	(0.003)
- Senior HS	0.067	0.059	0.025	0.063	0.062	0.007
	(0.043)	(0.037)	(0.037)	(0.039)	(0.043)	(0.587)
- College	0.054	0.052	0.014	0.055	0.053	0.010
	(0.037)	(0.033)	(0.171)	(0.037)	(0.037)	(0.370)
Prob of Work						
- Junior HS	0.495	0.515	0.063	0.476	0.523	-0.081
	(0.210)	(0.218)	(0.313)	(0.214)	(0.209)	(0.203)
- Senior HS	0.677	0.674	0.109	0.661	0.675	0.016
	(0.173)	(0.171)	(0.035)	(0.192)	(0.169)	(0.765)
- College	0.826	0.824	0.066	0.826	0.812	0.030
	(0.166)	(0.148)	(0.158)	(0.164)	(0.157)	(0.530)
Observations	638	473		759	494	

Notes: The "Corrected Difference" between the expectations of mother and youth corrects for sample selection by instrumenting for who responds to the expectation questions. As instruments we use variables that capture the timing of the interview, which are strong predictors for who responds (see First-stage Tables).

Table 4: Subjective Expectations of Future Earnings: Senior High School Graduates

Senior HS Graduates:		**Boys**			**Girls**	
Resp:	**Youth**	**Mother**	**Corr Diff (Y-M)**	**Youth**	**Mother**	**Corr Diff (Y-M)**
	Mean/(SE)		(P-Val)	Mean/(SE)		(P-Val)
Exp Log Earnings						
- Senior HS	7.618	7.648	0.075	7.525	7.637	-0.123
	(0.497)	(0.444)	(0.758)	(0.507)	(0.418)	(0.401)
- College	8.267	8.317	0.153	8.236	8.308	0.010
	(0.646)	(0.440)	(0.526)	(0.485)	(0.426)	(0.943)
Exp Return						
- College	0.648	0.669	0.078	0.711	0.671	0.133
	(0.359)	(0.365)	(0.681)	(0.407)	(0.364)	(0.277)
Std Dev of Log Earn						
- Senior HS	0.065	0.059	-0.015	0.063	0.060	0.010
	(0.039)	(0.036)	(0.463)	(0.039)	(0.038)	(0.431)
- College	0.054	0.053	-0.012	0.054	0.050	0.007
	(0.033)	(0.033)	(0.486)	(0.035)	(0.030)	(0.498)
Prob of Work						
- Senior HS	0.661	0.653	0.011	0.664	0.644	0.218
	(0.173)	(0.180)	(0.909)	(0.188)	(0.178)	(0.001)
- College	0.821	0.813	-0.054	0.825	0.813	0.065
	(0.161)	(0.154)	(0.512)	(0.161)	(0.156)	(0.201)
Observations	583	737		778	731	

Notes: The "Corrected Difference" between the expectations of mother and youth corrects for sample selection by instrumenting for who responds to the expectation questions. As instruments we use variables that capture the timing of the interview, which are strong predictors for who responds (see First-stage Tables).

Table 5: Differences between Mothers' and Youths' Expectations: Junior High School Graduates

Dep Var:	**Expected Log Earnings**					
	Boys			Girls		
	Jr HS	**Sr HS**	**College**	**Jr HS**	**Sr HS**	**College**
	Coeff/(SE)			Coeff/(SE)		
Youth Respondent (Instrumented)	-0.263	-0.075	-0.221	-0.705***	-0.543***	-0.333*
	(0.190)	(0.166)	(0.162)	(0.214)	(0.195)	(0.182)
GPA of Junior HS (0-100)	0.000	0.000	-0.001	-0.000	0.000	0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Obese	-0.050	-0.051	-0.037	-0.141	-0.128	-0.098
	(0.110)	(0.095)	(0.093)	(0.107)	(0.097)	(0.091)
Number of Brothers	-0.032**	-0.021*	-0.020	0.007	0.007	-0.001
	(0.014)	(0.013)	(0.012)	(0.015)	(0.014)	(0.013)
No Father in HH (Mother Single/Sep/Div)	0.025	0.057	0.049	0.007	-0.000	0.030
	(0.047)	(0.041)	(0.040)	(0.049)	(0.044)	(0.041)
Mother's Educ - Jr HS	-0.025	-0.036	0.015	-0.075*	-0.049	-0.030
	(0.041)	(0.036)	(0.035)	(0.042)	(0.039)	(0.036)
Father's Educ - Jr HS	-0.066	0.003	-0.030	0.080	0.064	0.072*
	(0.048)	(0.042)	(0.041)	(0.050)	(0.046)	(0.043)
Per cap Income - 5 to 10k	-0.019	0.017	0.037	0.042	0.080**	0.092***
	(0.039)	(0.033)	(0.033)	(0.040)	(0.036)	(0.034)
Per cap Income - more than 10k	-0.014	-0.005	0.030	0.084*	0.073*	0.108***
	(0.046)	(0.040)	(0.040)	(0.048)	(0.044)	(0.041)
Father's Occup - Employer	0.032	0.329*	0.382**	-0.135	-0.343*	-0.471**
	(0.202)	(0.175)	(0.172)	(0.222)	(0.203)	(0.188)
Observations	1111	1111	1111	1253	1253	1253
R-Squared	0.148	0.162	0.106	.	.	0.042
Adjusted R-Squared	0.128	0.143	0.085	.	.	0.023
First-stage F-Stat	12.14	12.14	12.14	10.90	10.90	10.90

Notes: Table displays coefficients and standard errors in brackets. * p<0.1 ** p<0.05 *** p<0.01. Excl. categories: size of locality less than 15k, not obese, father in household, mother's and father's education primary or less, father's occupation worker, per capita income less than 5000 pesos. All specifications include state dummies. The regressor "Youth Respondent" is instrumented for by variables characterizing the timing of the interview (Dummy for Weekday, Dummy for Week 41 or 42 (Beginning of October), Dummy for Week 45 or 46), which are strongly significant predictors of whether the youth is at home when the interviewer arrives and thus responds herself to the expectation questions.

Table 6: Differences between Mothers' and Youths' Expectations: Senior High School Graduates

Dep Var:	**Expected Log Earnings**			
	Boys		Girls	
	Sr HS	**College**	**Sr HS**	**College**
	Coeff/(SE)		Coeff/(SE)	
Youth Respondent (Instrumented)	-0.171	-0.090	-0.259*	-0.134
	(0.198)	(0.196)	(0.155)	(0.154)
GPA of Junior HS (0-100)	0.001	0.004**	0.003***	0.005***
	(0.002)	(0.002)	(0.001)	(0.001)
Obese	0.115	0.085	0.050	0.026
	(0.087)	(0.086)	(0.067)	(0.066)
Number of Female Siblings	0.024**	0.022**	-0.005	0.001
	(0.011)	(0.011)	(0.010)	(0.009)
No Father in HH (Mother Single/Sep/Div)	0.021	0.028	-0.061*	-0.061*
	(0.045)	(0.044)	(0.035)	(0.035)
Mother's Educ - Jr HS	0.003	0.013	-0.104***	-0.080**
	(0.037)	(0.037)	(0.035)	(0.035)
Mother's Educ - Univ	0.017	0.166	0.114	0.251
	(0.187)	(0.184)	(0.270)	(0.269)
Father's Educ - Univ	-0.130	0.007	0.251*	0.188
	(0.150)	(0.148)	(0.147)	(0.147)
Per cap Income - 5 to 10k	0.064**	0.008	0.035	0.008
	(0.032)	(0.031)	(0.029)	(0.029)
Per cap Income - more than 10k	0.082**	0.039	0.044	0.013
	(0.037)	(0.036)	(0.035)	(0.035)
Father's Occup - Employee	-0.050	-0.004	-0.049	-0.078**
	(0.039)	(0.039)	(0.034)	(0.034)
Father's Occup - Fam. Worker	-0.085	-0.115	-0.119	-0.158**
	(0.077)	(0.076)	(0.074)	(0.074)
Observations	1320	1320	1509	1509
R-Squared	0.116	0.091	0.099	0.065
Adjusted R-Squared	0.098	0.074	0.084	0.049
First-stage F-Stat	7.16	7.16	11.80	11.80

Notes: Table displays coefficients and standard errors in brackets. * p<0.1 ** p<0.05 *** p<0.01. Excl. categories: size of locality less than 15k, not obese, father in household, mother's and father's education primary or less, father's occupation worker, per capita income less than 5000 pesos. All specifications include state dummies. The regressor "Youth Respondent" is instrumented for by variables characterizing the timing of the interview (Dummy for Weekday, Dummy for Week 40 to 42 (October), Dummy for Week 45 or 46), which are strongly significant predictors of whether the youth is at home when the interviewer arrives and thus responds herself or whether the mother answers to the expectation questions.

