

The Perceived Returns to Education and the Demand for Schooling

Robert Jensen^{*}

Watson Institute for International Studies

Brown University

Abstract: Economists have long emphasized the link between the market returns to education and investments in schooling. While many studies estimate these returns with earnings data, it is the returns *perceived* by individuals that affect schooling decisions, and these perceptions may be inaccurate, due to limited or imperfect information. Using data from the Dominican Republic, we find that while the measured returns to schooling are high, the returns perceived by students are extremely low. Students provided with information on the higher measured returns reported increased perceived returns several months later. The least-poor of these students were also significantly less likely to drop out of school in subsequent years. However, there was little or no effect on schooling for the poorest students. Finally, we find some support for the hypothesis that students underestimate the returns to education in part because they rely heavily on information on the returns within their own community, which are downwards biased due to residential segregation by income.

^{*} I would like to thank Christopher Avery, Pascaline Dupas, Eric Edmonds, Andrew Foster, Alaka Holla, Nolan Miller, Kaivan Munshi, Meredith Pearson, Richard Zeckhauser and Jonathan Zinman for useful comments and discussions, and Eric Driggs, Jason Fiumara, Zachary Jefferson, Magali Junowicz, Yesilernis Peña, Louisa Ramirez, Rosalina Gómez, Alexandra Schlegel and Paul Wassenich for valuable research assistance. Assistance and financial support from the Fundación Global Democracia y Desarrollo (FUNGLODE) and President Leonel Fernández is gratefully acknowledged.

I. INTRODUCTION

How important are the returns to education in determining schooling decisions? Do students have accurate information about these returns when they choose whether to continue schooling? Becker's canonical model of human capital views education as an investment, where costs are compared to the discounted stream of expected future benefits, primarily in the form of greater wages. However, while there is a large literature estimating the returns to schooling with earnings data, as pointed out by Manski (1993), it is the returns *perceived* by students and/or their parents that will influence actual schooling decisions. Given the great difficulties in estimating the returns encountered even by professional economists using large data sets and advanced econometric techniques, it seems likely that the typical student makes their schooling decision on the basis of limited or imperfect information.¹ In this setting, there is little reason to expect the level of education chosen to be either individually or socially efficient.

This possibility is particularly important to consider for developing countries, where educational attainment remains persistently low, despite high measured returns (see Psacharopoulos and Patrinos 2004). For example, in the Dominican Republic, while 80 to 90 percent of youths today complete (compulsory) primary schooling, only about 25 to 30 percent complete secondary school. Yet the mean earnings of workers who complete secondary school is over 40 percent greater than workers who only complete primary school.² There are of course many potential explanations for this "puzzle," such as poverty and credit constraints, high discount rates or simply mis-measured returns on the part of the researcher (i.e., education is low because the true returns are low,³ or the returns are low for the subset of students not enrolled). However, if misperceived returns and the resulting reduced demand for schooling are limiting factors, simply providing accurate information may be the most cost-effective solution,

¹ The possibility that decision-makers may not be well-informed has been examined for other areas of economic behavior. For example, many studies find that individuals underestimate the costs of borrowing (see Stango and Zinman 2007 for examples) or are poorly informed of their own pension or social security benefits (ex., Mitchell 1988, Gustman and Steinmeier 2005, Chan and Stevens, forthcoming). Viscusi (1990) finds that individuals overstate the risks of lung cancer from smoking, and that these misperceptions reduce smoking behavior.

² Assuming a discount rate of .05, the net present value of expected lifetime earnings, including forgone wages and the direct costs of schooling, is over 15 percent greater with secondary schooling. Further, more educated workers are likely to receive greater non-wage benefits.

³ Though, for example Duflo (2001) finds high returns to schooling in Indonesia even when using variation in schooling induced by a policy experiment to alleviate concerns about potential omitted variables bias.

especially relative to more widely used programs such as enrollment-contingent cash transfers or private school vouchers that, while effective, are typically very expensive.

A handful of studies for the United States have found that high school seniors and college students are relatively well-informed of the returns to a college education (Smith and Powell 1990, Dominitz and Manski 1996, Betts 1996, Avery and Kane 2004, and Rouse 2004). However, to date there has been no evidence for a low-income country. And there is reason to believe that students and/or their parents in these countries may not be as well-informed as their American counterparts. For example, the decision to drop out of school is often made at a much younger age, when students have less information about the returns. And schools typically do not have guidance counselors to provide information about the returns.⁴ Further, in general there may just be less information available at all on earnings, because data may not be collected as regularly or comprehensively by governments or private organizations, or because the results may not be as widely disseminated.⁵ As a result, often the only data on earnings available to youths may be the individuals they can observe around them,⁶ which could lead to inaccuracies. For example, youths in remote, rural communities or small towns where few adults have any education may have little information from which to infer the returns, including the potential returns in the urban sector. A variant of this argument, which asserts not only that students are poorly informed, but in particular underestimate the returns, arises in the literature on segregation and neighborhood effects (see Durlauf 2004 for a review). Most prominently, Wilson (1987) argues that youths in urban ghettos with high unemployment rates underinvest in schooling because they have little or no salient evidence of a connection between education and gainful employment.⁷ In many developing countries, where much of

⁴ Betts (1994) for example finds that over 60 percent of the college seniors surveyed report using their school's career services center to obtain information about job prospects by field of study.

⁵ For example, for this study we had to conduct our own labor force survey to estimate the returns to education because no data were available at the time, nor were there any available published studies of the returns.

⁶ For example, over 70 percent of secondary school students in our survey reported their main source of information about earnings was the people they knew in their community. By contrast, Betts (1994) reports that the most widely used source of information on employment prospects among college students was newspapers and magazines.

⁷ Broadly similar arguments were also made by Handlin (1959) and Glazer and Moynihan (1963).

the population lives in rural areas, small towns or urban slums with few highly educated workers, a similar segregation effect may be even more widespread.

Using data from a panel survey of students in the Dominican Republic in the 8th grade, the last year of compulsory schooling and the point at which most students terminate their education, we find perceptions of the returns to secondary schooling are extremely low, especially relative to those measured with earnings data. Beyond providing evidence of perceptions for a developing country, the primary contribution of this paper is to test whether providing additional information can affect enrollment.⁸ Thus, students at a randomly selected subset of schools were informed of the returns estimated from earnings data. Relative to students not provided with this information, these students reported dramatically increased perceived returns when re-interviewed 4 to 6 months later. They were also nearly 4 percentage points (7 percent) more likely to be enrolled in school the next academic term, and 4 years later had completed on average about .20 more years of schooling; however, they were not statistically significantly more likely to complete secondary school. Splitting the sample by income, we find that the program had little or no effect for households below the median per capita income, but large and statistically significant effects on the likelihood of returning to school, years of schooling and finishing secondary school for students from households above the median. Finally, we find some suggestive evidence consistent with a version of Wilson's (1987) hypotheses, namely that students rely heavily on the earnings of workers in their own community in forming their expectations of earnings, and these local returns underestimate the population returns due to residential segregation by income.

The remainder of this paper proceeds as follows: Section II discusses the data and experimental design, and section III presents the results. Section IV explores the link between residential segregation by income and the perceived returns to education, and section V concludes.

⁸ A handful of studies have found that information can influence behavior in other areas. Dupas (2006) finds that providing age-disaggregated information on HIV prevalence rates affects the incidence of risky sexual behavior among girls in Kenya. Duflo and Saez (2003) find that retirement plan decisions respond to being given incentives to attend a session providing benefits information, and Hastings, Van Weelden and Weinstein (2007) find that school choices respond to being provided information sheets on average test scores and admissions probabilities.

II. DATA AND METHODOLOGY

II.A. Data

In order to first estimate the returns to education, we conducted a household-based income survey in January, 2001.⁹ The survey was conducted nationwide, but only in non-rural areas (comprising about two-thirds of the population) because of the greater difficulty in estimating earnings for agricultural households. The household sample was drawn in two stages. First, from the 30 largest cities and towns, we chose 150 sampling clusters at random,¹⁰ with the number of clusters chosen in each town approximately proportional to that town's share of the combined population of the 30 cities/towns.¹¹ A listing of all dwellings in the cluster was then made, and 10 households were drawn at random from each cluster. The questionnaire gathered information on education, employment and earnings, and background demographic and socioeconomic characteristics for all adult household members.

