Adolescents' demand for information

Marco Castillo[*] Mofioluwadasademi Ige[†‡]

May 2022

Abstract

Worldwide, adolescents make crucial career and schooling decisions in the face of significant information disparities. We investigate their demand for information and how it conforms with economic rationality by selling information on wages and college admissions in an RCT in Ibadan, Nigeria. We implement the RCT at a critical juncture when adolescents must decide to drop out or continue schooling. Over 70% of participants exhibit monotone preferences for the quantity of information. The intervention increased the high school dropout rate one year later by 3.8 percentage points from a baseline of 9.4 percent. We find that a 100% increase in willingness to pay (WTP) is associated with a 70% increase in the probability of discontinuing education and this is not statistically significantly different from a proportional relationship. We corroborate these results using belief data. Adolescents facing high-stakes decisions behave close to the economic paradigm.

JEL classifications: C93, J24, D83

Keywords: human capital accumulation, costly information acquisition, randomized controlled trial

[*]Department of Economics, Texas A&M University, Melbourne Institute and IZA, e-mail: marco.castillo@tamu.edu

[†]Freddie Mac, e-mail: Mofioluwasademi_ige@freddiemac.com

[‡]We gratefully acknowledge the support of the state government of Oyo, Nigeria for this research. The project was pre-registered at the AEA RCT Registry (AEARCTR-0004839) and received human subjects approval from Texas A&M University's IRB (IRB2018-0424D).

1 Introduction

At age 15 or earlier, about half of the adolescents across the world decide whether to continue their education and, if they decide to do so, along which path (e.g. academic, vocational).[1] These decisions require careful consideration of available options since mistakes can be very costly (Lai, Sadoulet and de Janvry, 2009; Lucas and Mbiti, 2012). Existing evidence about the effectiveness of information provision in increasing levels of education is mixed.[2] Understanding the effect, or lack thereof, of information on adolescents' human capital accumulation requires determining how they use this information and whether it conforms to economic rationality. Yet, we know very little how adolescents process economic information. This paper aims to fill this gap.

We develop a test for rational demand for information based on the intuition that willingness to pay (WTP) for information is a function of its anticipated influence on behavior. Blackwell (1951) offers the insight that data are valuable because they enable better decision making. The same idea is discussed by Hirshleifer (1971), who proposes that information is useful insofar as it influences choices, and de Lara and Gossner (2020) who have analyzed the implications of the analysis of the value of information. We exploit the fact that the vector of choice probabilities compatible with rational choice coincides with the gradient of the ex-ante expected utility (Sørensen and Fosgerau, forthcoming; Chiong, Galichon and Shum, 2016; McFadden, 1978; Rust, 1994) to derive a variational representation of the WTP for information as a function of the expected changes in behavior it generates.[3] In the binary choice case, e.g. continue education or not, this implies that the derivative of the WTP for information with respect to relative expected payoffs equals the expected change in behavior. This is a precise prediction that is testable.

Testing this implication of rational information acquisition requires observing the WTP for information and the actual effect of the information on decisions. To do this, we embed an incentive-compatible measure of adolescents' WTP for information in an information randomized control trial (RCT). Random assignments of information provide the needed counterfactual responses to information. The elicited WTP for information allows testing

[1] See https://data.worldbank.org/indicator/SE.COM.DURS.

[2] Jensen (2010) finds that information about the returns of education increase schooling. Other studies providing information about several aspects of educational investment report null as well as positive and negative results (see Ayaita, Spengler, Nagengast and Trautwein, 2020; Bergman, Denning and Manoli, 2019; Bettinger, Long, Oreopoulos and Sanbonmatsu, 2012; Busso, Dinkelman, Martinez and Romero, 2017; Goux, Gurgand and Maurin, 2017; Gurantz, Howell, Hurwitz, Larson, Pender and White, 2021; Hastings, Neilson and Zimmerman, 2015; Hoxby and Turner, 2015; Hyman, 2020; Loyalka, Liu, Song, Yi, Huang, Wei, Zhang, Shi, Chu and Rozelle, 2013).

[3] The assumptions needed for this result are given in Section 2 as well as its relationship to previous results in the literature.

if those who are affected by the provided information are also those who are willing to pay more for it, and by how much.

We conducted a field experiment in Ibadan, Nigeria that randomized information about wages and college admission rates to over 3,600 14-year-old adolescents who were deciding whether to continue to senior high school and, if so, which curriculum track to follow. As part of the study, we collected the adolescents' beliefs about earnings and their own future academic choices as well as their WTP for different amounts of information. This allowed us to test for monotonicity of preferences since we expect WTP to be (weakly) increasing in the amount of information (see Blackwell, 1951). For a subgroup of students, these beliefs were collected twice, before and after the decision to reveal the information was made. Our data collection allows us not only to conduct our proposed test using actual field outcomes, but also to investigate the same pattern using self-reported beliefs data. Our experiment then allows us to investigate if belief data reproduces field outcomes.

We find that 70 percent of participants exhibit monotone preferences for information. Overall, the information intervention led to a 3.8 percentage point decrease in education continuation rates a year later from a baseline of 91 percent.[4] We find that a 100% increase in WTP is associated with a 70% increase in the probability of discontinuing education. We cannot reject the hypothesis that the relation is exactly proportional. Conversely, the treatment effect is a 40 percent increase in dropout rates, and the willingness to pay for information of compliers is between 31 and 80 percent higher than the average depending of the information content. This relationship is closer to proportional if we control for incentivized measures of risk tolerance. We also find little evidence that WTP is explained observable characteristics including academic performance or measures of soft skills. This suggests WTP is not capturing unobservable characteristics related with human capital accumulation.

Since our study included a belief updating experiment, we are able to explore other aspects of adolescents' demand for information and, in a sense, to confirm that our results are robust. We find that participants updated their information on earnings at levels found for adults in the literature (Fuster, Perez-Truglia, Wiederholt and Zafar, forthcoming; Hjort, Moreira, Rao and Santini, 2021). Similarly, the participants' beliefs about their own future choices were consistent with partial sorting based on earnings (Arcidiacono, Hotz, Maurel and Romano, 2020). Finally, using the framework developed by Wiswall and Zafar (2015), and consistent with related studies (Delavande and Zafar, 2019; Haaland, Roth and Wohlfart, forthcoming; Wiswall and Zafar, 2018), we find that participants' choice elasticity with

[4] Contrary to Jensen (2010), who found that students underestimated the returns of education, the participants in our study overestimated these returns.

respect to earnings was relatively low (about 17 percent). These estimates are commensurate with the observed field behavior. Finally, we confirm that WTP is also increasing on reported changes in beliefs due to information.

Our paper contributes to the growing literature on decision-making by minors (Brocas and Carrillo, 2021; Brocas, Carrillo, Combs and Kodaverdian, 2019; Castillo, Ferraro, Jordan and Petrie, 2011; Castillo, Jordan and Petrie, 2018, 2019; Harbaugh, Krause and Berry, 2001; Sutter, Kocher, Glaetzle-Ruetzler and Trautmann, 2013). This literature concentrates on children's economic rationality and strategic sophistication and its relationship to life outcomes (see List, Petrie and Samek, forthcoming, for a review of this literature). Our paper opens a new area of inquiry by looking directly at how minors deal with costly acquisition of information (see Caplin, 2016; Caplin, Csaba, Leahy and Nov, 2020). Our results are consistent with minors rationally allocating their attention.

Our paper also contributes to the literature on the formation of human capital. First, we provide further evidence that minors are actors in their own development (Del Boca, Flinn, Verriest and Wiswall, 2019). Embedding behavioral measures in an RCT enables us to show the impact of students' information processing on outcomes. Second, our work is relevant to the literature on the measurement and importance of expected returns of education. That literature shows that heterogeneous beliefs and preferences regarding career choices are important; however, it retains the assumption that such beliefs are exogenous.[5] Our research shows that these beliefs are not independent of the responsiveness to those returns. This points to the need to model endogenous belief formation in the study of human capital accumulation.

Recent work has explored the demand for potentially valuable information (e.g., Allcott and Kessler, 2019; Fuster et al., forthcoming, and references therein). While the idea that information is valuable to the extent that it can influence behavior is old, such a connection is rarely exploited. An important exception is Chassang, Padro i Miquel and Snowberg (2012) who show that the willingness to pay to be treated in an experiment can be used to estimate actual and perceived treatment effects. Berry, Fischer and Guiteras (2020) apply this to sanitation interventions. Our approach applies these ideas to the context of information campaigns where the treatment being sold is information. In a related study, Fuster et al. (forthcoming) show that the WTP for information is increasing in experimental stakes. They find that WTP for information is 20 percent higher when stakes are 10 times higher.[6]

[5] This body of literature is too large to summarize here. Relevant papers utilizing belief elicitation include Arcidiacono et al. (2020), Delavande and Zafar (2019), and Wiswall and Zafar (2015).

[6] Beliefs, and choices, are affected by the information provided, but the study is not designed to test if WTP is increasing in changes in behavior. In Fuster et al. (forthcoming), rewards are based on the accuracy of beliefs. In our study, we observe behavior after information is provided. Bronchetti, Kessler, Magenheim,

Our paper appears at odds with the growing literature showing behavioral biases among the young (Brocas and Carrillo, 2021; Fe, Gil and Prowse, forthcoming), and its relationship with less human capital accumulation (Dynarski, Libassi, Michelmore and Owen, 2021; Cadena and Keys, 2015; Castillo et al., 2018, 2019). It would also appear at odds with the literature exploring the prevalence of behavioral biases among the global poor (Haushofer and Fehr, 2014; Kremer, Rao and Schilbach, 2019). However, our results are less unusual if we consider recent research showing that exchange anomalies are less prevalent among small-scale farmers in Zambia facing financial constraints (Fehr, Fink and Jack, forthcoming). Our study shows that Nigerian adolescents making high-stakes decisions in the face of significant information constraints use information economically. It would be important to determine if this pattern of behavior changes with fewer financial and information constraints.

The rest of this paper is organized as follows. Section 2 derives the main theoretical results used in the paper. Section 3 describes the RCT. Section 4 discusses the empirical strategy. Section 5 presents the main results. Section 6 discusses potential limitations of the study. Section 7 concludes the paper.

2 Theoretical framework

This section discusses a testable implication of costly information acquisition models that applies to our experimental setting. The basic intuition is that those who would benefit most from information should be willing to pay more for it. Since information has value insofar there are some states of the world in which behavior is affected, we expect that those who would change their behavior more have more reasons to acquire the information. The primary identification problem we face is that we cannot simultaneously observe a person with and without new information. We cannot estimate the individual treatment effect of information and its relationship to the willingness to pay for information. However, random assignment of information allows us to relate the treatment effect of information to the willingness to pay for it at the population level.

The formalisms in this section are based on the presentation in de Lara and Gossner (2020). To mimic the decision to continue to senior high school that we analyze in this paper, we concentrate on a two-action model. We represent the decision to continue to senior high school as action a being equal to 1 and the decision to not continue to senior high school as action a being equal to 0. Uncertainty about the returns to alternative career paths are

Taubinsky and Zwick (2020) use the framework developed by Caplin et al. (2020) to derive a test of rational inattention. They find that subjects undervalue reminders relative to the costly information acquisition benchmark.

represented as a finite set of possible states of the world $\Omega = \{\omega_1, \ldots, \omega_N\}$. Students are endowed with a utility function $u(a, \omega) = u(a) + e_a(\omega)$ that depends on the action a that is taken and the state of the world ω. The term $u(a)$ depends only on action a, and term $e_a(\omega)$ depends both on ω and option a. Decisions are made prior to the realization of ω, and students have a prior distribution $\mu \in \Gamma = \Delta(\Omega)$ over states of the world.

The set Π of information structures includes all of the functions $\pi : \Omega \to \Delta(\Gamma)$ that select measures with a finite support $\Gamma(\pi) \subset \Gamma$, and satisfy Bayes's law. Intuitively, an information structure provides a finite set of signals that agents can use to form posterior distributions of the state of the world, with different signals producing different posterior distributions. These posteriors are then used to decide which actions are optimal. The gross ex-ante payoff $G(\pi, u)$ for a student who uses information structure π optimally is:

$$G(\pi, u) = \sum_{\gamma \in \Gamma(\pi)} q(\gamma)[max_{a \in \{0,1\}} \sum_{\omega \in \Omega} \gamma(\omega)(u(a) + e_a(\omega))] \tag{1}$$

where $q(\gamma) = [\sum_{\omega \in \Omega} \mu(\omega)\pi(\gamma|\omega)]$. The equation above says that, given a state of the world ω, a person receives a signal that allows to form a posterior belief γ. This posterior beliefs is used to evaluate which action is optimal. The gross ex-ante payoff is the expected utility taking into account all the signals that are possible for each state of the world. More succinctly, information structures π can be represented as a distribution over a set of posterior distributions that average to the prior distribution.