Table 7: College Attendance Choice: Youth and Mother Expectations

Dependent Variable	**College Attendance Decision**			
	Youth Respondent		Mother Respondent	
	Boys	Girls	Boys	Girls
Expected Return - College	0.379**	0.175	-0.004	0.218
	(0.166)	(0.132)	(0.142)	(0.141)
Prob of Work - Sr HS	-0.550	0.296	-0.526	0.080
	(0.441)	(0.357)	(0.365)	(0.349)
Prob of Work - College	0.153	0.095	0.473	0.990**
	(0.469)	(0.411)	(0.416)	(0.388)
Obese	-0.375	-0.723*	-0.322	-0.364
	(0.578)	(0.388)	(0.331)	(0.298)
Number of Siblings	-0.007	0.009	0.024	0.015
	(0.037)	(0.034)	(0.031)	(0.032)
GPA of Jr HS (0-100)	0.029***	0.026***	0.021***	0.008*
	(0.007)	(0.006)	(0.005)	(0.004)
No Father in Household	0.406**	-0.063	0.187	-0.026
	(0.187)	(0.160)	(0.173)	(0.148)
Mother's Educ - Jr HS	0.288*	-0.012	0.371**	0.370***
	(0.166)	(0.159)	(0.148)	(0.143)
Mother's Educ - Sr HS	0.409	0.551*	0.932**	0.675**
	(0.473)	(0.323)	(0.390)	(0.286)
Father's Educ - Jr HS	0.060	0.353**	0.238	-0.027
	(0.191)	(0.164)	(0.157)	(0.154)
Father's Educ - Sr HS	0.528*	0.451	-0.008	0.521*
	(0.319)	(0.280)	(0.314)	(0.306)
Per cap Income - below 5k	0.121	-0.174	-0.213*	-0.164
	(0.147)	(0.132)	(0.120)	(0.128)
Per cap Income - above 10k	0.188	0.450***	0.269*	0.575***
	(0.176)	(0.151)	(0.146)	(0.141)
Father's Occup - Employee	0.191	0.067	-0.176	0.244*
	(0.170)	(0.141)	(0.150)	(0.144)
Father's Occup - Fam Worker	0.096	-0.472	-0.843**	-0.216
	(0.350)	(0.410)	(0.416)	(0.322)
Dist to Univ 20 to 40km	-0.351**	-0.256*	0.026	0.140
	(0.156)	(0.131)	(0.132)	(0.122)
Dist to Univ above 40km	-0.165	-0.518***	0.153	0.141
	(0.175)	(0.159)	(0.151)	(0.148)
Tuition more than 750 Pesos	0.047	-0.716***	-0.062	-0.235
	(0.195)	(0.180)	(0.169)	(0.172)
Observations	2829		2829	
Censored Obs	1468		1361	
Log Likelihood	-2515.239		-2720.632	
Sample Sel: Corr of Errors (P-Val)	0.149 (0.652)		-0.091 (0.772)	

Notes: Table displays coefficients and standard errors in brackets. * p<0.1 ** p<0.05 *** p<0.01. Excl. categories: not obese, father in household, mother's and father's education primary or less, father's occupation worker, per capita income between 5000 and 10000 pesos, distance to university less than 20km, tuition costs less than 750 pesos. All specifications include state dummies.

Table 8: College Attendance Choice: No Father in Household (Mother Single/Sep/Div)

Dependent Variable	**College Attendance Decision**			
	Youth Respondent		Mother Respondent	
	Boys	Girls	Boys	Girls
Expected Return College * Father in Household	0.259	0.041	-0.04	0.035
	(0.194)	(0.155)	(0.157)	(0.156)
Prob of Work Sr HS * Father in Household	-0.915*	0.012	-0.634	0.071
	(0.539)	(0.398)	(0.396)	(0.402)
Prob of Work College * Father in Household	0.067	-0.200	0.501	0.260
	(0.496)	(0.388)	(0.388)	(0.393)
Expected Return College * No Father in Household	1.297***	0.127	-0.451	0.519
	(0.395)	(0.306)	(0.424)	(0.346)
Prob of Work Sr HS * No Father in Household	-0.003	0.874	-1.204	0.746
	(0.896)	(0.927)	(1.037)	(0.833)
Prob of Work College * No Father in Household	-0.405	-0.533	0.526	2.296**
	(1.050)	(1.046)	(1.257)	(1.051)
Obese	-0.361	-0.663*	-0.352	-0.371
	(0.586)	(0.390)	(0.332)	(0.299)
GPA of Jr HS (0-100)	0.028***	0.027***	0.016***	0.010***
	(0.007)	(0.006)	(0.005)	(0.004)
No Father in Household	-0.424	-0.410	0.787	-2.496***
	(0.861)	(0.684)	(0.816)	(0.849)
Mother's Educ - Jr HS	0.287*	-0.002	0.363**	0.352**
	(0.168)	(0.159)	(0.148)	(0.145)
Mother's Educ - Sr HS	0.342	0.553*	0.951**	0.691**
	(0.472)	(0.330)	(0.390)	(0.287)
Father's Educ - Jr HS	0.099	0.375**	0.242	-0.023
	(0.193)	(0.165)	(0.157)	(0.154)
Father's Educ - Sr HS	0.613*	0.489*	-0.006	0.475
	(0.322)	(0.283)	(0.314)	(0.305)
Per cap Income - below 5k	0.154	-0.17	-0.205*	-0.131
	(0.149)	(0.133)	(0.121)	(0.129)
Per cap Income - above 10k	0.193	0.440***	0.271*	0.562***
	(0.180)	(0.151)	(0.146)	(0.142)
Father's Occup - Fam Worker	0.123	-0.463	-0.832**	-0.166
	(0.352)	(0.415)	(0.414)	(0.318)
Dist to Univ 20 to 40km	-0.378**	-0.252*	0.031	0.128
	(0.159)	(0.131)	(0.132)	(0.122)
Dist to Univ above 40km	-0.154	-0.505***	0.171	0.147
	(0.177)	(0.161)	(0.151)	(0.148)
Tuition more than 750 Pesos	0.022	-0.695***	-0.071	-0.213
	(0.199)	(0.179)	(0.170)	(0.172)
Observations	2829		2829	
Censored Obs	1468		1361	
Log Likelihood	-2509.603		-2719.807	
Sample Sel: Corr of Errors (P-Val)	0.114 (0.740)		-0.155 (0.626)	

Notes: Table displays coefficients and standard errors in brackets. * p<0.1 ** p<0.05 *** p<0.01. Excl. categories: not obese, father in household, mother's and father's education primary or less, father's occupation worker, per capita income between 5000 and 10000 pesos, distance to university less than 20km, tuition costs less than 750 pesos. All specifications include state dummies.

Table 9: High School Attendance Choice: Youth and Mother Expectations

Dependent Variable	**High School Attendance Decision**			
	Youth Respondent		Mother Respondent	
	Boys	Girls	Boys	Girls
Expected Return - Sr HS	-0.060	0.146	0.239	-0.078
	(0.204)	(0.190)	(0.246)	(0.281)
Expected Return - College	0.219	0.232	-0.005	-0.224
	(0.169)	(0.175)	(0.240)	(0.240)
Prob of Work - Jr HS	0.098	0.046	0.530	-0.817*
	(0.378)	(0.331)	(0.444)	(0.475)
Prob of Work - Sr HS	-0.047	-0.184	-0.884	-0.086
	(0.502)	(0.442)	(0.639)	(0.626)
Prob of Work - College	0.652	-0.205	0.869	1.075**
	(0.405)	(0.458)	(0.557)	(0.543)
Var of Log Earn - Jr HS	4.626	10.010	12.784	-7.731
	(6.506)	(8.017)	(8.250)	(8.328)
Var of Log Earn - Sr HS	12.180	-10.609	-9.004	-16.763*
	(9.484)	(8.181)	(11.111)	(9.021)
Var of Log Earn - College HS	5.770	7.270	-10.691	7.430
	(11.140)	(10.815)	(12.668)	(10.417)
Locality Size 15 to 50k	0.227	-0.089	0.385*	0.143
	(0.178)	(0.163)	(0.199)	(0.193)
Locality Size above 50k	-0.091	-0.110	0.266	0.411*
	(0.143)	(0.143)	(0.190)	(0.227)
Number of Male Siblings	-0.166***	-0.085*	0.031	0.028
	(0.054)	(0.048)	(0.060)	(0.068)
Number of Female Siblings	-0.119**	-0.052	-0.106*	0.056
	(0.052)	(0.048)	(0.054)	(0.061)
GPA of Jr HS (0-100)	0.027***	0.034***	0.025***	0.031***
	(0.006)	(0.007)	(0.007)	(0.007)
No Father in Household	-0.105	0.403**	0.248	-0.231
	(0.166)	(0.172)	(0.200)	(0.220)
Mother's Educ - Jr HS	0.253	0.407**	0.361*	0.548**
	(0.169)	(0.173)	(0.200)	(0.229)
Father's Educ - Jr HS	0.079	0.153	0.443*	-0.016
	(0.192)	(0.180)	(0.231)	(0.234)
Per cap Income - below 5k	0.165	-0.052	-0.205	-0.513***
	(0.153)	(0.148)	(0.176)	(0.187)
Father's Occup - Self-Empl	0.246	0.413**	-0.212	-0.121
	(0.220)	(0.204)	(0.220)	(0.277)
Father's Occup - Fam Worker	5.344	-0.091	0.289	-0.976*
	(5101.242)	(0.706)	(0.497)	(0.507)
Observations	2364		2364	
Censored Obs	967		1397	
Log Likelihood	-1983.456		-1838.484	
Sample Sel: Corr of Errors (P-Val)	-0.507 (0.053)		0.301 (0.414)	

