Compliers, defiers and the evaluation of information interventions

Marco Castillo* Mofioluwasademi Ige†‡

December 2023

Abstract

When beliefs are heterogeneous, equal information might lead to unequal responses, and the intent-to-treat estimate of the effect of an information intervention might underestimate its true impact. We show that, under standard behavioral assumptions, testing for a uniform response to information is equivalent to testing for instrument validity using prior beliefs. We use this fact to analyze an information provision RCT in Ibadan, Nigeria, implemented when adolescents had to decide whether to stop at junior high school or continue to senior high school. We reject a uniform response to information: the intent-to-treat estimate of the effect of the intervention (13.1 versus 9.4 percent dropout rate) underestimates it by at least a factor of two. If we account for their choice of major, the intervention affected one in four students.

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 participants approval from Texas A&M University's IRB (IRB2018-0424D).

1 Introduction

At age 15 or earlier, about half of adolescents worldwide decide whether to continue their education and, if 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 on the effectiveness of information provision in altering levels of education is mixed.[2] Why this is so is unclear. Intrinsic to information interventions is that a focus solely on average behavior can lead to underestimation of its total effect if information induces some to continue schooling and some to discontinue schooling. Indeed, the value of information is highest when a decision-maker is indifferent between alternatives (see de Lara and Gossner, 2020), implying that information is most likely to have nonuniform effects when it is valued the most. In this paper, we show how to use prior beliefs and field outcomes to nonparametrically test for a nonuniform response to information, identify treatment effects, and test for the instrumental value of information.

To illustrate the problem, and using results from our study, consider that an information intervention increased the dropout rate from senior high school by four percentage points from a baseline of nine. Four percentage points is the intention-to-treat estimator. It is consistent with the assumption that all students who responded to information dropped out. However, the experiment is also consistent with twenty-one percent of students being affected by the information provided. This would be the case if thirteen percent who planned to continue their education decided to drop out and nine percent who planned to discontinue their education decided to continue. Since an experiment reveals only the marginal distribution of treatment status and effects, we cannot determine which of the two estimates is correct without additional assumptions or data.

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

[2] Jensen (2010) finds that information about the returns of education increases schooling. Other studies providing information about several aspects of educational investment report null as well as positive and negative results (see Ajayi, Friedman and Lucas, 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; Barr, Bird, Castleman and Skimmyhorn, 2022).

Under random assignment, if beliefs are updated in the direction of signals, prior beliefs can be used to identify nonuniform responses to information. Intuitively, if responses to information are nonuniform, we should observe distributional changes in prior beliefs across treatment conditions conditional on treatment status. For participants with ex-ante optimistic beliefs about the returns to education, we should observe relatively fewer of them continuing education in the treated group, when they receive "bad news" that the returns are lower than expected than in the control group that does not receive news. Analogously, for participants with ex-ante pessimistic beliefs on the returns to education, we should observe relatively more of them deciding to continue education in the treated group when they receive "good news" than in the control group.

We build on the literature on testable implications of monotonicity (Balke and Pearl, 1997; Heckman and Vytlacil, 2005) and observe that, under standard assumptions, if prior beliefs determine how information affects choices, then analogous testable implications must hold for prior beliefs as well. This means that instrument validity tests (Kitagawa, 2015; Sun, 2023; Mourifie and Wan, 2017) can be used to directly test for a nonuniform response to information using prior beliefs and field choices.[3] This is necessary to analyze heterogeneous treatment effects in information interventions. Since prior beliefs could be orthogonal to observed individual characteristics, an instrument validity test using other covariates is not equivalent to the test we propose.[4]

The main insight of the paper is that changes in the distribution of prior beliefs, conditional on treatment status, by treatment assignment can be used to measure flows in and out of the decision of interest, i.e., continuing education. Since the procedure relies on distributional changes of priors only, identification does not require knowing the signals agents observed.[5] In the case of multiple choices, knowledge of flows in and out of each choice identifies the treatment effect of information but might not be enough to identify their joint distribution across treatments. However, the joint distribution can be point- or set-identified by matrix completion. This approach,

[3]In the case of multiple choices and multiple treatments, this requires testing for unordered monotonicity (Heckman and Pinto, 2018; Sun, 2023).

[4]For instance, Frandsen, Lefgren and Leslie (2023) propose using covariates independent of treatment assignment to test for average monotonicity in the case of "judge fixed effect" design.

[5]In practice, a researcher wants to observe priors, signals, and posteriors to test the validity of the underlying assumptions (see Haaland, Roth and Wohlfart, 2023, for a review). We provide such evidence in this paper.

together with the Fréchet bounds, can be used to derive bounds on the effect of information. Identifying the sets of beliefs for which uniform response to treatment holds makes feasible the estimation of treatment effects for different complier groups (e.g., Abadie, 2003; Heckman and Pinto, 2018).

Identifying if responses to information are heterogeneous is also needed to test for its instrumental value. Intuitively, the willingness to pay (WTP) for information should be higher among those who would change their choice after receiving it.[6] This hypothesis is difficult to test because it requires knowing who would or thought would benefit from the information. In other words, we would like to know the WTP of compliers. This is a straightforward task if the response to information is uniform (e.g., Marbach and Hangartner, 2020), but complicated when the response is not. Our approach allows us to test this hypothesis in the latter case.[7]

We use this approach to analyze a field experiment in Ibadan, Nigeria, that randomized information about wages and college admission rates to over 3,600 14-year-olds deciding whether to continue to senior high school. In Nigeria, as in other areas of the world, college admission is very selective; thus, choosing to continue to senior high school to go to college can be risky if a student is unsuccessful.[8] We subsequently observe whether they did or did not continue to senior high school and which track they followed. 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. For a subgroup of participants, these beliefs were collected twice, before and after randomization into groups that received information.[9]

[6]We develop a test based on the intuition that willingness to pay (WTP) for information is a function of its anticipated influence on behavior (Blackwell, 1951; Hirshleifer, 1971; de Lara and Gossner, 2020). The assumptions needed for this result are given in Section 2.

[7]Incidentally, we provide a new way to test for the instrumental value of information in field settings not requiring observing how the demand for information changes as the stakes of the decision vary (e.g., Caplin, 2016; Caplin, Csaba, Leahy and Nov, 2020; Chambers, Liu and Rehbeck, 2020). This is of practical importance since varying the stakes of real-life decisions is not always feasible.

[8]According to Nigeria's Joint Admissions and Matriculation Board, one in three students who apply to college is admitted (see Section 3).

[9]Participants updated their beliefs on earnings at levels found for adults (Fuster, Perez-Truglia, Wiederholt and Zafar, 2022; Hjort, Moreira, Rao and Santini, 2021). Similarly, the participants' beliefs about their future choices were consistent with partial sorting based on earnings (Arcidiacono, Hotz, Maurel and Romano, 2020; Wiswall and Zafar, 2021). Finally, using the framework developed by Wiswall and Zafar (2015a), and consistent with related studies (Delavande and Zafar, 2019; Haaland et al., 2023), we find that participants' choice elasticity with respect to earnings was relatively low (about 17 percent). These estimates are commensurate with observed field behavior. This suggests that our belief data is informative.

The information intervention led to an average 3.7 percentage point decrease in education continuation rates one year later, from a baseline dropout rate of 9.4 percent. We find evidence against a uniform response to information. Firstly, under this hypothesis, we would expect the WTP of compliers in the Control and Treated conditions to be the same.[10] We strongly reject this hypothesis, thus casting doubt on a uniform effect of information. Secondly, we directly test for the hypothesis of a uniform response to information and find evidence against it. For this, we use instrument validity tests (Kitagawa, 2015; Sun, 2023; Mourifie and Wan, 2017) to assess the statistical significance of the distributional changes in prior beliefs conditional on continuing education.[11]

Having established a nonuniform response to information, we estimate treatment effects on educational choices. For the decision to continue education or not, we estimate that the number of participants affected by information is twice as large as the intention-to-treat estimate suggests. For each participant deciding to continue education due to the information provided, two students decided to discontinue it. The intervention also affected the distribution of fields of study. Accounting for changes in fields of study, the intervention affected more than one in four students. By recovering the joint distribution matrix of choices, we show that some students must have switched fields of study, not just decided to discontinue education.

Finally, we use the proposed method to test if those reacting to the information were ex-ante more willing to pay for it as theory predicts. We find that those who discontinued education due to the intervention valued information more, but those who continued education due to the intervention valued information less. This is consistent with some participants valuing information instrumentally, while others either avoid it or fail to anticipate their behavior. This pattern suggests caution in interpreting WTP for information as purely instrumental or extrapolating from studies that do not allow for a nonuniform response to situations where this is possible.

The paper's main contribution is to provide a way to nonparametrically test and quantify nonuniform responses to information by combining field and belief data. The approach uses as inputs conditional density functions of prior beliefs across treat-

[10]Under the assumption of monotonicity and random assignment, we can estimate the characteristics of compliers in both the control and treatment groups (e.g., Heckman and Pinto, 2018).