Let $G(\mu, u)$ be the gross ex-ante payoff given prior μ. The WTP for information structure π is the number $W_\pi(u)$ such that:

$$G(\pi, u) - W_\pi(u) = G(\mu, u). \tag{2}$$

i.e., the difference between the gross ex-ante payoffs under π and μ. WTP is positive (see Chavas, 1993, for sufficient conditions for this result). de Lara and Gossner (2020) show that WTP is zero unless information structure π produces a signal at which optimal behavior under prior μ is no longer optimal. Moreover, WTP is increasing in the probability of receiving a signal leading to a change in behavior.[7]

Since the gross ex-ante payoff function is convex (see de Lara and Gossner, 2020), it is subdifferentiable. Sørensen and Fosgerau (forthcoming) show that this subdifferential coincides with the choice probability correspondence, i.e. $\partial G(\pi, u) = Pr(a = 0|\pi, u)$, and satisfies cyclical monotonicity. If the gross ex-ante payoff function is differentiable we have that $\frac{\partial G(\pi,u)}{\partial u(0)} = Pr(a = 0|\pi, u)$, which an instance of the Williams-Daly-Zachery theorem for

[7]These are Proposition 3.1 and Theorem 3.2 in de Lara and Gossner (2020)

additive random utility (McFadden, 1978; Rust, 1994). Given (2), we conclude that:

$$dW_\pi(u)/du(i) = Pr(a = i|\pi, u) - Pr(a = i|\mu, u) = \Delta Pr(a = i|u), i = 0, 1. \tag{3}$$

at any point of differentiability.[8]

Chassang et al. (2012) were the first to point out that WTP to be treated can be used to elicit actual and perceived treatment effects. We note, however, that if preferences are not separable, equation (3) can over(under)estimate the relationship between expected changes in behavior ($\Delta Pr(a = 0|u)$) and WTP.[9,10] Non-separable preferences are a common assumption in the literature on the value of information (e.g. Gould, 1974; Chavas, 1993; Eeckhoudt and Godfroid, 2000; Cabrales, Gossner and Serrano, 2017). We present evidence that WTP is increasing in risk tolerance. Importantly, the qualitative result that WTP is positive only if it has an expected effect on behavior does not depend on the separability assumption.

We illustrate these results with an example (see Figure 1). The example assumes that $u_0 = \Delta_0$ and $u_1 = 0$, where Δ_0 is the return to discontinuing senior high school. It also assumes that $e_0 = -4$ or $e_0 = 4$ with equal probability, and that e_1 is always equal to 0. That is, the are 2 states of the world that obtain with equal probability. In the absence of new information, students with $\Delta_0 < 0$ choose $a = 1$ and those with $\Delta_0 > 0$ choose $a = 0$. Any choice is optimal if $\Delta_0 = 0$. This obtains because the expected payoff to $a = 0$ equals $\frac{1}{2}(\Delta_0 + 4) + \frac{1}{2}(\Delta_0 - 4) = \Delta_0$. The optimal choice is illustrated in Figure 1 by a solid black line.

Consider now an information structure π that updates the probability that $e_0 = -4$ to $\frac{1}{4}$ with probability $\frac{1}{2}$ and that updates the probability that $e_0 = -4$ to $\frac{3}{4}$ with probability $\frac{1}{2}$. This is a valid information structure since the average of the posteriors equals the prior (i.e., $\frac{1}{2}$). A student in possession of this information will update her options accordingly. With probability $\frac{1}{2}$, the expected payoff to $a = 0$ equals $\frac{3}{4}(\Delta_0 + 4) + \frac{1}{4}(\Delta_0 - 4) = \Delta_0 + 2$, and with probability $\frac{1}{2}$, the expected payoff to $a = 0$ equals $\frac{1}{4}(\Delta_0 + 4) + \frac{3}{4}(\Delta_0 - 4) = \Delta_0 - 2$. A person with $\Delta_0 < 0$ will not change her behavior if the probability of $e_0 = -4$ increases, but could change her behavior if the probability of $e_0 = -4$ decreases. In particular, any student with $\Delta_0 \in (-2, 0)$ will switch from $a = 1$ to $a = 0$ when the positive signal is received. Since this switch in behavior happens with probability $\frac{1}{2}$, we obtain that the probability of choosing

[8]This equation holds for small changes in $\Delta Pr(a = i|u)$ which is our case. Closed-form solutions for function $W_\pi(u)$ require parametric assumptions as in (McFadden, 1978).

[9]Chambers, Liu and Rehbeck (2020) note that the prior and posterior distributions play a role similar to that of prices in consumer theory. Information structures are, by definition, mean preserving spreads of the prior distribution. We therefore expect that the certainty equivalent, and hence WTP, associated with an information structure is increasing in risk tolerance.

[10]The value of information itself depends on risk attitudes, however, incentive compatible measurement of this value of information is robust to risk attitudes.

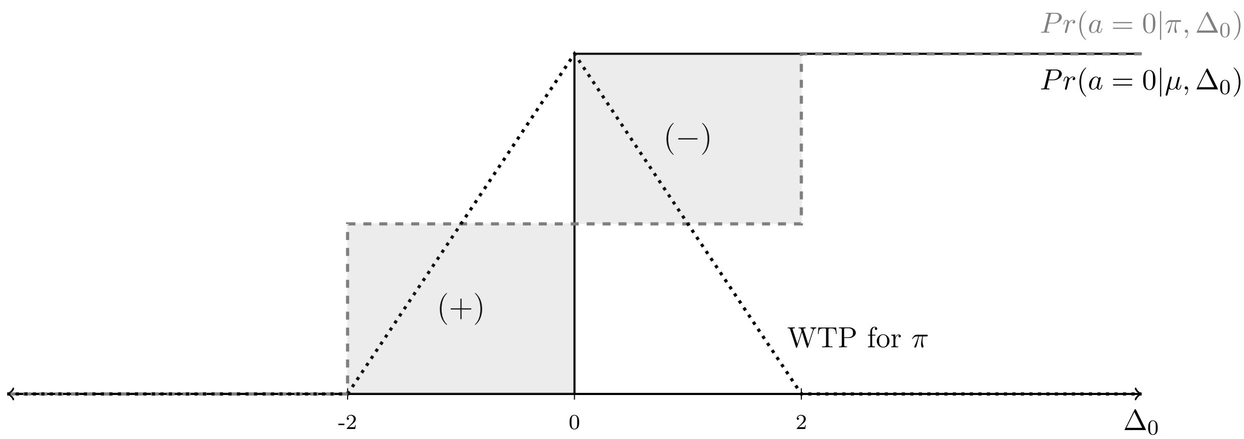


Figure 1: WTP for information as a function of returns to discontinuing education

Notes: The x-axis represents the return to discontinuing education, Δ_0. The y-axis represents the probability of discontinuing education and the willingness to pay for information structure π. $Pr(a = 0|\mu, \Delta_0)$, black solid line, denotes the best-response correspondence as a function of Δ_0 given prior μ. $Pr(a = 0|\pi, \Delta_0)$, gray dashed line, denotes the best-response correspondence as a function of Δ_0 given information structure π. The willingness to pay for information structure π is represented by the dotted black line. The marginal WTP for information structure π equals $Pr(a = 0|\pi, \Delta_0) - Pr(a = 0|\mu, \Delta_0)$ whenever this is single valued. It increases in the area marked with $(+)$, when a student might switch to $a = 0$ if information structure π is available, and decreases in the area marked (-), when a student might switch to $a = 1$ if information structure π is available. WTP is maximal when an agent is indifferent between options and decreases as choices become more certain.

$a = 0$ increases to $\frac{1}{2}$ when information structure π is available. For those with $\Delta_0 < -2$, the increase in probability of $e_0 = 4$ is not large enough to change their decisions. Those with $\Delta_0 = -2$ are indifferent between $a = 0$ and $a = 1$ when good news are received, so any choice is optimal. Their probability of choosing $a = 0$ is therefore between 0 and $\frac{1}{2}$ (half of the time $Pr(a = 0) = 0$ and half of the time $Pr(a = 0) \in [0, 1]$). Similar arguments imply that students with $\Delta_0 \in (0, 2)$ will decrease their probability of choosing $a = 0$ from 1 to $\frac{1}{2}$, and those with $\Delta_0 = 2$ will choose $a = 0$ with probability between $\frac{1}{2}$ and 1. Finally, those with $\Delta_0 = 0$ will choose $a = 0$ when the signal is good and choose $a = 1$ when the signal is bad. They choose $a = 0$ with probability $\frac{1}{2}$. Figure 1 represents these choices with a dashed gray line.

WTP for information depends on the optimal choices given information is accessed and the gains from these choices. Figure 1 shows that those with $\Delta_0 \notin (-2, 2)$ do not change their decisions if information structure π is available. So, their WTP for information is 0. Following the results above, the expected payoff for those in $(-2, 0)$ is $\frac{1}{2}(\Delta_0 + 2) + \frac{1}{2} \times 0$ under information structure π and 0 (since $a = 1$) under prior μ. Their WTP for information structure π is therefore $\frac{1}{2}\Delta_0 + 1 - 0$ and it is increasing in Δ_0. The expected payoffs for those in $(0, 2)$ is $\frac{1}{2}(\Delta_0 + 2) + \frac{1}{2} \times 0$ under information structure π and Δ_0 (since $a = 0$) under prior

μ. Their WTP for information structure π is therefore $\frac{1}{2}\Delta_0 + 1 - \Delta_0 = 1 - \frac{1}{2}\Delta_0$ which is decreasing in Δ_0. Figure 1 represents WTP by a dotted black line. We confirm that WTP for π is largest when indifference between alternatives is the smallest ($\Delta_0 = 0$).

Figure 1 makes clear that a change in beliefs is a necessary, but not sufficient, condition for a change in behavior (those with $\Delta_0 \notin (-2, 2)$ do not change behavior with new information). An information RCT might change the beliefs of two students in the same way, but it might change the behavior on only one person, both or neither one. Finally, $Pr(a = 0|\pi, \Delta u_0)$ shape reflects the number and distribution of signals. It will be a step function as long as the number of signals is finite.

Regarding testable implications of rational behavior, we cannot observe counterfactual behavior for each information structure and each potential signal. However, we can obtain consistent estimates of $Pr(a = 0|\pi, W_\pi)$ and $Pr(a = 0|\mu, W_\pi)$, that is, the conditional probability of choosing $a = 0$ given information structure π and prior μ and the WTP W_π, if individuals are randomly assigned the information structure π. Random assignment of information allows consistent estimates of these probabilities under the assumption that signals are themselves random and uncorrelated with access to information. We should note that this identification problem is not existent if one has expectation data on future behavior conditional on receiving information or not. That is, expectation data can be used to directly test the implication that willingness to pay for information is increasing in expected changes in behavior. We provide evidence consistent with this prediction later in the paper.

Figure 1 helps illustrate our main identification challenge when only actual behavior is available. Information structure π can both increase and decrease the proportion of those choosing $a = 0$. This implies that we might fail to detect the effect of information on average behavior even when information has an effect on individual behavior. In this case, WTP will be positive even if we do not observe behavioral changes on average. Our ability to test for rational use of information depends on the distribution of returns to education (Δ_0). If either most students are such that $\Delta_0 < 0$ or $\Delta_0 > 0$, we will be able to test that WTP is significantly correlated with $\Delta Pr(a = 0|\pi)$.[11] In this case, those who expect to change their behavior upon receiving information will be willing to pay for it. Moreover, for small changes in $\Delta Pr(a = 0|\pi)$, this relationship should be proportional.