Notes: Table displays coefficients and standard errors in brackets. * p<0.1 ** p<0.05 *** p<0.01. Excl. categories: size of locality less than 15k, not obese, father in household, mother's and father's education primary or less, father's occupation worker, per capita income between 5000 and 10000 pesos. All specifications include state dummies.

Table 10: High School Attendance Choice: No Father in Household (Mother Single/Sep/Div)

Dependent Variable	**High School Attendance Decision**			
	Youth Respondent		Mother Respondent	
	Boys	Girls	Boys	Girls
Expected Return Sr HS * Father in Household	-0.339	0.102	0.057	-0.156
	(0.240)	(0.215)	(0.294)	(0.338)
Expected Return College * Father in Household	0.024	0.190	-0.177	-0.446
	(0.191)	(0.209)	(0.283)	(0.286)
Prob of Work Jr HS * Father in Household	0.479	0.095	1.099**	-1.333**
	(0.479)	(0.402)	(0.541)	(0.615)
Prob of Work Sr HS * Father in Household	-1.000	-0.004	-1.179	0.504
	(0.665)	(0.552)	(0.757)	(0.873)
Prob of Work College * Father in Household	0.711	-0.261	0.877	0.188
	(0.476)	(0.473)	(0.606)	(0.715)
Var of Log Earn Jr HS * Father in Household	3.864	11.634	18.602*	-3.915
	(7.981)	(9.871)	(11.030)	(10.667)
Var of Log Earn Sr HS * Father in Household	11.937	-10.361	-10.887	-21.457**
	(11.070)	(9.422)	(12.721)	(10.943)
Var of Log Earn College * Father in Household	-3.520	8.452	-25.096	-2.866
	(11.505)	(13.546)	(17.250)	(13.338)
Expected Return Sr HS * No Father in Household	0.485	-0.096	0.135	-0.353
	(0.494)	(0.458)	(0.639)	(0.603)
Expected Return College * No Father in Household	0.923**	0.266	0.321	-0.011
	(0.441)	(0.395)	(0.683)	(0.506)
Prob of Work Jr HS * No Father in Household	-0.814	-0.463	-1.233	-0.042
	(0.796)	(0.755)	(1.298)	(1.068)
Prob of Work Sr HS * No Father in Household	1.645*	1.316	-2.110	-2.496*
	(0.929)	(1.034)	(2.177)	(1.416)
Prob of Work College * No Father in Household	0.268	-2.307*	1.066	2.189*
	(0.909)	(1.342)	(1.855)	(1.165)
Var of Log Earn Jr HS * No Father in Household	-0.146	15.584	18.473	-20.670
	(21.395)	(21.282)	(20.069)	(19.946)
Var of Log Earn Sr HS * No Father in Household	-0.365	14.842	33.416	-3.288
	(27.502)	(31.459)	(46.968)	(31.298)
Var of Log Earn College * No Father in Household	91.004*	-3.960	27.282	83.934
	(48.877)	(24.718)	(40.221)	(61.462)
Number of Male Siblings	-0.156***	-0.099**	0.009	0.033
	(0.055)	(0.048)	(0.063)	(0.071)
Number of Female Siblings	-0.142***	-0.066	-0.117**	0.050
	(0.053)	(0.049)	(0.057)	(0.064)
GPA of Jr HS (0-100)	0.032***	0.029***	0.022***	0.034***
	(0.006)	(0.006)	(0.007)	(0.007)
No Father in Household	-2.086**	1.489	1.219	-0.799
	(0.848)	(1.040)	(1.364)	(1.022)
Mother's Educ - Jr HS	0.293*	0.438**	0.406*	0.585**
	(0.177)	(0.174)	(0.208)	(0.241)
Father's Educ - Jr HS	0.108	0.133	0.433*	-0.038
	(0.195)	(0.183)	(0.236)	(0.244)
Per cap Income - below 5k	0.150	-0.073	-0.231	-0.558***
	(0.157)	(0.150)	(0.180)	(0.193)
Father's Occup - Self-Empl	0.211	0.399*	-0.337	-0.062
	(0.231)	(0.208)	(0.227)	(0.291)
Father's Occup - Fam Worker	7.357	-0.078	0.28	-1.092**
	.	(0.719)	(0.514)	(0.502)
Observations	2364		2364	
Censored Obs	967		1397	
Log Likelihood	-1977.423		-1822.286	
Sample Sel: Corr of Errors (P-Val)	-0.456 (0.083)		0.099 (0.794)	

Notes: Table displays coefficients and standard errors in brackets. * p<0.1 ** p<0.05 *** p<0.01. Excl. categories: size of locality less than 15k, not obese, father in household, mother's and father's education primary or less, father's occupation worker, per capita income between 5000 and 10000 pesos. All specifications include state dummies.

Table 11: College Attendance Choice: Marriage Market Proxy Conditional on Education

Dependent Variable	**College Attendance Decision**			
	Youth Respondent		Mother Respondent	
	Boys	Girls	Boys	Girls
Expected Return - College	0.726***	0.298*	0.167	0.382
	(0.241)	(0.157)	(0.223)	(0.239)
Prob of Work - Sr HS	0.046	0.393	-0.387	-0.112
	(0.564)	(0.455)	(0.475)	(0.546)
Prob of Work - College	-0.519	0.349	0.696	0.959
	(0.588)	(0.544)	(0.586)	(0.692)
Ratio of Unmarried Men to Women (Sr HS)	-0.148		0.103	
	(0.135)		(0.133)	
Ratio of Unmarried Women to Men (Sr HS)		0.174*		0.061
		(0.099)		(0.130)
GPA of Jr HS (0-100)	0.043***	0.030***	0.014**	0.011*
	(0.011)	(0.010)	(0.007)	(0.006)
No Father in Household	0.429*	-0.239	0.268	0.124
	(0.235)	(0.169)	(0.194)	(0.189)
Mother's Educ - Jr HS	0.177	0.065	0.227	0.448*
	(0.225)	(0.193)	(0.208)	(0.261)
Mother's Educ - Sr HS	-4.288	0.874**	0.228	0.648
	(1254.662)	(0.357)	(0.535)	(0.402)
Per cap Income - above 10k	-0.061	0.436**	0.317	0.570***
	(0.223)	(0.175)	(0.197)	(0.209)
Observations	1237		1237	
Censored Obs	564		564	
Log Likelihood	-1174.544		-1171.207	
Sample Sel: Corr of Errors (P-Val)	-0.421 (0.367)		-0.507 (0.359)	

Notes: Table displays coefficients and standard errors in brackets. * p<0.1 ** p<0.05 *** p<0.01. Excl. categories: father in household, mother's education primary or less, per capita income between 5000 and 10000 pesos.