For the student survey, for each of the 150 household sample clusters, we selected the school where students from that cluster attend 8th grade.¹² From each school, during April and May of 2001, we interviewed 15 randomly selected boys¹³ enrolled in 8th grade, the final year of primary school and therefore the point right before the very large declines in enrollment.¹⁴ The students were administered a survey gathering information on a variety of individual and household characteristics, as well as some simple questions on expected earnings by education (discussed below).

A second round survey of the youths from the student survey was conducted after the beginning of the next academic term (October, 2001), with respondents interviewed again (at home, school or work) about perceived returns to education and current enrolment status. In addition, at this time, parents were

⁹ At the time the study began, there was no publicly available micro-data on income available.

¹⁰ Cities and towns were divided into a set of clusters with the help of local experts, including community leaders and government officials.

¹¹ Though for greater geographic variation, we undersampled the capital, Santo Domingo. The city contains roughly 45 percent of the total population of the 30 cities/towns, but is only about 25 percent of our sample.

¹² In 6 cases, two clusters primarily used the same school; for these cases, we also chose the nearest alternate school.

¹³ We did not interview girls because of difficulties in eliciting expected earnings. Due to a low female labor force participation in the Dominican Republic (about 40 percent), in focus groups, most girls wouldn't estimate their expected earnings because they didn't think they would ever work.

¹⁴ Students were randomly selected from a list of currently enrolled students, and interviewed individually at the school. If a student was not present on the day of the interview, enumerators returned to the school the following day, and then contacted the student at home if they were still not available. 58 students were interviewed in their homes, primarily due to extended illness. Students were not compensated for their participation.

also interviewed to gather additional information on socioeconomic status. A third round, follow-up survey on schooling was also conducted in May-June of 2005, by which time students should have been finishing their last year of secondary school; for the approximately 120 students who were still enrolled in 2005 but were not yet through their final year of secondary school (due primarily to grade repetition), we conducted follow-ups for each of the next two years. For all follow-up surveys, if the respondent could not be found after two attempts, their parents, siblings or other relatives were interviewed about the youth's enrollment status. If these relatives also could not be located, neighbors were interviewed about the youth. Overall, we were able to obtain follow-up information in the October 2001 follow-up directly from 93 percent of youths, with another 2 percent from relatives and 5 percent from neighbors. By the 2005 survey, this had changed to 89 percent from youths, 4 percent from relatives, and 7 percent from neighbors. In all cases, we attempted to verify educational attainment by contacting the school students were reported to be attending or had attended. We were able to do so for 97 percent of students in the second round survey, and 91 percent in the third round.

II.B. The Measured and Perceived Returns to Schooling

Table 1 shows estimates of the differences in monthly earnings by education (estimates are in 2001 Dominican Pesos (RD\$), weighted to be representative of the 30 largest cities and towns).¹⁵ The data refer to the earnings of 30 to 40 year old men, since information regarding this age group is the basis for the experiment. The mean earnings (including both workers and non-workers)¹⁶ in column 1 are RD\$4,479 per month for those who completed secondary school (but not higher education) and RD\$3,180 for those who completed only primary school, yielding a difference of \$RD1,299. This difference represents an approximately 41 percent return to completing (just) secondary school over completing (just) primary school. Estimates of the returns are similar, and in fact slightly larger, when we apply a limited instrumental variables strategy to attempt to account for potential omitted variables bias

¹⁵ Our weighting scheme is based on year 2000 population estimates based on the 1993 census.

¹⁶ About 10-15 percent of both groups (with a slightly higher rate for the primary school group) reported they had no earnings in the past month. However, the gap in earnings by education is not substantially different if we focus only on employed workers. For the intervention, we chose to present the data not conditioned on employment.

and measurement error.¹⁷ While we cannot claim to have perfectly accounted for all econometric concerns, our best available estimates suggest that the returns to secondary schooling are high in the Dominican Republic.

By contrast, in pre-study focus groups, it was evident that few students perceived significant returns.¹⁸ Quantifying these perceptions is difficult, especially with young respondents.¹⁹ Therefore, the survey asked only some simple questions about perceived earnings, based on Dominitz and Manski (1996), though more limited. In particular, students were asked to estimate what they expected they themselves might earn under three alternative education scenarios:

“Suppose, hypothetically, you were to complete [this school year/ secondary school/ university], and then stop attending school. Think about the kinds of jobs you might be offered and that you might accept. How much do you think you will earn in a typical week, month or year when you are about 30 to 40 years old?”

Students were also asked to estimate the earnings of current 30 to 40 year old workers with different levels of education:

“Now, we would like you to think about adult men who are about 30 to 40 years old and who have completed only [primary school/ secondary school/ university]. Think not just about the ones you know personally, but all men like this throughout the country. How much do you think they earn in a typical week, month or year?”

While own-expected earnings are likely to be the relevant criteria for decision-making, this second set of questions was asked in order to arrive at estimates of earnings that are purged of any beliefs students may have about themselves, their household or their community, such as the quality of their school or self-

¹⁷ Using distance to primary and secondary school in childhood, proxies for the cost of schooling, as instruments for education as in Card (1995) and Kane and Rouse (1993) yields estimated returns to secondary school of \$RD1,433. The identifying assumption is that distance has no direct effect on adult earnings; this might be violated due to unobserved family or community characteristics. Like Card (1995), we find that the impact of distance on schooling is greatest for individuals predicted to have the lowest educational attainment, i.e., those from lower socioeconomic backgrounds. We therefore also estimated regressions including distance in the earnings regression and distance interacted with socioeconomic status in childhood as instruments. This strategy exploits the greater effect of distance on poorer children, with the slightly weaker (but still significant) identifying assumption that any effect of distance on earnings does not vary with family socioeconomic status. Using this approach, the returns to secondary schooling are \$RD1,552. The coefficient on distance in the earnings equation is small and not statistically significant. However, an important caveat is that the first stage predictive power is low, with F-statistics of about 5.

¹⁸ Though most students believed there were significant returns to completing primary school.

¹⁹ Students varied in age from 13 to 17, with variation due to late starts and grade repetitions.

perceived ability, or beliefs about factors such as race in determining earnings. They can thus be used to determine in part whether students' perceived earnings differ from those measured with earnings data because they have poor information, or because they have information or beliefs about themselves (correct or incorrect) that influence what they expect earnings will be for them personally.²⁰

These simple questions have several obvious and significant limitations. First, they are not precise in specifying the meaning of "expected" earnings, such as referring to the mean, median or mode.²¹ In addition, they do not elicit perceived uncertainty or the lifetime profile of earnings. Further, they ignore non-wage benefits such as health or pension plans, which are likely to be greater (though still not common) among secondary school educated workers. Finally, the questions deal with hypothetical situations, are stated in fairly formal language, and are slightly lengthy and complicated; as a result, about 10 percent of students did not provide responses to these questions, or responded "don't know." Given the ages of the students and their degree of math literacy, these various limitations could not be overcome. Thus, we do not view these as perfect measures of youths' decision-making criteria, nor will we rely on them for our primary analysis. We present these data simply as a way of quantifying as best as possible the impressions from the focus group discussions, in order to motivate the intervention.

Column 2 of table 1 shows that the average 8th grade boy reports that if they were to leave school at the end of the current year and not complete any more schooling, their (own) expected monthly wage would be RD\$3,516, which is greater than that actually measured in the household survey. There was considerable variation in the responses, with a standard deviation of RD\$846. Students on average expect monthly earnings of RD\$3,845 if they complete secondary school, which is much lower than that observed in the data (again, with considerable variation in responses). Thus, students overestimate the earnings of workers with primary schooling (by about \$RD330) and underestimate the earnings of

²⁰ For example, if students believe (perhaps correctly) they themselves will not gain from education because of labor market discrimination based on race or because they believe both education and employment opportunities are allocated by "connections," this should not be reflected in their perceptions of the earnings of other workers.

