Information, Heterogeneous Updating & Higher Education Decisions: Experimental Evidence from India

(JOB MARKET PAPER)

Tanvi Rao*

November 7, 2016

Abstract

I randomize information on population earnings to a sample of 12^{th} grade students, drawn from schools affiliated with a large public state university in India, who at the time were roughly six months away from making a decision regarding college attendance and college track (technical, academic, vocational) conditional on attendance. Upon the receipt of potentially new information, students revise beliefs regarding own-wages in the direction of the information, though the average extent of updating is small and masks substantial sub-group heterogeneity. Additionally, subsequent changes in enrollment intentions and intentions to borrow for higher education are in line with both the extent and direction of wage belief updating. A portion of the heterogeneity in wage belief updating can be explained by initial misperceptions regarding population earnings, and baseline relevance of earnings to enrollment intentions. Yet a large portion remains unexplained, consistent with substantial heterogeneity in updating heuristics, at the individual level. From a policy standpoint, these findings point to the limited capacity of information campaigns based on population-level aggregates to induce, on average, large changes in individual priors and help to rationalize a number of recent papers that find heterogeneous impacts of information provision on education outcomes.

JEL Classification: D83, I23, I26, J24

Keywords: information, subjective beliefs, higher education choices, framed field experiment, India.

---

*Correspondence: Dyson School of Applied Economics & Management, Cornell University, 434 Warren Hall, Ithaca, NY 14853. Email: tr254@cornell.edu; Website: www.tanvirao.com

I am extremely grateful to Ravi Kanbur, Jim Berry, Prabhu Pingali and David Just for advice and encouragement. I especially thank Victoria Prowse for several insightful conversations and Basit Zafar and Adeline Delavande for helpful comments and suggestions on research design and questionnaire construction. The smooth and timely completion of data-collection on ground would not have been possible without institutional support from the TATA-Institute of Social Sciences and in particular Bhaskar Mittra and Maya Nair. An excellent field-team led by Thangkhanlal Thangsing helped manage and implement the survey. Lastly, I thank the entire leadership and administration of Ranchi University and the enthusiastic student-subjects who participated in the survey. Funding from the TATA-Cornell Agriculture and Nutrition Initiative (TCi) is gratefully acknowledged. All errors are mine.

1 Introduction

Providing information on the returns to education in the labor market is generally seen as a powerful demand-side tool to encourage human capital accumulation. The hypothesis that parents and their children might underestimate sizeable returns to education in the labor market, and hence under invest in education, makes information provision a compelling and cost-effective intervention. Encouraging results from Nguyen (2008) & Jensen (2010), who find that randomizing information to a sample of students increased average schooling attainment and school attendance at a basic level of schooling, have spurred a significant literature examining information interventions in education. A substantial focus of the recent literature (Oreopoulos and Dunn (2013); Wiswall and Zafar (2015*b*); Hastings, Neilson and Zimmerman (2015); Pekkala Kerr et al. (2015)) has been in the context of post-secondary education, where premiums vary dramatically by degree and institution, and the focus of policy-makers has been to minimize the extent to which misinformation may lead to sub-optimal decisions, which includes enrollment on account of over estimating the net-returns to certain degree-college tracks vis-a-vis others.

By and large, the emergent literature on information interventions in education has been equivocal in it's findings. While some studies do find information to be effective on average[1], others document only a subset of the overall sample responding to the information in some manner, or not at all (Fryer Jr (2013); Loyalka et al. (2013); Avitabile and De Hoyos Navarro (2015); Pekkala Kerr et al. (2015); Bonilla, Bottan and Ham (2016)). Low average effectiveness of information interventions have mostly been explained by examining why certain groups of individuals may not be able to act on newly acquired information. For instance, Fryer Jr (2013) suggests that information on returns may have increased students intrinsic effort to do better but not real outcomes like test-scores because students may not know how to translate effort into output. In the context of a similar model, Avitabile and De Hoyos Navarro (2015) also find better learning outcomes among high-income individuals, on account of being exposed to information, as these individuals are plausibly in a better position to translate effort into output. Affordability/credit constraints have also been suggested to be binding in Bonilla, Bottan and Ham (2016) who find only a response on the intensive margin of enrollment (towards more selective tracks), driven by the relatively richer sub-group.

In this paper, I offer another explanation for why the provision of population-level information on returns may lead to highly heterogeneous (final) outcomes, by examining in detail the first-link in the causal chain that links population-level information to ed-

[1] For instance Oreopoulos and Dunn (2013) find that providing information on mean earnings differences between those who complete high school and those who have college degrees along with access to a financial aid calculator, leads students to revise upwards the expected earnings from college relative to high school and makes them more likely to state college attendance. Hastings, Neilson and Zimmerman (2015) provide highly customized degree and institution specific information to individuals, adapted to their individual enrollment intentions, and find low SES individuals to enroll in degrees with higher net earnings.

ucation outcomes, that is, the extent to which individuals update own-earnings beliefs in response to receiving information about population-level averages. Heterogeneity in this case depends on (a) the extent to which individuals are misinformed about population earnings to begin with and (b) the degree to which new information on population earnings is *relevant* to individual's own earnings beliefs. By and large papers in the literature either do not systematically collect information on the impact of information on own belief updating (Fryer Jr (2013);Pekkala Kerr et al. (2015)) or do not establish heterogeneous outcomes between sub-groups over and above homogeneous updating in own-beliefs (Loyalka et al. (2013); Avitabile and De Hoyos Navarro (2015); Bonilla, Bottan and Ham (2016)).

In a framed field experiment (Harrison and List (2004)), I study the role of information provision on own-earnings beliefs, stated enrollment and borrowing intentions, in a setting which abstracts away from both affordability and eligibility constraints. More specifically, I examine the impact of an experiment that provides information to high-school students on the distribution of post-secondary, track-specific population earnings. The impact of the experiment is measured by students' updating of own wage beliefs contingent on pursuing each post-secondary track, and their stated probability of enrollment across tracks. Borrowing intentions are measured for higher-education attendance and are not elicited as a track-specific decision. Wage beliefs and enrollment intentions were elicited for (potentially) hypothetical choice-sets where all tracks were available to all individuals. Surveyed students were 5-9 months away from making an actual post-secondary education enrollment decision and were therefore likely to be thinking more actively about their post-secondary education status. The experiment was carried out with 12^{th} grade students drawn from constituent schools of a large, public, state university in India and provided earnings information conditional on three post-secondary tracks- technical, general, vocational- and information conditional on not pursuing post-secondary education.

Students in this sample have substantially biased beliefs about population earnings at baseline[2]. Moreover, at baseline, earnings are a statistically important determinant of enrollment intentions. Despite this the average impact of information provision is small. This is the case even when commonly cited binding constraints like credit availability and eligibility for enrollment are not directly constraining students. In the current setting, the small average impact of information on own-wage belief updating and hence subsequent decisions, stems from highly heterogeneous updating of own-wage beliefs. Heterogeneity is examined by students' current subject stream of study, an important predictor of post-secondary education in the Indian context and also the dimension along which the survey sample was stratified. For one group of students (students in Arts/Humanities) own-wage belief updating, changes in enrollment intentions and treatment-control differences in borrowing intentions are all statistically insignif-

[2]While I do not examine the formation of earnings expectations here, Maertens (2011) indicates that in her sample from rural India, individuals' information sets regarding earnings are positively influenced by the frequency with which information is received via media outlets and from schools, on the number of educated people known and on the respondent's own education.

icant. The second group of students (students in the Commerce stream), revise own-earnings beliefs for attendance tracks, relative to non-attendance, *downwards* by a large and highly statistically significant magnitude. This is driven by a large, downward revisions for wage-beliefs conditional on attendance tracks and no statistically significant revision for beliefs conditional on the non-attendance track. This group of students state lower likelihood of enrollment relative to non-enrollment and treatment individuals are not any more likely to borrow than control ones. The final group of students (students in the Science stream), revise own-earnings beliefs for attendance tracks, relative to non-attendance, *upwards*, and this is driven by a large, downward revision for the non-attendance track and no statistically discernible updating for the attendance tracks. This group of students state higher likelihood of enrollment relative to non-enrollment and treatment individuals are much more likely to borrow than control ones. Therefore, two sub-groups of students, Commerce and Science, update attendance earnings, relative to non-attendance, in opposite directions. These patterns are indicative of systematic updating behavior because enrollment and borrowing intentions for the two sub-groups are in line with the direction of wage belief updating. Moreover, within these sub-groups, the same set of individuals seem to be driving both wage belief updating and revisions regarding enrollment intentions.