Table 12: College Attendance Choice: Marriage Market Proxy Not Conditional on Education

Dependent Variable	**College Attendance Decision**			
	Youth Respondent		Mother Respondent	
	Boys	Girls	Boys	Girls
Expected Return - College	0.699***	0.295*	0.172	0.379
	(0.241)	(0.154)	(0.226)	(0.241)
Prob of Work - Sr HS	0.069	0.376	-0.382	-0.134
	(0.554)	(0.446)	(0.480)	(0.551)
Prob of Work - College	-0.406	0.325	0.705	0.938
	(0.577)	(0.531)	(0.599)	(0.701)
Ratio of Unmarried Men to Women	-0.562		0.039	
	(0.449)		(0.338)	
Ratio of Unmarried Women to Men		0.595**		0.119
		(0.284)		(0.322)
GPA of Jr HS (0-100)	0.049***	0.027***	0.015**	0.010*
	(0.012)	(0.010)	(0.007)	(0.006)
No Father in Household	0.404*	-0.261	0.264	0.126
	(0.232)	(0.166)	(0.197)	(0.191)
Mother's Educ - Jr HS	0.169	0.133	0.224	0.456*
	(0.222)	(0.185)	(0.211)	(0.266)
Mother's Educ - Sr HS	-4.105	0.866**	0.253	0.640
	(1682.852)	(0.353)	(0.545)	(0.403)
Per cap Income - below 5k	0.350*	-0.285*	0.05	-0.374
	(0.210)	(0.170)	(0.190)	(0.297)
Per cap Income - above 10k	-0.045	0.431**	0.315	0.574***
	(0.220)	(0.174)	(0.199)	(0.210)
Observations	1237		1237	
Censored Obs	564		564	
Log Likelihood	-1174.980		-1173.440	
Sample Sel: Corr of Errors (P-Val)	-0.494 (0.263)		-0.475 (0.414)	

Notes: Table displays coefficients and standard errors in brackets. * p<0.1 ** p<0.05 *** p<0.01. Excl. categories: father in household, mother's education primary or less, per capita income between 5000 and 10000 pesos.

Table 13: High School Attendance Choice: Marriage Market Proxy Conditional on Education

Dependent Variable	**High School Attendance Decision**			
	Youth Respondent		Mother Respondent	
	Boys	Girls	Boys	Girls
Expected Return - Sr HS	-0.059	0.116	0.179	-0.127
	(0.204)	(0.189)	(0.249)	(0.281)
Expected Return - College	0.230	0.235	-0.044	-0.232
	(0.168)	(0.174)	(0.246)	(0.239)
Prob of Work - Jr HS	0.084	0.027	0.360	-0.779*
	(0.376)	(0.329)	(0.451)	(0.474)
Prob of Work - Sr HS	-0.011	-0.158	-0.918	-0.152
	(0.499)	(0.439)	(0.646)	(0.626)
Prob of Work - College	0.688*	-0.246	1.220**	0.965*
	(0.403)	(0.457)	(0.571)	(0.542)
Var of Log Earn - Jr HS	4.772	9.719	15.009**	-7.091
	(6.477)	(7.951)	(7.585)	(8.312)
Var of Log Earn - Sr HS	12.558	-10.321	-12.144	-17.524*
	(9.481)	(8.106)	(11.159)	(9.214)
Var of Log Earn - College HS	5.669	7.767	-7.106	5.751
	(11.103)	(10.857)	(13.080)	(10.251)
Ratio of Unmarried Men to Women (Jr HS)	0.011		-0.340***	
	(0.102)		(0.115)	
Ratio of Unmarried Women to Men (Jr HS)		0.464*		0.436*
		(0.256)		(0.256)
Locality Size 15 to 50k	0.225	-0.121	0.258	0.102
	(0.179)	(0.163)	(0.203)	(0.195)
Locality Size above 50k	-0.090	-0.183	0.103	0.315
	(0.145)	(0.148)	(0.191)	(0.226)
Number of Brothers	-0.161***	-0.081*	0.048	0.028
	(0.054)	(0.048)	(0.061)	(0.068)
Number of Sisters	-0.116**	-0.046	-0.112**	0.047
	(0.051)	(0.048)	(0.055)	(0.061)
GPA of Jr HS (0-100)	0.029***	0.031***	0.031***	0.029***
	(0.006)	(0.007)	(0.007)	(0.007)
No Father in Household	-0.097	0.409**	0.291	-0.223
	(0.166)	(0.171)	(0.206)	(0.219)
Mother's Educ - Jr HS	0.254	0.401**	0.300	0.540**
	(0.169)	(0.172)	(0.202)	(0.228)
Father's Educ - Jr HS	0.081	0.125	0.479**	-0.071
	(0.192)	(0.179)	(0.235)	(0.234)
Per cap Income - below 5k	0.171	-0.066	-0.268	-0.558***
	(0.152)	(0.148)	(0.180)	(0.190)
Father's Occup - Self-Empl	0.248	0.464**	-0.156	-0.025
	(0.219)	(0.206)	(0.218)	(0.280)
Father's Occup - Fam Worker	5.709	-0.085	0.249	-0.970*
	(2.0e+04)	(0.690)	(0.498)	(0.509)
Observations	2364		2364	
Censored Obs	967		1397	
Log Likelihood	-1978.935		-1826.721	
Sample Sel: Corr of Errors (P-Val)	-0.541 (0.040)		0.320 (0.374)	

Notes: Table displays coefficients and standard errors in brackets. * p<0.1 ** p<0.05 *** p<0.01. Excl. categories: size of locality less than 15k, not obese, father in household, mother's and father's education primary or less, father's occupation worker, per capita income between 5000 and 10000 pesos. All specifications include state dummies.

Table 14: College Attendance Choice: Rationalization of Choices

	Youth Resp		**Mother Resp**	
	Boys	**Girls**	**Boys**	**Girls**
	P-Val of KS-Test		P-Val of KS-Test	
Exp Log Earnings				
- Senior HS	0.557	0.650	0.939	0.002
- College	0.972	0.415	0.030	0.076
Exp Return				
- College	0.412	0.134	0.007	0.002
Prob of Work				
- Senior HS	0.604	0.034	0.884	0.662
- College	1.000	0.300	0.954	0.827
Observations				
(Sen HS Grads/Grade 12)	583/212	778/257	737/298	731/375

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

Figure 1: Comparing Expectations of Senior High School Graduates with a One-Year Younger Cohort: Youth Respondents

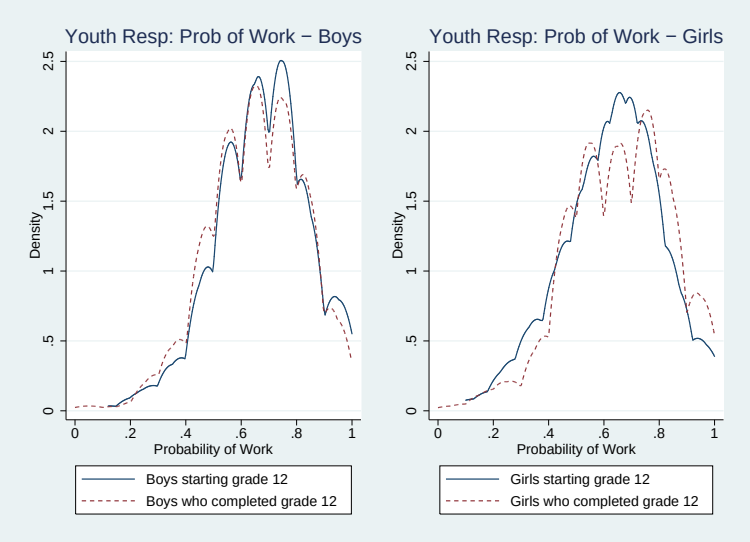


Figure 2: Comparing Expectations of Senior High School Graduates with a One-Year Younger Cohort: Mother Respondents