²¹ Though even if these more precise definitions could have been elicited, it is unclear which quantity students actually use in decision-making. The wording was intended to elicit as best as possible the level of earnings students expect or associate with different levels of schooling.

workers with secondary schooling (by about RD\$700). While they were not directly asked for the expected difference in earnings or the expected returns to schooling, the average implied perceived return is RD\$329 (9 percent), which is only one-quarter as large as the estimate from the earnings data. About 42 percent of students report no difference in own-expected earnings for the two levels of education, while 12 percent had implied returns that exceeded those measured in the data. Using these expectations, if we assume students expect to work until they are 65, and have a discount rate of .05, even if there were no direct costs of schooling, the implied net present value of the lifetime expected stream of earnings without secondary school is 11 percent *greater* than with secondary school. Thus, unless there are high non-wage returns, completing secondary school would only be worthwhile for students with these expectations if they were extremely patient (i.e., had a discount rate of .005 or less).²²

As stated above, any discrepancy between students' expected earnings and those measured among current workers could arise because students feel they have information about themselves that influences where they will fall in the earnings distribution, for example because they attend a poor quality school or because of some other important omitted variable. Thus, column 3 presents data on students' perceived wages of current adult workers aged 30-40. The means here are lower than own-expected earnings for both levels of education, consistent with a general optimism bias. About 55-60 percent of students reported the same mean for current workers as they expected for themselves for both levels of schooling, with about 25-30 percent expecting higher wages for themselves and 10-15 percent expecting lower wages. As with own-expected earnings, the implied perceived returns to a secondary school degree are much lower than what was measured in the earnings data (and lower than what they expect for themselves). This is not to say that students are unaware that some high school educated workers do enjoy a large return to their schooling; they just do not believe the effect on average is large. And again, the expectation questions can be scrutinized on a variety of fronts, so we do not take these estimates as the

²² We omit analysis of the additional effect of tertiary education. Less than 10 percent of adult males have a university degree; outside of the capital Santo Domingo, it is less than 5 percent. In our survey, only 13 percent of students reported planning to attend college, and by the final survey, only 6 percent had actually enrolled.

precise criteria students use in making schooling decisions.²³ However, these results are consistent with the focus group evidence and strongly suggest that despite the high measured returns, students do not believe there are large returns to secondary education. This raises the possibility that providing information may improve schooling.

II.C. The Intervention

Our study was motivated by a specific hypothesis regarding perceived returns to schooling, namely that students are poorly informed of the earnings of workers outside of their immediate community (this will be discussed in more detail in section IV).²⁴ In this context, providing them with information on the earnings of workers throughout the country provides a better information set. Therefore, at the end of the student survey, each respondent at a randomly selected subset of schools²⁵ was given information on the estimated earnings from the household survey, and the absolute and percent return implied by those estimates:

“Before we end, I would like to provide you with some information from our study. In January, we interviewed adults living in this community and all over the country. We asked them about many things, including their earnings and education. We found that the average earnings of a man 30 to 40 years old with only a primary school education was about 3,200 pesos per month. And the average income of a man the same age who completed secondary school, but did not attend university, was about 4,500 pesos per month. So the difference between workers with and without secondary school is about 1,300 pesos per month; workers who finish secondary school earn about 41 percent more than those who don’t. And people who go to university earn about 5,900 pesos per month, which is about 85 percent more than those who only finish primary school.”

²³ Though these perceptions do predict actual schooling. In regressions, an additional RD\$1,000 implied expected return increases the likelihood of returning to school the next year by about .11. However, these regressions may be plagued by omitted variables bias (ex., those with low perceived returns may come from lower quality schools) or reverse causality (ex., a “sour grapes effect” whereby those who don’t want to go to school or who want to but are constrained from doing so by poor grades or low income, report low returns, or students who want to continue schooling report high expected returns for self-justification). Instrumenting for perceived returns using the assignment of the intervention yields a larger (.15), and statistically significant, effect on schooling.

²⁴ As mentioned above, over 70 percent of students in our survey reported their primary source of information on earnings was the people they knew in their community.

²⁵ Assignment of the treatment was done at the school-level rather than for individual students within schools, since students within schools are likely to communicate, which would contaminate the control group. We cannot rule out that communication across schools occurred, though to the extent that such contamination took place, the true effect of the treatment would likely be even greater than what we estimate.

While the statement is again perhaps a bit lengthy, formal and complicated, the training of enumerators stressed that it was essential to emphasize the key elements of the statement, namely the earnings levels by education and the difference between them, by repeating them a second time after the statement was read, in order to make sure students understood the findings (students were then also invited to ask any questions about the data and results that they might have). It should also be noted that given the limitations in estimating the market returns to education from survey data and in eliciting students' perceptions of the returns, our intervention does not per se rely on estimating either correctly. The expected effect does depend on whether the estimates provided are above or below the returns perceived by students; but again, focus groups consistently revealed that most students believed there were little or no returns to schooling, so this was not a major concern for the study.²⁶

III. RESULTS

III.A. Basic Results

Table 2 provides data on key variables for the treatment and control groups. As expected given randomization, in the initial interview, there was little difference between the two groups in (own) expected earnings with or without a secondary school degree in round 1, and thus little difference in the implied expected returns. However, in the follow-up survey 4 to 6 months later, the treatment group reported on average greater expected earnings associated with secondary school completion, and lower expected earnings with only primary school. For the control group, there was an increase in expected

²⁶ Though we were concerned about providing misleading information, such as grossly overstating the returns, especially if they vary by race, region, or family background (such as political connections). However, the intervention was justified on the grounds of simply providing students the best available information, as well as informing them of the methodology and its limitations (as best as possible), and making it clear that the earnings data were national averages, not necessarily what they could expect for themselves: *"We also used statistical methods to try to account for the fact that different kinds of people get different amounts of education; the results were similar. However, no method is perfect, and people differ in many ways that affect their earnings, and statistics can't always capture those differences. And of course, there is no way to predict anyone's future, so our results don't signify that this is what you yourself will earn, these are only averages over the population."* While the returns may vary by, say, race, so the returns are not as great for some students in our sample, we would only believe the intervention was potentially harmful to those students if we believed their current level of schooling was efficient, which we find unlikely. We also view our intervention as consistent with the numerous efforts under way to increase education for all students, but especially for the most disadvantaged groups.

earnings for both levels of schooling, though more so for secondary.²⁷ Thus the treatment group experienced a large relative decrease (RD\$284) in expected earnings with only primary school and a smaller relative increase in expected earnings with secondary school (RD\$80). Based on a simple differences-in-differences calculation, the intervention on average differentially raised perceived returns by RD\$364; regressions in columns 1 and 2 of table 3 show that this estimate is statistically significant, and unchanged when controlling for other household characteristics. Overall, 54 percent of the treatment group had increased implied own-expected returns between the two rounds, compared to about 27 percent for the control group. However, there was heterogeneity in response to the treatment. About 28 percent of the treatment group had increased implied returns of RD\$1,000 or above, compared to 7 percent for the control group.

It is these large changes in the expected returns that we predict will affect schooling behavior. It is also worth noting that because the change in the expected return is driven to a great extent by a decline in expected earnings with only primary schooling, the intervention not only increased the expected future wage gap, but also lowered the opportunity cost of schooling, which is borne much sooner and thus not reduced as much through discounting. Thus we might expect a bigger effect than if the increase in implied expected returns was driven more by an increase in expected earnings with secondary schooling.

As stated earlier, since schooling is only compulsory through the 8th grade, the students in our sample were not required to return to school in the academic year following the first survey. The bottom three rows of table 2 provide data on subsequent school attainment; for now, we present data on reported schooling (by the student, their family or neighbors); we examine differences when using only verified schooling data below. The table shows that the treatment group was about 4 percentage points (7 percent) more likely to be attending school the following year, and 2 percentage points (7 percent) more likely to

²⁷ While there may just have been an overall general increase in expected earnings due to changes in labor market or macroeconomic factors or because students grew older between the rounds, sample selection is also likely to cause an increase in the mean implied expected return to schooling for both treatment and controls. Students who returned to school in round 2 (and thus who presumably had higher expected returns to schooling) were slightly more likely to be interviewed in that round than students who did not return, and thus we are more likely to have second-round data on expected earnings for these students.

have completed secondary school by 2005. And by 2005, the treatment group completed on average about .18 more years of schooling. Table 3 presents regression estimates of the effects of the treatment on these outcomes. The third column simply replicates the results in table 2; the coefficient is positive but not statistically significant at conventional levels (the p-value is .12). In column 2, the effect of the treatment is just slightly increased by the inclusion of other controls such as household socioeconomic status²⁸ and (teacher-assessed) 8th grade school performance (on a scale of 1 to 5 (much worse than average, worse than average, average, above average, much better than average)).²⁹ The coefficient is also estimated slightly more precisely in this specification, and as a result it is now statistically significant at the 10 percent level (p-value of .08). The other variables have the expected sign, with higher socioeconomic status (income and father's education) and better school performance associated with increases in the likelihood of continuing schooling, though the latter is not statistically significant.