What explains this differential updating on account of the receipt of the same information? Arts students have at baseline a small and statistically insignificant elasticity of enrollment to wage beliefs. At baseline they also make smaller errors, on average, with regards to beliefs about population earnings. However, these factors do not explain differential updating on the part of Commerce & Science students. Ex-ante, these students have a statistically important and similar in magnitude elasticity of enrollment to wage beliefs- therefore earnings likely play an important role in their decisions for future education. Importantly, I establish that differential updating for these two groups of students cannot be established on account of differences in baseline errors, regarding population wages, between the groups. Both groups of students have statistically identical baseline errors for all four tracks. This indicates that at the individual level heuristics relating beliefs regarding population earnings to own-earnings are highly varied and undermine the extent to which information campaigns based on population aggregates might be effective on average. Consistent with predictions of belief-based models of Bayesian updating, I find some suggestive evidence to support that individuals with stronger likelihoods to pursue a track are less likely to update earnings beliefs for that track, compared to individuals with weaker likelihoods at baseline. However, in the absence of data on individuals' variance of their prior beliefs, I cannot rule out that a portion of the non-response to information may also be non-Bayesian in nature. However, some insights from the literature indicate that this is a possibility. Wiswall and Zafar (2015*b*), examine the extent to which individual-level updating of beliefs deviates from the Bayesian benchmark. Given each individual's prior belief and variance of their prior belief, they construct a Bayesian benchmark for every individual and then use data on their actual posteriors to classify deviants from the benchmark. They document a wide range of updating heuristics among respondents; nearly a fifth of their sample comprises of "Non-Updaters", in the non-Bayesian sense. Among those who update, while the most common heuristic is within the band of Bayesian updating, a substantial portion of the sample is more Conservative in their updating and up to 12-19 percent of the respondents update in the Opposite ("Contrary") direction.

To summarize, in this paper, I establish that heterogeneity in the updating of own-wage beliefs in response to population-level information is important and drives significant differences in decision-making between sub-groups in the sample. A large part of this updating is unexplained by initial differences in misperceptions regarding population earnings. Some suggestive evidence indicates that this differential behavior is consistent with predictions of the Bayesian model, but we cannot rule the extent to which non-Bayesian behavior may account for these findings. However the literature suggests that the latter is likely to play an important role and merits further investigation. In this paper, reference to differential updating heuristics indicates both variation in updating consistent with the Bayesian model (i.e. on account differential variance of prior beliefs) and non-Bayesian updating.

Campaigns designed to provide earnings information based on population level aggregates are attractive to policy-makers. For instance, based on Jensen (2010)'s influential study, the Mexican Secretariat of Public Education implemented an intervention to provide students entering 10^{th} grade with information about the returns to high-school and tertiary education, with an aim to improve on-time graduation and learning outcomes (Avitabile and De Hoyos Navarro (2015)). A major appeal to information interventions are also their potential for being cost-effective[3]. For a policy-maker looking to implement an information campaign to encourage more optimal education decision making, these findings are not encouraging because they highlight that the heterogeneity by which individuals apply population-level information as relevant to themselves is important. Therefore, even if a policy-maker has accurate knowledge about the direction and magnitude of baseline errors regarding population wages, for a particular group, they may not be able to induce large changes, on average, in individual's beliefs about themselves. More detailed data on earnings conditional on different types of education (as is available in Chile (Hastings, Neilson and Zimmerman (2015)), Finland (Pekkala Kerr et al. (2015)) & Colombia (Bonilla, Bottan and Ham (2016)) and for different groups of individuals (like the National Survey of College Graduates in the U.S.) may help in the design of more specific information interventions which may provide more informative signals to different groups of individuals. However, currently, in most developing countries, such information is not systematically collected.

The primary contribution of this study is to the literature that evaluates the potential of information policies and campaigns to influence education decision-making. However, it's findings can also throw light in other areas of economics where aggregate in-

[3]Cost-effectiveness analysis accessed on J-PAL's website at: https://www.povertyactionlab.org/policy-lessons/education/improving-student-participation show that information provision is more cost-effective than merit scholarships and cash transfers. However, only one study on information provision is used as reference

formation is used to influence individual decision-making. Some examples include the use of information in impacting occupation choice (Osman (2014)), sexual-behaviors (Dupas (2011)) and migration decisions (Shrestha (2016)). The results on the demand for borrowing as a result of information provision in this paper, speak to an important strand in the large literature on higher education access, which focuses on borrowing constraints (Stinebrickner and Stinebrickner (2008); Delavande and Zafar (2014); Kaufmann (2014)). In principal, individuals who would like to borrow at the going interest-rate but are unable to gain credit are classified to be credit-constrained. However, being credit-constrained depends on each individuals net-return calculation from education, which itself may suffer from information gaps. In my study, the fraction of the sample which revises expected earnings from post-secondary attendance, relative to non-attendance, upwards is also more likely to state an increased intentions to borrow. Therefore, to the extent that the link between information provision and own-wage beliefs is present, information can affect behavior to lift more binding barriers to education access.

The rest of the paper is organized as follows: section 2 briefly discusses aspects of post-secondary education in India, section 3 outlines a conceptual framework to motivate sources of heterogeneity in own-wage belief updating, section 4 discusses data and experimental details, section 5 is devoted to results, section 6 entails a discussion on factors that explain (or fail to) heterogeneous updating in the sample and the final section concludes.

2 Post-Secondary Education in India

I study the decision-making of students between three post-secondary tracks and the non-attendance alternative. After the completion of high school, students choose whether or not to attend post-secondary education and what type of post-secondary education to enroll in. I classify post-secondary education into three "attendance tracks"- technical/professional degrees, general academic degrees and vocational diploma or certificate courses. This is also how the Government of India classifies higher education tracks in its collection of post-secondary education data as part of the National Sample Surveys (NSS) and this is the lowest level of aggregation at which nationally representative earnings data is available in India.

Each higher education track studied in this paper lies at distinct points of the net-return spectrum from post-secondary education in India. The three attendance track are also distinct in the type of educational content they impart and have distinct labor-market implications. Technical degree courses include professional degrees in fields like medicine, engineering and architecture as well as "job-oriented" degrees like Bachelors of Computer Application, Business Administration, Information-Technology (IT), Pharmacy or Hotel Management. These courses are offered both by government and private institutions and are regulated by the "All-India Council for Technical Education (AICTE)". General degree courses are non-technical and award a bachelor's degree in either the arts, sciences or commerce, further categorized according to subject. Mostly, these are offered by the government via central or state level universities and colleges. Vocational courses are not academic and focus on imparting a set of skills (rather than broader academic knowledge) targeted towards employment in a specific sector. They are offered by both government and private institutes. Under the government, these courses are offered either by Industrial Training Institutes/Centers (ITI/ITC) or by Polytechnics.

Recent reports of the NSS 71^{st} round on education expenses, estimates the average yearly costs for technical/professional degrees to be a little over 60,000 rupees (approx. 1,000 dollars) with the expenditure on private institutions being 1.5-2.5 times the cost of government institutes. Average yearly expenses for a general education, in contrast, were found to be around 7,000 rupees (approx. 100 dollars) and for vocational courses, around 30,000 rupees (approximately 450 dollars). Measured wage premiums for technical degrees are more than a 100% of the wages of those who complete high school. Despite the fact that vocational training is more expensive than general degree courses, wage premiums for vocational courses (42% of high-school wage) are around 8 percentage points lower than the wage premiums for general courses.