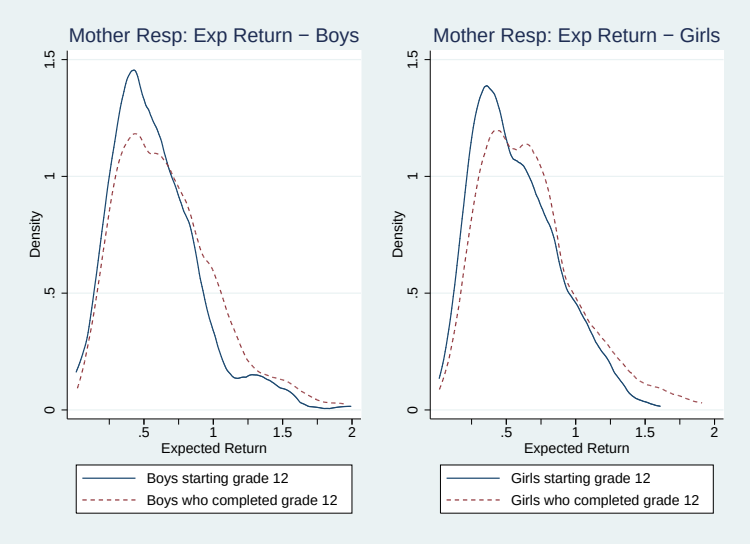


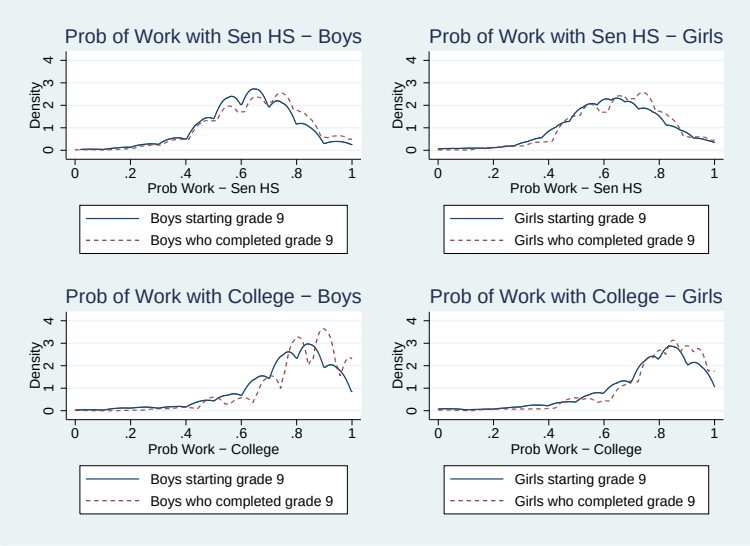
Table 15: High School Attendance Choice: Rationalization of Choices

	Youth Resp		**Mother Resp**	
	Boys	**Girls**	**Boys**	**Girls**
	P-Val of KS-Test		P-Val of KS-Test	
Junior HS Graduates				
Exp Log Earnings				
- Junior HS	0.767	0.693	0.000	0.000
- Senior HS	0.532	0.454	0.000	0.002
- College	0.314	0.117	0.000	0.071
Exp Return				
- Senior HS	0.089	0.358	0.089	0.598
- College	0.090	0.290	0.554	0.086
Prob of Work				
- Junior HS	0.608	0.605	0.134	0.616
- Senior HS	0.310	0.045	0.001	0.008
- College	0.978	0.299	0.000	0.089
Observations				
(Jun HS Grads/Grade 9)	683/119	759/141	473/267	494/197

Notes: Table displays the p-values of Kolmogorov-Smirnov tests of equality of distributions. The null hypothesis is that the cross-sectional distribution of -for example- expected returns is the same for the sample of junior high school graduates (whose schooling decision we are analyzing) and the sample of a cohort that is one year younger and just starting grade 9 (who have thus not decided yet about whether to enrol in senior high school or not).

Figure 3: Comparing Expectations of Junior High School Graduates with a One-Year Younger Cohort: Youth Respondents - Boys Without Father

Figure 4: Comparing Expectations of Junior High School Graduates with a One-Year Younger Cohort: Mother Respondents



Appendix B: Robustness Checks

First-Stage Results

Table 16: First-stage Regression for Whether the Youth Responds Herself: Young Cohort

Dep Var:	**Youth Respondent: Yes/No**		
	Instruments 1	Instruments 2	Instr and Controls
	Coeff/(SE)	Coeff/(SE)	Coeff/(SE)
Interview on a Weekday	-0.081***	-0.093**	-0.045
	(0.029)	(0.039)	(0.029)
Interview in Week 41 or 42	0.153***	0.161***	0.141***
	(0.021)	(0.029)	(0.021)
Interview in Week 45 or 46	-0.127***	-0.089**	-0.081***
	(0.032)	(0.043)	(0.031)
Female		0.040	-0.082
		(0.057)	(0.215)
Interview on a Weekday * Male		0.029	
		(0.058)	
Interview in Week 41 or 42 * Male		-0.017	
		(0.043)	
Interview in Week 45 or 46 * Male		-0.081	
		(0.063)	
Sex Ratio (Jr HS) * Male			-0.055**
			(0.023)
Sex Ratio (Jr HS) * Female			0.044
			(0.057)
Locality Size above 50k * Male			0.023
			(0.035)
Locality Size above 50k * Female			0.061*
			(0.033)
GPA of Jr HS (0-100) * Male			-0.003*
			(0.002)
GPA of Jr HS (0-100) * Female			-0.003
			(0.002)
Father's Occup - Employee * Male			0.064*
			(0.036)
Father's Occup - Employee * Female			-0.045
			(0.035)
Number of Siblings	No	No	Not Sign
Dummy for No Father in HH	No	No	Not Sign
Dummies for Mother's Educ	No	No	Not Sign
Dummies for Father's Educ	No	No	Not Sign
Dummies for Parental Income	No	No	Not Sign
State Dummies	No	No	Yes
Observations	2364	2364	2364
F-Statistik	32.452	14.595	8.441
R-Squared	0.040	0.042	0.144
Adjusted R-Squared	0.038	0.039	0.127

Notes: Table displays coefficients and standard errors in brackets. * p<0.1 ** p<0.05 *** p<0.01. Excl. categories: not obese, mother's and father's education primary or less, father's occupation worker, per capita income between 5000 and 10000 pesos. The last specification includes state dummies, which are also interacted with the gender dummy.

Table 17: First-stage Regression for Whether the Youth Responds Herself: Old Cohort

Dep Var:	**Youth Respondent: Yes/No**		
	Instruments 1 Coeff/(SE)	Instruments 2 Coeff/(SE)	Instr and Controls Coeff/(SE)
Interview on a Weekday	-0.069*** (0.026)	-0.087** (0.034)	-0.065** (0.026)
Interview in Week 40 to 42	0.115*** (0.023)	0.129*** (0.031)	0.122*** (0.023)
Interview in Week 45 or 46	-0.055** (0.022)	-0.072** (0.030)	-0.071*** (0.023)
Female		0.114** (0.053)	-0.258 (0.190)
Interview on a Weekday * Male		0.046 (0.051)	
Interview in Week 40 to 42 * Male		-0.031 (0.046)	
Interview in Week 45 or 46 * Male		0.039 (0.044)	
GPA of Jr HS (0-100) * Male			-0.005*** (0.002)
GPA of Jr HS (0-100) * Female			-0.000 (0.001)
Number of Siblings * Male			-0.010 (0.008)
Number of Siblings * Female			-0.019** (0.008)
No Father in Household * Male			0.136*** (0.042)
No Father in Household * Female			-0.001 (0.038)
Father's Occup - Employee * Male			0.103*** (0.037)
Father's Occup - Employee * Female			0.063* (0.034)
Dummies for Mother's Educ	No	No	Not Sign
Dummies for Father's Educ	No	No	Not Sign
Dummies for Parental Income	No	No	Not Sign
Dummies for Dist to College	No	No	Not Sign
Dummy for Tuition Costs	No	No	Not Sign
State Dummies	No	No	Yes
Observations	2829	2829	2829
F-Statistik	18.284	10.376	5.114
R-Squared	0.019	0.025	0.081
Adjusted R-Squared	0.018	0.023	0.065

Notes: Table displays coefficients and standard errors in brackets. * p<0.1 ** p<0.05 *** p<0.01. Excl. categories: not obese, mother's and father's education primary or less, father's occupation worker, per capita income less than 5000 pesos, distance to university less than 20km, tuition costs less than 750 pesos. The last specification includes state dummies, which are also interacted with the gender dummy.

Descriptive Statistics: Earnings Expectations and Census Earnings

Concerning the comparison of expectations to Census earnings, we are particularly interested in the pattern for the different respondents, that is for boys and girls and for youth and mother respondents. We use Census data of the year 2000 and compute average log earnings (by gender and by schooling degree) in the municipality of residence for each youth in the sample. While the expectations questions refer to earnings at age 25, we use earnings of individuals who are between 25 and 30 years old to get a sufficiently large sample size for each municipality.

While the comparison between the subjective expectations data and the Census data is certainly informative, if nothing else to check whether the subjective expectations data are roughly of the same order of magnitude, a direct comparison and a formal test of equality between the two would be misleading. There are many reasons why the Census data and the subjective expectations would be different. First, the former refer to a specific year (2000) and are therefore affected by specific aggregate shocks that might have been relevant in that year. To test some version of Rational expectations, one would need several years of realizations to average out aggregate shocks. Second, the data refer to individuals who were between 25 and 30 in 2000 and therefore belong to a different cohort from the individuals whose expectations were elicited in our survey (who are around 15 or 18 in 2005, and who were asked about their expected earnings when they are 25). Finally, the Census data report realizations for individuals who self-selected into a specific education level and do not contain "counterfactual" earnings, which are instead elicited in the expectations questions.

From these considerations, it should be clear that to establish whether the elicited expectations are 'rational' is probably impossible. And, in a sense, it is not too important: for modeling education choices what matters is whether the elicited expectations reflect the subjective expectations that people base their decisions on.