A significant correlation is strong evidence in favor of rational behavior since changes in average behavior across treatments underestimates changes in individual behavior under the null hypothesis of rational behavior. In sum, our ability to test for rational behavior depends

[11] The identification challenge persists in the case in which the outcome variable is not binary, e.g. when several career paths are available. The formula for the marginal WTP also extends to this case *mutatis mutandis*.

on the information RCT having a uniform effect across students. In the language of causal inference models, this is a monotonicity assumption.[12] Those who change from $a = 1$ to $a = 0$ due to the information are *compliers* (Imbens and Angrist, 1994). The characteristics of compliers, including their WTP for information, can then be identified using standard techniques (Imbens and Rubin, 1997; Marbach and Hangartner, 2020). We conclude that our test is viable when LATE estimates of the treatment effect of discontinuing high school are valid. Conversely, WTP for information can be used to test if a monotonicity assumption for an information RCT holds.

3 Experimental design and implementation

In this section we provide background on the study setting and details about the data collection.

3.1 Background

The secondary education system in Nigeria is divided into junior high school and senior high school. After spending six years in primary school, students attend three years in junior high school and can spend three years in senior high school. At the end of junior high school, students have an opportunity to choose between an academic or a vocational curriculum. Students wishing to proceed with the academic option in senior high school have three curriculum options: arts, commercial/social science, and science. The vocational track also provides students with different areas for specialization. The diversity of the course curricula in senior high school gives students the opportunity to choose their future career paths. Students are exposed to both mainstream academic courses and vocational courses. At the end of junior high school, or grade 9, students take a statewide examination—the Basic Education Certificate Examination (BECE)—which allows the transition to the next level of schooling.

3.2 Overview

The present study recruited students in their last year of junior high school who had to decide whether to continue to senior high school, go to vocational school/take an apprenticeship, or drop out of school entirely. The study took place in the city of Ibadan, the capital of the state of Oyo, Nigeria and Nigeria's third most populous city (3.2 million). The study was

[12] Whenever students can choose from strictly more than two career options, the test will underestimate the presence of rational behavior even assuming unordered monotonicity (see Heckman and Pinto, 2018).

conducted with the approval of the State of Oyo's Ministry of Education, Science, and Technology. The experiment had five stages: recruitment, baseline data collection, information provision in treated schools, collection of endline data, and collection of administrative data on educational choices.

In the first stage, schools were recruited and consent was obtained. Students were assigned to three experimental conditions: treatment, impure control and pure control. All students were asked to fill out a baseline survey that collected basic demographic information, attitudes towards schooling, a proposed curriculum/track choice, career aspirations, and subjective expectations. Students in the treatment group and impure control group were then asked to respond to three distinct information elicitation tasks. Following this, students in the treatment group were provided with information on average earnings for the different tracks and college admission probabilities. Then students in the treatment group and the impure control group were asked to respond to the same expectation questions a second time. This design follows Wiswall and Zafar (2015)'s design for college choice. The survey ended by collecting information about (hypothetical) time and (paid) risk preferences.

3.3 Measures of expectations

The questions regarding self-beliefs were questions about educational outcomes, i.e., the students' chances of ending their education with junior high school, going to a vocational school or apprenticeship, dropping out of senior high school, finishing senior high school, dropping out of college, and finishing college together with a curriculum track. These questions also included predicted probabilities of working full-time at a job related to a specific major and earnings after finishing schooling, at ages 30 and 50. We also asked for their estimated probabilities of earning at least 50,000 Nigerian Naira (N50,000), N100,000, and N200,000. Similar questions were asked regarding their beliefs about the population, with reference to a typical student. Figure 2 provides the wording of one of the self-belief elicitation tasks regarding educational attainment.

What are the chances that you will	**Number**
go to art class?	
go to science class?	
go to commercial/social science class?	
go to vocational school after JSS3?	
drop out of school after JSS3?	
TOTAL: THE TOTAL SHOULD ADD UP TO 100	

Figure 2: Example of self-belief elicitation

The information treatment included statistics about the earnings and labor supply in Nigeria and population-level college acceptance rates and college choices. This information came from the Joint Admissions and Matriculation Board (JAMB) and Stutern (2018).[13] To the best of our knowledge, this is the most up-to-date information on admissions, graduation, and earnings for recent graduates in Nigeria.

Section 3.5 presents the information provided in the study. The instrument used in this study was a simplified version of the approach of Wiswall and Zafar (2015), who used this kind of information to estimate human capital accumulation models; see Haaland et al. (forthcoming) for a review of the approach. Importantly for us, the belief information we collect allows testing if students update their beliefs when the information is provided. Verifying that the students update their beliefs is a necessary condition for the rational use of information.

3.4 Elicitation of willingness to pay for information

We elicited the WTP for different pieces of information using a multiple price list (MPL) that is a discretization of the Becker-DeGroot-Marshack (BDM) incentive-compatible mechanism. The MPL included prices from N0 to N200 in increments of N25. Participants were asked to respond to 3 MPLs. The first one asked for their WTP for information on college admissions, the second asked for their WTP for information on wages by curriculum track, and the third MPL asked for their WTP for both pieces of information. This was done to test for differences in WTP for the different types of information and to check for adherence to the law of demand: more information should be valued (weakly) more.

To embed the WTP elicitation in the information RCT, we drew prices from the binary set $\{0, 250\}$. These draws were fixed at the classroom/school level in order to avoid the expected spillover effects if randomization were done at the individual level. As a consequence, either all of the students in a school were assigned to the information treatment group or they were all were assigned to a no-information group. Since we could not ask students to pay for information with their out-of-pocket money, we provided all the students in our study with N200 that they could use in different experimental tasks, including the WTP tasks. They were told that payments would be calculated based on their choices in one of the tasks chosen at random.[14] Providing subjects with money is consistent with common practice

[13]A total of 5,219 Nigerian graduates who graduated during the years 2013–2017 completed the survey. The data collection took place between February 8 and May 15, 2018. The survey was hosted using Google Forms, and Stutern.com recruited respondents via email and social media sites. To account for graduates in marginalized locations, an offline version of the survey was conducted in five states (Edo, Enugu, Ibadan, Imo, and Kaduna).

[14]In practice, one of the three MPLs was chosen at random to determine the cost of information. If the

in experimental economics. Note, however, that this imposed a budget constraint on the subjects. The amount that we provided to each student is about 1.1% of Nigeria's minimum wage and enough to cover a student lunch. Our study balances the need for the salience of payoffs and the risk associated with transferring money to minors.

Previous research using the BDM shows that both the distribution of prices (Mazar, Koszegi and Ariely, 2014) and the upper bound of the distribution of prices (Bohm, Linden and Sonnegard, 1997) can affect elicited valuations. To avoid these issues in our experiment, we indicated to participants that the prices could take values as low as N0 and higher than N200. However, we cannot test whether the set of prices that was provided altered the average valuation of information. Different approaches to eliciting valuations that would further minimize these issues (e.g., Allcott, Braghieri, Eichmeyer and Gentzkow, 2020; Mosquera, Odunowo, McNamara, Guo and Petrie, 2020) were not feasible because students needed to be provided with money to participate.

Whenever price randomization is implemented at the individual level, instead of at the cluster level as in our experiment, it is possible to disentangle the effect on behavior of the treatment status and the price paid (see Berry et al., 2020). That is, it allows to identify screening and sunk-cost effects separately. Sunk-cost bias is unlikely in our experiment since students paid a price of zero for information. Also, as we will show later, the behavior of those not receiving information is not affected by their WTP being elicited. We note that our study does not have treatment where information is given and WTP is not elicited. This treatment would allow to test if elicitation itself has an added effect on behavior beyond information. Given that Berry et al. (2020) find no evidence of sunk-cost effects in their experiment, we interpret our results as mainly identifying screening effects. Finally, students might value information even if it does not affect their behavior. This might be due to risk attitudes (e.g. Kreps and Porteus, 1978) or curiosity. Our results depend on the assumption that compliers with our treatment are not more intrinsically curious or differ in their preference for early resolution of uncertainty. While these alternative explanation are testable, they are beyond the scope of our study.

3.5 Randomization and implementation

Information intervention: The intervention provided information to students in randomly selected schools. Students in the treatment schools received information about average wages, the percent working full time, and the percent earning more than N50,000 and N100,000 for

price was $0, we provided all pieces of information. We did this because we would not be able to detect the effects of different sets of information since the treatments were assigned at the school level to avoid potential contamination.

the different curricula/tracks, as well as the proportion of males (females) who applied and were admitted to college across the three tracks. Figure 3 shows how the information was presented; it was done in this way to make it easy to understand. We consulted with the State's Ministry of Education to ensure that the information was accurate.

In the state of Oyo, there are over 600 public secondary schools. The sample of schools includes 115 coeducational junior high schools out of a universe of 133 in Ibadan city. These schools are evenly distributed in four areas of the city across the 5 local government areas in the city. The subset of schools randomized into the study were visited by enumerators carrying official letters from the government, our IRB approval letter, a study overview/permission letter to obtain permission to visit the school, and a proposed date and time to visit.

The pre-registered study planned for 32 schools: 16 in the treatment group and 16 in the control group (planned number of observations = 5,200).[15] We divided the control group so that half would be asked the belief questions and the other half would not. This resulted in a pure control group, an impure control group, and a treatment group, which allowed us to test for rationality as well as whether asking belief questions to those who did not receive information would affect their behavior. We planned for 16 schools in the treatment group, 8 schools in the pure control group, and 8 schools in the impure control group. The intervention was done on students who took their exams in June/July 2020 and needed to decide which track to choose by the beginning of senior high school in September 2020. The study was designed to detect a 5 percentage point change in dropout rates with a power of 0.8 at a 5% significance level.[16]

The first stage of the study was conducted between November 8 and December 3, 2019. After assessing potential implementation failures, we decided to recruit additional schools. We retain the 36 schools that completed the data collection. There were 18 schools in the treatment group (N = 1,925), 6 in the impure control group (N = 658), and 12 in the pure control group (N = 1,054). The main deviation from the original protocol was a reduction in the time that schools allowed for implementing the study.[17] The student examinations were delayed until August 2020 due to Covid-19. Figure 4 presents the design and implementation of the study in a graphic form.[18]

Table 1 presents basic statistics for the sample and its comparability across the three

[15] AEA RCT Registry number AEARCTR-0004839.

[16] We used administrative data on dropout rates to calculate the intraclass correlation (0.02).

[17] We reclassified three schools from the impure control to the control group since belief data were not collected due to time constraints.

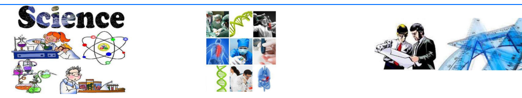
[18] We conducted a pilot test in 2018 to assess the feasibility of WTP elicitation techniques. We visited three schools for a total of 195 students. Two of these schools were single-sex and not included in this study. The WTP elicitations from the pilot are comparable to those in this study. Results are available from the authors upon request.

Figure 3: Information provided

Admissions information

Arts

- 15% of girls applied to Arts in university
- 10% of boys applied to Arts in university
- Of the girls that applied, 32% got admission
- Of the boys that applied, 30% got admission



Science

- 53% of girls applied to science in university
- 60% of boys applied to science in university
- Of the girls that applied, 26% got admission
- Of the boys that applied, 29% got admission

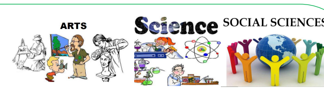
Social-sciences

- 33% of girls applied to social-science in university
- 30% of boys applied to social-science in university
- Of the girls that applied, 26% got admission
- Of the boys that applied, 27% got admission

Source: JAMB and CINFORES (2017)

Salary information

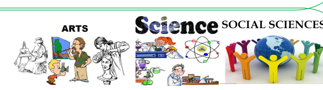
Female



This information is from a survey of university graduates in Nigeria. Among recent ***female*** graduates from a university with a Bachelor's degree in each of the above fields:

	Arts	Science	Social Sciences
• The percentage that are working full time is	**44.6%**	**45.6%**	**45.9%**
• The average monthly salary of those that are working full time is	**₦59,158**	**₦68,740**	**₦64,696**
• The percentage of those who are working full time that earn more than N50,000 monthly	**45.74%**	**56.2%**	**52.23%**
• The % of those who are working full time that earn more than N100,000 monthly is	**5.26%**	**7.85%**	**4.86%**

Male



This information is from a survey of university graduates in Nigeria. Among recent ***male*** graduates who just graduated from a university with a Bachelor's degree in each of the above fields:

	Arts	Science	Social Sciences
• The percentage that are working full time is	**48.54%**	**50.4%**	**49.35%**
• The average monthly salary of those that are working full time is	**₦53,100**	**₦77,849**	**₦71,924**
• The % of those who are working full time that earn more than N50,000 monthly is	**44%**	**60%**	**53.16%**
• The % of those who are working full time that earn more than N100,000 monthly is	**6%**	**25.4%**	**8.42%**

Among recent university graduates who received a Bachelor's degree in the above fields

The percentage of those who are women is	**65.52%**	**35.64%**	**56.62%**

Source: JAMB and CINFORES (2017)

groups.[19] The average age of participants was 14 years. There were slightly fewer females in the sample than males. About 40 percent declared themselves to be Christian and had roughly four siblings. Almost 80 percent of the students lived with both of their parents. Around 10 percent of the students declared that they had repeated at least one grade. The three groups are balanced in all of the variables we checked, except for the number of females. The pure control group had slightly more females than the information treatment group.