Columns 7 and 8 focus on the effect of the program on whether students completed secondary schooling. Having received information about the returns to schooling increases the likelihood of completing secondary school by about 2 percentage points; however, the coefficient is not statistically significant, with or without the additional controls. Thus, despite providing information specifically on the returns to completing secondary school, the experiment had no overall effect on actual secondary school completion. However, columns 9 and 10 show that the intervention resulted in a statistically significant increase of about .20 years of completed schooling for the treatment group.

One reason students may not have completed the full four years of secondary schooling, and indeed a potential limiting factor for the effectiveness of the intervention, is that poverty and credit constraints may prevent some youths who want to continue schooling from doing so (i.e., demand alone is not sufficient). Table 4 presents separate regressions for youths in households above ("less poor") and below ("poor") the median household income. Cases where the student's family was not interviewed in

²⁸ Income was gathered from parents in the second survey round; data on income are missing for 139 observations, almost evenly split between treatment and control groups. We assign the median income to these observations. Dropping these observations from the regression instead does not change the results appreciably.

²⁹ Unfortunately, more objective data such as national test scores were not available. The data used here are useful for comparing students within schools, but comparing performance across schools is likely to be less informative.

round 2 lack income data and are excluded from this analysis (however, re-classifying households with missing data as either all poor or all less poor does not change the results appreciably). Focusing on the top panel, for the poorest households the effect of the treatment is small and not statistically significant for all three measures of schooling. This is despite the fact that in column 4, the treatment appears to have had a large effect on perceived returns to schooling for these students. By contrast, for youths from wealthier households, the effects are large and statistically significant at the 10 percent level or better for all three education measures. For this group, the intervention increased the likelihood of returning the next year by 7 percentage points (11 percent, from a base of 56 percent) and the likelihood of completing secondary school by 5 percentage points (13 percent, relative to a base of 40 percent). The net effect on total years of schooling completed was .34.³⁰ The fact that the increase in implied expected returns to schooling for this group is comparable to the increase for the poorer students suggests that providing information on the returns may have stimulated increased demand for schooling among both groups, but poverty and credit constraints limited investment in schooling even when there was demand. Overall though, for students above the median income, the effects are large and striking. The magnitudes compare favorably to the effects found for large-scale programs implemented elsewhere, such as Mexico's Progresa, which provided direct cash incentives to increase school attendance.³¹ However, many of these

³⁰ Unfortunately, we are unable to separate how much of the gain in secondary school completion or years of schooling is among students newly induced to enter secondary school by the treatment, versus an increased likelihood of continuation or completing secondary schooling among those who would have returned for 9th grade anyway, but would have dropped out before finishing schooling. Doing so would require being able to identify which specific students were induced to enter secondary school by the treatment, which is not possible.

³¹ Progresa, whose payments also were conditioned on other requirements, and also provided other benefits, increased enrolments for 9th grade boys from 60 to 66 percentage points (Schultz 2004), close to what was found here for wealthier students. For other comparisons, Duflo (2001) finds that a program in Indonesia that built approximately 61,000 primary schools (effectively doubling the stock), resulted in a .25-.40 increase in years of schooling, or, .12-.19 years (comparable to the results found here for the full sample) for each additional school built per 1,000 students. Angrist, Bettinger and Kremer (2006) find that a large voucher program in Colombia increased secondary school completion rates by 5 to 7 percentage points (a 15 to 20 percent gain), similar to what we find for the wealthier students. Of course, these results are not directly comparable; for example, Indonesia was in 1973 (and still is) a much poorer country than the Dominican Republic today, the Progresa program started from a much higher enrolment base, and both it and the Colombian voucher program targeted only the poorest students, so improvements in schooling may have been harder to achieve in these other cases.

other programs are extremely expensive,³² whereas in the present case, information could potentially be provided at little or no cost.

III.B. Interpretation and Robustness

One issue we consider is whether just by asking students to form their expectations of earnings for various levels of schooling, they acquire information or begin to think about the schooling decision in a way they would not have otherwise; alternatively, there may just be an effect of being interviewed by a research team as part of a project from an American university. Since both treatment and controls were administered the same survey except for whether they were provided with information on returns at the end, this does not affect our interpretation of the effect of the treatment.³³ However, one issue to consider is whether the control group was influenced by the interview. Therefore, in column 5 of table 3, we compare the full-sample control group to a “shadow” control group of 15 randomly selected students at each of 30 randomly selected non-sample schools.³⁴ These students were identified but not interviewed until the second round (unfortunately, they were not followed after this round). However, we only gathered data on the enrollment status for this group, so in the regression we only include an indicator for being in the control group that was interviewed (the lack of other data also prevents us from splitting the analysis by income as in table 4). The results show that the original “sample” or interviewed control group experienced no differential change in enrolment relative to the non-interviewed control group; the coefficient is positive, but small and not statistically significant. Thus, the provision of information on the returns to schooling appears to be the critical factor for achieving schooling gains.

³² For example, Progresa cost nearly .2% of Mexico’s GDP to provide benefits to about one-ninth of all Mexican households. Indonesia’s program cost about 1.5% of 1973 GDP, or about 750 million dollars in 2007. And the Colombian vouchers came at a cost of about \$190 per year of attendance (though for the government some of the cost would likely be offset by savings in expenditures for public schools). There are of course other interventions that have also been shown to be very cost-effective, such as the de-worming program studied by Miguel and Kremer (2004), which achieves gains at a cost of about \$3.50 per additional year of schooling.

³³ Unless we believe that the intervention would not have been effective without students first going through the interview, or without the presence of our research team.

³⁴ These schools were chosen to obtain approximately the same population distribution as the original student sample; for example, 6 of the 30 shadow schools (20 percent) were in Santo Domingo, compared to 35 of the 150 original schools (23 percent).

To this point, we have used data on reported education. The primary concern is that students may inflate the amount of education they achieved, especially if they received the treatment. A second concern is a general decline in accuracy when students or their relatives could not be interviewed (typically because the family had moved), and schooling data was obtained from neighbors. As stated, we attempted to verify schooling data for all students, but were unable to do so for 3 percent of students in the second round and 9 percent in the third round. Most of the cases where data could not be verified were due to obtaining information from neighbors or more distant relatives, since they often did not know which school the youth attended. Therefore, in the bottom panel of table 4, we consider only the “higher quality” observations where schooling could be directly verified. Before turning to these results, we make two observations. First, there were very few cases (27) where the youth reported enrolment that differed from that reported by the school. This is largely because students were typically interviewed during the daytime on school days (at home, work or school), so students not in school should be less likely to mis-report that they do attend school. Second, to an extent, the top panel of table 4 already eliminated many of the non-verified households, since if a neighbor had to report on the youth’s schooling, we would also not have income data for that household and they would have been dropped from the analysis. However, the overlap is not perfect, as there are some households where neighbors provided data that could be verified.

The bottom panel of table 4 reveals that using only the verified data reduces the sample sizes slightly, but does not change the results dramatically. The effect of the treatment for wealthier households is still positive for all three measures of education, though slightly smaller for years of schooling and having completed secondary school; and in the latter case, the significance level declines (p-value of .12) so that it no longer falls within conventional levels. However, in terms of both returning for 9th grade and total years completed, the results suggest that the schooling gains were real, rather than a reporting bias induced by the intervention. However, we must maintain the assumption that enrollment among students whose data could not be verified is not negatively correlated with the treatment.³⁵

³⁵ For example, if we make the strong assumption that all control students whose data could not be verified were enrolled whereas all treatment students whose data could not be verified were not enrolled, the treatment effects

Finally, while the results suggest that the increased schooling was due to the observable impact of the intervention on perceived returns to schooling, we are unable to rule out that some of the effect was due to other factors, such as reducing the uncertainty of students' estimates,³⁶ or that when providing information on the returns, enumerators provided additional information or encouragement to students to remain in school. In addition, we are unable to directly link at the individual level the increases in perceived returns to schooling and actual schooling, since there were large increases in perceived returns even for some members of the control group (perhaps due to measurement error, uncertainty or other new outside information).