Table 18 presents mean expected earnings based on the distribution of earnings and Census earnings for different schooling degrees. Mean earnings are presented separately for boys and girls and mother and youth respondent. To correct for sample selection we present estimates that are based on the Heckman selection correction (again using the timing of the interview as an exclusion restriction). Estimates for the old (young) cohort are in the top (bottom) part of the table.

For the young cohort, all respondents expect lower junior high school earnings than observed in the Census. Youth respondents expect slightly lower senior high school earnings and similar college earnings. Mothers in general expect higher earnings and thus their expectations are closer to actually observed senior high school earnings but higher than currently observed college earnings. The old cohort expects similar high school earnings and higher college earnings. This pattern is consistent with a recent trend of a decrease of junior high school earnings in real terms, stagnating senior high school earnings and a rise in college earnings.

Girls responding themselves expect lower earnings than boys. The expected gender gap is similar to the one observed in Census earnings in 2000 with the exception of junior high school earnings, for which the actual gender gap is larger than expected. Also mothers expect their daughters to earn less than their sons (with one exception for the young cohort for senior high school earnings).

In terms of returns, girls' and boys' expectations are similar for returns to high school, while

girls expect higher returns to college. According to current Census earnings, returns are indeed larger for girls for both schooling degrees. While girls' return expectations are slightly higher than observed, boys' expectations are much larger than observed returns.

Table 18: Expected Earnings of Mothers and Youths and Realized Census Earnings (With Heckman Selection Correction)

	Boys		**Girls**	
Respondent:	**Youth**	**Mother**	**Youth**	**Mother**
	Mean	Mean	Mean	Mean
Old Cohort				
Exp Log Earnings				
- Senior HS	7.829	7.746	7.377	7.371
	(0.201)	(0.151)	(0.142)	(0.134)
- College	8.506	8.406	8.254	8.206
	(0.200)	(0.151)	(0.137)	(0.129)
Log Census Earnings				
- Senior HS	7.720	7.508	7.647	7.195
	(0.119)	(0.121)	(0.101)	(0.138)
- College	8.294	8.200	8.054	8.143
	(0.128)	(0.114)	(0.075)	(0.102)
Observations	507	605	585	542
Young Cohort				
Exp Log Earnings				
- Junior HS	6.982	7.089	6.539	7.133
	(0.108)	(0.121)	(0.138)	(0.136)
- Senior HS	7.580	7.522	7.242	7.641
	(0.095)	(0.108)	(0.122)	(0.127)
- College	8.132	8.200	7.898	7.917
	(0.090)	(0.101)	(0.117)	(0.146)
Log Census Earnings				
- Junior HS	7.580	7.328	6.973	7.178
	(0.058)	(0.085)	(0.066)	(0.106)
- Senior HS	7.907	7.682	7.566	7.324
	(0.039)	(0.061)	(0.046)	(0.096)
- College	8.179	7.961	8.153	7.799
	(0.051)	(0.083)	(0.040)	(0.088)
Observations	599	437	631	424

Notes: In this table we use Census data of the year 2000 and compute average log earnings (by gender and by schooling degree) in the municipality of residence for each youth in the sample. For this reason average Census earnings can differ between youth and mother sample, because the individuals in the two samples come from different municipalities. This table differs in the number of observations compared to the main tables, because for some municipalities in which junior and senior high school graduates live, we do not have Census data. Therefore we drop observations with missing information on actual earnings in this table (to compare expectations with actual earnings for the same municipalities), while we keep those observations for the main analysis.

Main Results: Tables With Marginal Effects

Table 19: College Attendance Choice: Youth and Mother Expectations - Marginal Effects

Dependent Variable	**College Attendance Decision**			
	Youth Respondent		Mother Respondent	
	Boys	Girls	Boys	Girls
Expected Return - College	0.091*	0.042	-0.001	0.063
	(0.047)	(0.032)	(0.041)	(0.042)
Prob of Work - Sr HS	-0.132	0.071	-0.151	0.023
	(0.115)	(0.087)	(0.115)	(0.100)
Prob of Work - College	0.037	0.023	0.136	0.285**
	(0.113)	(0.098)	(0.125)	(0.129)
Obese	-0.074	-0.115**	-0.080	-0.089
	(0.092)	(0.057)	(0.078)	(0.067)
GPA of Jr HS (0-100)	0.007***	0.006***	0.006***	0.002*
	(0.003)	(0.002)	(0.002)	(0.001)
No Father in Household	0.117*	-0.015	0.000	-0.007
	(0.061)	(0.037)	(0.063)	(0.042)
Mother's Educ - Jr HS	0.079	-0.003	0.121**	0.121**
	(0.055)	(0.038)	(0.058)	(0.053)
Mother's Educ - Sr HS	0.118	0.167	0.340**	0.238**
	(0.162)	(0.125)	(0.156)	(0.115)
Father's Educ - Jr HS	0.015	0.100*	0.075	-0.008
	(0.049)	(0.059)	(0.057)	(0.044)
Per cap Income - above 10k	0.050	0.132**	0.085	0.198***
	(0.053)	(0.059)	(0.052)	(0.058)
Father's Occup - Self-Empl	-0.064*	-0.026	-0.051	0.013
	(0.039)	(0.035)	(0.037)	(0.040)
Father's Occup - Fam Worker	0.024	-0.087	-0.160**	-0.057
	(0.092)	(0.064)	(0.077)	(0.078)
Dist to Univ 20 to 40km	-0.070**	-0.054*	0.008	0.043
	(0.034)	(0.032)	(0.039)	(0.039)
Dist to Univ above 40km	-0.036	-0.093**	0.047	0.043
	(0.036)	(0.039)	(0.050)	(0.046)
Tuition more than 750 Pesos	0.012	-0.115**	-0.018	-0.061
	(0.049)	(0.048)	(0.047)	(0.047)
Observations	2829		2829	
Censored Obs	1468		1361	
Log Likelihood	-2515.239		-2720.632	
Sample Sel: Corr of Errors (P-Val)	0.149 (0.652)		-0.091 (0.772)	

Notes: Table displays marginal effects (regressors are evaluated at their median values) and standard errors in brackets. * p<0.1 ** p<0.05 *** p<0.01. Excl. categories: not obese, father in household, mother's and father's education primary or less, father's occupation worker, per capita income between 5000 and 10000 pesos, distance to university less than 20km, tuition costs less than 750 pesos. All specifications include state dummies.

Table 20: College Attendance Choice: No Father in Household - Marginal Effects

Dependent Variable	**College Attendance Decision**			
	Youth Respondent		Mother Respondent	
	Boys	Girls	Boys	Girls
Expected Return College * Father in Household	0.08	0.013	-0.013	0.011
	(0.062)	(0.048)	(0.051)	(0.051)
Prob of Work Sr HS * Father in Household	-0.281	0.004	-0.207	0.023
	(0.180)	(0.122)	(0.139)	(0.131)
Prob of Work College * Father in Household	0.021	-0.061	0.164	0.085
	(0.152)	(0.122)	(0.133)	(0.127)
Expected Return College * No Father in Household	0.399***	0.039	-0.147	0.169
	(0.154)	(0.094)	(0.142)	(0.121)
Prob of Work Sr HS * No Father in Household	-0.001	0.269	-0.393	0.243
	(0.276)	(0.290)	(0.348)	(0.277)
Prob of Work College * No Father in Household	-0.125	-0.164	0.171	0.749**
	(0.324)	(0.323)	(0.411)	(0.377)
Obese	-0.096	-0.152*	-0.101	-0.106
	(0.132)	(0.079)	(0.089)	(0.078)
GPA of Jr HS (0-100)	0.009***	0.008***	0.005***	0.003**
	(0.003)	(0.003)	(0.001)	(0.001)
No Father in Household	-0.109	-0.106	0.298	-0.262**
	(0.183)	(0.151)	(0.325)	(0.126)
Mother's Educ - Jr HS	0.097	-0.001	0.130**	0.126**
	(0.063)	(0.049)	(0.059)	(0.054)
Mother's Educ - Sr HS	0.117	0.198	0.361**	0.260**
	(0.177)	(0.135)	(0.147)	(0.113)
Father's Educ - Jr HS	0.032	0.129*	0.084	-0.008
	(0.064)	(0.067)	(0.061)	(0.050)
Father's Educ - Sr HS	0.222*	0.173	-0.002	0.174
	(0.130)	(0.113)	(0.102)	(0.122)
Per cap Income - above 10k	0.063	0.154**	0.095*	0.208***
	(0.064)	(0.063)	(0.055)	(0.056)
Father's Occup - Fam Worker	0.04	-0.117	-0.192**	-0.051
	(0.117)	(0.090)	(0.088)	(0.092)
Dist to Univ 20 to 40km	-0.099**	-0.070*	0.01	0.043
	(0.045)	(0.040)	(0.044)	(0.043)
Dist to Univ above 40km	-0.045	-0.125***	0.059	0.050
	(0.049)	(0.048)	(0.055)	(0.051)
Tuition more than 750 Pesos	0.007	-0.157***	-0.023	-0.065
	(0.062)	(0.061)	(0.054)	(0.053)
Observations	2829		2829	
Censored Obs	1468		1361	
Log Likelihood	-2509.603		-2719.807	
Sample Sel: Corr of Errors (P-Val)	0.114 (0.740)		-0.155 (0.626)	

Notes: Table displays marginal effects (regressors are evaluated at their median values) and standard errors in brackets. * p<0.1 ** p<0.05 *** p<0.01. Excl. categories: not obese, father in household, mother's and father's education primary or less, father's occupation worker, per capita income between 5000 and 10000 pesos, distance to university less than 20km, tuition costs less than 750 pesos. All specifications include state dummies.