Information treatment (N = 1,925, 18 schools):

Background survey → Beliefs → WTP → Info → Beliefs → Preferences → Field outcomes

Impure control (N = 658, 6 schools):

Background survey → Beliefs → WTP → Beliefs → Preferences → Field outcomes

Pure control (N = 1,054, 12 schools):

Background survey → Beliefs → Preferences → Field outcomes

Figure 4: Study implementation

Notes: A subset of the expectation questions and preferences were collected from all participants. Preferences were collected at the end of the survey. Students were provided with N200 to be used in the willingness to pay (if applicable) and preference elicitations.

3.6 Outcomes collected

The study has two main outcome variables: self-beliefs about educational choices and actual educational choices. The self-beliefs were collected before and after the information intervention, and the educational outcomes were collected more than a year after the baseline survey. We present the outcomes below.

Probability of enrolling in school: During the baseline study, we asked students to report their estimated probability of enrolling in school beyond junior high school. We also asked for their estimated probability of choosing different senior high school and college tracks (arts, commerce, science, or vocational education) as well as dropping out of school after junior high school, during senior high school, and during college. These data were used to investigate the way students updated their self-beliefs upon receiving information.

Attendance/dropout rates: We obtained administrative data from schools on attendance and enrollment. In particular, we recorded whether a student took a junior high

[19] All classrooms in the last year of junior high school of the participating schools were visited by enumerators.

Variable	(1) Cntrl (0)	(2) Info (1)	(3) Imp. Cntrl (2)	(4) (1)-(0)	(5) (2)-(0)	(6) (1)-(2)
age	13.978	14.018	14.189	-0.041	0.083	0.007
	(1.591)	(1.398)	(1.549)	(0.141)	(0.183)	(0.203)
female	0.498	0.448	0.465	-0.037	-0.029	-0.029
	(0.500)	(0.497)	(0.499)	(0.028)	(0.050)	(0.039)
mother_edu_years	12.511	12.484	12.644	-0.079	0.443	-0.457
	(3.677)	(3.540)	(3.216)	(0.264)	(0.359)	(0.310)
christian	0.394	0.364	0.394	-0.030	0.021	-0.052
	(0.489)	(0.481)	(0.489)	(0.058)	(0.097)	(0.074)
siblings	3.985	3.892	3.930	-0.072	-0.075	-0.030
	(3.118)	(2.636)	(2.290)	(0.187)	(0.203)	(0.205)
twoparents	0.786	0.785	0.757	0.010	-0.014	0.007
	(0.411)	(0.411)	(0.429)	(0.019)	(0.023)	(0.026)
Grades	0.617	0.616	0.605	0.004	-0.005	0.001
	(0.144)	(0.141)	(0.135)	(0.011)	(0.016)	(0.013)
repeated	0.097	0.093	0.156	-0.027	0.057	-0.048
	(0.296)	(0.290)	(0.363)	(0.022)	(0.042)	(0.033)
Observations	1,054	1,925	658	3,637	1,528	2,767

Table 1: Characteristics by treatment group

Notes: Columns 1-3 show the average by treatment group. The numbers in parentheses are used to indicate pairwise comparisons between the groups in columns 4-6. Missing data items are replaced with the mean of the variable over the entire sample.

school exit exam, the grade obtained on the exam, and whether they registered for senior high school. For students to qualify for admission to a senior high school and higher education, nationwide examinations are held each year. Because exam scores determine a student's future educational choices, schools tend to stress memorization of facts rather than creative problem-solving. Students are required to pass at least six subjects to proceed to senior high school at the same institute or a different institute.

Curriculum choice: We obtained administrative data from schools on high school track choices. The curriculum tracks include arts, commerce, and science. All senior secondary students are required to study English, mathematics, one science, and one Nigerian language course.[20] The remaining subjects are electives and are selected based on the students' interest in either the sciences, the social sciences, or the arts. We note that it is too early in the study to know the actual career paths taken by students who registered in senior high school. We are also not aware of any data showing a correlation between curriculum choices in senior high school and career paths.

The state of Oyo does not have a centralized system with all students' data. In order to minimize potential biases due to non-response, we visited each high school in the study to collect information on registration. This allowed us to cross-check whether students changed schools after junior high school as well as their decisions. We also conducted a phone survey for all students who were not found in any of the school records. This procedure allowed us to determine the outcomes for over 95% of the sample. We do not find significant differences

[20]Science is not required for non-science tracks.

in missing data across the treatments.

4 Empirical strategy

To analyze the information treatment effects, we estimate the following empirical model:

$$y_i = \alpha + X_i\beta_i + \sum_k \gamma_k \mathbf{1}[Z_i = k] + \varepsilon_i,$$

where y_i denotes an outcome variable, e.g., not enrolling in senior high school, X_i is a set of covariates, e.g., household structure, age, etc., and $\mathbf{1}[A]$ is an indicator function that equals 1 when statement A is true, and 0 otherwise. The variable Z_i denotes the treatment condition for student i, and ε_i is an unobserved random component. The estimates assume that this random component is correlated at the level of the school.

We set $y_i = 1$ if a student does not enroll in senior high school, and 0 otherwise. Following Imbens and Angrist (1994), a response function is a mapping between the treatment assignment Z_i and outcome y_i, that is, $y_i : Z_i \to \{0, 1\}$. Non-compliers are those subjects who always choose either 0 or 1. Compliers are those whose decision to continue to senior high school depends on Z_i. If the treatment is binary, a monotonicity condition can be introduced wherein all compliers react to the treatment assignment, Z_i, similarly. When treatments are non-binary and unordered, a monotonicity condition can be imposed for each pair of treatment assignments (see Heckman and Pinto, 2018). We will show that the behavior in the impure and pure control groups are not statistically significantly different, and therefore the analysis can be conducted as if the treatment is binary.

We can use random assignment to treatment conditions to characterize compliers (see Marbach and Hangartner, 2020). Using the Law of Iterated Expectations, we have, for any variable X:

$$E[X] = E[X|NT]Pr(NT) + E[X|AT]Pr(AT) + E[X|C]Pr(C)$$

where NT denotes never-takers, i.e., those who continued to senior high school whether they they were provided with information or not, AT denotes always-takers, i.e., those who did not continue to senior high school whether they were provided with information or not, and C denotes compliers, i.e., those who switched from continuing to senior high school to not continuing when they were provided with information. We can estimate $E[X]$, $E[X|NT]$, and $E[X|AT]$ directly, since the data identify them. $E[X|C]$ can be estimated using the above equation. An alternative approach to identifying counterfactuals is given in Heckman and Pinto (2018). The counterfactual characteristic of compliers can be estimated by regressing $X_i\mathbf{1}[y_i = 1]$ on y_i using Z_i as an instrument and that of non-compliers by regressing $X_i\mathbf{1}[y_i = 0]$ on $1 - y_i$ using Z_i as an instrument.

5 Results

5.1 The demand for information

This section describes the demand for information corresponding to the 2,583 subjects in the information treatment and impure control conditions. Eighty-three percent of the answers to the MPLs have no switch-backs. This is comparable to the 95% consistency rate in Fuster et al. (forthcoming) and the 98% consistency rate in Allcott et al. (2020), who both use only one MPL per subject. Eighty percent of the subjects are consistent on each of the three MPLs. Preferences for information are monotonic if $WTP_{\text{college \& wage info}} \geq max\{WTP_{\text{college info}}, WTP_{\text{wage info}}\}$. Seventy percent of the subjects satisfy monotonicity.

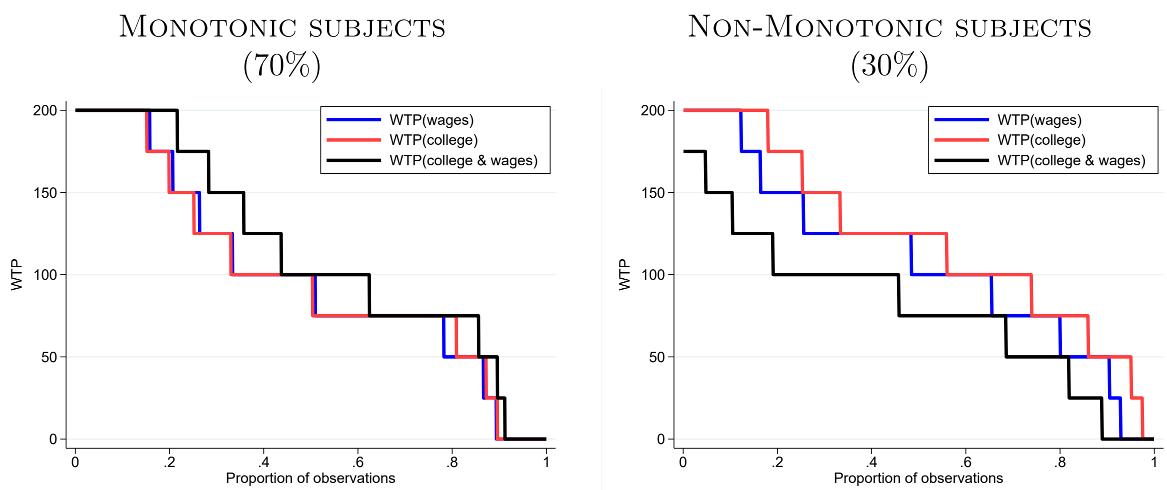


Figure 5: The demand for information

Notes: A subject is classified as having monotonic preferences if her valuation of all pieces of information is the (weakly) largest. WTP is defined as the largest amount of money a subject would pay for information. The WTP of subjects who have switch-backs is calculated as the WTP of the closest pattern, in monetary terms, of behavior without switch-backs. There were 2,389 usable observations of WTP for both pieces of information, 2,409 for salary information, and 2,435 for admissions information.

Figure 5 presents the demand curves for the monotonic and non-monotonic subjects. The x-axis in each panel gives the percent of subjects whose WTP is less than the number on the y-axis. We define a subject's WTP as the maximum price for which a subject would be willing to purchase the information. In the case where a subject has a switch-back, we use the WTP that minimizes the absolute distance to actual choices weighted by the price at which decisions are reversed.

In order to make our results comparable to those reported for developed countries, here we calculate the WTP in US dollars. The average WTP is \$0.25, which is equivalent to 0.6% of the monthly minimum wage (\$43.4), or 0.2% of the monthly wage of liberal arts graduates (\$143.6). For comparison, Allcott and Kessler (2019) found that subjects were willing to pay \$3 on average for home energy reports, and Fuster et al. (forthcoming) found that people were willing to pay \$4 for information on home prices. This is roughly equivalent to 0.2% of the monthly salary of \$2160 (\$12 per hour).

Table 2 presents the relationship between WTP and individual covariates from the rich set of baseline variables we collected. We include questions from Castillo et al. (2019), who study dropout rates in a US sample, as well as proxy variables used in that study. We find that, with the exception of risk attitudes, there is little, or no consistent, correlation between WTP and observable covariates.[21] This is consistent with the WTP providing new information about students' decision-making. While these results do not exclude the possibility of omitted variables bias, it helps allay the concern that WTP is correlated with unobservable participants' characteristics rather than their demand for information itself.

We conclude this section by discussing the implication of observing a consistent correlation between risk tolerance and WTP for information. The regression implies that a participant investing all her endowment in the paid lotteries is willing to pay about 30% above the average WTP. As discussed in the theoretical framework, if the average participant is risk averse, this implies that WTP will be less responsive to expected changes in behavior. In the extreme, expected changes in behavior might have little impact on WTP. Importantly, the correlation between risk tolerance and WTP cautions against interpreting lack of a proportional relationship between changes in behavior and WTP as a failure of rational behavior. This confirms that qualitative nature of our test.