[11]We implement the test by constructing a measure of lifetime earnings from continuing and discontinuing education using elicited prior beliefs and measuring distributional changes conditional on choices.

ments. These estimates can recover the joint distribution of choices across treatment conditions and identify treatments in different complier groups. Rewriting the problem in this fashion makes it straightforward to test for nonuniform responses and to deal with multidimensional problems and situations where signals are not observed.[12] These are conditions frequent in the education setting we study.[13] The importance of prior beliefs in the analysis of information interventions is well known (e.g., Thornton, 2008; Jensen, 2010; Hoxby and Turner, 2015; Kendall, Nannicini and Trebbi, 2015; Cantoni, Yang, Yuchtman and Zhang, 2019; Bailey, Davila, Kuchler and Stroebel, 2019; Bursztyn, Gonzalez and Yanagizawa-Drott, 2020); however, we are not aware of a formal test for a uniform response to information. A distinct advantage of the proposed approach is using prior beliefs as surrogates of yet-to-observe outcome variables for testing heterogeneous responses to information. Since rejecting a uniform response to information implies that the intent-to-treat estimator is a lower bound of the treatment effect, this procedure can warn policymakers early about the importance of information barriers in educational choices.

The second main contribution is empirical. We provide direct evidence that the effect of information on education choices can be significantly underestimated. We estimate that more than one in four students changed their decisions due to the intervention. As we show in the paper, underestimation can be due to the assumption of uniform treatment response and conditioning on too small a set of prior beliefs when decisions depend on multiple alternatives. We also provide direct evidence for and against the instrumental value of information. As we discuss in the paper, assuming uniform response to treatment would have biased our test of theory. This implies that our approach affords to uncover violations of theory in different subpopulations. A researcher assuming uniformity will derive conclusions on a subpopulation, perhaps not representative of those affected by the intervention.

Our study speaks to the design and analysis of experiments. We find there is value in exploring the reaction to information across a wide spectrum of prior beliefs. Our results advise waiting for a broader set of life outcomes to test whether the intervention

[12] Knowledge of the signals subjects observed is needed to implement split sample tests of heterogeneous responses to information in survey experiments (see Haaland et al., 2023).

[13] Initial beliefs affect treatment selection producing a non-separable treatment choice equation as discussed in Heckman, Urzua and Vytlacil (2006) section 7. Heterogeneity in beliefs in the educational context has been reported before (Jensen, 2010; Stinebrickner and Stinebrickner, 2014; Wiswall and Zafar, 2015b) and is confirmed in our data.

was beneficial or harmful when uncertain about its full effect on educational choices. If information aids self-selection, treatment effects might be detectable in the long run (e.g., earnings, life satisfaction), even if we observe small intent-to-treat effects on educational choices in the short run.[14] The analysis presented here would be consistent with larger intent-to-treat effects in the long run than in the short run. The study also highlights the importance of testing theoretical implications across different populations. While we expected symmetric results on the value of information for those continuing and discontinuing education, we found the opposite. This discovery would not have been possible had we only implemented an intervention to change beliefs in one direction.

Kirkeboen, Leuven and Mogstad (2016) show that, in decision contexts like ours, knowledge of the best-next option and access to as many instruments as alternatives can be used to identify treatment effects for different subpopulations. We show that measures of prior beliefs can have identifying power even with one binary instrument. Since knowledge of signals received by subjects is not strictly needed for the proposed approach, it might be useful when information varies randomly across identifiable groups, and prior beliefs are available. The ability of beliefs to aid identification ultimately depends on their relevance in decisions and data quality.

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 presents the main results. Section 5 concludes the paper. A series of appendices provide additional results.

2 Theoretical framework

2.1 Decision model

This section discusses the decision framework used in the analysis. Students choose from $K+1$ choices $0, 1, \dots, K$. The utility they derive from choice k can be decomposed into two additive parts: U_k+e_k, where U_k represents the lifetime expected utility associated with k, while e_k represents idiosyncratic shocks across choices. A student chooses option k^* if $U_{k^*} + e_{k^*} \geq max_k\{U_k + e_k\}$. We assume that $U_0 = e_0 = 0$. We define the "social surplus function" (McFadden, 1981; Sørensen and Fosgerau, 2022),

[14] See Manski (2007) for bounds on treatment effects under outcome optimization.

which is the expected utility obtained from the choice problem:

$$\mathcal{W}(U) = E[max_k\{U_k + e_k\} - max_k e_k|U] \tag{1}$$

where $U = (u_1, \ldots, u_K)$. We define the conditional choice correspondence, $\mathcal{P}(U)$, as the probability of choosing each option given U is consistent with maximization. $\mathcal{P}(U)$ is a correspondence because without further assumptions ties and different tie-breaking rules are possible. Sørensen and Fosgerau (2022) prove that $\mathcal{W}(U)$ is finite, convex and everywhere subdifferentiable, and its subdifferential coincides with the conditional choice probability correspondence, i.e. $\partial\mathcal{W}(\cdot|U) = \mathcal{P}(\cdot|U)$. Moreover, the conditional choice correspondence $\mathcal{P}(U)$ is cyclic monotone. Shi, Shum and Song (2018) show that if the distribution of e_k's is absolutely continuous with respect to the Lebesgue measure and independent of U, then $\mathcal{W}(U)$ is differentiable.

The implication of cyclic monotonicity can be seen most clearly in comparing the conditional choice probabilities given two possible vectors of expected utilities U and U'. In particular, cyclic monotonicity requires that, for all (U', U), we must have that $(\mathcal{P}(U') - \mathcal{P}(U))'(U' - U) \geq 0$.[15] In particular, if there are only two options $k = \{0, 1\}$ with $k = 0$ representing dropping out of school and $k = 1$ representing continuing education, cyclical monotonicity implies that if the relative utility gain from continuing education, U', is larger than U, then the observed probability of continuing education must weakly increase. The model's predictions depend on the assumptions of how information affects expected utilities. If we have a proxy of changes in expected utility, cyclical monotonicity can help identify the effects of information on behavior consistent with maximization. The following section discusses the needed assumptions to implement this approach using belief data as proxies of utilities.

This framework can be used to evaluate the value of information. For instance, let $U(s_j), j = 1, \ldots, M$ be the vector of expected utilities if signal s_j is received. Signal s_i is distributed according to a finite probability distribution π such that $E_\pi[U(s)] = U$.[16] For a given status quo U, we can define the willingness to pay for this information, WTP, as the solution to the following equation:

$$E_\pi[\mathcal{W}(U(s))] - \text{WTP} = \mathcal{W}(U) \tag{2}$$

[15] This condition must hold for any selection $p \in \mathcal{P}(U)$.

[16] Information structures can be represented as a distribution over a set of posterior distributions that average to the prior distribution.

The case in which $k \in \{0, 1\}$ illustrates the usefulness of this result. Differentiating equation (2), we obtain that $d\text{WTP}/dU = E_\pi[\mathcal{P}(U(s))] - \mathcal{P}(U)$, where U is the return to continuing education.[17] The slope of the WTP is positive for those who would ex-ante increase the chances of continuing education upon receipt of the information, and negative for those who would ex-ante decrease the chances of continuing education with information. The WTP is maximal among those who are ex-ante indifferent between the options and therefore have the most to gain from acquiring information (see de Lara and Gossner, 2020, for results in greater generality). This result is useful because it implies that if the information has instrumental value, then behavioral changes should be larger among those with higher WTP for information.

2.2 Identification

We build on the previous framework to develop a simple procedure to identify non-monotone responses to information. We let $D_i = 1$ be the decision of individual i to continue education and $D_i = 0$ be the decision to discontinue education. $Y_{i,1}$ denotes i's prior belief in lifetime earnings of continuing education and $Y_{i,0}$ denotes i's prior belief in lifetime earnings of discontinuing education.

Following our experiment, we denote the updated beliefs after receiving a signal S by $Y_i' = (Y_{i,0}', Y_{i,1}')$. We assume that beliefs are not updated if no new information is available, i.e. $Y_{i,j}' = Y_{j,i}, j = 0, 1$.[18] Finally, let Z_i equal 1 if i receives signal S and Z_i equal 0 if i receives no signal. We let variable V_i be a barrier to continuing education not observed by the researcher. We make the following assumption.

Assumption 1: (i) $sgn((Y_{i,1}' - Y_{i,0}') - (Y_{i,1} - Y_{i,0})) = sgn(S - Y_{i,1})$, (ii) $D_i = \mathbf{1}[\nu(Y_{i,1}' - Y_{i,0}') \geq V_i]$ for a non trivial increasing function $\nu(\cdot)$ of $Y_{i,1}' - Y_{i,0}'$, (iii) Z_i is jointly independent of $(Y_{i,0}, Y_{i,1}, V_i)$.

Assumption 1(i) states that beliefs on returns to education update in the direction of the signal. Expected returns to education are updated upwards if the prior belief of lifetime earnings of continuing education is below the signal. Expected returns to education are updated downwards if the prior belief of lifetime earnings of continuing education is above the signal. Assumption 1(i) is satisfied if prior beliefs and signals are distributed normal, and signals on returns to education do not affect the belief

[17] We assume here that $U(s) = U + \eta(s)$, where $\eta(s)$ are adjustments to priors given signal s.