Another feature of higher education in India is that students study in a specific subject-stream during 11^{th} and 12^{th} grades. This is the case in the current sample and is also true nationally. Typically, there are three subject streams: (1) Arts/Humanities, (2) Commerce and (3) Science. As is discussed subsequently, students' current stream of study is expected to be strongly correlated with future post-secondary education choice on account of preferences, with regards to eligibility for specific courses or degrees within tracks and also, and also on account of ability (measured by test-scores) and socio-economic status (SES). In this paper, we discuss heterogeneity in the impact of the treatment, by students' current stream of study, because a-priori, we expect that students from different streams would have different baseline intentions of pursuing different tracks. In section 5.2, we further discuss correlates of students current stream of study to frame our findings.

3 Conceptual Framework

I discuss here a simple model of belief updating proposed in Wiswall and Zafar (2015*b*) which is useful to frame the set-up, analysis and findings of this paper. The model highlights that students update beliefs about their own-earnings, upon receiving information about population earnings if (1) they are misinformed about populations earnings and (2) their beliefs about their own earnings are linked to their beliefs about population earnings. Additionally, the function that links population earnings beliefs to own earnings beliefs, known as the updating function, varies at the level of the individual, and matters in determining both the direction and extent to which individuals update beliefs.

Let X_{it} be individual $i's$ expectation at time t about her own earnings at some future date, denoted X and let Ω_{it} denote $i's$ information set at time t. Prior to receiving information, in the pre-stage, respondent i reports her beliefs about self earnings as:

$$X_{it} = E(X|\Omega_{it}) = \int X dG_i(X|\Omega_{it}) = f_i(\Omega_{it}) \tag{1}$$

where $G_i(X|\Omega_{it})$ is individual $i's$ belief about the distribution of future earnings conditional on the information Ω_{it}. $f_i(.)$ is the updating function that provides the mapping between the individual's information set to beliefs about own-earnings at some future date.

An individual's information set has two parts: $\Omega_{it} = \{I_{it}, B_{it}\}$. Here, let I_{it} be individual $i's$ current belief about the information we provide in our treatment- average track-specific post-secondary education earnings in the Indian population. B_{it} contains all other elements of an individual's information set which includes both other population-level information and private information available only to the individual like her perceived ability to succeed in a particular track. After the provision of information, in the post-stage, the individual's information set is Ω_{it+1}. At this stage, we also elicit her beliefs about her own-earnings at a future X_{it+1}, again.

Two conditions are necessary for an individual to update their beliefs about their own earnings and for $X_{it} \neq X_{it+1}$:

1. $I_{it} \neq I_{it+1}$ and the individual should not already know the information that we provide. Therefore the information should be new and also accepted by the individual as credible.

2. $f_i(\Omega_{it}) \neq f_i(\Omega_{it+1})$ and the individual should consider the population-level information as relevant to themselves.

If we observe that individuals at baseline have beliefs about population earnings that are substantially different from the information we provide, then condition (1) is met and the information provided is likely "new" to the individuals in the sample.

If individuals do not update own-earnings beliefs, despite the information being new to them, i.e. $\frac{\partial f_i(\Omega_{it+1})}{\partial I_{it}} = 0$, then this non-updating can be Bayesian or non-Bayesian in nature. Theoretically, Bayesian updating is the benchmark model for updating beliefs about a quantity regarding which information is provided. In our case, a modification of Bayesian updating applies because individuals receive information over one variable (population earnings) and update beliefs regarding a separate variable (own earnings). In this Quasi-Bayesian case, updated earnings (X^B_{it+1}) are a linear combination of an individual's prior beliefs (X_{it}) and new information about population earnings (I_{it+1}), wherein the relative weight placed on the new information is the variance of the prior belief $V(X_{it})$ divided by the variance of the new information $V(I_{it+1})$, i.e. $\frac{V(X_{it})}{V(I_{it+1})}$. Therefore, a core prediction of the Bayesian model of updating is that individuals are more responsive to information regarding a quantity that they have weaker priors about (i.e. higher $V(X_{it})$) (DellaVigna and Gentzkow (2010)). In the context of this paper, $V(I_{it+1})$ is the same for all individuals and the same track-specific information is provided to everyone. Therefore, to the extent that updating is Bayesian, variation in responsiveness to information stems for differences in the variance of prior beliefs.

Non-Bayesian updating denotes deviations from the Bayesian benchmark; that is the extent to which individuals posteriors differ from their Bayesian benchmark which is defined given the variance of their prior beliefs. A body of literature (Kahneman and Tversky (1972); Wiswall and Zafar (2015*b*)) in economics and psychology documents that updating can be (i) Alarmist (response being more exaggerated relative to the Bayesian benchmark), (ii) Conservative (less than the Bayesian benchmark) and (iii) Non-updater (no response even though the Bayesian benchmark predicts response). Ideally, for information interventions[4] to be effective on average and induce significant updating for the complete sample or by direction and magnitude of baseline error, we would want individual-level updating heuristics to be less variable. In other words, the direction and magnitude of baseline population errors should be predictive of the direction and magnitude of own-belief updating. As discussed above, ex-ante, we should not expect this to be the case and whether or not this is the case is an empirical question. The results in this paper emphasize that the heterogeneity in individual-level updating heuristics is important. We show that this is the case by highlighting differential updating of own-earnings beliefs for sub-groups with statistically identical baseline-errors. The documented differential updating of own-earnings beliefs also has important consequence for the manner in which these sub-groups update enrollment probabilities and borrowing intentions- decisions that take updated own-earnings beliefs as inputs.

4 Data & Experiment Details

4.1 Data collection & Timing

The data for this study was collected from a sample of 1525 students across nine public schools in the East Indian state of Jharkhand. All nine schools are constituent units of a large state university and the students, at the time of the survey, were studying in in their 12^{th} grade.[5] Four of the nine schools are situated in the capital city of Ranchi, one in a rural block of Ranchi district and four others are in surrounding rural districts. The survey was conducted between October 2014 and February 2015, five-nine months

[4] I refer here to the class of information interventions that provide information on a population average to perturb beliefs about an individual-level quantity.

[5] Specifically, these students were studying in the final year of their "intermediate degree" in what are known as "intermediate colleges". After completing 10^{th} grade, students chose between attending either an "intermediate" college, for two years of higher-secondary schooling, or attending a high school which offers 11^{th} and 12^{th} grades. Public "intermediate" colleges, like the ones surveyed here, are often co-located with public colleges offering undergraduate degrees. Since intermediate education is equivalent to higher secondary education, I refer to these students as being in 12^{th} grade, throughout the paper, to avoid confusing terminology as most people think of colleges as referring to only post-secondary education.

prior to the time when students make actual decisions regarding enrollment in post-secondary education.

Figure A1 highlights the structure of the survey. Half of the complete sample was randomly assigned to the information treatment group and the other half to the control group. The sample was also stratified by gender and current stream of study to ensure equal representation of the two sets of groups across treatment and control (Duflo, Glennerster and Kremer (2007)). We drew, approximately, an equal number of students from each school. Further, within each school, students were randomly assigned to survey-sessions of 15 students each. Survey sessions were either a control session or a treatment session, with the latter differing only on account of the feature that it included an approximately 20 minute long information session, at the end of the collection of baseline data and disbursal of "loan cards". For a given survey-session, round 2 of data collection was conducted the day after the first round.[6] In every school, both rounds of all control sessions were conducted before the treatment survey-sessions, in order to prevent students from the treatment group to share information with students in the control group, in a manner that can influence the results of this paper. Both sets of students answered exactly the same round 1 and round 2 questions. Survey sessions were conducted in classrooms within the students' school and were led by a team of two enumerators. Students answered the questions, posed by the enumerators, on android tablets. The questionnaires were fielded using Open Data Kit (ODK) software.