IV. LOCAL INFORMATION, RESIDENTIAL SEGREGATION AND PERCEIVED RETURNS

We view the results of the intervention as the primary contribution of this paper; students appear to have low perceived returns to education and providing them with information on the higher measured returns increases schooling. However, we briefly explore one potential explanation for why students might underestimate the returns. While we are unable to undertake a complete test of all potential explanations, we consider a simple model of residential segregation by income. The model is akin to the argument of Wilson (1987), and is based on anecdotal impressions suggesting a high degree of segregation³⁷ and substantial residential mobility in the Dominican Republic, with successful workers often moving to better neighborhoods (both within and across cities/towns), especially to access higher quality public services such as electricity, water or schools. While some of Wilson's argument has been formalized elsewhere (for example, Streufert 2000 and Moizeau, Tropeano and Vergnaud 2004), we provide a simple discussion here to present and clarify (a slight variant of) the basic argument, and to derive additional predictions not discussed elsewhere that will be valuable in exploring the hypothesis.

would be smaller and no longer statistically significant. While we have no reason to believe non-verified treatment students are less likely to be enrolled than non-verified control students, this assumption is not testable.

³⁶ For example, if students were initially more uncertain of their estimates for earnings with secondary school than their estimates for primary school, reduced uncertainty due to the treatment might in itself have had an independent effect on the decision to stay in school, even if estimates of the returns were unchanged.

³⁷ Unfortunately, there are no empirical studies of the extent of income segregation in the Dominican Republic, nor is there sufficient data to undertake such an analysis.

The main intuition is easily described; a slightly more formal presentation is in the Appendix. Consistent with the survey results, we assume that the only information available to a given youth is the education and earnings of the set of workers they can observe in their own community. For simplicity, we consider schooling that takes on only two states, “unschooled” and “schooled.” Earnings depend on education and a random shock; panel A of figure 1 plots the two earnings distributions, where all individuals live in one community (or there are multiple communities, with identical distributions of earnings). Provided they have access to an unbiased sample of workers, students in principle have enough information to form an unbiased estimate of the mean earnings for the two groups, and thus can accurately estimate the returns to education.

Now consider residential segregation by income, where all households with income greater than some threshold live in a “rich” neighborhood, while all others live in a “poor” neighborhood. Assume students are unaware of the segregation mechanism, and continue to have information only on workers in their community.³⁸ Panel B shows that for both levels of education, workers living in the poor neighborhood will be those with the least favorable income shocks. This truncation reduces the means of both distributions, but much more so for the schooled workers. For example, it may be that all unschooled workers live in the poor neighborhood, so the local, truncated mean is close to the population mean, but only those schooled workers with the very worst income shocks are found there, leading to a greatly truncated mean, and thus a smaller difference in the local, within-community means by education. Thus, it is not that there are no educated workers in the poor community, but that these workers are negatively selected on income; the only schooled workers that youths in poor communities observe do not earn much more than the unschooled, giving the impression that education does not increase earnings substantially.

Panel C shows that the differences in the population means will also be underestimated in the rich neighborhood; while the truncated distributions overestimate the (full-population) earnings of both schooled and unschooled workers, the difference between them is underestimated because the rich

³⁸ Alternatively, students may observe “success stories” (for example, knowing some people have a nicer house or car or live in a nicer area), but still underestimate their earnings, or the role that education played in their success.

neighborhood contains most of the distribution of schooled workers but only those unschooled workers who had very good income shocks. Thus, while the emphasis is often on the effects of segregation on “underclass social isolation,” it may adversely affect those living in wealthier communities as well.

Because we are unable to measure the extent of residential segregation with our small cluster sizes, a full test of this hypothesis is not possible. However, a few observations are consistent with the hypothesis playing at least a partial role in depressing perceived returns to schooling. First, the simple model predicts not just lower perceived returns in both rich and poor communities, but in particular that students in poor communities will underestimate the population average earnings for both groups, whereas in rich communities, students will overestimate both. Columns 4 to 9 of table 1 split the sample into the 75 poorest and wealthiest (though still relatively poor) sample clusters (based on cluster median income). The data are consistent with the predictions; for both themselves and for all other workers, mean earnings are underestimated (relative to the full population means) in the poor neighborhoods, but dramatically more so for high school educated workers, and vice versa for the rich neighborhoods. Further, the estimate of the mean earnings of workers with only primary school by students in the wealthier communities exceeds the estimates of the mean earnings of high school educated workers by students in the poor communities.³⁹ In addition, the estimates of earnings by education within rich and poor communities are closer to those measured in the data for those communities.⁴⁰ These results suggest that students’ perceptions of earnings are heavily influenced by information/workers within their own communities. Of course, it might be argued that students should only use local information in forming expectations, since there may be differences across communities in important factors that influence expected earnings, such as school quality. However, again, table 1 reveals that the pattern of expected

³⁹ While it may seem unlikely that students in rich communities think the earnings of low educated workers are so high, it is important to keep in mind that the estimates are for workers who completed primary school. Wealthy students may observe poorer workers, but believe most of them have less than primary education. For all of these results, it is also important to keep in mind that responses were for the “typical worker,” so, for example, students in rich areas may know there are some low education workers who earn much less than what they report for the typical worker, and students in poor areas may know that some workers with high education earn significantly more.

⁴⁰ We cannot undertake this analysis at a more disaggregated level, such as clusters. Due to the small number of households per cluster, few contain more than one or two workers with exactly 8 or 12 years of schooling, and even fewer have more than one or two of both. Thus, estimates of earnings by education within clusters are very noisy.

earnings holds for youths' perceptions of workers throughout the Dominican Republic, not just in their own expected earnings; provided students interpreted the questions correctly, consideration of local factors such as school quality should have influenced only their own expected earnings.

Second, the simple model of residential segregation and local information would predict that a mean-preserving increase in the variance of the distribution of earnings in a community should lead to an increase in the perceived returns to schooling in that community, as the truncated distributions come to more closely approximate the population distributions. In other words, less segregated communities should result in students having higher perceived returns to schooling; dividing the distributions in figure 1 into an increasing number of communities, in the limit income segregation is complete and all workers in a given community, regardless of education, have identical incomes, so the local returns are zero. Bearing in mind the concerns about eliciting expected earnings, table 5 shows regressions for the correlates of implied perceived returns to schooling in the first round. Controlling for median cluster earnings,⁴¹ a RD\$100 increase in the standard deviation of within-cluster earnings (the mean standard deviation is 476) is associated with a RD\$8 increase in students' perceived returns.⁴² While we cannot rule out that there are other characteristics of communities with varying degrees of earnings dispersion that also affect the actual or perceived returns to schooling (such as within-community variation in school quality), these results are again consistent with an over-reliance on local information, coupled with residential income segregation.

Finally, though perhaps most weakly, table 5 also suggests that factors that may influence the information students have about the returns to schooling are correlated with implied perceived returns. Older students (conditional on still being in the 8th grade), who are likely to have more information, perhaps from visiting outside the community, have higher (and, while not explicitly shown in the table, more accurate) estimates of the returns to schooling. Similarly, students who watch more television have

⁴¹ Here we use wage earnings; similar results hold when using household income per capita.

⁴² Unfortunately, we cannot include in the regressions a variable measuring the returns to schooling in the local community. As noted above, the small cluster size means few have more than one or two observations with exactly primary or secondary schooling, and even fewer have both, making estimates of the local returns very noisy.

higher estimates of the returns to education; for students living in poor communities with little visible evidence of the returns, television may provide salient examples of the wealth and lifestyles of more highly educated workers (such as doctors or lawyers) depicted on television. However, these results may be driven by other unobserved variables, so the interpretation is again speculative.