[18] This assumption is made for exposition. Proposition 1 only requires that those receiving information react to it more strongly.

on earnings of discontinuing education too strongly. In this case, $(Y'_{i,1} - Y'_{i,0}) - (Y_{i,1} - Y_{i,0}) = \frac{\sigma_1^2 - \sigma_{1,0}}{\sigma_1^2 + \sigma_\varepsilon^2}(S - Y_{i,1})$ and $sgn(Y'_{i,1} - Y_{i,1}) = sgn(S - Y_{i,1})$ if $\sigma_1^2 - \sigma_{1,0} > 0$, where σ_i^2 is the variance of Y_i, $\sigma_{1,0}$ their covariance, and σ_ε the variance of the signal. Bayesian updating and normality are not necessary conditions for Assumption 1 to hold (see Benjamin, 2019). Assumption 1(ii) states that students self-select into education based on their expected returns (see Willis and Rosen, 1979).[19,20] Assumption 1(iii) states that priors and selection into education are independent of treatment assignment.

Proposition 1: Under Assumption 1, $Pr(D_i = 1, Y_{i,1}|Z_i = 1) \geq Pr(D_i = 1, Y_{i,1}|Z_i = 0)$ if $Y_{i,1} < S$ and $Pr(D_i = 1, Y_{i,1}|Z_i = 1) \leq Pr(D_i = 1, Y_{i,1}|Z_i = 0)$ if $Y_{i,1} > S$.

Proof: Fix prior $(Y_{i,0}, Y_{i,1}) = (y_{i,0}, y_{i,1})$. For those receiving the signal, i.e., $Z_i = 1$, Assumption 1(i) implies that $Y'_{i,1} - Y'_{i,0} \geq Y_{i,1} - Y_{i,0}$ if $Y_{i,1} < S$ and $Y'_{i,1} - Y'_{i,0} \leq Y_{i,1} - Y_{i,0}$ if $Y_{i,1} > S$. For those not receiving the signal, i.e., $Z_i = 0$, $Y'_{i,1} - Y'_{i,0} = Y_{i,1} - Y_{i,0}$. Assumption 1(ii) implies that $E[\mathbf{1}[\nu(Y'_{i,1} - Y'_{i,0}) \geq V_i]|Y_{i,0} = y_{i,0}, Y_{i,1} = y_{i,1}, Z_i = 1] \geq E[\mathbf{1}[\nu(Y'_{i,1} - Y'_{i,0}) \geq V_i]|Y_{i,0} = y_{i,0}, Y_{i,1} = y_{i,1}, Z_i = 0]$ if $y_{i,1} < S$, and $E[\mathbf{1}[\nu(Y'_{i,1} - Y'_{i,0}) \geq V_i]|Y_{i,0} = y_{i,0}, Y_{i,1} = y_{i,1}, Z_i = 1] \leq E[\mathbf{1}[\nu(Y'_{i,1} - Y'_{i,0}) \geq V_i]|Y_{i,0} = y_{i,0}, Y_{i,1} = y_{i,1}, Z_i = 0]$ if $y_{i,1} > S$. Assumption 1(iii) then implies that $E[\mathbf{1}[\nu(Y'_{i,1} - Y'_{i,0}) \geq V_i]|Y_{i,1} = y_{i,1}, Z_i = 1] \geq E[\mathbf{1}[\nu(Y'_{i,1} - Y'_{i,0}) \geq V_i]|Y_{i,1} = y_{i,1}, Z_i = 0]$ if $y_{i,1} < S$, and $E[\mathbf{1}[\nu(Y'_{i,1} - Y'_{i,0}) \geq V_i]|Y_{i,1} = y_{i,1}, Z_i = 1] \leq E[\mathbf{1}[\nu(Y'_{i,1} - Y'_{i,0}) \geq V_i]|Y_{i,1} = y_{i,1}, Z_i = 0]$ if $y_{i,1} > S$. To see this, note that the inequalities hold pointwise for all $Y_{i,0}$. Since the distribution of $Y_{i,0}$ conditional on $(Y_{i,1}, Z_i)$ is invariant due to random assignment, the result follows from the monotonicity of the integral operator. Those with pessimistic beliefs $(Y_{i,1} < S)$ reconsider discontinuing education, and those with optimistic beliefs $(Y_{i,1} > S)$ consider discontinuing education.

Proposition 1 makes a nonuniform response to information a testable hypothesis. If the response to information is not uniform, we should find that an instrument validity test (Kitagawa, 2015; Sun, 2023; Mourifie and Wan, 2017) using priors as

[19] Assumption 1(ii) is stronger than required for Proposition 1. The previous subsection shows that if preferences are separable, any conditional choice correspondence consistent with maximization will satisfy cyclic monotonicity. We do not need to assume a tie-breaking rule as in Assumption 1(ii). The crucial assumption is that students behave as if belief changes correspond to expected utility changes and beliefs update in the direction of the signal.

[20] Assumption 1(ii) can be written as a function of Z_i and V_i instead of $Y'_{i,1} - Y'_{i,0}$ and V_i if we allow for random coefficients or non-separability in V_i as in Heckman et al. (2006) (Section VII). However, we want to highlight that beliefs mediate treatment effects.

surrogates of outcomes is rejected. Instrument validity tests build on Balke and Pearl (1997)'s observation that, under the monotonicity assumption (Imbens and Angrist, 1994), the density functions of outcomes conditional on treatment status can be ordered by treatment assignment. We show that a similar pattern should be observed on covariates determining how agents react to treatment. If the decision to continue education depends on additional beliefs, e.g., earnings from alternative options, the appropriate test is equivalent to an instrument validity test with multi-valued outcomes (see Kitagawa, 2015, footnote 8). This is true whether an agent receives new information on other aspects relevant to a decision or not.

Proposition 1 also provides a method to estimate how many participants are affected by the information campaign. Let $q_d(Y_d) = f(Y_d, D = d|Z = 0)$ and $p_d(Y_d) = f(Y_d, D = d|Z = 1)$ be the joint densities of prior beliefs Y_d and $D = d$ given $Z = z$. Under Assumption 1, we can estimate the proportion of participants switching from discontinuing education to continuing education by $\int_{Y_1} max\{p_1(Y_1) - q_1(Y_1), 0\}dY_1$, and the proportion of participants switching from continuing education to discontinuing education as $\int_{Y_1} max\{q_1(Y_1) - p_1(Y_1), 0\}dY_1$. The first expression is the number of participants continuing education, and the second is the number of participants discontinuing education.[21] If decisions depend on multiple beliefs, this procedure provides a lower bound on the effect of information. This follows from the max function being a convex function and the definition of marginal densities via Jensen's inequality.

To illustrate the use of Proposition 1 to identify heterogeneous responses to information treatments, Figure 1 shows functions $q_1(Y_d)$ and $p_1(Y_d)$ using simulated data. We assume that $\bar{Y_1} = 11.34$, $\bar{Y_0} = 9.87$, $\rho(Y_0, Y_1) = 0.19$, $S = 10.8$, $\theta = 0.8$ and $Pr(D = 1|Y_1(S), Y_0) = (1 + exp(-1 - 5(Y_1' - Y_0')))^{-1}$.[22] Consistent with Proposition 1, $q_1(Y_1) \geq p_1(Y_1)$ if $Y_1 > S$ and $q_1(Y_1) \leq p_1(Y_1)$ if $Y_1 < S$. The average difference in dropout rates for those receiving signal S is 4 percentage points. However, the proportion deciding to stop education due to the signal is 9 percentage points and the proportion deciding to continue education due to the signal is 3 percentage points.

[21] Dahl, Huber and Mellace (2017) show that treatment effects and characteristics of compliers and defiers can be identified if a local monotonicity condition holds. Local monotonicity states that conditional on a potential outcome, compliers and defiers do not co-exist. The authors propose a method to identify different regions of the potential outcome distribution to estimate treatment effects on compliers and defiers and to characterize them. This condition is violated here since, for each set of beliefs, individuals can choose to continue or discontinue education.

[22] The values chosen are based on the data we collected and describe in the next sections.

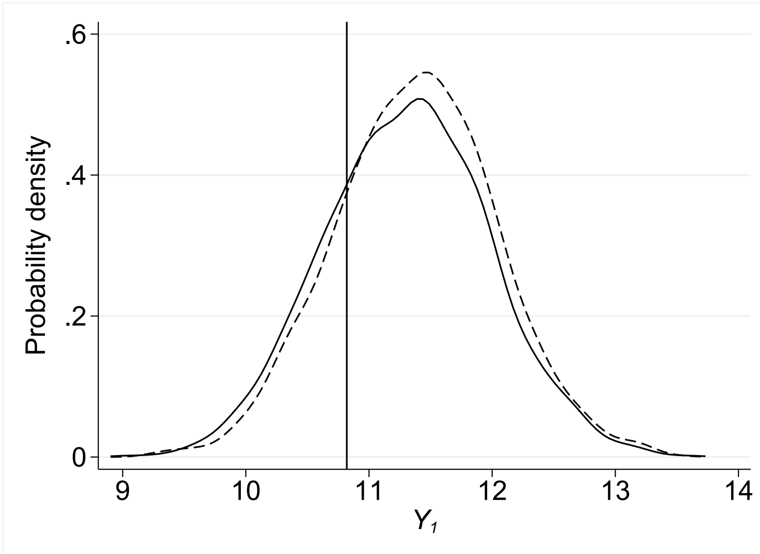


Figure 1: Simulated belief densities under Proposition 1. Prior beliefs on earnings of continuing education by treatment conditional on continuing education. Solid lines for treatment, dotted lines for control.