4.2 Survey Questionnaire & Information Treatment

Round 1 of the survey consisted of questions on (i) socio-economic details including gender, caste, religion, a "household assets" module, parental education and occupation, older sibling gender and education, scores on previous centralized board examinations and history of grade repetition and (ii) baseline beliefs contingent on each higher education alternative i.e. technical/professional degrees, general degrees, vocational diplomas/certificate courses and the fourth alternative of not attending further education after 12^{th} grade. While the three post-secondary education tracks were constructed to maintain consistency with education data collected by the country's National Sample Survey (NSS), the categories are broad and encompass a variety of courses of study. Therefore, data collection was preceded by a detailed explanation of possible courses/degrees that are part of every category. Since a majority of the beliefs questions were either probabilistic in nature or required students to express responses on a scale of 0-100, the

[6] The short time-period between round 1 and round 2 of the survey follows from the research designs of Wiswall and Zafar (2015*a*) and Osman (2014). The short-time span between rounds allows one to be sure that all other factors in an individuals utility function that are correlated with earnings beliefs remain invariant and also limits the time that students in the control group have to acquire information from other sources, over time. The drawback is that I cannot comment on the process of expectations formation over the long term or on the persistence of the effects of information provision. However, Wiswall and Zafar (2015*a*), who collect revised beliefs both instantaneously and over the long term, find the effects of information provision to be strongly persistent two years after the provision of information.

baseline beliefs module was also preceded by a discussion (with examples) on answering probabilistic questions[7].

In the baseline beliefs module, stated probabilities of enrollment were elicited for (i) all four higher-education alternatives[8] and (ii) only for higher education alternatives that comprise an individual's affordable choice-set. Next, individuals were asked about certain non-pecuniary and pecuniary beliefs conditional on each education alternative. Pecuniary beliefs included data on expected probability of employment and expected average monthly earnings. These pecuniary beliefs were collected both for individuals' perceptions regarding their own expected labor market outcomes and outcomes they believe apply to an average individual in the population[9]. Non-pecuniary beliefs included questions regarding enjoyment of coursework, parental approval of education track and likelihood of graduation. In this paper, we are interested in disentangling the effect of information-gaps in one particular aspect of individuals' information sets- their knowledge about the distribution of population (public) earnings by post-secondary track. Therefore, the only belief our experiment manipulates (and which we collect post-treatment data on) are individuals' beliefs regarding track-specific expected average monthly earnings. We utilize beliefs about expected average monthly earnings in this paper, and the other aforementioned belief variables are not analyzed in detail here. Secondly, since the focus of this paper is on recovering the elasticity of enrollment intentions to earnings beliefs, we utilize data on all four higher education alternatives hypothetically "available" to an individual. Restricting estimation to only affordable alternatives biases our parameter of interest, as individuals constrained by costs might appear falsely unresponsive to the information intervention.

Another part of the data-collection focused on measuring the demand for higher-education loans in our sample. Therefore, additionally, at the end of round 1, all students were given a "loan-card" which had two questions related to borrowing for higher education which they had to think about at home and discuss with their family members. The two questions were- a) whether the individual would like to accept a loan, offered at a

[7]We ensured that answers to all probabilistic questions sum to 100 by placing the total as a constraint in the questionnaire, without fulfilling which, the survey would not proceed to the subsequent question.

[8]The exact wording of the question used to elicit enrollment probabilities for potentially hypothetically choice sets of individuals was: *"Think ahead to next year when you have completed (sic) intermediate. Imagine that you have passed your (sic) intermediate examinations and are able to secure admission in one degree/course belonging to each of the options 1, 2 and 3. Option 4 is also available to you. Suppose that you are provided with financial aid such that all your expenses (tuition, boarding, room, etc.) are paid for at a private/government institute for a course belonging to each options 1, 2 and 3. State the percent chance that you would enroll in each of the following?"* This statement was followed by the four education options among which students had to allocate probabilities.

[9]For e.g. the question used to elicit own earnings beliefs was: *Consider the situation where you graduate from a degree belonging to the alternative "insert track". Look ahead to when you will be 30 years old. Think about the types of jobs associated with degree/course. How much do you think YOU would earn per MONTH on AVERAGE, if you completed a degree of this type?* "You" was replaced with the phrasing "Typical person" to elicit beliefs about earnings of an average person in the population.

fair interest rate, for attending higher education, to be repaid only after completion of their studies- "yes" or "no" b) If "yes", keeping in mind the length of their desired degree, how much would they like to borrow on a yearly basis?

The 20 minute information session discussed the average and the 25^{th} and 75^{th} percentile of the monthly earnings distribution of men and women who have completed each higher education alternative. This data was calculated from two latest rounds of National Sample Survey (NSS) data[10]. Individuals part of the information treatment group also took home a sheet of paper with a graph and some statistics that summarized the contents of the information session that they were part of. The script of the information-session is reproduced in Appendix 1 and the "information sheet" taken home by the students is given in Figure A2. The loan card taken home by the students is given in Figure A3, along with the accompanying loan script.

The next day, for round 2, students were (i) re-asked about their stated enrollment probabilities for all four higher education alternatives, (ii) expected average monthly earnings for each higher education alternative. In addition, their response to the questions posed in the "loan card" were also recorded.

5 Results

5.1 Covariate Balance

Table A1 in Appendix 2 summarizes the key background variables of sample individuals and checks for balance in these characteristics across control and treatment groups, for the full sample. Control and treatment individuals do not differ statistically on account of almost all relevant socio-economic and demographic characteristics, at baseline. However, we see that control individuals are more likely to own land (p-value=0.04) and individuals in the treatment group have a slightly higher index of household assets (p-value=0.10). Nevertheless, one other variable that is also indicative of the individual's household's well-being, namely the "HH Facility Index", does not statistically differ between control and treatment groups. More importantly, baseline differences in land ownership and household assets, do not manifest in statistically different baseline enrollment probabilities. Figure 4-Figure 7 also show that there are no pre-existing differences between the two groups in the distributions of track-specific own-wage beliefs.

Table A2 breaks down the sample by current stream of study. Here, the Arts and the Science streams seem to be balanced on baseline variables, but there are some imbalances in the commerce stream (5 out of 20 variables). However, some of the variables go in opposite directions. For instance, treatment group individuals have a higher asset index but are less likely to own land. Therefore, in Table A3, I complement Tables A1 & A2 by using the F-test approach to testing for balance. Here, for the commerce stream, the p-value for the F-test that all coefficients are zero is around 0.30.

[10]Refers to two latest rounds of "employment-unemployment" data; i.e. - NSS 66th round (2009-10) and NSS 68th round (2011-12)

5.2 Correlates of Current Stream of Study

Sub-group heterogeneity in this paper is examined by students' current stream of study, a dimension along which the survey sample was stratified. Typically, higher-secondary education in India (grades 11^{th} and 12^{th}) entails students to be enrolled in one of three streams of study- (1) Arts (humanities), (2) Commerce or (3) Science.[11] Students' current stream of study is strongly correlated with their future post-secondary education choices, in part because it often determines eligibility for future study. That is, which particular degrees/courses a student could potentially study within the three attendance tracks discussed in this paper, is to a large extent determined by their stream of study at the higher-secondary level. Accordingly, students' preferences for post-secondary study and eventual occupations are taken into account by them when they choose their stream of study in 11^{th} grade. Hence, a-priori, the impact of track-specific information is expected to differ according to students' current educational stream. Students belonging to a particular stream take classes together and hence can also be expected to have correlated information sets at baseline and to further develop common proclivities towards type of future study.

Current stream of study is correlated with expected post-secondary enrollment. This can be seen in the last four rows of Table A4. Students in the Science stream are nearly 18 percentage points (pp.) more likely to want to enroll in technical tracks as compared to Arts students and 12 pp. more likely as compared to Commerce students. Arts and Commerce students are more likely than Science students to enroll in general tracks. Arts students state the highest non-attendance probabilities, followed by Commerce and then Science students.