5.2 The use of information

In this section we describe how students use the provided information. We first discuss the extent to which students update their beliefs. Then we evaluate whether their career choices reflect a selection based on earnings. Finally, we discuss the effect of the provided information on beliefs about these career choices and the effect of the provided information on dropout rates.

[21] We find little correlation between WTP and other baseline data as well. Results are avaiable from the authors upon request.

	Non-monotonic	WTP by type of information		
		Admissions	Wages	Both
Female	-0.028	-1.455	3.758	-1.073
	(0.023)	(4.085)	(3.788)	(3.318)
Age	-0.008	0.948	1.208	1.956
	(0.011)	(1.156)	(1.296)	(1.226)
First born	-0.023	-5.344	-3.588	-3.903
	(0.022)	(4.124)	(3.969)	(3.680)
Two-parent household	0.011	2.012	2.503	2.498
	(0.020)	(2.509)	(3.317)	(3.030)
Mother's education level	0.009	1.023	1.418	0.447
	(0.016)	(2.181)	(3.224)	(2.444)
% invested in lotteries	0.001	0.343**	0.294*	0.288*
	(0.001)	(0.147)	(0.147)	(0.151)
Discount rate	-0.029	11.079	13.362	9.575
	(0.055)	(9.495)	(8.913)	(9.656)
Nobody at home helps with homework	0.021	-4.902	-2.055	-6.429
	(0.024)	(4.536)	(5.333)	(4.774)
Repeated a grade	0.010	-0.229	3.995	2.564
	(0.033)	(4.115)	(3.808)	(3.969)
Average grades	-0.030	5.085	24.099*	11.356
	(0.105)	(12.277)	(13.771)	(11.605)
Received a suspension	-0.013	-18.144*	-13.143	-14.493
	(0.078)	(8.856)	(7.813)	(10.255)
Has extra lessons	-0.018	-2.502	-2.210	-2.291
	(0.020)	(2.448)	(2.939)	(2.996)
Currently attends an apprenticeship	0.030	-0.685	-2.180	-4.433
	(0.036)	(4.566)	(3.603)	(3.993)
Best school subject is math	0.015	-2.218	1.726	-0.088
	(0.029)	(4.009)	(3.453)	(4.002)
I pay attention in class	-0.006	4.135	0.359	2.549
	(0.019)	(2.486)	(2.424)	(2.230)
I like being at school	0.027*	-1.871	1.000	-1.909
	(0.016)	(1.825)	(2.397)	(1.910)
I get in trouble at school	0.026**	2.024	1.617	1.187
	(0.011)	(1.206)	(1.324)	(1.388)
Observations	2100	2165	2142	2126
Adj R2	0.000	0.022	0.016	0.016

* $p<0.10$, ** $p<0.05$, *** $p<0.010$. Errors clustered at the school level. Estimates are based on cases with complete data. A subject is classified as having monotonic preferences if her valuation of all pieces of information is weakly largest. WTP is defined as the largest amount of money a subject would pay for information. WTP of subjects who have switch-backs is calculated as the WTP of the closest, in monetary terms, pattern of behavior without switch-backs.

Table 2: Correlates of the WTP for information

5.2.1 Information updating

We use a Bayesian learning model to evaluate information updating. In this model, a student has a normally distributed prior belief $prior \sim N(\mu_0, \sigma_0^2)$. The variance of the prior belief, σ_0^2, captures the uncertainty of this belief. A signal is drawn from the true distribution of the variable that is distributed: $signal \sim N(\mu, \sigma^2)$. A Bayesian agent will update her prior according to the following formula:

$$posterior = (1 - \alpha)prior + \alpha signal,$$

where $\alpha = \frac{\sigma_0^2}{\sigma_0^2 + \sigma^2}$. Rearranging terms, we obtain:

$$posterior - prior = \alpha(signal - prior).$$

This provides an empirical framework for assessing the confidence students place on their beliefs. More confident students will update information less or will have lower values of α.

Our design allows testing if the information updating is due to new information that has been received or due to a reversion to the mean. For instance, students might correct reported expectations after noticing that they made a mistake or simply upon reflection. We can address this issue by comparing information updating between the information treatment and impure control conditions.

Table 3 presents estimates of a regression of the change in the log of earnings beliefs[22] as a function of the difference between the log of the information provided on earnings and the log of the first set of earnings beliefs.[23] We interact this variable with an indicator of having received the information. We estimate these regressions for the subset of beliefs closest to the data provided to students. The parameters are identified because male and female students have different relevant information. The results are similar if the regression is pooled across all questions to provide additional variation. We observe that the net effect of information on beliefs ranges from 0.12 to 0.24. This is comparable to the estimates in Hjort et al. (2021), who found the effect of signals to be between 0.26 and 0.37 (see Table 3 of their paper). The estimates on information updating for those who did not receive any information at all suggest that the belief data in this population are measured with a significant amount of error.[24]

[22]We did not collect information on beliefs of being admitted to college, so we cannot conduct a similar analysis for these beliefs.

[23]We winsorized the data at 1% to avoid extreme reports. The log of the earnings beliefs is close to a normal distribution.

[24]Fuster et al. (forthcoming) also observe this phenomenon, although to a lesser extent, when analyzing the effect of information on beliefs about housing prices.

Table 4 reproduces this analysis on the set of probability beliefs. This provides a second test of the ability of students to process the information that is provided. We find that the probability beliefs are less responsive to information.[25] The parameters associated with the signal are a fraction of those estimated for beliefs about earnings. The estimation, however, reproduces the pattern consistent with measurement error that was found for measurement updating.

	Salary at 21			Salary at 30		
	Arts	Comm	Science	Arts	Comm	Science
(signal-prior)	0.559***	0.526***	0.496***	0.517***	0.448***	0.575***
	(0.040)	(0.039)	(0.038)	(0.037)	(0.039)	(0.036)
(signal-prior) × Info treatment	0.212***	0.198***	0.239***	0.234***	0.245***	0.126***
	(0.046)	(0.045)	(0.045)	(0.043)	(0.046)	(0.043)
Observations	2207	2206	2187	2228	2226	2202
Adj R2	0.364	0.340	0.340	0.369	0.308	0.348

* p<0.10, ** p<0.05, *** p<0.010

Table 3: Expectation updating

Notes: Salary at 21 for Arts refers to the answer to the question: "Imagine that you enrolled in the arts track and studied one of the arts courses at university, and tell me how much you think you would be paid monthly if you have just graduated and now work full time?" The labels for the other columns follow the same pattern. The dependent variable is the difference between the belief elicited the second time and the belief elicited the first time. Beliefs are winsorized at 1% and expressed in logs. "Info treatment" equals 1 if the subject was provided with information, and 0 otherwise.

	Pr(Salary at 30<50,000)			Pr(Salary at 30<100,000)		
	Arts	Comm	Science	Arts	Comm	Science
(signal-prior)	0.518***	0.614***	0.637***	0.603***	0.603***	0.487***
	(0.044)	(0.043)	(0.041)	(0.043)	(0.044)	(0.040)
(signal-prior) × Info treatment	0.146***	0.083*	0.065	0.180***	0.152***	0.205***
	(0.049)	(0.048)	(0.047)	(0.049)	(0.050)	(0.046)
Observations	2133	2115	2121	2115	2108	2105
Adj R2	0.324	0.345	0.369	0.391	0.364	0.334

* p<0.10, ** p<0.05, *** p<0.010

Table 4: Threshold probability updating

Notes: The first column for Arts refers to the answer to the question: "What is the percentage chance that if you were working full time you would earn at least N50,000 monthly if you graduated from the arts track at university and were 30 years old?" The labels of the other columns follow the same pattern. The dependent variable is the difference between the belief elicited the second time and the belief elicited the first time. "Info treatment" equals 1 if the subject was provided with information, and 0 otherwise.

[25] This could partly be due to the fact that probability beliefs are not necessarily distributed normally, and therefore the learning model is inadequate for these data.

5.2.2 The perceived returns of education

Information RCTs providing earnings estimates assume that this information is relevant to students. However, recent research shows that career choices are only partially driven by concerns about earnings (e.g. Arcidiacono et al., 2020; Delavande and Zafar, 2019; Wiswall and Zafar, 2018). It is therefore important to assess whether students' beliefs about their career choices correspond to their expected earnings.

Since the students were asked what they thought their earnings would be for alternative career choices, we can estimate what students think the returns of education are. We follow Arcidiacono et al. (2020)'s methodology for measuring perceived treatment effects. These effects can be calculated because students reported what their earnings would be for every possible alternative. This means that we can measure the perceived treatment effect of any possible combination of individual choices. The treatment effects of interest are the treatment effect on the treated, i.e., the treatment effect on those choosing a particular alternative, and the treatment effect on the untreated, i.e., the treatment effect for each possible alternative that is not chosen. In particular, a measure of the perceived treatment effect on the treated for a particular career choice is the weighted mean of earnings at, say, 30 relative to junior high school. The weight is the probability that a student would choose such a career path. Analogously, the perceived treatment effect on the untreated is the weighted mean of earning at, say, 30 relative to junior high school, where the weight is the probability of not choosing such a career path.[26]

Table 5 shows these estimates using beliefs on earnings at 30 for different career paths (senior high school, arts in college, commerce in college, and science in college) compared to ending education with junior high school.[27] The table presents separate estimates for the first set of elicited beliefs and the second set of elicited beliefs. We only use data from the information treatment group. First, we observe a clear ordering on the perceived returns of career choices. Senior high school is ranked lowest and science is ranked highest. Second, we observe that the perceived returns of education decreased significantly after the provision of information. For instance, the perceived returns of a career in a science field drop from N121,804 to N65,946. This is almost half of the initially perceived returns for this choice relative to junior high school. The same pattern repeats for other educational choices. Third, we observe that in most cases, estimates of the treatment effect on the treated are larger than estimates of the treatment effect on the untreated. This is consistent with selection based on earnings. Finally, we observe that students perceive that earnings in some careers that they have not chosen are large. Indeed, the difference between the maximum earnings a student

[26] We use probabilities of choosing or not choosing an alternative as a proxy for actual choices.

[27] We bound probabilities to be between 0.01 and 0.99 since some students' beliefs were exactly 0 or 1.

could obtain given her beliefs and the expected earnings according to her expected choices is N75,000 (median N30,000). This is consistent with significant perceived nonpecuniary benefits of the chosen careers or significant barriers to education.

Estimates prior to receiving information

	Treatment on the treated		Treatment on the untreated	
	mean	s.e.	mean	s.d
Senior HS	10,953	941	11,883	1,020
Arts	68,723	2,365	58,552	2,071
Commerce	86,969	2,719	79,803	2,804
Science	121,804	3,478	98,283	3,383

Estimates after receiving information

	Treatment on the treated		Treatment on the untreated	
	mean	s.e.	mean	s.d
Senior HS	23,629	1,889	25,640	1,903
Arts	31,025	2,040	21,620	1,963
Commerce	50,508	2,573	43,354	2,821
Science	65,946	3,254	56,537	3,160

Table 5: Perceived treatment effects of educational choices

Notes: Treatment effect of the treated = weighted mean of earnings at 30 relative to junior high school (weight = probability of choosing the career path). Treatment effect on the untreated = weighted mean of earnings at 30 relative to junior high school (weight = probability of not choosing the career path). Calculations are based on data from the information treatment group ($N = 1925$). The top panel uses data from the first round of belief elicitation and the bottom panel uses data from the second round of belief elicitation. Earnings expectations are winsorized at 1%. Probability beliefs are bounded between 0.01 and 0.99.

5.3 The effect of information

The analysis so far shows that information is used for updates and that career choices reflect earnings considerations. We now discuss the effect of information on perceived future decisions and actual decisions.

To measure the effect of information on career choices we follow Wiswall and Zafar (2015), who estimate changes in beliefs about career choices as a function of changes in relative earnings. Specifically, let $\pi_{k,i}$ be student i's belief that she will choose career k. We can define the log-odds of choosing k over $\tilde{k}$ as $ln\pi_{k,i} - ln\pi_{\tilde{k},i}$. Let $w_{k,i}$ be students i's belief about the earnings associated with career k. We can define the relative earnings with respect to option $\tilde{k}$ as $ln\pi_{k,i} - ln\pi_{\tilde{k},i}$.