Table 21: High School Attendance Choice: Youth and Mother Expectations - Marginal Effects

Dependent Variable	**High School Attendance Decision**			
	Youth Respondent		Mother Respondent	
	Boys	Girls	Boys	Girls
Expected Return - Sr HS	-0.017	0.042	0.09	-0.029
	(0.059)	(0.057)	(0.094)	(0.104)
Expected Return - College	0.063	0.067	-0.002	-0.084
	(0.051)	(0.055)	(0.090)	(0.088)
Prob of Work - Jr HS	0.028	0.013	0.199	-0.306*
	(0.109)	(0.096)	(0.168)	(0.172)
Prob of Work - Sr HS	-0.014	-0.053	-0.332	-0.032
	(0.145)	(0.128)	(0.242)	(0.235)
Prob of Work - College	0.188	-0.059	0.326	0.403*
	(0.126)	(0.130)	(0.215)	(0.214)
Var of Log Earn - Jr HS	1.332	2.883	4.794	-2.899
	(1.908)	(2.428)	(3.116)	(3.146)
Var of Log Earn - Sr HS	3.508	-3.055	-3.376	-6.286*
	(2.855)	(2.456)	(4.170)	(3.360)
Var of Log Earn - College HS	1.662	2.094	-4.009	2.786
	(3.224)	(3.195)	(4.765)	(3.968)
Locality Size 15 to 50k	0.059	-0.027	0.132**	0.052
	(0.044)	(0.050)	(0.067)	(0.071)
Locality Size above 50k	-0.027	-0.033	0.094	0.140**
	(0.044)	(0.044)	(0.060)	(0.065)
Number of Male Siblings	-0.048***	-0.025*	0.012	0.011
	(0.018)	(0.013)	(0.023)	(0.026)
Number of Female Siblings	-0.034**	-0.015	-0.040*	0.021
	(0.016)	(0.013)	(0.021)	(0.024)
GPA of Jr HS (0-100)	0.008***	0.010***	0.009***	0.012***
	(0.003)	(0.003)	(0.003)	(0.003)
No Father in Household	-0.032	0.097**	0.088	-0.090
	(0.052)	(0.044)	(0.068)	(0.084)
Mother's Educ - Jr HS	0.065	0.097**	0.124*	0.178**
	(0.042)	(0.043)	(0.067)	(0.075)
Mother's Educ - Sr HS	0.085	0.210***	0.362*	-0.023
	(0.091)	(0.070)	(0.185)	(0.252)
Father's Educ - Jr HS	0.022	0.041	0.149**	-0.006
	(0.052)	(0.046)	(0.075)	(0.088)
Father's Educ - Sr HS	0.138*	0.095	0.244*	0.077
	(0.077)	(0.074)	(0.135)	(0.177)
Per cap Income - below 5k	0.044	-0.015	-0.079	-0.201***
	(0.039)	(0.044)	(0.068)	(0.072)
Father's Occup - Self-Empl	0.064	0.099**	-0.082	-0.046
	(0.053)	(0.048)	(0.084)	(0.106)
Father's Occup - Fam Worker	0.210***	-0.027	0.102	-0.371**
	(0.070)	(0.218)	(0.164)	(0.177)
Observations	2364		2364	
Censored Obs	967		1397	
Log Likelihood	-1983.456		-1838.484	
Sample Sel: Corr of Errors (P-Val)	-0.507 (0.053)		0.301 (0.414)	

Notes: Table displays marginal effects (regressors are evaluated at their median values) and standard errors in brackets. * p<0.1 ** p<0.05 *** p<0.01. Excl. categories: size of locality less than 15k, not obese, father in household, mother's and father's education primary or less, father's occupation worker, per capita income between 5000 and 10000 pesos. All specifications include state dummies.

Table 22: High School Attendance Choice: No Father in Household - Marginal Effects

Dependent Variable	**High School Attendance Decision**			
	Youth Respondent		Mother Respondent	
	Boys	Girls	Boys	Girls
Expected Return Sr HS * Father in Household	-0.085	0.026	0.014	-0.037
	(0.064)	(0.055)	(0.071)	(0.081)
Expected Return College * Father in Household	0.006	0.048	-0.042	-0.106
	(0.048)	(0.055)	(0.070)	(0.085)
Prob of Work Jr HS * Father in Household	0.12	0.024	0.262	-0.318
	(0.125)	(0.101)	(0.185)	(0.219)
Prob of Work Sr HS * Father in Household	-0.251	-0.001	-0.281	0.120
	(0.179)	(0.139)	(0.234)	(0.217)
Prob of Work College * Father in Household	0.179	-0.066	0.209	0.045
	(0.129)	(0.117)	(0.183)	(0.176)
Var of Log Earn Jr HS * Father in Household	0.971	2.923	4.441	-0.935
	(2.028)	(2.575)	(3.413)	(2.614)
Var of Log Earn Sr HS * Father in Household	2.999	-2.603	-2.599	-5.122
	(2.889)	(2.446)	(3.300)	(3.477)
Var of Log Earn College * Father in Household	-0.884	2.124	-5.991	-0.684
	(2.903)	(3.458)	(5.038)	(3.167)
Expected Return Sr HS * No Father in Household	0.122	-0.024	0.032	-0.084
	(0.127)	(0.115)	(0.154)	(0.149)
Expected Return College * No Father in Household	0.232*	0.067	0.077	-0.003
	(0.127)	(0.101)	(0.167)	(0.121)
Prob of Work Jr HS * No Father in Household	-0.205	-0.116	-0.294	-0.010
	(0.206)	(0.192)	(0.337)	(0.255)
Prob of Work Sr HS * No Father in Household	0.413	0.331	-0.504	-0.596
	(0.258)	(0.273)	(0.585)	(0.452)
Prob of Work College * No Father in Household	0.067	-0.58	0.254	0.523
	(0.229)	(0.365)	(0.456)	(0.383)
Var of Log Earn Jr HS * No Father in Household	-0.037	3.915	4.410	-4.935
	(5.375)	(5.443)	(5.290)	(5.372)
Var of Log Earn Sr HS * No Father in Household	-0.092	3.729	7.977	-0.785
	(6.909)	(7.944)	(11.780)	(7.492)
Var of Log Earn College * No Father in Household	22.865*	-0.995	6.513	20.038
	(13.522)	(6.208)	(10.008)	(17.606)
Number of Male Siblings	-0.039**	-0.025**	0.002	0.008
	(0.017)	(0.012)	(0.015)	(0.018)
Number of Female Siblings	-0.036**	-0.017	-0.028	0.012
	(0.016)	(0.012)	(0.020)	(0.018)
GPA of Jr HS (0-100)	0.008***	0.007***	0.005	0.008*
	(0.003)	(0.002)	(0.004)	(0.005)
No Father in Household	-0.701***	0.161**	0.143	-0.260
	(0.180)	(0.063)	(0.119)	(0.392)
Mother's Educ - Jr HS	0.063*	0.087**	0.078	0.100
	(0.037)	(0.039)	(0.057)	(0.071)
Mother's Educ - Sr HS	0.061	0.168***	0.155	-0.001
	(0.085)	(0.063)	(0.126)	(0.156)
Father's Educ - Sr HS	0.127**	0.074	0.119	0.043
	(0.059)	(0.064)	(0.095)	(0.107)
Per cap Income - below 5k	0.035	-0.019	-0.062	-0.169*
	(0.035)	(0.041)	(0.056)	(0.089)
Father's Occup - Self-Empl	0.048	0.081**	0.000	-0.015
	(0.048)	(0.041)	(0.073)	(0.073)
Father's Occup - Fam Worker	0.168***	-0.020	0.000	-0.376*
	(0.063)	(0.194)	(0.096)	(0.193)
Observations	2364		2364	
Censored Obs	967		1397	
Log Likelihood	-1977.423		-1822.286	
Sample Sel: Corr of Errors (P-Val)	-0.456 (0.083)		0.099 (0.794)	

Notes: Table displays marginal effects (regressors are evaluated at their median values) and standard errors in brackets. * p<0.1 ** p<0.05 *** p<0.01. Excl. categories: size of locality less than 15k, not obese, father in household, mother's and father's education primary or less, father's occupation worker, per capita income between 5000 and 10000 pesos. All specifications include state dummies.