For the sample under study the selection of stream in 11^{th} grade was not purely based on preferences but followed a cut-off system where students apply to study in a given stream, and admission is based on points scored in the 10^{th} grade board examination. The highest cut-offs were for Science, followed by Commerce and then Arts. Table A4 confirms that students in the Science stream scored, on average, nearly 11 pp. more than students in the Arts stream and roughly 6 pp. more than students in the Commerce stream, in the 10^{th} grade. The Science group also has more males- almost 22 pp. more males than Arts and 16.5 pp. more males than Commerce. Other correlates are as expected. Students in the Science stream have the highest Asset/Household facility indices, are least likely to be lower caste (i.e. Scheduled Tribe), most likely to belong to the majority religion (Hinduism), and have the most educated parents and older siblings. These measures are least favorable to students in the Arts stream and Commerce students are in the middle.

[11]The arts stream includes subjects like history, political science, psychology, sociology, languages; the commerce stream includes the study of accounts, business, business mathematics and the science stream includes subjects such as physics, chemistry, computer science and mathematics. Economics can typically be studies across the three streams.

5.3 Baseline Relationship Between Expected Earnings & Enrollment Intentions

The experiment in this paper follows from the premise that college enrollment decisions are based on *perceived* net benefits from college (Manski (1993)) and that subjective expectations of future earnings are important determinants of current education decisions (Arcidiacono, Hotz and Kang (2012);Wiswall and Zafar (2015*a*)). In Table A5 I provide *prima facie* evidence on the relationship between expected earnings and enrollment intentions, at baseline, in my sample. Here, I regress the probability of enrollment of individual i for track j, on individual and track-specific non-pecuniary and pecuniary beliefs. I control for individual (student) fixed effects and exploit only within-individual variation in beliefs and enrollment intentions, to estimate the importance of earnings as a determinant of intended enrollment. Nevertheless, these estimates are only suggestive and not causal because unobserved track-specific beliefs that are correlated with earnings beliefs, and predict intented enrollment, are not accounted for.

Consistent with previous literature (Delavande and Zafar (2014); Zafar (2013)) these estimates imply that expected earnings are small but statistically significant determinants of enrollment intentions and that non-pecuniary factors are generally more important in the decision-making process of students. The regression function underlying these results is linear-log in wages, therefore a 1% increase in wage beliefs regarding track j imply an increase of 0.023% increase in probability of enrolling in track j (col. 1 of Table A5). Beliefs regarding expected enjoyment of coursework is the most important correlate of enrollment intentions (also the case in Zafar (2013)), with a 1% increase in the probability of enjoying coursework being associated with a 0.45% increase in the probability of enrolling in track j[12].

Additionally, it is relevant to note that the coefficient on "log own wage" is more than an order of magnitude smaller for Arts students as compared to the other two groups and statistically insignificant (col. 2 and 6 of Table A5). To the extent that it is costlier for individuals to process or pay attention to information not relevant to their decision-making process (Hanna, Mullainathan and Schwartzstein (2014);Sims (2003)), we might expect these students to not update their own-earnings beliefs in response to information provision. If the baseline elasticity of enrollment to earnings is strongly correlated with the experimental elasticity, we might expect these students to not update enrollment intentions/borrowing decisions despite updating own earnings beliefs.

[12]If I estimate this relationship using a log-odds specification, comparable to the reduced form model of Wiswall and Zafar (2015*a*) estimated with cross-sectional data, then my estimates imply that a 1% increase in beliefs about own-earnings in a track (relative to own-earnings for non-attendance) increase the log-odds of enrolling in that track by 0.2%. This estimate is much smaller than their estimated elasticity of 1.6%.

5.4 Baseline Beliefs Regarding Population Earnings

Figure 1 plots the distribution of logged wage beliefs held by males and females in the sample, for an average person in the population. In Indian Rupee (U.S. Dollar) terms, the "true" (measured from nationally representative data) average monthly earnings for working age individuals having completed a technical education track is Rs. 22,071 (328 USD) per month for males and Rs. 16,453 (245 USD) per month for females. Average monthly earnings for having completed a general education track is Rs. 15,280 (227 USD) per month for males and Rs. 12,750 (190 USD) per month for females and average monthly for having completed a vocational education track is Rs. 14,495 (216 USD) per month for males and Rs. 12,210 (182 USD) per month for females. Average monthly earnings of those who don't pursue post-secondary is Rs. 9,973 (148 USD) per month for males and Rs. 8,907 (132 USD) per month for females. Thus, measured college premiums for completing post-secondary education are high and range from around 121 (85) percent for technical tracks to 45 (37) percent for vocational tracks, for males (females).

Figure 1 indicates that a majority of males seem to substantially over-estimate population wages for the three attendance tracks and for the non-attendance alternative. A majority of females also seem to over-estimate population wages[13]. However, for the three attendance tracks, the proportion of over-estimators is smaller for females as compared to males and for the non-attendance alternative the proportion of over-estimators is smaller as compared to the attendance tracks. In investigating the role of information gaps with regards to college attendance, we are interested in the errors that individuals make for the attendance tracks, relative to the non-attendance alternative. In this regard, males can be said to have more accurate beliefs at baseline as compared to females. Table 1 tabulates the percentage of students who over-estimate earnings in all four tracks. The "Full Sample" panel of Table 1 indicates that 70% of males over-estimate population earnings for the non-attendance alternative. This proportion is higher by 2, 4 and 13 percentage points for technical, general and vocational tracks, respectively. In contrast, 49% of females overestimate population earnings for the non-attendance alternative. For girls, this proportion is higher by 21, 13 and 17 percentage points for technical, general and vocational tracks, respectively.

Figure 2 plots the distribution of logged wage beliefs, for an average person in the population, broken down by students' current stream of study. Three facts are apparent in these figures. One, for all tracks, the extent of over-estimation is higher for students in the Commerce and Science streams, as compared to students in the Arts (Humanities) stream; two, the distributions of population-beliefs for Commerce and Science students closely overlay each other; three at least for students in the Commerce & Science

[13]Interestingly, Bonilla, Bottan and Ham (2016) and Gamboa and Rodríguez Lesmes (2014) also find that their respective samples of high school students substantially overestimate the wages of college graduates in Colombia. Therefore, even within low-income populations, among the demographic of high school students, underestimation of college earnings does not seem to be a serious impediment to individuals under-investing in college-level education.

streams, the extent of over-estimation is higher for the three attendance tracks as compared to the non-attendance alternative. Stream-specific panels in Table 1 further illustrate these points. While for students in the Arts stream there is no dominant direction in which baseline errors prevail (especially when attendance tracks are compared to the non-attendance alternative), for Commerce and Science students a majority of students (1) overestimate earnings for all tracks and (2) overestimate attendance earnings to a larger extent than non-attendance earnings. In addition, for both of these streams, a much larger proportion of females, as compared to males, over-estimate attendance earnings relative to the proportion that over-estimates non-attendance earnings.

Figure 3 gives an idea of the relative magnitude of overestimation versus underestimation in the sample. It is a scatter plot of the percentile mean of baseline population errors where "error" is defined as (perceived population wages)$_{ij}$-(true population wages)$_{ij}$ for individual i and track j, and measured in true-wage units. Looking at errors on either side of the zero-error line, and with added focus on errors within 1 true-wage unit of zero-error, we see that for the three attendance tracks there are substantially more individuals one-unit above the true wage than there are below, though, within this range, individuals are more evenly distributed for the non-attendance track.

5.5 Impact on Own-Wage Beliefs

The impact of the treatment on own-wage beliefs is measured first for each track separately (equation 2) and then for each attendance track relative to non-attendance (equation 3):

$$log(W_{ijt}) = \alpha + \beta_1 Post + \beta_2 T + \beta_3 (Post \times T) + \theta X_{it=1} + u_{ijt} \tag{2}$$

$$log(W_{ijt}) - log(W_{iJt}) = \alpha + \beta_1 Post + \beta_2 T + \beta_3 (Post \times T) + \theta X_{it=1} + u_{ijt} \tag{3}$$

Where $log(W_{ijt})$ is the log of own-wage belief of individual i, conditional on enrollment in track j, at either $t = 1$ (pre-treatment) or $t = 2$ (post-treatment). $Post$ is a dummy variable which equals 1 for post-treatment data, T is a dummy variable which equals 1 for individuals in the treatment group. β_3, our coefficient of interest, measures the average effect of the treatment on updating of own-wage beliefs. $X_{it=1}$ denotes baseline controls and u_{ijt} is a mean zero error term. $log(W_{iJt})$ is the log of own-wage beliefs of individual i for the non-attendance track.