A simple regression of $ln\pi_{k,i} - ln\pi_{\tilde{k},i}$ on $ln\pi_{k,i} - ln\pi_{\tilde{k},i}$ is likely to be biased, since career preferences and abilities are likely not independent. However, we can take advantage of the fact that the information intervention did change expectations to estimate the following regression:

$$(ln\pi'_{k,i} - ln\pi'_{\tilde{k},i}) - (ln\pi_{k,i} - ln\pi_{\tilde{k},i}) = \beta_0 + \beta_1[(lnw'_{k,i} - lnw'_{\tilde{k},i}) - (lnw_{k,i} - lnw_{\tilde{k},i})] + \nu_k + \Delta\epsilon_{k,i},$$

where $'$ indicates beliefs elicited after the information is provided, ν_k is a fixed effect for choice k, and $\Delta\epsilon_{k,i}$ captures changes that are uncorrelated with beliefs about earnings. The first-difference regression eliminates unobservable differences across students.

Table 6 provides estimates of this regression using data from the information treatment group.[28] We use beliefs about earnings at 30 years of age since these are closest to the information provided. We provide estimates using both all of the subjects in the information treatment group and only the subjects in the information treatment group whose WTP was monotonic on the amount of information acquired. We find a moderate estimate of career choice elasticity. The elasticity is 0.06 for the overall population and 0.10 for the population with monotone preferences. The estimates increase to 0.165 and 0.252 when we correct for measurement error.[29] For comparison, Wiswall and Zafar (2015)'s estimate for a sample of New York University college students is 0.275 (see Table 6 in their paper). Given that the perceived returns of continuing education after information is obtained are about half of the initially perceived returns, these estimates suggest a 2.5 percentage point decrease in each educational choice other than junior high school if preferences for these alternatives are uniformly distributed.[30] Equivalently, they predict a 2.5 percentage point increase in senior high school dropout rates.

The analysis so far indicates that students use the information that is provided in expected ways. Beliefs about earnings are updated if the information is provided, and beliefs about career paths are updated once the earnings beliefs are updated. Next, we look at the

[28]We use junior high school as the reference group and consider senior high school, arts, commerce, and science as alternatives. Wages are winsorized at 1% to eliminate extreme values. Beliefs about career choices are bounded between 0.01 and 0.99. The results are qualitatively similar if we use different bounding thresholds.

[29]To correct for measurement error, we calculate the reliability of changes in relative wages using data collected in the impure control group. Specifically, let x be the change in relative wages. We calculate the reliability coefficient as $\frac{Var(x|Control)}{Var(x|Information)}$, where $Var(x|Control)$ is the variance of x in the impure control condition and $Var(x|Information)$ is the variance of x in the information treatment condition.

[30]This calculation assumes that the pre-intervention probability of choosing junior high school is 10% and the probability for the other four alternatives is 22.5%.

	Accounts for measurement error?			
	No		Yes	
	All	Monotonic	All	Monotonic
Log of wage at 30 (rev.)	0.060**	0.101***	0.165*	0.252**
	(0.030)	(0.037)	(0.087)	(0.098)
Observations	4726	3227	4726	3227

* p<0.10, ** p<0.05, *** p<0.010

+ clustered at the individual level

Table 6: Changes in self-beliefs about career choices as a result of information provision

Notes: The dependent variable is the change in the log odds of choosing a career path between the first and second round of belief elicitation. The choices considered are senior high school, arts, commerce, and science. The baseline choice is junior high school. The explanatory variable is the change in relative earnings at 30 years of age between the first and second rounds of belief elicitation. The baseline comparison is junior high school. Dummies for each choice are included. The estimation excludes self-beliefs that add up to less than 95 or more than 105 in either the first or second round of belief elicitation. The results are less precise if these data are included, but the qualitative results are the same.

effect that the information intervention had on field outcomes.

Table 7 shows linear probability models for the effects of the group treatments on not pursuing senior high school. We find that students in the pure and impure control group are respectively 3.7 and 3.8 percentage points less likely to not pursue senior high school. This difference is not significant (p-value = 0.9442). The effect is 3.7 if we combine both control groups into one control group. The estimated effect is similar if we account for non-responses.[31] This effect is within the 90 percent confidence interval of the predictions using the estimates in Table 6. The estimated effect of information on education is large. The fraction of students not continuing to senior high school is about 9.4 percentage points in the control groups. We conclude that information has a significant effect on the reported and actual decisions of students in our study. Importantly, estimates using the elicited beliefs and the field outcomes are compatible.

5.4 The instrumental value of information

Section 2 shows that if students acquire information rationally, we should observe a relationship between the WTP for information and the size of the information treatment effect. Similarly, we should observe that those who react to the intervention (compliers) value the information more. We should also expect that willingness to pay for information is increasing

[31]The estimated treatment effect is 0.0343 (s.e. 0.0215) using an inverse probability weighting correction. We predict attrition using age, sex, mother's years of schooling, Christianity, average grades, and indicators for two-parent households and having repeated a grade.

	No return	No return
Impure control	-0.039	
	(0.030)	
Pure control	-0.037*	
	(0.020)	
Information treatment		0.038*
		(0.021)
Observations	3473	3473

* p<0.10, ** p<0.05, *** p<0.010

Standard errors in parentheses. Clustered at the school level.

Table 7: Treatment effect of dropping out of senior high school (pp.)

Notes: The dependent variable is equal to 1 if a student does not register for senior high school, and 0 otherwise. The estimates are marginal effects obtained from a Probit regression. Outcome data are available by request for 95.5% of the participants (3,473 out of 3,637).

in self-reported expected changes in behavior.

5.4.1 Beliefs

Equation (3) in Section 2 provides a variational representation of the value of information. As discussed there, this equation cannot be evaluated directly because we never observe counterfactual behavior with and without information.[32] Our survey, however, does provide those counterfactuals since they are elicited before and after the information is released. It stands to reason that WTP for information should be larger for those whose beliefs about career choices change the most.

Table 8 estimates the relationship between several measures of belief change of career choice on the willingness to pay for information. We measure changes in beliefs using the L1 norm, the L2 norm, and the Kullback-Leibler divergence. We find that both the L2 norm and the Kullback-Leibler divergence are positively related to the willingness to pay for information. This is consistent with the theoretical prediction that those who update their choices the most are willing to pay more for information.

Since the relationship between the value of information and changes in behavior depend on priors, individual preferences and risk attitudes, we take this result as encouraging.[33] Indeed, this makes clear that access to data on beliefs and expected choices as well as willingness

[32]Equation (3) in Section 2 generalizes to the case of more than two choices.

[33]Frankel and Kamenica (2019) show that no metric can be a valid measure of information. While the Kullback-Leibler is a valid measure of information, it is not clear it is appropriate in this context. As shown in Section 2 the value of information varies with expected changes in behavior under risk neutrality. We take these results as qualitative evidence that willingness to pay for information is related to expected changes in behavior. Their paper focus on ex-post measures of information.

to pay for information can be used to improve structural estimations of human capital accumulation models via equation (3) in Section 2. This is a topic of current research.[34]

	WTP (1)	WTP (2)	WTP (3)
L1	3.836		
	(3.489)		
L2		5.928*	
		(2.880)	
KL			2.576*
			(1.228)
Observations	1209	1209	1209
schools	18	18	18
R2	0.042	0.044	0.044

* p<0.10, ** p<0.05, *** p<0.010

+ clustered at the school level

Table 8: Belief changes and the value of information

Notes: The dependent variable is equal to WTP for all pieces of information. Each column estimates the effect of a measure in changes in prior and posterior beliefs about career choices. The definition of career choices are the same as in Table 6. The first column uses the L1 norm to measure changes in prior and posterior beliefs of career choices. The second column uses the L2 norm and the third column uses the Kullback-Leibler divergence. All regressions include a dummy for females, age, a dummy for a two-parent household, mother's education, average investment in the lotteries and discount rate. The sample is restricted to those receiving information and those whose beliefs answers add up to 100±5.

5.4.2 Field behavior

Table 9 tests whether the size of the information treatment effect varies with the WTP for information. The regression uses the WTP for both pieces of information as a moderator. Figure 6 graphically shows the implied information treatment effects of the model (using estimates from Table 9, column 2). As theoretically predicted, those who were more willing to pay for information reacted to the information treatment more strongly. While the effect on the treated is 3.7 percentage points, the effect on those willing to pay N200 for both pieces of information is almost 10 percentage points. Column (3) in the Table shows that the result is robust to the inclusion of additional moderating variables. Table 9 allows us to

[34]As shown in Section 2, WTP has a closed form solution if one is willing to make assumptions about the uncertainty faced by subjects. For instance, if one assume extreme value errors, the ex-ante gross payoff function has a known expression (McFadden, 1978). Importantly, since information structures are equivalent to mean preserving spreads of prior beliefs, this implies that those willing to pay more for information should also be more responsive to information. This implies that the elasticities estimated in Table 6 should be increasing in WTP. The belief data is consistent with this prediction. The elasticity of those with WTP above 100 is 0.36 (s.e. 0.16), and the elasticity of those WTP below 100 is 0.26 (s.e. 0.13).

calculate how much WTP is likely to increase due to an increase in expected behavior. We have that a 100% increase in WTP is associated with a 70% increase in the probability of discontinuing education. Inversely, a doubling of the probability of discontinuing education increases WTP for information by 143%.[35]

	No return		
	(1)	(2)	(3)
Info treatment	0.038*	-0.037	-0.008
	(0.021)	(0.030)	(0.060)
WTP/100		-0.047***	-0.062***
		(0.009)	(0.011)
WTP/100 × Info treatment		0.070***	0.087***
		(0.014)	(0.017)
Observations	3473	2279	2135

* p<0.10, ** p<0.05, *** p<0.010

Standard errors in parentheses. Clustered at the school level.

Table 9: Treatment effects and the value of information

Notes: The dependent variable is equal to 1 if a student does not register for senior high school and 0 otherwise. The estimates are marginal effects obtained from a Probit regression. The second column is estimated using data from the information treatment and impure control groups (24 instead of 32 schools). The third column includes a dummy for female, two-parent household, being suspended, repeating a grade, averages grades, average investment in the paid lotteries and their interactions with the information treatment. The estimates including additional covariates are qualitatively similar.

We can confirm this pattern by analyzing the characteristics of those responding to the information intervention (compliers). Figure 7 shows the WTP for each type of information for the different response types. We can see that compliers have a higher WTP for each type of information. Table 10 shows the characteristics of the different response types for a larger set of variables. Since the complier group is relatively small ($\sim$ 4%), a potential concern is that the results are due to chance. To test whether the differences are significant, we calculate the mean characteristics of compliers and non-compliers using 5,000 bootstrap samples. The penultimate column in the table shows the percentage of times the mean of a variable for the complier population was larger than the mean of the variable for the non-complier population. A two-sided test of significance at the 10% level corresponds to these proportions being less than 5% or above 95%. We find that the mean of the WTP is significantly larger for compliers for all measures except the WTP for salary information, which is only one-sided significant. Compliers are also more likely to be monotone. This adds to the evidence in favor of the instrumental value of information. We also find that

[35]We reject the hypothesis that the coefficient on the interaction term of information treatment and WTP is equal to 0.1 (p-value = <0.001).

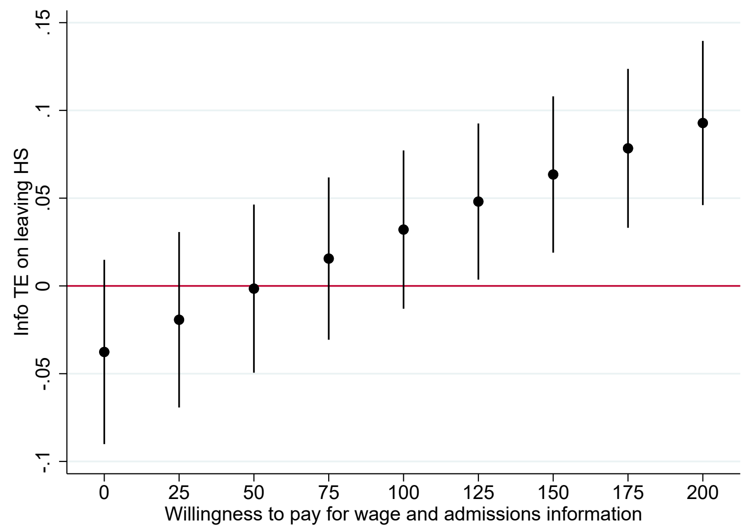


Figure 6: Treatment effects by WTP

Notes: The treatment effects for each value of WTP are calculated using the estimates in column (2) of Table 9. The dependent variable is equal to 1 if a student does not register for senior high school, and 0 otherwise.

compliers are more likely to have repeated a grade and are more likely to declare themselves Christian. Overall, we find that WTP is a good predictor of the effect of information on behavior. The last column of Table 10 test the hypothesis that the increase in WTP of compliers with respect to non-compliers is proportional to the increase in dropout rates. We cannot reject the hypothesis that this relation is proportional for information on salaries and college admission, but we can reject it for the WTP for both pieces of information (p-value = 0.09). We note that the WTP of compliers is likely to over-estimate the relationship between WTP and expected changes in behavior since it concentrates on those who ex-post change their behavior not those who would change behavior ex-ante.