Figure 1 makes the analogy of the proposed approach to instrumental validity tests evident.

Multiple choices and signals. Information interventions might provide information on several educational alternatives. For instance, Jensen (2010) provided information on primary, secondary, and tertiary education, and our study provided information on alternative education tracks: Arts, Commerce, and Science. Information is likely to have a uniform effect if the signal of one of the alternative education options was higher than the subjects' beliefs and the signals of all other alternatives were lower than the subjects' beliefs. Information is likely to have a nonuniform effect whenever subjects have mixed signals. To generalize the approach, let $Y = (Y_0, Y_1, \ldots, Y_K)$ be a $K + 1$ vector of prior beliefs on earnings and define $q_k(Y) = f(Y, D = k|Z = 0)$ and $p_k(Y) = f(Y, D = k|Z = 1)$. $q_k(Y)$ and $p_k(Y)$ are the joint density functions of prior beliefs given an educational choice k for the control and treatment group. We can estimate the proportion of participants switching to k by $\int_Y max\{p_k(Y) - q_k(Y), 0\}dY$, and the proportion of participants switching away from k as $\int_Y max\{q_k(Y) - p_k(Y), 0\}dY$.

We provide a set of assumptions under which the proposed procedure has the desired interpretation. Let Y' be a vector of posterior beliefs, Y a vector of prior beliefs, S a set of signals, $D \in \{0, 1, \ldots, K\}$ a choice set, and $\eta_i, i = 1, 2$ vectors of parameters of unknown dimension. Vector η_1 is a vector that determines how posteriors affect educational choices and η_2 is a vector that determines how signals affect posteriors. We assume that $\eta_i, i = 1, 2$ are measurable.

Assumption 2: (i) $Y', Y, S, D, \eta_1, \eta_2 \perp\!\!\!\perp Z$, (ii) $Pr(D = k|Y', Y, S, \eta_1, \eta_2) = Pr(D = k|Y', \eta_1)$, (iii) $Y' = b(Y, S, \eta_2) = \{b_k(Y, S, \eta_2)\}_k$ where $b_k(\cdot)$ are functions of (Y, S, η_2).

Assumption 2(i) says that the relationship between beliefs, signals, and decisions is independent of treatment assignment. Assumption 2(ii) is an exclusion restriction implying that treatment effects are mediated only by their effect on (posterior) beliefs. Assumption 2(iii) says that knowledge of (Y, S, η_2) is enough to determine Y'.

Proposition 2: Under Assumption 2, the proportion of participants switching to k is $\int_Y max\{p_k(Y) - q_k(Y), 0\}dY$, and the proportion of participants switching away from k is $\int_Y max\{q_k(Y) - p_k(Y), 0\}dY$.

Proof: We have that $f(Y, D = k|Z = z) = Pr(D = k|Y, Z = z)f(Y|Z = z)$. We have that $Pr(D = k|Y', Y, S, \eta_1, \eta_2) = Pr(D = k|b(Y, S, \eta_2), \eta_1) = Pr(D = k|Y, Z = k, \eta_2, \eta_1)$. The first equality follows from assumptions 2(ii) and 2(iii). The second inequality follows from the fact that S is fixed for each $Z = z$. So, we have that $Pr(D = k|Y, Z = k) = \int_{\eta_1}\int_{\eta_2} Pr(D = k|Y, Z = k, \eta_2, \eta_1)dF_{\eta_1}dF_{\eta_1} = \int_{\eta_1}\int_{Y'=b(Y,S,\eta_2)} Pr(D = k|Y', \eta_1)dF_{\eta_1}dF_{\eta_1}$. Finally, under assumption 2(i), $f(Y|Z = 1) = f(Y|Z = 0)$. This implies that $max\{f(Y, D = k|Z = 1) - f(Y, D = k|Z = 1), 0\}$ is the probability density of those switching into k and $max\{f(Y, D = k|Z = 0) - f(Y, D = k|Z = 1), 0\}$ is the probability density of those switching out of k.

Proposition 2 shows that distributional changes can be used to measure flows in and out of educational choices even if students are inconsistent with the behavioral model in Proposition 1 or Section 2. The assumptions do not impose rational information processing. It imposes that posterior beliefs capture all the information needed to decide and that the experiment influences choice only through beliefs. Assumption 2(ii) is testable since it implies that choices conditional on posterior beliefs are independent of treatment assignment. This condition can be violated if the set of posterior beliefs is incomplete. Proposition 2 implies that testing for a nonuniform response to information with multiple choices is equivalent to testing for unordered monotonicity (Heckman and Pinto, 2018) using multiple prior beliefs (Sun, 2023).

We can use this approach to recover the joint distribution of treatment status in the treatment and control groups. Table 1 uses data from our experiment to show

how this can be done. For instance, the sum of $a_{1,2}+a_{1,3}$ is the proportion of students who switch from dropping out to continuing education, and the sum of $a_{2,1}+a_{3,1}$ is the proportion of students who switch from continuing education to discontinuing it. The sum of other off-diagonal terms can be similarly estimated. We can recover all the terms of the matrix by exploiting additional restrictions on probabilities. This approach only partially identifies this joint probability if four or more options exist. Additional moment conditions for identification can be derived from joint events like $D \in \{k, j\}$. This is the approach we follow in the paper by combining two educational tracks: Commerce and Science. Since the proposed approach identifies the sets of beliefs where monotonicity holds for different alternatives, we can potentially exploit this knowledge to recover the treatment effects on different complier groups (e.g., Abadie, 2003). We do this when we estimate the WTP for information among those who switched to discontinue education and those who switched to continue education. These results demonstrate the usefulness of belief data in identifying treatment effects.

Table 1: Finding the joint distribution of choices

	Treatment			
Control	Dropout	Arts	Comm & Sci	Total
Dropout	$a_{1,1}$	$a_{1,2}$	$a_{1,3}$	0.09
Liberal Arts	$a_{2,1}$	$a_{2,2}$	$a_{2,3}$	0.39
Commerce & Science	$a_{3,1}$	$a_{3,2}$	$a_{3,3}$	0.53
Total	0.13	0.36	0.51	1.00

3 Experimental design and implementation

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 of junior high school and can spend three years in senior high school. Students can choose between an academic or a vocational curriculum at the end of junior high school. Students wishing to proceed with the academic option in senior high school have three curriculum options: arts, commercial/social sciences, and science.

The vocational track also provides students with different areas for specialization. The diversity of the course curricula in senior high school allows students 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 occurred in Ibadan, the capital of Oyo, Nigeria, and Nigeria's third most populous city (3.2 million). The study was 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, and pure control. All students were asked to complete a baseline survey that collected basic demographic information, attitudes toward schooling, a proposed curriculum/track choice, career aspirations, and participant expectations. Students in the treatment and impure control groups were then asked to respond to three distinct information elicitation tasks. Following this, students in the treatment group were given information on average earnings for the different tracks and college admission probabilities. Then, students in the treatment and impure control groups were asked to respond to the same expectation questions a second time. This design follows Wiswall and Zafar (2015a)'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 student's 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. Table 2 provides the wording of one of the self-belief elicitation tasks regarding educational attainment.

Table 2: Example of self-belief elicitation

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	

The information treatment included statistics about Nigeria's earnings and labor supply and population-level college acceptance rates and choices. This information came from the Joint Admissions and Matriculation Board (JAMB) and Stutern (2018).[23] 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 (2015a), which used this kind of information to estimate human capital accumulation models; see Haaland et al. (2023) 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 necessary for rational use of information. It is also a direct way to test the behavioral assumptions introduced in Section 2.2.

[23] 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. An offline survey was conducted in five states (Edo, Enugu, Ibadan, Imo, and Kaduna) to account for graduates in marginalized locations.

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 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. 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 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 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[24] 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.[25] Providing participants with money is consistent with common practice in experimental economics.

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 price set altered the average valuation of information.[26]

[24]This amount is about 1.1% of Nigeria's minimum wage and enough to cover a student's lunch. Our study balances the need for the salience of payoffs and the risk associated with transferring money to minors.

[25]In practice, one of the three MPLs was chosen randomly to determine the cost of information. If the price was \$0, we provided all the information. We did this because we could not detect the effects of different sets of information since the treatments were assigned at the school level to avoid potential contamination.

[26]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.

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 the different curricula/tracks, as well as the proportion of males (females) who applied and were admitted to college across the three tracks. Figure 2 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 accurate information.