The latter specification is important because we analyze enrollment decisions in a log-odds framework and interpret attendance log-odds relative to a base-case of non-attendance.

5.5.1 Full Sample

Figure 4-Figure 7 plot pre and post treatment distributions of own-wage beliefs for control and treatment groups, by track. To focus on the bulk of the distribution and to avoid stretching out the densities to the extremes, I plot densities in the 1-99 percentile

range of the data. The presented Kolmogorov-Smirnov p-values for equality of distributions are however based on the full data. For all four tracks, there are no baseline differences in the respective distributions. Post-treatment, the distribution for the three attendance tracks shifts leftward (a downward revision) and this shift is statistically significant. The shift is perceptibly larger for general and vocational tracks. There is no statistically discernible shift in the distribution of non-attendance earnings.

Table 2 examines the effect of the information treatment on updating of own-wage beliefs in a regression. Panel A looks at each track separately and Panel B presents updating for each of the three attendance tracks relative to updating for the non-attendance alternative. For all four tracks, the treatment is associated with a downward revision in own-wage beliefs, but the revision is statistically significant for only one track (vocational). Relative to non-attendance, the overall effect of the treatment on updating of attendance-track wage beliefs for all three tracks is not statistically significant. The effect-sizes imply an upward revision of earnings-beliefs of around 6% (technical track) and downward revisions ranging from around 3 to 7.3 percent (general and vocational tracks).

5.5.2 By Baseline Error

Table 3 examines the effect of the treatment on own wage-belief updating, separately for those who under and over estimate track-specific population wage beliefs at baseline[14]. Panels A and C look at updating for all four tracks separately, for under and over-estimators, respectively. Panels B and D present updating relative to the non-attendance track. In this case baseline under (over) estimators are also defined as those who under (over) estimate attendance-track population earnings relative to non-attendance[15].

Focusing on Panels A and C, in general, under-estimators seem to revise wage beliefs upwards and over-estimators seem to revise wage beliefs downwards (3 out of 4 tracks in each case). Wage beliefs revisions for baseline under-estimators are relatively smaller and statistically indistinguishable from zero as compared to wage belief revisions for over-estimators. This can partly be explained by differences in the extent of baseline errors between the two-groups, shown in Figure 3. Overall, there are fewer under-estimators than over-estimators (20-30% of the full sample, depending on track), and over-estimators are farther away from the zero error line than are under-estimators. This is true, even if we restrict our attention to 1 true-wage unit below and above zero. Among over-estimators, individuals revise wage-beliefs downwards for all four tracks, with the magnitudes being largest (and statistically significant) for the general and vocational tracks. However, the magnitude of downward revision for the non-attendance track is also quite large. As can be seen in Panel D it is not possible to reject the hypothesis that wage-

[14] Baseline under estimators are those for whom (perceived population wages)$_{\mathrm{ij}}$-(true population wages)$_{\mathrm{ij}}$ < 0. Over-estimators are conversely defined.

[15] Therefore, for this specification, baseline under estimators are those for whom $\{[(\text{perceived population wages})_{\mathrm{ij}}\text{-}(\text{true population wages})_{\mathrm{ij}}]\text{-}[(\text{perceived population wages})_{\mathrm{iJ}}\text{-}(\text{true population wages})_{\mathrm{iJ}}]\} < 0$

revisions, relative to non-attendance, on account of the treatment are zero. The same holds for baseline under-estimators in Panel B.

5.5.3 By Current Stream of Study

Table 4 examines the effect of the treatment on own wage-belief updating, by students' current stream of study, an important determinant of post-secondary education decisions. For each of the three streams- Arts/Humanities, Commerce and Science I first examine updating for all four tracks (Panels A, C, and E) and then updating for the three attendance tracks relative to non-attendance (Panels B, D, and F). For students in the Arts stream, the impact of the information treatment on own wage revisions is statistically insignificant. Individuals revise own-wage beliefs downwards by similar magnitudes for general, vocational and non-attendance tracks, implying small (1.5-2.2 percent) downward revisions for general and vocational tracks relative to non-attendance.

In contrast, wage belief updating for attendance tracks relative to the non-attendance alternative (Panels D & F), are large and statistically significant for Commerce and Science students, but run in opposite directions. Students in the Commerce stream, strongly revise own-wage beliefs downward for all three attendance tracks, relative to the non-attendance alternative (Panel D). Here, the magnitudes of wage belief updating indicate downward revisions of the magnitude 20-30 percent in treatment relative to control groups[16]. This is driven by large downward revisions for the attendance tracks (specifically general and vocational tracks) and no statistically discernible (but upward) updating for the non-attendance track (Panel C).

For the third group of students, students in the Science stream, the treatment induces them to revise relative attendance earnings upward for all three tracks (the effect is statistically significant for technical and general tracks). This is driven entirely by a strong downward revision in wage-beliefs for the non-attendance alternative and no systematic updating for the three attendance tracks looked at separately (Panel E).

I also find that the pattern of relative wage belief updating established in Table 4 is largely driven by females in the sample. A part of the explanation for this could be that females, as compared to males, perceive the returns to attendance, relative to non-attendance, more inaccurately. This is shown in Table 5. Downward revision is much stronger for females in the Commerce stream (the differential ranges from 24 to 33 percent) and the upward revision in the Science stream is driven entirely by females (Panel C). Females in the Science stream revise own-wage beliefs upwards by magnitudes of 43-66 percent in treatment relative to control groups. Table 6 confirms that this pattern holds even if we restrict the sample for these two streams to only baseline over-estimators. Therefore, differential updating between Commerce & Science groups exists holding fixed the direction of baseline error.

Differential updating by Commerce and Science students is further apparent in Table 7

[16] I use the formula prescribed in Kennedy et al. (1981) to interpret interaction terms as the estimating equation is of log-linear form and the independent variables for interest are dummy variables.

where we pool the three attendance tracks and examine own-wage revision as a function of the magnitude of baseline error[17]. A statistically significant "treatment x error" coefficient implies that own-wage revisions are a systematic function of baseline error. A negative coefficient on the interaction term implies that larger amounts of baseline under-estimation are associated with larger upward revisions in treatment relative to control group. Focusing on columns (4)-(6), which include the full set the baseline controls, we see that Commerce students systematically respond to information presented for the attendance-tracks and Science students respond to information presented for the non-attendance alternative.

Track by stream regressions in which wage-revisions cannot be systematically established as a consequence of the treatment indicate that the track-specific information provided was not relevant to the individuals of a given stream. In section 6, I discuss that differences in updating between Commerce and Science students is not a mechanical consequence of sub-group differences in baseline errors. I also provide some evidence that non-updating is consistent with the Bayesian model but cannot rule out the extent to which non-Bayesian updating constitute these findings.

5.6 Impact on Enrollment Intentions

Next, I examine whether updating of own-wage beliefs in response to information on public earnings leads to updating of track-specific enrollment intentions among the sample of students. I examine the impact of the treatment on enrollment in a multinomial logit framework, elaborated in equation 5.

$$\begin{aligned} \eta_{ijt} = log(\frac{\pi_{ijt}}{\pi_{iJt}}) = \alpha + \beta_1 Post + \beta_2 T + \beta_3 Track + \beta_4 (Post \times T) \\ + \beta_5 (T \times Track) + \beta_6 (Post \times Track) + \beta_6 (Post \times T \times Track) + \theta X_{it=1} + u_{ijt} \end{aligned} \tag{5}$$

Here, $log(\pi_{ijt})$ is the stated probability at round t of individual i enrolling in track j and $log(\pi_{iJt})$ is the stated probability of non-attendance. η_{ijt} denotes the odds of choosing in track j as opposed to non-attendance. β_4 and β_6 are coefficients of interest, where β_4 measures the average effect of the treatment on the log-odds that an individual chooses the technical track as opposed to non-attendance and β_6 measures the differential log-odds (relative to the base track technical) of choosing general and vocational tracks.