6 Discussion

We discuss some limitations of our study. We cannot exclude the possibility that the most curious students are also the ones more likely to gain from discontinuing secondary education. However, we consider this case to be of similar interpretation since it is equivalent to minors valuing information that might lead to changes in behavior. We acknowledge it is important to determine the non-instrumental value of information since it affects the interpretation of null effects of information campaigns. If information has non-instrumental value, willingness to pay for it does not imply that treatment effects are not uniform in this case. We em-

	All	Always-Taker	Never-Taker	Compliers	Pr(Complier $\geq$ Non-Complier)	Pr($\frac{WTP_{Complier}}{WTP_{Non-Complier}} \leq \frac{Pr(dropout\|T=1)}{Pr(dropout\|T=0)}$
Percent of population	1.00	0.09	0.87	0.04		
	Mean of each characteristic					
Age	14.03	14.27	13.99	14.51	0.82	
Female	0.47	0.52	0.45	0.83	0.89	
Mother ed. (years)	12.54	13.04	12.51	11.81	0.38	
Christian	0.38	0.35	0.36	0.87	0.98	
Siblings (No.)	3.92	4.03	3.89	4.41	0.58	
Two-parent HH	0.78	0.78	0.79	0.49	0.09	
Grades	0.62	0.62	0.62	0.56	0.16	
Repeated a grade	0.10	0.11	0.09	0.41	0.99	
Discount factor	0.50	0.48	0.51	0.34	0.10	
Lottery investment	40.41	38.07	41.16	27.16	0.11	
Monotonic	0.69	0.60	0.69	1.07	0.95	
WTP admissions	108.31	105.36	105.73	167.91	0.99	0.22
WTP salaries	103.83	106.47	102.16	137.77	0.93	0.47
WTP both	105.60	101.34	102.44	192.75	0.99	0.09

Notes: This analysis is done at the student level. Row 1 shows the share of each compliance group in the sample. The shares of each compliance group are slightly different for the data on WTP since this was collected only for the non pure control conditions. The remaining rows show the means of each student or school characteristic across the different subgroups. The last column shows the probability that the mean characteristic of the complier group is larger than the mean characteristic of non-compliers. P-values are calculated using 5000 bootstrap samples. The mean characteristic uses the approach suggested by Marbach and Hangartner (2020).

Table 10: Characteristics of response types

phasize that we find little evidence that willingness to pay for information is correlated with attitudinal survey questions that are predictive of human capital accumulation.

Our study provides a limited amount of information, and it is possible that a larger set of variables might lead to larger or different changes in behavior. However, we have little knowledge of the kind of information adolescents' value. Our findings suggest that time spent searching could be used to study the use and value of information. In this revealed preference approach, distinguishing between intrinsic and instrumental value of information is paramount.

The main analysis is restricted to one field outcome: discontinuing education. The outcomes registered for the study also include senior high school graduation, choice of major, and post-secondary education. The framework developed in Section 2 extends to non-binary

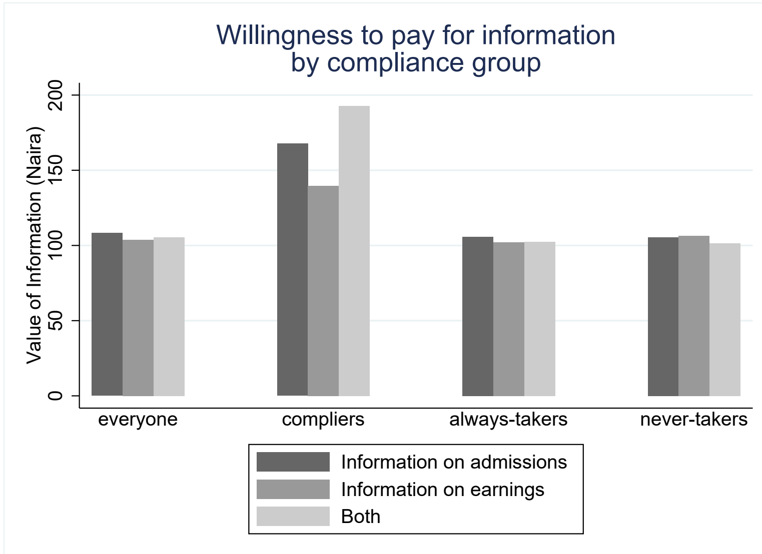


Figure 7: WTP by response type

Notes: Estimates of characteristics for each information treatment response type are calculated using Marbach and Hangartner (2020)'s approach. The graph includes data from the information treatment and impure control groups.

outcomes and therefore our results can be evaluated using a categorical variable. We do not implement this approach presently because those outcomes are not yet available. However, we show that willingness to pay for information is also correlated with changes in expectation of career choices using the belief elicitation data. This provides further support for our findings and shows that willingness to pay for information is an important source of identification in structural models using belief data.

Data collection does not include beliefs about the likelihood of being admitted to college. In retrospect, this is likely an important variable in the Nigerian case because of limited supply of post-secondary education alternatives.

7 Conclusions

This study investigates the instrumental value of information for adolescents. A simple test of rational information acquisition is implemented in a large sample of 14-year-old Nigerians who are deciding whether to continue on to senior high school or dropout of formal education. We find clear evidence that adolescents value information in accordance with economic theory. We cannot reject the hypothesis that WTP increases proportionally with changes in behavior. WTP for information is a strong predictor of who might benefit from an information intervention. This theoretical prediction is confirmed using belief data on career choices before and after receiving the information. Belief data provides the counterfactuals that the field intervention does not provide.

These results have implications for the way we model decisions made by the young, and in particular, regarding human capital accumulation. Heterogeneity of beliefs about returns of human capital accumulation might reflect endogenous information acquisition. WTP for information might reveal barriers to human capital accumulation. Our study highlights the importance of collecting expectation data in field experiments. These data can be useful in understanding the mechanisms leading to field outcomes as well help improve structural estimates of human capital accumulation models.

The study contributes to the knowledge of the decision-making of minors. Previous research has concentrated mainly on the study of their preferences and strategic behavior. Our study reveals adolescents' agency on human capital accumulation decisions and their ability to use information efficiently when the stakes are high. The study also contributes to the analysis of information interventions. We show that WTP for information is higher among compliers precisely when information has a uniform effect on subjects. This is because average response to information underestimates the effect of information on behavior. In the extreme, when different subjects respond to information in opposite ways, WTP might be positive while estimated treatment effects are nil. This situation reveals a failure of monotonicity rather than the uselessness of information. We show that this identification problem can be remedied using belief data. If WTP for information correlates with treatment responsiveness ex-ante, it can be used to design more efficient experiments and information interventions.

References

Allcott, H., Braghieri, L., Eichmeyer, S., Gentzkow, M., 2020. The Welfare Effects of Social Media. American Economic Review 110, 629–676.

Allcott, H., Kessler, J.B., 2019. The Welfare Effects of Nudges: A Case Study of Energy Use Social Comparisons. American Economic Journal-Applied Economics 11, 236–276.

Arcidiacono, P., Hotz, V.J., Maurel, A., Romano, T., 2020. Ex Ante Returns and Occupational Choice. Journal of Political Economy 128, 4475–4522.

Ayaita, A., Spengler, M., Nagengast, B., Trautwein, U., 2020. Field of study and earnings: The role of abilities, personality, and family background .

Bergman, P., Denning, J.T., Manoli, D., 2019. Is Information Enough? The Effect of Information about Education Tax Benefits on Student Outcomes. Journal of Policy Analysis and Management 38, 706+.

Berry, J., Fischer, G., Guiteras, R.P., 2020. Eliciting and Utilizing Willingness to Pay: Evidence from Field Trials in Northern Ghana. Journal of Political Economy 128, 1436–1473.

Bettinger, E.P., Long, B.T., Oreopoulos, P., Sanbonmatsu, L., 2012. The Role of Application Assistance and Information in College Decisions: Results from the H & R Block Fafsa Experiment. Quarterly Journal of Economics 127, 1205–1242.

Blackwell, D., 1951. Comparison of experiments, in: In Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability. volume Berkeley, CA: University of California Press of *Annual Review of Economics*, pp. 93–102.

Bohm, P., Linden, J., Sonnegard, J., 1997. Eliciting reservation prices: Becker-DeGroot-Marschak mechanisms vs. markets. Economic Journal 107, 1079–1089.

Brocas, I., Carrillo, J.D., 2021. Steps of Reasoning in Children and Adolescents. Journal of Political Economy 129, 2067–2111.

Brocas, I., Carrillo, J.D., Combs, T.D., Kodaverdian, N., 2019. The development of consistent decision-making across economic domains. Games and Economic Behavior 116, 217–240.

Bronchetti, E.T., Kessler, J.B., Magenheim, E.B., Taubinsky, D., Zwick, E., 2020. Is attention produced rationally?.

Busso, M., Dinkelman, T., Martinez, C.A., Romero, D., 2017. The effects of financial aid and returns information in selective and less selective schools: Experimental evidence from Chile. Labour Economics 45, 79–91.

Cabrales, A., Gossner, O., Serrano, R., 2017. A normalized value for information purchases. Journal of Economic Theory 170, 266–288.

Cadena, B.C., Keys, B.J., 2015. Human Capital and the Lifetime Costs of Impatience. American Economic Journal-Economic Policy 7, 126–153.

Caplin, A., 2016. Measuring and Modeling Attention, in: Aghion, P and Rey, H (Ed.), Annual Review of Economics. volume 8 of *Annual Review of Economics*, pp. 379–403.

Caplin, A., Csaba, D., Leahy, J., Nov, O., 2020. Rational inattention, competitive supply, and psychometrics. Quarterly Journal of Economics 135, 1681 – 1724.

Castillo, M., Ferraro, P.J., Jordan, J.L., Petrie, R., 2011. The today and tomorrow of kids: Time preferences and educational outcomes of children. Journal of Public Economics 95, 1377–1385.

Castillo, M., Jordan, J.L., Petrie, R., 2018. Children's rationality, risk attitudes and field behavior. European Economic Review 102, 62–81.

Castillo, M., Jordan, J.L., Petrie, R., 2019. Discount Rates of Children and High School Graduation. Economic Journal 129, 1153–1181.

Chambers, C.P., Liu, C., Rehbeck, J., 2020. Costly information acquisition. Journal of Economic Theory 186.

Chassang, S., Padro i Miquel, G., Snowberg, E., 2012. Selective Trials: A Principal-Agent Approach to Randomized Controlled Experiments. American Economic Review 102.

Chavas, J., 1993. Oon the demand for information. Economic Modelling 10, 398–407.

Chiong, K.X., Galichon, A., Shum, M., 2016. Duality in dynamic discrete-choice models. Quantitative Economics 7, 83–115.

Del Boca, D., Flinn, C.J., Verriest, E., Wiswall, M., 2019. Actors in the child development process.

Delavande, A., Zafar, B., 2019. University Choice: The Role of Expected Earnings, Nonpecuniary Outcomes, and Financial Constraints. Journal of Political Economy 127, 2343–2393.

Dynarski, S., Libassi, C.J., Michelmore, K., Owen, S., 2021. Closing the gap: The effect of reducing complexity and uncertainty in college pricing on the choices of low-income students. American Economic Review 111, 1721–1756.

Eeckhoudt, L., Godfroid, P., 2000. Risk aversion and the value of information. Journal of Economic Education 31, 382–388.