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. The subset of schools randomized into the study was 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).[27] We divided the control group so 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 and 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 in the pure control group, and 8 in the impure control group. The intervention was implemented with students who would take 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.[28]

The first stage of the study was conducted between November 8 and December 3, 2019. To account for potential attrition of schools from the experiment, 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

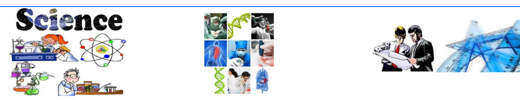
[27] AEA RCT Registry number AEARCTR-0004839.

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

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

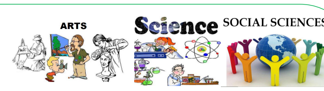
Social-sciences

- 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

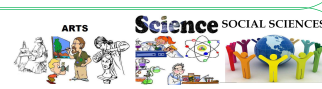
Female



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:

• 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:

• 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)

Figure 2: Information provided

impure control group ($N = 658$), and 12 in the pure control group ($N = 1{,}054$). The main deviations from the original protocol were a reduction in the time that schools allowed for implementing the study[29] and the delay in student examinations until August 2020 due to Covid-19. Figure 3 presents the design and implementation of the study in a graphic form.[30]

Table 3 presents basic statistics for the sample and its comparability across the three groups.[31] 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 they had repeated at least one grade. The three groups are balanced in all 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):

$\boxed{\text{Background survey}} \rightarrow \boxed{\text{Beliefs}} \rightarrow \boxed{\text{WTP}} \rightarrow \boxed{\text{Info}} \rightarrow \boxed{\text{Beliefs}} \rightarrow \boxed{\text{Preferences}} \rightarrow \boxed{\text{Field outcomes}}$

Impure control (N = 658, 6 schools):

$\boxed{\text{Background survey}} \rightarrow \boxed{\text{Beliefs}} \rightarrow \boxed{\text{WTP}} \rightarrow \boxed{\text{Beliefs}} \rightarrow \boxed{\text{Preferences}} \rightarrow \boxed{\text{Field outcomes}}$

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

$\boxed{\text{Background survey}} \rightarrow \boxed{\text{Beliefs}} \rightarrow \boxed{\text{Preferences}} \rightarrow \boxed{\text{Field outcomes}}$

Figure 3: 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.

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

[30] 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.

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

Table 3: Characteristics by treatment group

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's years of education	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)
No. of 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)
Two-parent household	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 a grade	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

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 variable's mean over the entire sample.

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.

Earning beliefs: We asked students about their beliefs on earnings for different educational choices. We also asked them to estimate the probability that earnings would exceed predetermined thresholds. Table A.1 tests if prior beliefs on earnings are balanced across the information and impure information treatments. Only earning beliefs for high school at ages 21 and 30 are larger in the impure control. In the Result section, we will return to the issue of balanced prior beliefs and how we deal with it.

Probability of enrolling in school: During the baseline study, we asked students to report their estimated probability of enrolling 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) and dropping out after junior high school, senior high school, and college. These data were used to investigate how students updated their self-beliefs upon receiving information.

Attendance/dropout rates: We obtained school administrative data on attendance and enrollment. In particular, we recorded whether a student took a junior high school exit exam, the grade obtained 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 must pass at least six subjects to proceed to senior high school at the same or different institutes.

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 must study English, mathematics, one science, and one Nigerian language course.[32] The remaining electives are selected based on the student's interest in the sciences, the social sciences, or the arts.

The state of Oyo does not have a centralized system with all students' data. 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 and their decisions. We also conducted a phone survey for all students not found in the school records. This procedure allowed us to determine the outcomes for over 95% of the sample. We do not find significant differences in missing data across the treatments.

4 Results

4.1 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: $signal \sim N(\mu, \sigma^2)$. A Bayesian agent will update her prior according to the following formula:

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

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

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

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

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 θ.

Figure A.1 shows the average prior by educational choice and age. Numbers 20, 30, and 50 refer to the age at which expected earnings are reported. For each educational choice, prior beliefs on earnings increase with age. Prior beliefs are also increasing in educational choice; earning beliefs are lowest for Junior High School and largest for those with a college degree in sciences. Arts, Commerce, and Science refer to earnings to different college degrees. Figure A.2 shows the density functions of the difference between the log of earning beliefs at 30 years of age and the log of the gender-specific signal provided. The densities correspond to the earnings if a major in Arts, Science, or Commerce is followed. We see a high degree of heterogeneity in beliefs and a majority of participants over-estimating earnings. This confirms that the necessary condition for a nonuniform response to treatment is satisfied in our setting.

Our design allows testing if 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 4 presents estimates of a regression of the change in the log of earnings beliefs 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.[33] We interact this variable with an indicator of having received the information and estimate these regressions for the subset of beliefs closest to the data provided to students (i.e., 21 and 30 years of age). 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 (Science at 30) to 0.24 (Commerce at 30). This is comparable to the

[33] We winsorized the data at 1% to avoid extreme reports. The log of the earnings beliefs is close to a normal distribution. We did not collect information on beliefs about being admitted to college, so we cannot conduct a similar analysis for these beliefs.

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).[34]

Table A.2 reproduces this analysis on the set of probability beliefs. This provides a second test of the ability of students to process the information provided. We find that beliefs on the probability of earning a certain salary are less responsive to information.[35] The parameters associated with the signal are a fraction of those estimated for beliefs about earnings. We confirm that participants update information consistent with Assumption 1(i).

Table 4: Expectation updating

	Salary at 21			Salary at 30		
	Arts	Comm	Science	Arts	Comm	Science
(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

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 participant was provided with information and 0 otherwise.

4.2 Treatment effects on dropout rates

So far, the analysis indicates that students use the information 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 effect of information intervention on field outcomes.

Table 5 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

[34]The estimates on information updating for those who did not receive any information suggest that the belief data in this population are measured with error. Fuster et al. (2022) also observe this phenomenon, although to a lesser extent, when analyzing the effect of information on beliefs about housing prices.

[35]This could partly be because probability beliefs are not necessarily distributed normally, and therefore the learning model is inadequate for these data.

control group are respectively 3.9 and 3.7 percentage points less likely not to 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.[36] The estimated effect is similar if we account for non-responses.[37] This effect is within the 90 percent confidence interval of the predictions using the estimates in Table **??**. The estimated effect of education information is large. The percentage of students not continuing to senior high school is 9 percentage points in the control groups. We conclude that information significantly affects students' reported and actual decisions in our study. Importantly, estimates using the elicited beliefs and the field outcomes are compatible.

Columns (3) and (4) of Table 5 test whether the size of the information treatment effect varies with the WTP for information. The regressions use the WTP for both pieces of information as a moderator. Those 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 (4) in the Table shows that the result is robust to including additional moderating variables.

4.3 Evidence against a uniform response to information

We divide the population between always-takers (AT), never-takers (NT), compliers (C), and defiers (F). Always takers drop out of school in the treatment and control conditions. Never-takers continue education in the treatment and control group. Since the treatment effect increased dropout rates, we assume compliers dropped out of school in the information treatment and remained in school in the control group. Defiers continued education in the information treatment and dropped out of school in the control group.

We start by describing the characteristics of response types under the assumption of monotonicity (see Table A.4). Compliers' willingness to pay for information (WTP)

[36]We implement the randomization test proposed by Canay, Romano and Shaikh (2017) to account for possible biases due to a small number of clusters. The test requires that treatment effects be estimated in each cluster, so we grouped adjacent clusters into one to implement the test. This reduced the number of clusters from 36 to 18. The estimated treatment effect is significant at 10 percent (p-value = 0.077). For robustness, we estimate the p-values for 1,000 random pairings of clusters. The average p-value is 0.077 (s.e. 0.033.)

[37]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.

Table 5: Treatment effects and the value of information continuing to senior high school

	(1)	(2)	(3)	(4)
Impure control	0.038			
	(0.028)			
Pure control	0.037			
	(0.020)			
Info treatment		-0.037	0.031	0.003
		(0.020)	(0.033)	(0.061)
WTP/100			0.038	0.049
			(0.007)	(0.011)
WTP/100 $\times$ Info treatment			-0.063	-0.075
			(0.014)	(0.017)
Constant	0.869	0.906	0.863	0.891
	(0.017)	(0.010)	(0.026)	(0.031)
Observations	3473	3473	2279	2135

Standard errors in parentheses. Clustered at the school level.

Notes: The dependent variable equals 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 households, being suspended, repeating a grade, average grades, average investment in the paid lotteries, and their interactions with the information treatment. The estimates, including additional covariates, are qualitatively similar.

significantly differs from those of always-takers and never-takers. However, under monotonicity, it is possible to identify the characteristics of compliers both in the Control and in the Treated group (e.g. Heckman and Pinto, 2018). If the condition holds, we would expect that the characteristics of compliers will be similar in both groups due to random assignment. Table 6 provides such a comparison. We observe that estimates of WTP are significantly higher for compliers in the Control group than in the Treated group. The significant difference in estimated WTP for compliers across treatment conditions and the violation of the bounds of WTP for information cast doubt on the monotonicity assumption.