[17] The regression specification here is:

$$(W_{ijt=2} - W_{ijt=1})^T = \alpha + \beta_1 error^T + \beta_2 T + \beta_3 (error^T \times T) + \theta X_{it=1} + u_{ijt} \tag{4}$$

Here, the dependent variable, W_{ijt}^T which measures own-wage revision and "$error^T$" which measures the difference between perceived and true population wage beliefs, can both take on negative or positive values. Therefore, both variables are transformed using an inverse hyperbolic since transformation, which behaves and is interpreted like a log-transformation, but allows keeping zero and negative values Bellemare, Barrett and Just (2013).

Based on the log-odds regression in (5), the probability of enrolling in each track j is given by:

$$\Pi_{ijt} = \frac{\exp\{\eta_{ijt}\}}{\sum_{j=1}^{J} \exp\{\eta_{ijt}\}} \tag{6}$$

I use equation (6) to compute the predicted probability of enrollment in each track for control and treatment groups in rounds 1 and 2. The marginal effect of the treatment on the predicted probability of enrollment in each track is given by the difference $\{\hat{\Pi}_{jt=2} - \hat{\Pi}_{jt=1}\}_T - \{\hat{\Pi}_{jt=2} - \hat{\Pi}_{jt=1}\}_C$. Where the subscripts T and C are for treatment and control groups, respectively. $\hat{\Pi}_{jt}$ denotes the predicted probability of enrollment in track j, calculated using parameter estimates from (5).

This allows me to track how students allocate probability of enrollment across tracks at baseline and post-treatment.

5.6.1 Full Sample

Table 8 (A) presents the effect of the treatment on the log-odds of pursuing each of the three attendance tracks relative to the base-case of non-attendance. However, we are interested not only in the relative likelihood of enrolling in each track relative to non-enrollment, but in the absolute probability of choosing each track which depends on how the three separate relative effects balance out each other. Therefore, in Table 8 (B), I present the marginal effect of the treatment on the absolute probability of choosing each track. Together, both tables establish that the overall effect of the treatment on enrollment is small and statistically insignificant. This is consistent with the small effect of the treatment on the updating of wage beliefs for the full sample. This result does not importantly change when the sample is broken down by gender or by baseline error.[18]

5.6.2 By Current Stream of Study

In Table 9 (A), I examine the effect of the treatment on the log-odds of pursuing each of the three attendance tracks relative to the base-case of non-attendance, by current stream of study. In Table 10, I present these log-odds separately for males and females.

For students in Arts/Humanities, the effect of the treatment on enrollment log-odds is small, and statistically insignificant and the marginal effect of the treatment on the predicted probability of enrollment is also small Table 9 (B). As can be seen in Table 10, for both males and females in this stream, we cannot reject the null that the effect of the treatment on enrollment log-odds, relative to non-enrollment, for each track is zero. This is consistent with the impact of the treatment on own-wage belief updating for these students.

[18]These result are omitted for brevity but available upon request

For students belonging to the Commerce stream, there is a decrease in enrollment log-odds relative to non-attendance for technical and general streams (Table 9 (A)), and this effect is entirely driven by females (compare panel B to A of Table 10), for whom there is a decrease in enrollment log-odds for all three attendance tracks relative to non-attendance. This is consistent with the fact that for this group of students, the overall effect of the treatment was a downward revision in own-wage beliefs for each of the three attendance tracks, relative to non-attendance. As can be seen in Table Table 9 (B), which takes into account the relative size of track-specific treatment effects from Table 9 (A), for Commerce students, there is a decrease the probability of enrollment in technical tracks by around 4.7 pp. and increase probability of enrollment in general tracks (1.45 pp.), vocational tracks (2.77 pp.) and for non-attendance (0.477 pp.).

For students belonging to the Science stream, there is an increase in enrollment log-odds relative to non-enrollment for all three attendance tracks Table 9 (A). A comparison of panel B to panel A in Table 10 shows that this effect of the treatment on enrollment log-odds is driven by females. This too is consistent with the fact that for this group of students, the overall effect of the treatment was an upward revision in own-wage beliefs for each of the three attendance tracks, relative to non-attendance. As can be seen in Table 9 (B), for Science students there is an increase the probability of enrollment in technical tracks by around 3.45 pp. and decrease probability of enrollment in general tracks (0.505 pp.), vocational tracks (2.28 pp.) and for non-attendance (0.668 pp.).

For both Science and Commerce students, and specifically for females, the direction of updating of enrollment intentions for each attendance track relative to non-attendance, is broadly consistent with the direction of updating of wage beliefs for each attendance track relative to non-attendance. However, the overall effect of the treatment for Commerce students is to induce a movement away from technical tracks and towards other tracks (more so vocational tracks) and on the contrary the effect of the treatment for Science students is to induce a movement towards technical tracks and away from other tracks (more so vocational tracks). This fact cannot be explained by the difference in magnitude of wage-belief revision between attendance tracks, which do not statistically differ from each other, within the groups of Commerce and Science students[19].

5.7 Impact on Borrowing

Recall that borrowing intentions for higher education were measured only post-treatment. Also, unlike stated enrollment intentions, the intent to borrow was not elicited as a track-specific decision. The average impact of the information treatment on borrowing intentions is given by β_1 in equation (7) below:

$$Y_{it=2} = \alpha + \beta_1 T + \theta X_{it=1} + u_{it} \tag{7}$$

[19] I refer here to panels D & F of Table 4. In a variant of these regressions, when tracks are included as an interaction term, I confirm that the magnitude of wage-belief updating for each attendance track does not statistically differ from the magnitude of updating for the other two tracks.

Where Y is a binary variable measuring whether or not the individual would like to accept a loan offer towards higher education enrollment. We may be concerned that the answer to this survey question might not be reflective of what would happen if a loan were actually made available, because individuals might not fully internalize the costs of borrowing while answering this question. While it is not possible to fully allay such concerns in this setting, it should be noted that the control group mean of the fraction of individuals wanting to borrow is only around 56 percent. This is despite the fact that on average only 1.04 tracks (out of 3) are thought to be affordable by individuals. The number of tracks affordable to individuals who do not want to borrow is 1 and this number is 1.08 tracks for individuals who do want to borrow.

Table 11 presents results on the impact of the information treatment on borrowing intentions, which was the only other primary outcome that we measured in our survey. Overall, treated individuals state a higher probability of intent to borrow, which is higher by 6.7 pp. By stream, this effect is largest and statistically significant only for students in the Science stream who have roughly a 15 pp. higher probability of wanting to borrow when compared with students in the control group. This represents a 25 percent increase in demand for borrowing, relative to the control group mean. This result is further consistent with an upward revision for this group of students in wage beliefs for attendance tracks relative to non-attendance. Conditional on loan acceptance, the amount of that individuals would like to borrow does not statistically differ on account of the treatment.

6 Can Heterogeneity in Updating be Explained?

In this section I discuss that while the wage belief updating of Arts students vis-a-vis the other two streams can be attributable to two observed patterns in the data, the differential updating between Science & Commerce students remains unexplained. This points to the existence of substantial heterogeneity in updating heuristics in the sample.