V. CONCLUSIONS

We find that while the measured returns to education in the Dominican Republic are high, the returns perceived by students are very low. An intervention that provided information on the measured returns both increased perceived returns and improved schooling outcomes. The effects are large and striking; there are few examples of policies or interventions that result in a 5 to 10 percent increase in the likelihood of school enrolment or a .20 year increase in schooling, much less interventions that are as inexpensive as this one. Provided information can be credibly and effectively delivered with existing teaching and administrative resources, programs such as this, or career counseling, mentorship programs or “career days,” are likely to be among the most cost-effective strategies available for improving education. Given the large and striking results found here, further replication studies in other settings are worth consideration.

The results also suggest that while addressing supply-side issues such as distance to school or fees and other costs may be worthwhile public policy, the returns to and the demand for education play an important role as well; more students will attend school if they believe there to be a sufficiently large return to doing so. In fact, information-based programs may ultimately prove more beneficial than other programs, because they stimulate the demand for schooling itself, rather than for, say cash incentives to be obtained through attendance, so students may be more committed to school and provide greater effort. However, it is important to keep in mind that the intervention had no effect for poorer households, suggesting that demand alone is not sufficient. For these households, a combination of stimulating demand coupled with lowering the barriers to attendance may be more effective.

In addition, though we are unable to rule out alternative hypotheses, the results are consistent with Wilson's (1987) model of residential segregation by income, suggesting that low perceived returns and low levels of education are in part an informational problem, which the experiment helped remedy. These results suggest that economic segregation may harm lower (and, potentially, higher) income communities; this could lead to the intergenerational transmission of poverty, as poor children underestimate the returns to education and therefore underinvest in education. On a larger scale, this could result in social inefficiency and a "development trap," as the relative skill composition demanded by the labor market is not transmitted to youths in the form of greater perceived returns, resulting in an undersupply of skilled labor which in turn inhibits the development of domestic skill-intensive industries or the ability to attract foreign direct investment.

Finally, an important limitation in understanding perceived returns and the demand for schooling is the difficulty in eliciting the returns for youths of this age. But based on the results of this analysis, improving the methodology for measuring these perceptions is worth further exploration.

APPENDIX: RESIDENTIAL SEGREGATION AND THE PERCEIVED RETURNS TO EDUCATION

We consider a very basic model to illustrate the central point described in the text, rather than a more complete model of all the factors that affect education decisions. Consider a basic set-up akin to Manski (1993), in which students infer the returns by observation of current workers. Assume the present value of lifetime earnings, Y , depends on ability, A , schooling, s , and a random shock, $Y = \alpha_Y + \beta_Y A + \gamma_Y S + \varepsilon_Y$. Students have a “distaste” for schooling c , which can be expressed in monetary terms as a function of ability and a taste factor ε_c , $c = \alpha_c + \beta_c A + \varepsilon_c$, $\beta_c < 0$. We assume that ability, the income shock and the taste factor are independently distributed normal random variables. For simplicity, we consider schooling that takes on only two values, 0 (unschooled) or 1 (schooled). The decision on schooling compares whether expected lifetime income with schooling minus the costs exceeds expected lifetime income without schooling, so that the child goes to school if and only if $E^*[Y | A, s=1] - c > E^*[Y | A, s=0]$. In forming expectations, if students observe the ability,⁴³ earnings and schooling of all current workers, for a given A , the mean incomes for schooled and unschooled workers are then $E[Y | A, s=1] = \alpha_Y + \beta_Y A + \gamma_Y$ and $E[Y | A, s=0] = \alpha_Y + \beta_Y A$. The difference between the two, γ_Y , is an accurate estimate of the true returns. Students choose schooling if the cost of schooling is less than the return, so those with the smallest taste shock attend school, $\varepsilon_c \leq \alpha_c + \beta_c A - \gamma_Y$.

Now we add residential segregation by income, where all households with income $Y \geq Y^*$ live in a rich neighborhood, while all those with $Y < Y^*$ live in a poor neighborhood. We assume students are unaware of the segregation mechanism, and compute the returns to education as the difference in means for schooled and unschooled persons in their neighborhood, conditional on ability. Educated workers are found in the poor neighborhood if $\alpha_Y + \beta_Y A + \gamma_Y + \varepsilon_Y < Y^*$, i.e., those with the lowest income shocks, $\varepsilon_Y < Y^* - \alpha_Y - \beta_Y A - \gamma_Y$. Mean income among the educated in the poor neighborhood, given A , is then,

⁴³ If students do not observe ability, the difference between perceived and actual differences in the returns will depend on whether the segregation effect overwhelms the unobserved ability bias; adding in classical measurement error for observations of education further leads to underestimates of the returns to education.

$$E[Y | A, s=1, \varepsilon_Y < Y^* - \alpha_Y - \beta_Y A - \gamma_Y] = \alpha_Y + \beta_Y A + \gamma_Y - \sigma_Y \left(\frac{\phi\left(\frac{Y^* - \alpha_Y - \beta_Y A - \gamma_Y}{\sigma_Y}\right)}{\Phi\left(\frac{Y^* - \alpha_Y - \beta_Y A - \gamma_Y}{\sigma_Y}\right)} \right)$$

A similar selection occurs among the unschooled, so that mean income given A for this group is

$$E[Y | A, s=0, \varepsilon_Y < Y^* - \alpha_Y - \beta_Y A - \gamma_Y] = \alpha_Y + \beta_Y A - \sigma_Y \left(\frac{\phi\left(\frac{Y^* - \alpha_Y - \beta_Y A}{\sigma_Y}\right)}{\Phi\left(\frac{Y^* - \alpha_Y - \beta_Y A}{\sigma_Y}\right)} \right)$$

Thus in both cases, the community-mean earnings by education are lower than the corresponding population means. The difference between these two, the estimate of the return to schooling is,

$$\gamma_Y + \sigma_Y \left(\frac{\phi\left(\frac{Y^* - \alpha_Y - \beta_Y A}{\sigma_Y}\right)}{\Phi\left(\frac{Y^* - \alpha_Y - \beta_Y A}{\sigma_Y}\right)} \right) - \sigma_Y \left(\frac{\phi\left(\frac{Y^* - \alpha_Y - \beta_Y A - \gamma_Y}{\sigma_Y}\right)}{\Phi\left(\frac{Y^* - \alpha_Y - \beta_Y A - \gamma_Y}{\sigma_Y}\right)} \right) < \gamma_Y$$

with the last inequality following from the fact that $\phi(\cdot)/\Phi(\cdot)$ is monotone decreasing because ε_Y is drawn from a log-concave distribution (Bagnoli and Bergstrom 2005).⁴⁴ Thus, students in the poor neighborhood underestimate the returns to schooling (because the truncation does not lower the means for both groups equally, instead lowering it more for the high educated worker than the lower educated workers).

Similarly, the perceived returns to education in the rich neighborhood will be,

$$E[Y | A, s=1, \varepsilon_Y \geq Y^* - \alpha_Y - \beta_Y A - \gamma_Y] - E[Y | A, s=0, \varepsilon_Y \geq Y^* - \alpha_Y - \beta_Y A - \gamma_Y] =$$

$$\gamma_Y + \sigma_Y \left(\frac{\phi\left(\frac{Y^* - \alpha_Y - \beta_Y A - \gamma_Y}{\sigma_Y}\right)}{1 - \Phi\left(\frac{Y^* - \alpha_Y - \beta_Y A - \gamma_Y}{\sigma_Y}\right)} \right) - \sigma_Y \left(\frac{\phi\left(\frac{Y^* - \alpha_Y - \beta_Y A}{\sigma_Y}\right)}{1 - \Phi\left(\frac{Y^* - \alpha_Y - \beta_Y A}{\sigma_Y}\right)} \right) < \gamma_Y$$

with the last inequality following from the fact that the hazard ratios in brackets are monotone increasing. Thus, students in the rich neighborhood also underestimate the returns to schooling; here, they overstate the population means for both groups, but even more so for the low educated workers. These results follow more generally from the idea that if we truncate two otherwise identical (log-concave) distributions with different means at a common point, whether from above or below, the difference in means (or any percentile of the distribution) are less than the difference for the full distributions.