To fully assess if these results are due to a non-uniform response to information, we follow the procedure suggested in Section 2.2. We now describe how we construct variable Y_1, the prior beliefs on life earnings from continuing education. The experiment collected prior beliefs on the earnings for ages $k =\{$at graduation, at 30 years

Table 6: Willingness to pay for information by compliers under monotonicity

	In Control	In Treatment	p-value	Combined
Admission & Earnings	345.80	132.71	0.01	239.26
Admission rates	292.33	132.21	0.02	212.27
Earnings	251.04	95.43	0.02	173.24

of age, at 50 years of age} corresponding to dropping out of senior high school, finishing senior high school, and graduating with a degree in arts, science or commerce. We construct a measure of earnings corresponding to continuing education, Y_1, as the natural log of the arithmetic mean of the winsorized earnings beliefs of finishing senior high school, graduating with a degree in arts, science, or commerce at graduation, 30 years of age, and 50 years of age. While we collected data on the likelihood that each one of these alternative education paths is taken and the probability of full employment by educational choice, the data is incomplete. We use an average of all earnings beliefs to reduce measurement error. As mentioned in Section 2.2, there is no exact correspondence between the signal provided in the experiment and prior beliefs. Therefore, we are agnostic about the exact threshold each participant uses. However, Proposition 1 should hold upon aggregating across individuals unless the distribution of individual threshold beliefs is sufficiently distinct from the signals. This measure of earnings is balanced across treatments.[38] We follow a similar procedure to construct a measure of Y_0, earnings if discontinuing education.

We implement the tests proposed by Kitagawa (2015) and its modification proposed by Sun (2023).[39] We should remark that these tests are based on the joint hypothesis that monotonicity holds for those who continue to senior high school and those who are discontinuing education. We implement the tests in the subsample of those continuing education since Proposition 1 applies to this subsample only.[40]

Figure 4 shows the pattern predicted by Proposition 1 using data from our experiment. Table 7 provides formal tests of instrument validity.[41] Table 7 also shows test

[38] Neither differences in means (p-value = 0.172) nor distributions (p-value = 0.205) are significant.

[39] We thank Zhenting Sun for providing code to implement these tests.

[40] As discussed in Section 2.2, the procedure can be modified to include prior beliefs conditional on discontinuing education.

[41] Sun (2023) suggests a modification of the test by Kitagawa (2015) that is more powerful.

results for measures of earning for different ages and the sub-sample with WTP for information above and below N100.[42] According to Section 2, the evidence against a uniform response to information must be stronger in this case. Finally, we implement the test using prior beliefs for continuing and discontinuing education.[43]

Table 7 shows that monotonicity cannot be rejected in the whole sample using a one-dimensional test. This is true if we use the aggregate measure of earnings or the measures disaggregated by age. We can reject monotonicity for several trimming constants for the subsample of subjects with a WTP above 100. This is consistent with not all subjects equally paying attention to information. It is also evidence that not all subjects respond to information uniformly. More importantly, Section 2.2 discusses that detecting a nonuniform response to information diminishes if decisions depend on multiple priors rather than only one. Therefore, we test for monotonicity using earning beliefs for continuing and discontinuing education. We find that monotonicity is strongly rejected when we account for both these earnings.[44] In sum, we reject the hypothesis that responses to information are uniform.

4.4 Response types

Section 2.2 provides a method to estimate the proportion of compliers and defiers by inspecting distributional changes in prior beliefs among those continuing education in the treatment and control conditions. Table 8 shows those probabilities estimated for the whole population. We obtain estimates by first re-balancing the sample to satisfy the conditions in Proposition 1.[45] Table 8 presents the mean estimates and their standard deviations. The estimates of those switching in and out of education differ from zero. The top panel presents estimates using prior beliefs on earnings if contin-

[42] The tests suggested by Mourifie and Wan (2017) are in Table A.3.

[43] We follow the suggestion by Kitagawa (2015), footnote 8, to extend the test to multidimensional outputs. In particular, we create a grid on each prior and create a class of rectangles over which the supremum is evaluated. We use a grid rather than the full support of the variables because the dimension of the problem makes it computationally infeasible. Table 7 shows results for the unidimensional test of both Kitagawa (2015) and Sun (2023) using the same grid as the one used for the two-dimensional test. The results are similar using a seventy equally spaced grid.

[44] Results using a reverse order, i.e., information increased education, produce identical results. Table A.3 confirms this result using Mourifie and Wan (2017) testing procedure.

[45] We binned beliefs and samples from the treated group to match the proportions in the control group. We use samples for which we find no significant difference in the distribution of prior beliefs. We then obtain kernel density estimates for these samples using Matlab's mvksdensity command. The results use Silverman's rule for bandwidths and Epanechnikov kernel. The results are robust using alternative kernel methods.

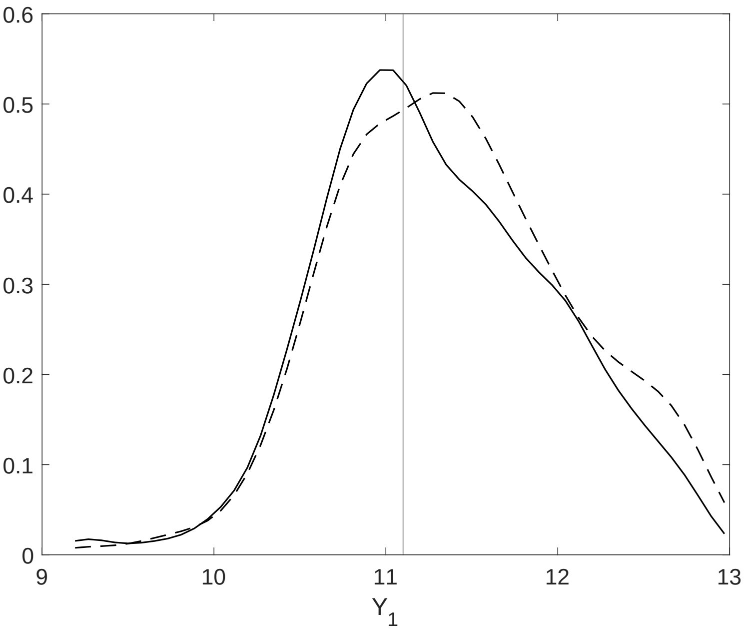


Figure 4: Kernel (Gaussian) density estimates. Prior beliefs on earnings of continuing education by treatment conditional on continuing education. Solid lines for treatment, dotted lines for control. Densities are re-weighted to account for the probability of being assigned to treatment.

Table 7: Monotonicity test using prior beliefs

	Trimming constant (ξ)									
	0.07	0.1	0.13	0.16	0.19	0.22	0.25	0.28	0.3	1
	One-dimensional test: Y_1									
All (Sun)	0.504	0.302	0.248	0.361	0.332	0.246	0.191	0.157	0.209	0.194
All (Kitagawa)	0.517	0.314	0.267	0.375	0.345	0.255	0.202	0.167	0.217	0.218
All (Sun)[1]	0.262	0.302	0.303	0.265	0.229	0.203	0.180	0.234	0.264	0.136
All (Kitagawa)[1]	0.287	0.325	0.345	0.306	0.261	0.229	0.205	0.264	0.292	0.162
Earnings at graduation (Sun)	0.051	0.104	0.159	0.123	0.210	0.406	0.588	0.627	0.684	0.773
Earnings at graduation (Kitagawa)	0.057	0.120	0.179	0.136	0.228	0.433	0.622	0.666	0.732	0.802
Earnings at 30 (Sun)	0.622	0.536	0.869	0.922	0.862	0.776	0.695	0.614	0.572	0.219
Earnings at graduation (Kitagawa)	0.639	0.551	0.878	0.935	0.874	0.794	0.722	0.648	0.609	0.255
Earnings at 50 (Sun)	0.459	0.428	0.526	0.613	0.513	0.598	0.675	0.787	0.803	0.889
Earnings at 50 (Kitagawa)	0.467	0.439	0.534	0.622	0.520	0.611	0.684	0.792	0.811	0.895
WTP>100 (Sun)	0.233	0.182	0.095	0.054	0.034	0.023	0.014	0.013	0.016	0.223
WTP>100 (Kitagawa)	0.252	0.195	0.103	0.060	0.039	0.027	0.018	0.015	0.018	0.240
WTP<100 (Sun)	0.148	0.203	0.245	0.133	0.153	0.159	0.124	0.091	0.104	0.046
WTP<100 (Kitagawa)	0.152	0.209	0.254	0.139	0.158	0.163	0.129	0.093	0.106	0.047
	Two-dimensional test: (Y_0, Y_1)									
All (Sun)[1]	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
All (Kitagawa)[1]	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

[1] Due to the dimensionality of the problem, the sets over which the test is conducted are based on a grid of fifty equally spaced points on the domain of prior beliefs instead of all possible sets. Results are similar using a seventy-five equally spaced grid.

uing education. The bottom panel presents estimates using prior beliefs on earning if continuing and discontinuing education.[46] The measured effect of information on behavior is twice as large using both priors. Of each participant deciding to continue education, two decided to discontinue education. Table 8 shows the estimated effect of information is significantly below the Fréchet upper bound.