Table 23: College Attendance Choice: Marriage Market Proxy Conditional on Education - Marginal Effects

Dependent Variable	**College Attendance Decision**			
	Youth Respondent		Mother Respondent	
	Boys	Girls	Boys	Girls
Expected Return - College	0.167**	0.068*	0.063	0.143*
	(0.081)	(0.038)	(0.083)	(0.083)
Prob of Work - Sr HS	0.011	0.09	-0.145	-0.042
	(0.129)	(0.102)	(0.177)	(0.204)
Prob of Work - College	-0.119	0.08	0.261	0.359
	(0.142)	(0.126)	(0.216)	(0.238)
Ratio of Unmarried Men to Women (Sr HS) (Loc)	-0.034	0.04	0.039	0.023
	(0.038)	(0.026)	(0.052)	(0.048)
GPA of Jr HS (0-100)	0.010***	0.007***	0.005**	0.004*
	(0.004)	(0.003)	(0.002)	(0.002)
No Father in Household	0.120*	-0.048	0.104	0.047
	(0.065)	(0.040)	(0.081)	(0.074)
Mother's Educ - Jr HS	0.044	0.015	0.088	0.176*
	(0.061)	(0.049)	(0.080)	(0.095)
Mother's Educ - Sr HS	-0.146	0.283**	0.088	0.254*
	(0.105)	(0.142)	(0.205)	(0.152)
Per cap Income - below 5k	0.091	-0.051	0.017	-0.124*
	(0.070)	(0.047)	(0.071)	(0.073)
Per cap Income - above 10k	-0.014	0.122**	0.123	0.224***
	(0.048)	(0.061)	(0.076)	(0.077)
Observations	1237		1237	
Censored Obs	564		673	
Log Likelihood	-1174.544		-1171.207	
Sample Sel: Corr of Errors (P-Val)	0.172 (0.585)		-0.507 (0.359)	

Notes: Table displays marginal effects (regressors evaluated at their median values) and standard errors in brackets. * p<0.1 ** p<0.05 *** p<0.01. Excl. categories: father in household, mother's education primary or less, per capita income between 5000 and 10000 pesos.

Table 24: College Attendance Choice: Marriage Market Proxy Not Conditional on Education - Marginal Effects

Dependent Variable	**College Attendance Decision**			
	Youth Respondent		Mother Respondent	
	Boys	Girls	Boys	Girls
Expected Return - College	0.168**	0.071*	0.064	0.141*
	(0.077)	(0.038)	(0.084)	(0.083)
Prob of Work - Sr HS	0.017	0.09	-0.142	-0.050
	(0.132)	(0.105)	(0.179)	(0.205)
Prob of Work - College	-0.098	0.078	0.263	0.349
	(0.142)	(0.128)	(0.222)	(0.241)
Sex Ratio of Unmarried Men to Women	-0.135		0.015	
	(0.120)		(0.126)	
Sex Ratio of Unmarried Women to Men		0.143*		0.044
		(0.075)		(0.118)
GPA of Jr HS (0-100)	0.012***	0.007***	0.006**	0.004
	(0.004)	(0.002)	(0.002)	(0.003)
No Father in Household	0.116*	-0.055	0.000	0.048
	(0.065)	(0.041)	(0.083)	(0.075)
Mother's Educ - Jr HS	0.044	0.034	0.000	0.179*
	(0.062)	(0.054)	(0.081)	(0.096)
Mother's Educ - Sr HS	-0.157	0.287**	0.000	0.251
	(0.107)	(0.138)	(0.209)	(0.153)
Per cap Income - below 5k	0.098	-0.059	0.019	-0.127*
	(0.071)	(0.047)	(0.072)	(0.072)
Per cap Income - above 10k	-0.010	0.125**	0.122	0.225***
	(0.051)	(0.060)	(0.077)	(0.077)
Observations	1237		1237	
Censored Obs	564		673	
Log Likelihood	-1174.980		-1173.440	
Sample Sel: Corr of Errors (P-Val)	-0.494 (0.263)		-0.475 (0.414)	

Notes: Table displays marginal effects (regressors evaluated at their median values) and standard errors in brackets. * p<0.1 ** p<0.05 *** p<0.01. Excl. categories: father in household, mother's education primary or less, per capita income between 5000 and 10000 pesos.

Table 25: High School Attendance Choice: Marriage Market Proxy Conditional on Education - Marginal Effects

Dependent Variable	**High School Attendance Decision**			
	Youth Respondent		Mother Respondent	
	Boys	Girls	Boys	Girls
Expected Return - Sr HS	-0.022	0.039	0.077	-0.056
	(0.065)	(0.062)	(0.094)	(0.104)
Expected Return - College	0.076	0.075	-0.010	-0.101
	(0.057)	(0.060)	(0.092)	(0.088)
Prob of Work - Jr HS	0.036	0.007	0.143	-0.315*
	(0.120)	(0.105)	(0.171)	(0.171)
Prob of Work - Sr HS	-0.006	-0.049	-0.319	-0.063
	(0.159)	(0.140)	(0.245)	(0.237)
Prob of Work - College	0.219	-0.082	0.435**	0.391*
	(0.138)	(0.145)	(0.220)	(0.210)
Var of Log Earn - Jr HS	1.581	3.117	5.554*	-2.539
	(2.113)	(2.644)	(2.867)	(3.118)
Var of Log Earn - Sr HS	3.994	-3.278	-4.924	-6.660**
	(3.140)	(2.663)	(4.247)	(3.351)
Var of Log Earn - College HS	1.656	2.410	-2.284	2.653
	(3.555)	(3.535)	(4.993)	(3.929)
Sex Ratio (Jr HS)	-0.001	0.142	-0.122***	0.122
	(0.033)	(0.093)	(0.041)	(0.095)
Locality Size above 50k	-0.031	-0.059	0.046	0.121*
	(0.049)	(0.052)	(0.070)	(0.071)
Number of Male Siblings	-0.050***	-0.027**	0.012	-0.007
	(0.019)	(0.014)	(0.023)	(0.025)
Number of Female Siblings	-0.036**	-0.016	-0.050**	0.005
	(0.017)	(0.014)	(0.021)	(0.023)
GPA of Jr HS (0-100)	0.009***	0.010***	0.011***	0.011***
	(0.003)	(0.003)	(0.003)	(0.003)
No Father in Household	-0.035	0.109**	0.1	-0.111
	(0.057)	(0.049)	(0.071)	(0.084)
Mother's Educ - Jr HS	0.074	0.109**	0.112	0.176**
	(0.047)	(0.048)	(0.069)	(0.072)
Mother's Educ - Sr HS	0.098	0.249***	0.401**	-0.097
	(0.102)	(0.083)	(0.179)	(0.251)
Father's Educ - Jr HS	0.026	0.038	0.149*	-0.023
	(0.058)	(0.053)	(0.076)	(0.089)
Father's Educ - Sr HS	0.156*	0.101	0.255*	0.084
	(0.091)	(0.087)	(0.136)	(0.175)
Father's Occup - Self-Empl	0.077	0.121**	-0.051	-0.014
	(0.059)	(0.055)	(0.084)	(0.107)
Father's Occup - Fam Worker	0.249***	-0.037	0.107	-0.309*
	(0.083)	(0.238)	(0.169)	(0.181)
Observations	2364		2364	
Censored Obs	967		1397	
Log Likelihood	-1979.491		-1833.455	
Sample Sel: Corr of Errors (P-Val)	-0.525 (0.048)		0.390 (0.263)	

Notes: Table displays marginal effects (regressors are evaluated at their median values) and standard errors in brackets. * p<0.1 ** p<0.05 *** p<0.01. Excl. categories: size of locality less than 15k, not obese, father in household, mother's and father's education primary or less, father's occupation worker, per capita income less than 5000 pesos. All specifications include state dummies.