Fe, E., Gil, D., Prowse, D., forthcoming. Cognitive skills, strategic sophistication, and life outcomes. Journal of Political Economy .

Fehr, D., Fink, G., Jack, K., forthcoming. Poor and Rational: Decision-making under Scarcity. Journal of Political Economy .

Frankel, A., Kamenica, E., 2019. Quantifying information and uncertainty. American Economic Review 109, 3650–3680.

Fuster, A., Perez-Truglia, R., Wiederholt, M., Zafar, B., forthcoming. Expectations with Endogenous Information Acquisition: An Experimental Investigation. Review of Economic and Statistics .

Gould, J., 1974. Risk, stochastic preferences, and value of information. Journal of Economic Theory 8, 64–84.

Goux, D., Gurgand, M., Maurin, E., 2017. Adjusting Your Dreams? High School Plans and Dropout Behaviour. Economic Journal 127, 1025–1046.

Gurantz, O., Howell, J., Hurwitz, M., Larson, C., Pender, M., White, B., 2021. A National-Level Informational Experiment to Promote Enrollment in Selective Colleges. Journal of Policy Analysis and Management 40, 453+.

Haaland, I., Roth, C., Wohlfart, J., forthcoming. Designing Information Provision Experiments. Journal of Economic Literature .

Harbaugh, W., Krause, K., Berry, T., 2001. GARP for kids: On the development of rational choice behavior. American Economic Review 91, 1539–1545.

Hastings, J., Neilson, C.A., Zimmerman, S.D., 2015. The effects of earnings disclosure on college enrollment decisions.

Haushofer, J., Fehr, E., 2014. On the psychology of poverty. Science 344, 862–867.

Heckman, J.J., Pinto, R., 2018. Unordered Monotonicity. Econometrica 86, 1–35.

Hirshleifer, J., 1971. Private and social value of information and reward to inventive activity. American Economic Review 61, 561–574.

Hjort, J., Moreira, D., Rao, G., Santini, J.F., 2021. How Research Affects Policy: Experimental Evidence from 2,150 Brazilian Municipalities. American Economic Review 111, 1442–1480.

Hoxby, C.M., Turner, S., 2015. What High-Achieving Low-Income Students Know About College. American Economic Review 105, 514–517.

Hyman, J., 2020. Can Light-Touch College-Going Interventions Make a Difference? Evidence from a Statewide Experiment in Michigan. Journal of Policy Analysis and Management 39, 159+.

Imbens, G., Angrist, J., 1994. Identification and Estimation of Local Average Treatment Effects. Econometrica 62, 467–475.

Imbens, G., Rubin, D., 1997. Estimating outcome distributions for compliers in instrumental variables models. Review of Economic Studies 64, 555–574.

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

Kremer, M., Rao, G., Schilbach, F., 2019. Behavioral development economics, in: Bernheim, B., DellaVigna, S., Laibson, D. (Eds.), HANDBOOK OF BEHAVIORAL ECONOMICS - FOUNDATIONS AND APPLICATIONS 2. Handbooks in Economics, pp. 345–458.

Kreps, D., Porteus, E., 1978. Temporal Resolution of Uncertainty and Dynamic Choice Theory. Econometrica 46, 185–200.

Lai, F., Sadoulet, E., de Janvry, A., 2009. The adverse effects of parents' school selection errors on academic achievement: Evidence from the Beijing open enrollment program. Economics of Education Review 28, 485–496.

de Lara, M., Gossner, O., 2020. Payoffs-beliefs duality and the value of information. SIAM Journal on Optmization 30, 464–489.

List, J., Petrie, R., Samek, A., forthcoming. How experiments with children inform economics. Journal of Economic Literature 135, 1681 – 1724.

Loyalka, P., Liu, C., Song, Y., Yi, H., Huang, X., Wei, J., Zhang, L., Shi, Y., Chu, J., Rozelle, S., 2013. Can information and counseling help students from poor rural areas go to high school? Evidence from China. Journal of Comparative Economics 41, 1012–1025.

Lucas, A.M., Mbiti, I.M., 2012. The Determinants and Consequences of School Choice Errors in Kenya. American Economic Review 102, 283–288.

Marbach, M., Hangartner, D., 2020. Profiling Compliers and Noncompliers for Instrumental-Variable Analysis. Political Analysis 28, 435–444.

Mazar, N., Koszegi, B., Ariely, D., 2014. True Context-dependent Preferences? The Causes of Market-dependent Valuations. Journal of Behavioral Decision Making 27, 200–208.

McFadden, D., 1978. Modelling the Choice of Residential Location.. Elgar Reference Collection. Pioneering Papers of the Nobel Memorial Laureates in Economics, vol. 3. Cheltenham, U.K. and Northampton, Mass.: Elgar, Unlisted. pp. 409 – 430.

Mosquera, R., Odunowo, M., McNamara, T., Guo, X., Petrie, R., 2020. The economic effects of Facebook. Experimental Economics 23, 575–602.

Rust, J., 1994. Structural estimation of Markov decision processes. Handbook of Econometrics 4, 3081–3143.

Sørensen, J., Fosgerau, M., forthcoming. How McFadden met Rockafellar and learned to do more with less. Journal of Mathematical Economics .

Stutern, 2018. Nigerian graduate report: 2018. https://www.kaggle.com/stutern/nigerian-gradaute-report-2018.

Sutter, M., Kocher, M.G., Glaetzle-Ruetzler, D., Trautmann, S.T., 2013. Impatience and Uncertainty: Experimental Decisions Predict Adolescents' Field Behavior. American Economic Review 103, 510–531.

Wiswall, M., Zafar, B., 2015. Determinants of college major choice: Identification using an information experiment. The Review of Economic Studies 82, 791–824.

Wiswall, M., Zafar, B., 2018. Preference for the workplace, investment in human capital, and gender. Quarterly Journal of Economics 133, 457–507.

8 Appendix: Willingness to pay for information - Instructions

Let's play the following game! From time to time, we will stop during our explanation and allow you to ask questions. The purpose of this game is to help us understand how much value you place on certain information. Remember, if any part of the game makes you feel uncomfortable, you can talk to your school counselor or principal about it. They will be able to help you.

Consider that I want to sell information to you. The information can help you to make better decisions about choosing between science, arts or commercial class in SS1. This type of information is important because it can affect what you become in future. The information tells you the chances of getting admitted into higher institutions based on the type of class you choose or information on average salary for different professions and chances of working full time. For example, you will know the percentage of boys and girls that apply to study Arts, commercial and science courses and what percentage of them get admitted. Imagine you have N200, I would like to know if you will be willing to exchange the money for information. I will offer an amount of money as shown in the table below.

You will play three versions of the game, but only one will be used to pay you. Once you have made the decision for each of the three rounds, we will choose a number from one to three by randomly choosing from numbered balls in a bag. One of your classmates will be the one to pick the ball. The number on the ball chosen will determine which of one of the three versions of the game will be chosen to pay you. Next, I will present the class with 10 cards in a bag which represents prices drawn from N0 to more than N200, and someone in your class will be asked to pick one (the person will not know which card represents what price and I also do not know). The price on the card chosen will be used to determine if you get the information or not. If the price that is drawn from the bag is less than what you select as your value for the information, you will pay the drawn price and receive the information. If, however, the drawn price is strictly greater than what you choose as your valuation, then you do not get the information but keep your money. Think carefully about each decision.

Please listen carefully to the following example of this game: Dele is willing to buy the information at N125 and no more. So, he chooses "yes" for prices N0 - N125 (rows A-F) and chooses "no" for prices 150 and above (options G through I). We present him with a bag that has cards drawn from N0 to more than N200, he puts his hand in the bag and chooses price N0. Since the price N0, is always less than any amount he could have chosen, he will receive the information and keep his N200. And if N250 is chosen, he will not get the information since it is larger than any amount he could have paid.

Do you accept the price?	Yes	No
A: Price: N0 (means you pay: N0) You receive the information AND keep a payment of N200	X	
B: Price: N25 (means you pay: N25) You receive the information AND keep a payment of N175	X	
C: Price: N50 (means you pay: N50) You receive the information AND keep a payment of N150	X	
D: Price: N75 (means you pay: N75) You receive the information AND keep a payment of N125	X	
E: Price: N100 (means you pay: N100) You receive the information AND keep a payment of N100	X	
F: Price: N125 (means you pay: N125) You receive the information AND keep a payment of N75	X	
G: Price: N150 (means you pay: N150) You receive the information AND keep a payment of N50		X
H: Price: N175 (means you pay: N175) You receive the information AND keep a payment of N25		X
I: Price: N200 (means you pay: N200) You receive the information AND keep a payment of N0		X

Do you have any questions?

Consider that I want to sell information to you.

Information to sell: The chances of people getting admitted into university based on the type of class they choose in SS1. For example, you will know the percentage of boys and girls that apply to study Arts, commercial and science courses in university and what percentage of them get admitted. This information is from JAMB.

Imagine you have N200, I would like to know how much of the N200 you would like to use to buy the information. Think carefully on how much you value this information and respond by marking X in the relevant column.

Do you accept the price?	Yes	No
A: Price: N0 (means you pay: N0) You receive the information AND keep a payment of N200		
B: Price: N25 (means you pay: N25) You receive the information AND keep a payment of N175		
C: Price: N50 (means you pay: N50) You receive the information AND keep a payment of N150		
D: Price: N75 (means you pay: N75) You receive the information AND keep a payment of N125		
E: Price: N100 (means you pay: N100) You receive the information AND keep a payment of N100		
F: Price: N125 (means you pay: N125) You receive the information AND keep a payment of N75		
G: Price: N150 (means you pay: N150) You receive the information AND keep a payment of N50		
H: Price: N175 (means you pay: N175) You receive the information AND keep a payment of N25		
I: Price: N200 (means you pay: N200) You receive the information AND keep a payment of N0		

Again, consider that I want to sell information to you.

Information to sell: The average salary for people that go to different classes in SS1 and the chances that they are working full time. For example, you will know the average salary for boys and girls who studied different courses in university and are now working and the chances that they are working full time.

Imagine you have N200, I would like to know how much of the N200 you would like to use to buy the information. Think carefully on how much you value this this information and respond by marking X in the relevant column.

Do you accept the price?	Yes	No
A: Price: N0 (means you pay: N0) You receive the information AND keep a payment of N200		
B: Price: N25 (means you pay: N25) You receive the information AND keep a payment of N175		
C: Price: N50 (means you pay: N50) You receive the information AND keep a payment of N150		
D: Price: N75 (means you pay: N75) You receive the information AND keep a payment of N125		
E: Price: N100 (means you pay: N100) You receive the information AND keep a payment of N100		
F: Price: N125 (means you pay: N125) You receive the information AND keep a payment of N75		
G: Price: N150 (means you pay: N150) You receive the information AND keep a payment of N50		
H: Price: N175 (means you pay: N175) You receive the information AND keep a payment of N25		
I: Price: N200 (means you pay: N200) You receive the information AND keep a payment of N0		

Finally, consider that I want to sell information to you.

Information to sell: The chances of people getting admitted into university based on the type of class they choose in SS1 and the average salary for people that go to different classes in SS1 and the chances that they are working full time. For example, you will know the percentage of boys and girls that apply to study Arts, commercial and science courses in university and what percentage of them get admitted. This information is from JAMB. Also, you will know the average salary for boys and girls who studied different courses in university and are now working and the chances that they are working full time. It combines the two types of information I previously sold to you.

Imagine you have N200, I would like to know how much of the N200 you would like to use to buy the information. Think carefully on how much you value this this information and respond by marking X in the relevant column.

Do you accept the price?	Yes	No
A: Price: N0 (means you pay: N0) You receive the information AND keep a payment of N200		
B: Price: N25 (means you pay: N25) You receive the information AND keep a payment of N175		
C: Price: N50 (means you pay: N50) You receive the information AND keep a payment of N150		
D: Price: N75 (means you pay: N75) You receive the information AND keep a payment of N125		
E: Price: N100 (means you pay: N100) You receive the information AND keep a payment of N100		
F: Price: N125 (means you pay: N125) You receive the information AND keep a payment of N75		
G: Price: N150 (means you pay: N150) You receive the information AND keep a payment of N50		
H: Price: N175 (means you pay: N175) You receive the information AND keep a payment of N25		
I: Price: N200 (means you pay: N200) You receive the information AND keep a payment of N0		