Table 8: Distribution of behavioral types

	One dimensional				
	Switched in	Switched out	Out minus In	Out plus In	Fréchet UB
Mean	0.009	0.058	0.049	0.067	0.212
SE	0.004	0.006	0.008	0.006	0.008

	Two dimensional				
	Switched in	Switched out	Out minus In	Out plus In	Fréchet UB
Mean	0.049	0.097	0.049	0.146	0.212
SE	0.006	0.007	0.008	0.010	0.008

4.5 Willingness to pay for information

Table 9 investigates the willingness to pay for information of those switching decisions after receiving information. We take advantage of our estimations allowing us to detect the distribution of participants switching in and out of education based on their prior beliefs. The first column is the average WTP for information of those who continue education in the control group and have prior beliefs as those who switched to continue education. The second column is the average WTP for information on those who continue education in the treatment group and have prior beliefs as those who switched to continue education. The only difference between these two groups is the participants who switched to continuing education due to the intervention. The third column presents the implicit willingness to pay for information of those switching.[47] The second column is lower than the first, implying that those who switched behavior are less willing to pay for information. We do not impose boundary conditions in the estimation and obtain negative values. The fourth column presents the willingness

[46]The method presented in Section 2.2 affords estimations using the full set of prior beliefs. We refrain from that approach due to the curse of dimensionality.

[47]We use the law of total probability to estimate this value.

to pay for information of those who continue education in the control group and have beliefs similar to those who discontinue education due to the intervention. The fifth column presents the willingness to pay for information of those who continue education in the treatment group and have beliefs similar to those who discontinue education due to the intervention. Since the difference between these two groups is those who switched behavior, we conclude that those who switched must be more willing to pay for information.

Table 9: WTP by prior beliefs, treatment, and response type

	AT	AT+In	Switched-In	AT+Out	AT	Switched-Out			
	(1) $Z = 0$	(2) $Z = 1$	(3)	(4) $Z = 0$	(5) $Z = 1$	(6)	(7) H0:(1)=(2)	(8) H0:(4)=(5)	(9) H0:(3)=(6)
All information:									
Mean	112.4	106.0	-97.9	113.5	101.8	212.7	0.083	< 0.001	< 0.001
SD	1.7	3.1	48.4	0.9	2.1	22.0			

We find a significant difference in willingness to pay for information between those who switched in and out of education due to the intervention. We note that this conclusion is not due to boundary violations. The willingness to pay for information for those who discontinue education is significantly larger than zero (and the mean), which would be an alternative estimate of the willingness to pay for information for those who switched to continue education. This is evidence consistent with information avoidance or myopia in some sub-populations and warrants caution in interpreting WTP solely as a measure of the instrumental value of information.

4.6 Flows across educational choices

The method proposed in Section 2.2 identifies flows in and out of several educational paths. The information intervention likely altered beliefs of several options relative to discontinuing education and relative to each other. Since subjects with similar beliefs are expected to react similarly to information, we can assess flows across choices by measuring distributional changes in the joint distribution of beliefs across treatment and control groups.

To implement this approach, we estimate the joint density function of the priors on earning for each choice (Arts, Commerce, and Science) and earning for junior high school for those choosing each option. For example, we estimate the joint density function of the priors on earnings for Arts, priors for junior high school, and choosing

Arts conditional on being in the Control and Treated group.[48] We estimate the probability of moving into Arts due to the intervention as the integral of the difference between the density in the Treated and Control group truncated at zero. We estimate the probability of moving out of Arts due to the intervention as the integral of the difference between the density in the Control and Treat group truncated at zero. This procedure is repeated for the decision to enroll in either Commerce or Science. We also estimate the flows in and out of senior high school by aggregating priors as in the previous section.[49] This approach does not double-count movements in and out of educational paths since they are constructed using marginal densities on mutually exclusive choices. For instance, when we estimate the flows in and out of senior high school, we do not account for movements across fields of study. When we estimate the flows in and out of Arts, we account for flows into other fields and in and out of senior high school.

We use linear programming to minimize the probability that the intervention generated flows across choices. Table 10 provides estimates using this approach. The last column is the marginal distribution of choices for the Control group and the last row is the marginal distribution of choice for the Treatment group. We estimate ninety-five confidence intervals using Bootstrap. The estimates on the flows in and out of Junior High School are slightly different from those in Table 8 because we estimate the model on the subsample with prior beliefs for all these options. Table 10 estimates that twenty-seven (the sum of the off-diagonal terms) of subjects change their decision due to the intervention.[50] Importantly, we find that the intervention also changed track choices. For instance, we estimate that seven percent of those who planned to enroll in Commerce or Science switched to Arts. Given these findings, we expect to observe larger treatment effects on earnings.

[48] That is, $f(Y_0, Y_{Arts}, Arts = 1|Z = 0)$ and $f(Y_0, Y_{Arts}, Arts = 1|Z = 1)$.

[49] We use Matlab mvksdensity to estimate these densities over a 50×50 grid. We use a Gaussian kernel with Silverman's rule of thumb bandwidth. We use plug-in formulas instead of cross-validation since they produce the most conservative estimates of the effect of information. We use the trapezoidal rule for numerical integration. We rebalance the sample before analyzing to minimize spurious effects due to a lack of balance in small samples. We split the data in 4×4 bins and resample to obtain balance on these bins. We estimate confidence intervals using clustered bootstrap.

[50] Heinesen, Hvid, Kirkeboen, Leuven and Mogstad (2022) show that aggregation across fields of study might lead to biases. Estimates using disaggregated majors produce a larger effect of information on choices and similar flow patterns. Our estimates are, therefore, conservative.

Table 10: Treatment effects on the distribution of educational choices

Control	Treated: Dropout	Arts	Commerce & Science	Total
Dropout	0.02	0.00	0.06	0.08
	[0.00,0.05]	[0.00,0.03]	[0.03,0.09]	[0.05,0.12]
Arts	0.06	0.29	0.03	0.38
	[0.00,0.13]	[0.21,0.34]	[0.00,0.09]	[0.35,0.41]
Commerce & Science	0.05	0.07	0.42	0.53
	[0.00,0.12]	[0.04,0.12]	[0.34,0.47]	[0.49,0.58]
Total	0.13	0.36	0.51	1.00
	[0.10,0.17]	[0.31,0.41]	[0.46,0.55]	

5 Conclusions

Despite large disparities in access to information on returns to education and the long-term consequences associated with poor decisions, the evidence on the effectiveness of information interventions in education is mixed. A possible interpretation of this mixed evidence is that information is of secondary importance in addressing gaps in human capital accumulation. This paper shows that standard analysis of information interventions likely underestimates their full effect and misdiagnoses the severity of informational barriers.

We develop a method to detect nonuniform responses to information and, therefore, the overall effect of information on educational choices. We show that instrument validity tests (e.g., Kitagawa, 2015; Sun, 2023; Mourifie and Wan, 2017) can be extended to a special set of covariates if those covariates determine how subjects respond to treatment. We find that the conventional measure of average treatment effects underestimates the effect of information by at least a factor of two. Information constraints are not trivial in this population. Standard experimental design advises sample sizes inversely related to expected effect sizes. If small effect sizes are due to nonuniform response to treatment, our results suggest that richer data might be a better alternative than larger samples.

References

Abadie, A., 2003. Semiparametric instrumental variable estimation of treatment response models. Journal of Econometrics 113, 231–263.

Ajayi, K.F., Friedman, W.H., Lucas, A.M., 2020. When information is not enough: Evidence from a centralized school choice system.

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

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

Bailey, M., Davila, E., Kuchler, T., Stroebel, J., 2019. House price beliefs and mortgage leverage choice. Review of Economic Studies 86, 2403–2452.

Balke, A., Pearl, J., 1997. Bounds on treatment effects from studies with imperfect compliance. Journal of the American Statistical Association 92, 1171–1176.

Barr, A., Bird, K.A., Castleman, B.L., Skimmyhorn, W.L., 2022. Can information and advising affect postsecondary participation and attainment for non-traditional students? evidence from a large-scale experiment with the u.s. army.

Benjamin, D.J., 2019. Errors in probabilistic reasoning and judgment biases, in: Bernheim, B., DellaVigna, S., Laibson, D. (Eds.), Handbook of Behavioral Economics - Foundations and Applications 2. Handbooks in Economics, pp. 69–186.

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+.

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.

Bursztyn, L., Gonzalez, A.L., Yanagizawa-Drott, D., 2020. Misperceived social norms: Women working outside the home in saudi arabia. American Economic Review 110, 2997–3029.

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.

Canay, I.A., Romano, J.P., Shaikh, A.M., 2017. Randomization tests under an approximate symmetry assumption. Econometrica 85, 1013–1030.

Cantoni, D., Yang, D.Y., Yuchtman, N., Zhang, Y.J., 2019. Protest as strategic games: Experimental evidence from hong kong's antiauthoritarian movement. Quarterly Journal of Economics 134, 1021–1077.