⁴⁴ Bagnoli-Bergstrom show $\phi(x)/(1-\Phi(x))$ is monotone increasing. But then $\phi(-x)/(1-\Phi(-x))$ is monotone decreasing, and since the normal is symmetric, $\phi(-x) = \phi(x)$ and $1-\Phi(-x) = 1-(1-\Phi(x)) = \Phi(x)$, so $\phi(-x)/(1-\Phi(-x)) = \phi(x)/\Phi(x)$.

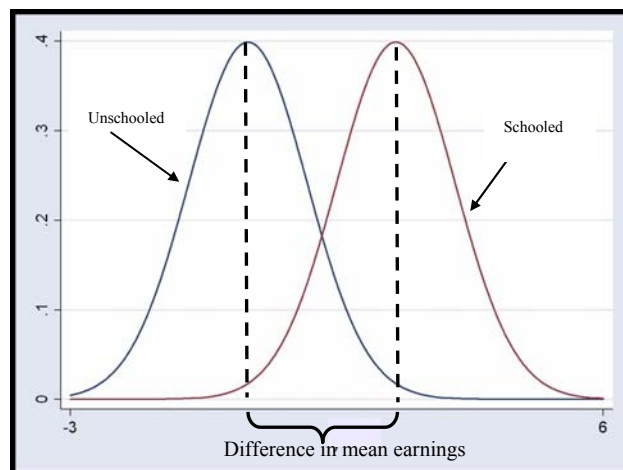
REFERENCES

- Angrist, Joshua, Eric Bettinger, and Michael Kremer (2006). "Long-Term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia," *American Economic Review*, 96(3), p. 847-862.
- Avery, Christopher and Thomas J. Kane (2004). "Student Perceptions of College Opportunities: The Boston COACH Program," in Hoxby, Caroline M., ed., College Choices: The Economics of Where to Go, When to Go, and How to Pay for It. Chicago: University of Chicago Press.
- Bagnoli, Mark and Ted Bergstrom (2005). "Log-concave Probability and its Applications," *Economic Theory*, 26, p. 445-469.
- Becker, Gary (1967). Human Capital and the Personal Distribution of Income. Ann Arbor Michigan: University of Michigan Press.
- Betts, Julian (1996). "What Do Students Know About Wages? Evidence From a Survey of Undergraduates," *Journal of Human Resources*, 31(1), p. 27-56.
- Card, David (1995). "Using Geographic Variation in College Proximity to Estimate the Return to Schooling," in Louis N Christofides, E. Kenneth Grant and Robert Swidinsky, eds., Aspects of Labour Market Behaviour. Toronto: University of Toronto Press.
- Chan, Sewin and Ann Huff Stevens (forthcoming). "What You Don't Know Can't Help You: Pension Knowledge and Retirement Decision-Making," *Review of Economics and Statistics*.
- Dominitz, Jeff and Charles F. Manski (1996). "Eliciting Student Expectations of the Returns to Schooling," *Journal of Human Resources*, 31(1), p. 1-26.
- Duflo, Esther (2001). "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment," *American Economic Review*, 91(4), p. 795-813.
- - and Emmanuel Saez (2003). "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment," *Quarterly Journal of Economics*, 118(3), p. 815-842.
- Dupas, Pascaline (2006). "Relative Risks and the Market for Sex: Teenagers, Sugar Daddies and HIV in Kenya," mimeo, Harvard University.
- Durlauf, Steven N. (2004). "Neighborhood Effects," in, Henderson, J.V. and J.F. Thisse, eds., Handbook of Regional and Urban Economics, Volume 4, Amsterdam: North Holland.
- Glazer, Nathan and Daniel P. Moynihan (1963). Beyond the Melting Pot. Cambridge, MA: MIT Press.
- Gustman, Alan L. and Thomas L. Steinmeier (2005). "Imperfect Knowledge of Social Security and Pensions," *Industrial Relations*, 44(2), p. 373-397.
- Handlin, Oscar (1959). The Uprooted. Philadelphia, Pa.: University of Pennsylvania Press.
- Hastings, Justine S., Richard Van Weelden and Jeffrey Weinstein (2007). "Preferences, Information and Parental Choice Behavior in Public School Choice," National Bureau of Economic Research Working Paper 12995, Cambridge, MA.
- Kane, Thomas J. and Cecilia Rouse (1993). "Labor Market Returns to Two- and Four- Year Colleges: Is a Credit a Credit and Do Degrees Matter?" NBER Working Paper No.4268.
- Manski, Charles F. (1993). "Adolescent Econometricians: How Do Youth Infer the Returns to Education?" in, Clotfelter, Charles T. and Michael Rothschild, eds., Studies of Supply and Demand in Higher Education, Chicago, Ill: University of Chicago Press.
- Miguel, Edward and Michael Kremer (2004). "Worms: Identifying Impacts of Education and Health in the Presence of Treatment Externalities," *Econometrica*, 72(1), p. 159-217.
- Mitchell, Olivia (1988). "Worker Knowledge of Pension Provisions," *Journal of Labor Economics*, 6(1), p. 21-39.
- Moizeau, Fabien, Jean-Philippe Tropeano and Jean-Christophe Vergnaud (2004). "Neighborhood Informational Effects and the Organization of the City," mimeo.
- Psacharopoulos, George and Harry Anthony Patrinos (2004). "Returns to Investment in Education: A Further Update," *Education Economics*, 12(2), p. 111-134.

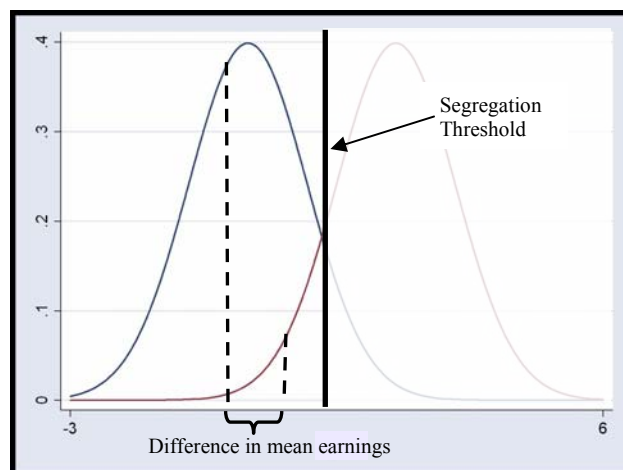
- Rouse, Cecelia Elena (2004), "Low-Income Students and College Attendance: An Exploration of Income Expectations," *Social Science Quarterly*, 85(5), p. 1299-1317.
- Schultz, T. Paul (2004). "School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program," *Journal of Development Economics*, 74, p. 199-250.
- Smith, Herbert L. and Brian Powell (1990). "Great Expectations: Variations in Income Expectations Among College Seniors," *Sociology of Education*, 63(3), p. 194-207.
- Stango, Victor and Jonathan Zinman (2007). "Fuzzy Math and Red Ink: When the Opportunity Cost of Consumption is Not What it Seems," mimeo, Dartmouth College.
- Struefert, Peter (2000). "The Effect of Underclass Social Isolation on Schooling Choice," *Journal of Public Economic Theory*, 2(4), p. 461-82.
- Viscusi, W. Kip (1990). "Do Smokers Underestimate Risks?" *Journal of Political Economy*, 98(6), p. 1253-169.
- Wilson, William Julius (1987). The Truly Disadvantaged. Chicago: University of Chicago Press.

FIGURE 1. RESIDENTIAL SEGREGATION AND MEAN EARNINGS BY EDUCATION

A. Full Population



B. Income Segregation—Poor Neighborhood



C. Income Segregation—Rich Neighborhood

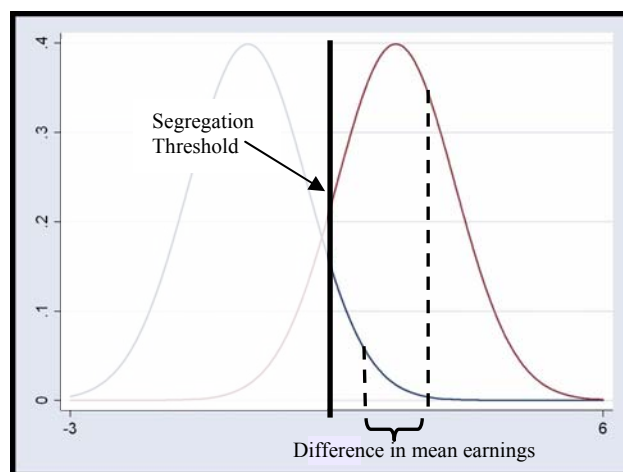


TABLE 1. MEASURED AND PERCEIVED MONTHLY EARNINGS

	POOR COMMUNITIES						RICH COMMUNITIES		
	(1) Measured Mean	(2) Perceived (self)	(3) Perceived (others)	(4) Measured Mean	(5) Perceived (self)	(6) Perceived (others)	(7) Measured Mean	(8) Perceived (self)	(9) Perceived (others)
Not Complete Secondary	3,180 [1400]	3,516 [846]	3,478 [829]	2,610 [1098]	2,760 [426]	2,740 [433]	4,021 [1153]	4,176 [578]	4,122 [554]
Complete Secondary	4,479 [1432]	3,845 [1003]	3,765 [957]	3,028 [990]	2,985 [502]	2,939 [433]	4717 [1332]	4,595 [756]	4,485 [719]
(Sec) – (Not Sec)	1,299	329 [406]	287 [373]	318	225 [302]	199 [271]	696	419 [469]	363 [439]

Notes: All figures in 2001 Dominican Pesos (RD\$). Standard deviations in bracket. Rich and poor communities are defined as above vs. below the 50th percentile of community-level median income.

TABLE 2. EFFECT OF THE INTERVENTION ON EXPECTED RETURNS AND SCHOOLING

	<u>ROUND 1</u>		<u>ROUND 2</u>		<u>ROUND 3</u>	
	Control	Treatment	Control	Treatment	Control	Treatment
Expected Earnings (Self):						
Primary (Only)	3,550 (117)	3,481 (121)	3,583 (118)	3,230 (92)		
Secondary (Only)	3,890 (134)	3,802 (144)	4,001 (132)	3,995 (114)		
Estimated Returns	339 (25)	321 (28)	418 (24)	765 (34)		
Returned to School?			.55 (.02)	.59 (.02)		
Completed Secondary School?					.30 (.02)	.32 (.02)
Years of schooling completed					9.75 (.070)	9.93 (.073)

Notes: All figures in 2001 Dominican Pesos (RD\$). Standard errors, corrected for clustering at the school-level, in parentheses. The number of observations for returning to school or completing secondary school is 1125 for both treatment and control groups; for years, it is 1033 for the controls and 1041 for the treatment group (most observations in the third round reported by parents or neighbors indicated whether the youth completed secondary school, but not how many years of schooling they had completed if they did not finish). The number of observations for expected earnings in round 1 is 1003 for the control group and 1022 for the treatment group; in round 2, the number of observations is 922 for the control group and 977 for the treatment group.

TABLE 3. EFFECTS OF THE INTERVENTION ON SCHOOLING

	<i>Δ Implied Return (Self)</i>		<i>Returned Next Year</i>			<i>Completed Secondary</i>		<i>Years of Schooling</i>	
	(1)	(2)	(3)	(4)	(5)	(7)	(8)	(9)	(10)
Treatment	366 (29)	366 (29)	.039 (.025)	.041 (.023)		.020 (.024)	.023 (.020)	.18 (.098)	.20 (.083)
Log (income per capita)		30.0 (48)		.075 (.042)			.21 (.044)		.75 (.16)
School Performance		1.1 (13)		.011 (.010)			.019 (.008)		.085 (.035)
Father's education		-26 (33)		.082 (.029)			.061 (.029)		.28 (.12)
Interviewed					.014 (.027)				
R ²	.087	.087	.002	.015	.0002	.001	.036	.003	.040
Observations	1,859	1,859	2,250	2,250	1,575	2,250	2,250	2,074	2,074

Notes: Heteroskedasticity-consistent standard errors accounting for clustering at the school-level in parentheses. *Δ Implied Return (Self)* is the change (round 2 minus round 1) in the implied returns to schooling the student expects for themselves (earnings with secondary completion minus earnings with primary only) in 2001 Dominican pesos (no adjustment made for inflation between the two rounds). School Performance is teacher assessment of the student's performance, on a scale of 1 to 5 (much worse than average, worse than average, average, above average, much better than average).

TABLE 4. EFFECTS OF THE INTERVENTION ON SCHOOLING BY SOCIOECONOMIC STATUS

	<u>POOR HOUSEHOLDS</u>				<u>RICH HOUSEHOLDS</u>			
	<i>Returned</i> (1)	<i>Completed School</i> (2)	<i>Years Education</i> (3)	<i>Δ Perceived Returns (self)</i> (4)	<i>Returned</i> (5)	<i>Completed School</i> (6)	<i>Years Education</i> (7)	<i>Δ Perceived Returns (self)</i> (8)
Treatment	.007 (.035)	-.001 (.026)	.041 (.11)	344 (41)	.072 (.038)	.054 (.031)	.34 (.12)	384 (41)
Log (income per capita)	.064 (.065)	.25 (.06)	.71 (.23)	202 (83)	.046 (.12)	.10 (.13)	.51 (.44)	31 (134)
School Performance	.001 (.014)	.015 (.012)	.065 (.048)	-9.2 (13.3)	.025 (.013)	.024 (.012)	.10 (.048)	8.6 (22)
Father's education	.055 (.045)	.019 (.043)	.16 (.18)	-29.7 (62)	.091 (.039)	.096 (.039)	.35 (.14)	-5.1 (40)
R ²	.003	.019	.013	.091	.019	.020	.029	.088
Observations	1055	1055	1007	920	1056	1056	1002	939

<u>VERIFIED SCHOOLING DATA ONLY</u>								
	<u>POOR HOUSEHOLDS</u>				<u>RICH HOUSEHOLDS</u>			
	<i>Returned</i> (1)	<i>Completed School</i> (2)	<i>Years Education</i> (3)	<i>Δ Perceived Returns (self)</i> (4)	<i>Returned</i> (5)	<i>Completed School</i> (6)	<i>Years Education</i> (7)	<i>Δ Perceived Returns (self)</i> (8)
Treatment	.010 (.035)	-.016 (.027)	.039 (.11)	344 (41)	.072 (.038)	.047 (.030)	.31 (.12)	385 (41)
Log (income per capita)	.059 (.065)	.23 (.065)	.71 (.23)	202 (83)	.044 (.12)	.010 (.14)	.49 (.45)	29 (134)
School Performance	.001 (.014)	.012 (.012)	.056 (.047)	-9.1 (13.3)	.024 (.013)	.026 (.013)	.10 (.048)	8.5 (22)
Father's education	.056 (.045)	.040 (.046)	.17 (.18)	-29.8 (62)	.091 (.039)	.099 (.039)	.38 (.14)	-5.5 (40)
R ²	.003	.017	.012	.090	.0219	.021	.029	.088
Observations	1052	994	983	919	1055	991	978	938

Notes: Heteroskedasticity-consistent standard errors accounting for clustering at the school-level in parentheses. Youths are split according to whether they live in a household that is below (poor) or above (rich) the median household income per capita. School performance is teacher assessment of the student's performance, on a scale of 1 to 5 (much worse than average, worse than average, average, above average, much better than average). Verified Schooling Data refers to only data where school enrollment and completion status could be verified at the school the child had attended.

TABLE 5. CORRELATES OF IMPLIED PERCEIVED RETURNS

	<i>Implied Perceived Returns (Self)</i>
Standard Deviation Wages	.083 (.029)
Median Wage	.087 (.018)
Amount of TV Watched	23.8 (5.0)
School Performance	18.8 (7.9)
Father's Education	27.7 (23.5)
Student's Age	33.5 (11.5)
R ²	.11
Observations	2,025

Notes: Heteroskedasticity-consistent standard errors accounting for clustering at the school-level in parentheses. *Implied Returns (Self)* is the Round 1 implied returns to schooling the student expects for themselves (earnings with secondary completion minus earnings with primary only) in 2001 Dominican pesos. Standard Deviation Wages is the standard deviation of wages in the cluster the student lives in and Median Wage is the cluster median wage. Amount of TV Watching is the hours of television watched on a typical day, scored on a scale of 0 to 5 (0: don't watch at all; 1: less than one hour; 2: 1-2 hours; 3: 2-3 hours; 4: 4-5 hours; 5: more than 5 hours).