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.

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

Chernozhukov, V., Kim, W., Lee, S., Rosen, A.M., 2015. Implementing intersection bounds in stata. Stata Journal 15, 21–44.

Dahl, C., Huber, M., Mellace, G., 2017. It's never too late: A new look at local average treatment effects with or without defiers. Discussion Papers on Business and Economics No. 2/2017.

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

Frandsen, B., Lefgren, L., Leslie, E., 2023. Judging judge fixed effectst. American Economic Review 113, 253–277.

Fuster, A., Perez-Truglia, R., Wiederholt, M., Zafar, B., 2022. Expectations with endogenous information acquisition: An experimental investigation. Review of Economic and Statistics 104, 1059–1078.

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., 2023. Designing Information Provision Experiments. Journal of Economic Literature 61, 3–40.

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

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

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

Heckman, J.J., Urzua, S., Vytlacil, E., 2006. Understanding instrumental variables in models with essential heterogeneity. Review of Economic and Statistics 88, 389–432.

Heinesen, E., Hvid, C., Kirkeboen, L.J., Leuven, E., Mogstad, M., 2022. Instrumental variables with unordered treatments: theory and evidence from returns to field of study.

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.

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

Kendall, C., Nannicini, T., Trebbi, F., 2015. How do voters respond to information? evidence from a randomized campaign. American Economic Review 105, 322–353.

Kirkeboen, L.J., Leuven, E., Mogstad, M., 2016. Field of study, earnings, and self-selection*. Quarterly Journal of Economics 131, 1057–1111.

Kitagawa, T., 2015. A test for instrument validity. Econometrica 83, 2043–2063.

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.

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.

Manski, C.F., 2007. Identification for Prediction and Decision. Cambridge and London:.

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., 1981. Econometric models of probabilistic choice, in: Manski, C., McFadden, D. (Eds.), Structural Analysis of Discrete Data with Econometric Applications. MIT Press: Cambridge, MA., pp. 198–272.

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

Mourifie, I., Wan, Y., 2017. Testing local average treatment effect assumptions. Review of Economic and Statistics 99, 305–313.

Shi, X., Shum, M., Song, W., 2018. Estimating semi-parametric panel multinomial choice models using cyclic monotonicity. Econometrica 86, 737–761.

Sørensen, J., Fosgerau, M., 2022. How McFadden met Rockafellar and learned to do more with less. Journal of Mathematical Economics 100.

Stinebrickner, R., Stinebrickner, T.R., 2014. A major in science? initial beliefs and final outcomes for college major and dropout. Review of Economic Studies 81, 426–472.

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

Sun, Z., 2023. Instrument validity for heterogeneous causal effects. Journal of Econometrics 237.

Thornton, R.L., 2008. The demand for, and impact of, learning hiv status. American Economic Review 98, 1829–1863.

Willis, R., Rosen, S., 1979. Education and self-selection. Journal of Political Economy 87, S7–S36.

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

Wiswall, M., Zafar, B., 2015b. How do college students respond to public information about earnings? Journal of Human Capital 9, 117–169.

Wiswall, M., Zafar, B., 2021. Human capital investments and expectations about career and family. Journal of Political Economy 129, 1361–1424.

APPENDICES INTENDED FOR ONLINE PUBLICATION

A ADDITIONAL MATERIAL

Table A.1: (Log) Earning beliefs by career track and age

Variable	(1) Info (1)	(2) Imp. Cntrl (2)	(3) (1)-(2)
Junior HS at 21	9.170	9.405	0.235
	(1.289)	(1.289)	(0.153)
Junior HS at 30	9.652	9.847	0.195
	(1.187)	(1.160)	(0.117)
Junior HS at 50	10.014	10.145	0.131
	(1.171)	(1.158)	(0.098)
Senior HS at 21	9.822	10.054	0.232
	(1.062)	(1.064)	(0.119)
Senior HS at 30	10.202	10.382	0.180
	(0.957)	(0.969)	(0.079)
Senior HS at 50	10.515	10.661	0.146
	(0.988)	(1.035)	(0.103)
Arts at 21	10.702	10.794	0.092
	(0.775)	(0.843)	(0.121)
Arts at 30	11.097	11.184	0.087
	(0.826)	(0.911)	(0.140)
Arts at 50	11.511	11.522	0.011
	(0.877)	(0.976)	(0.145)
Commerce at 21	10.891	10.804	-0.087
	(0.842)	(0.884)	(0.099)
Commerce at 30	11.229	11.229	-0.000
	(0.877)	(0.911)	(0.146)
Commerce at 50	11.606	11.560	-0.046
	(0.892)	(0.961)	(0.150)
Science at 21	11.036	11.105	0.068
	(0.975)	(1.035)	(0.148)
Science at 30	11.378	11.397	0.019
	(0.960)	(1.064)	(0.156)
Science at 50	11.671	11.675	0.004
	(1.046)	(1.136)	(0.195)
Observations	1,925	658	2,583

A.1 Characterizing compliers assuming monotonicity

This section reports the characteristics of those responding to the information intervention (compliers) under monotonicity (Table A.4). Since the complier group is relatively small ($\sim 4\%$), 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

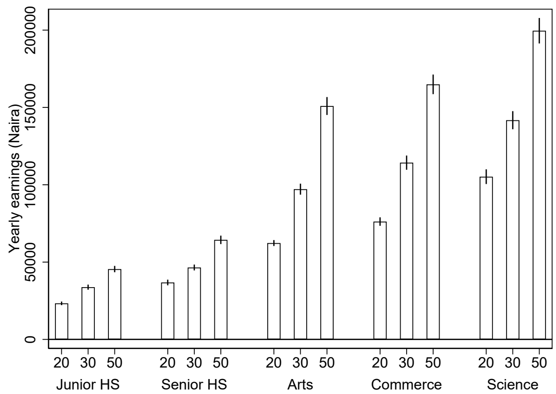


Figure A.1: Prior beliefs by educational choice and age

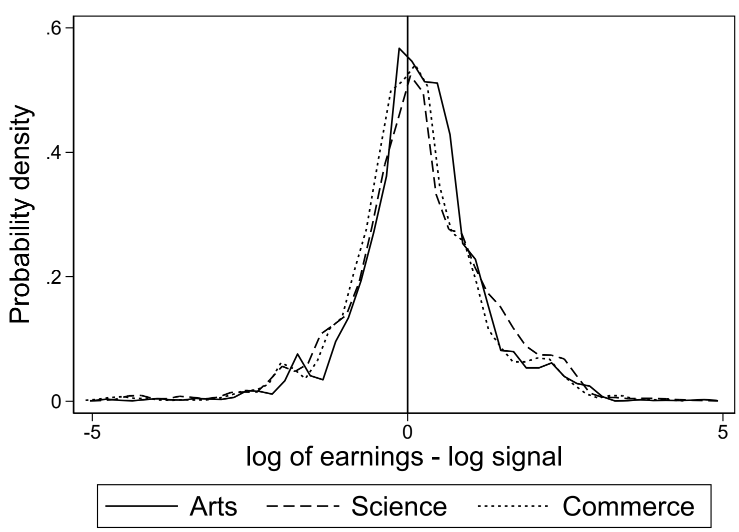


Figure A.2: Distribution of earnings beliefs at 30 around signal

Table A.2: Threshold probability updating

	Pr(Salary at 30<50,000)			Pr(Salary at 30<100,000)		
	Arts	Comm	Science	Arts	Comm	Science
(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

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 participant was provided with information and 0 otherwise.

Table A.3: Monotonicity test using baseline beliefs (Mourifie and Wan, 2017)

	significance level		
	10%	5%	1%
	Conditioning on Y_1		
One restriction ($D = 1$)	R	NR	NR
Two restrictions	R	NR	NR
	WTP > 100		
One restriction ($D = 1$)	R	R	NR
Two restrictions	R	NR	NR
	WTP < 100		
One restriction ($D = 1$)	NR	NR	NR
Two restrictions	NR	NR	NR
	Conditioning on (Y_0, Y_1)		
One restriction ($D = 1$)	R	R	R
Two restrictions	R	R	R
	WTP > 100		
One restriction ($D = 1$)	R	R	R
Two restrictions	R	R	R
	WTP < 100		
One restriction ($D = 1$)	NR	NR	NR
Two restrictions	NR	NR	NR

R: rejection; NR: no rejection. Estimates use Stata's clrtest default settings, i.e., a linear specification (see Chernozhukov, Kim, Lee and Rosen, 2015).

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 (no switch-backs in the elicitation task). We also find that compliers are more likely to have repeated a grade and are more likely to declare themselves Christian.

Table A.4: CHARACTERISTICS OF RESPONSE TYPES

	All	Always-Taker	Never-Taker	Compliers	Pr(Complier ≥ Non-Complier)
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
WTP salaries	103.83	106.47	102.16	137.77	0.93
WTP both	105.60	101.34	102.44	192.75	0.99

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).

B 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?

[QUESTION 1. Other questions are similar.]

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		