In section 5.3, and with reference to Table A5, we already established that ex-ante, we might expect Arts students to be less responsive to the information treatment and that is indeed what we find. This ex-ante prediction is based on the finding that for these students, at baseline, the enrollment elasticity of earnings beliefs is close to zero and statistically insignificant. Therefore, to the extent that it is costlier for individuals to process or pay attention to information not relevant to their decision-making process, we might expect these students to not update their own-earnings beliefs in response to information provision. Across tracks, the updating of own wage beliefs for Arts students, on account of the treatment, is statistically insignificant and we cannot reject the null that it is zero (first two panels of Table 3). The same analysis indicates that earnings beliefs seem to be important predictors of enrollment intentions for both Commerce & Science students, and the coefficient on log wages for both streams is of roughly equal magnitude. Experimentally, these students update relative own-earnings and also enrollment and borrowing intentions, in a manner consistent with the updating of own-wage beliefs. Therefore differences in the baseline relevance of earnings provide us with one explanation for why we may see differential updating for Arts as compared to Commerce & Science students. It cannot explain why the two sets of students, Commerce & Science, respond to different pieces of information and hence update relative earnings in opposite directions.

Therefore, next I examine whether differential updating can be rationalized on account of differences in baseline errors, regarding population earnings, between sub-groups. In Figure 2 we had established that the distributions of population wage beliefs for Commerce and Science students closely overlay each other. On the other hand, on average, Arts students seem to make smaller errors for all four tracks. This is further evident in Table A6. Here, we look at differences in baseline errors[20] between the three sub-groups in a regression framework, using OLS (col. 1) and quantile regressions (col. 2-6). Focusing our attention on col.1, it is evident that for all tracks, for students in the Arts stream, mean error is statistically significantly smaller (closer to zero error) than for students in the Science stream. However, the mean error for Commerce students does not statistically differ from that of Science students, for any track. Quantile regression results (col. 2-6) also consistently point to the fact that for several points along the distribution of baseline errors, Arts students differ consistently in their perceptions about population earnings when compared to Science students, while Commerce students do not. Therefore, while we cannot rule out the possibility that Arts students update own-earnings differently than students from the other two streams on an additional account of initial differences in perceptions regarding population earnings, this cannot explain differential updating between Commerce & Science streams. Therefore, it is evident that a substantial amount of updating heterogeneity in the sample is, most likely, not a mechanical consequence of differences in baseline errors regarding population earnings.

Next, I test whether the differential updating between Commerce & Science students can be explained by a core prediction of the Bayesian model. That is, I test whether individuals in the sample are more responsive to information regarding a track that they are less likely to enroll in (DellaVigna and Gentzkow (2010); Oreopoulos and Dunn (2013)), and hence have weaker priors about. In Table A7, I establish that Commerce students have higher baseline likelihoods of non-attendance. This is the case both when all four tracks are in an individual's choice set, regardless of affordability (unconstrained choice set), and when only affordable tracks enter an individual's choice set (constrained choice set)[21]. In the unconstrained case, Commerce students state 1.4-2 pp. higher likelihoods of non-attendance compared to the reference category of Science students (who state about a 5.3% likelihood of non-attendance). It is apparent that when the cost constraint is removed and all tracks are hypothetically made available to all individuals, students state generally small probabilities of non-attendance. With regards to data on individuals constrained choice sets, wherein unaffordable options are assigned zero

[20]Measured, as described in footnote 14, using an inverse hyperbolic sine transformation.

[21]In this case tracks unaffordable to an individual are assigned a zero probability of attendance.

probability of attendance, Commerce students are about 6.6-9.2 pp. more likely to not-attend compared to Science students whose probability of non-attendance is about 50%.

In Table A8 I test whether individuals with a higher baseline probability of enrolling in a track are less likely to update earnings beliefs in response to the information treatment. I regress the absolute value of wage belief revision on levels and an interaction of the treatment dummy with the baseline probability of enrollment. While the effect of the treatment does not vary with the baseline probability of enrollment for the unconstrained case, the "treatment x baseline enrollment probability" interaction term is negative and significant (at the 5% level) when affordability of tracks is taken into account. Therefore, the fact that Commerce students are less likely to attend and more responsive to information on attendance tracks and Science students are more likely to attendance and less responsive to information on attendance tracks, is consistent with the core prediction of Bayesian updating bearing out in the data. Previous work in the literature leads us to believe that a portion of non-updating may also be non-Bayesian in nature (i.e. not explained by the variance of priors), but unfortunately in the absence of further data, I cannot comment on the extent to which that may be the case in this sample.

7 Conclusion

In this paper I present results from an information experiment, which randomized information on population earnings, for three post-secondary education tracks- technical, general, vocational and the final alternative of not pursuing post-secondary education. The experiment was carried out with 1525 12^{th} grade students, across nine affiliated schools of a large, non-selective public state university in the Indian state of Jharkhand.

The impact of information provision is measured by students' updating of own wage beliefs contingent on pursuing each post-secondary track, their stated probability of enrollment across tracks and borrowing intentions for higher education enrollment. Average impact of the treatment on the updating of own-wage beliefs and subsequent changes in enrollment intentions is small, though the impact on borrowing intentions is positive and statistically important. Average results mask considerable sub-group heterogeneity, defined by the current subject stream of students. For two out of three sub-groups of students, the impact of the treatment on relative own-earnings beliefs, is statistically important. For these two sub-groups, own-wage belief updating is stronger for females, a pattern which may partly be on account of the fact that females perceive relative returns to attendance more inaccurately than males, at baseline. However, I also find that females are more responsive to the information treatment controlling for the size of baseline error (result omitted for brevity). Interestingly, Wiswall and Zafar (2015*b*) also find this to be the case in their sample of New York University (NYU) students. Within sub-group, the odds of enrollment, relative to non-enrollment, is consistent with direction of wage belief updating and is, reassuringly, stronger for the group (females) with larger wage belief updating. The effect of the treatment on borrowing intentions is also in line it's effect on wage-belief updating.

For the two sub-groups (Commerce & Science) for whom the impact of the treatment on wage beliefs for attendance tracks, relative to non-attendance, is statistically important, the updating takes place in opposite directions. This pattern is on account of the fact that individuals in the groups systematically respond to different pieces of track-specific information. Students in the Commerce sub-group revise wage beliefs downwards for attendance tracks only, and this pattern carries over to a downward revision in earnings beliefs for attendance tracks, relative to non-attendance. Science students revise wage beliefs downwards for the non-attendance track only, which translates into attendance earnings being relatively more attractive for this group, post-treatment.

A combination of factors may explain why we see the first sub-group, students in the Arts stream, not revise wage beliefs in response to the treatment. Ex-ante, these students have a low elasticity of enrollment to wage beliefs and at baseline they make smaller errors, on average, with regards to beliefs about population earnings. However, these factors do not explain differential updating on the part of Science & Commerce students. Ex-ante, these students have a statistically important and similar in magnitude elasticity of enrollment to wage beliefs- therefore earnings likely play an important role in their decisions for future education. These students also make nearly identical errors with regards to population wage beliefs at baseline.

As discussed in the conceptual framework (section 3) of the paper, non-updating in response to "new" information implies that the piece of information provided was not relevant to individuals. Differential relevance of track-specific information to different sub-groups of individuals in the sample drives the heterogeneous impacts of information provision on own-wage belief updating. Sub-groups of individuals with equally biased information sets vary in their response to information significantly, depending on the extent to which their beliefs about population earnings are linked to their beliefs about their own. Track-specific non-updating can be Bayesian or "rational" or Non-Bayesian. Suggestive evidence implies that variation in updating may be attributable to the variance of individuals' priors (consistent with Bayesian updating), but without further data on individual-level distributions of own-wage beliefs, we cannot quantify the extent to which non-updating is Bayesian. However, recent evidence (Wiswall and Zafar (2015*b*)) establishes that individuals do indeed deviate significantly from the Bayesian benchmark.

Recent papers in the literature which find information provision to have small average impacts on outcomes like test-scores or enrollment decisions state the presence of other binding constraints like credit constraints or lack of knowledge of the education production function as explanations. This paper offers another explanation for why the provision of population-level information on returns may lead to highly heterogeneous outcomes, by examining in detail the first-link in the causal chain that links population-level information to education outcomes, which is, the extent to which individuals update own-earnings beliefs in response to receiving information about population-level averages.

My study is a framed field experiment wherein participants deal with a subject of interest outside the experiment (their own education) but not in an environment where they would naturally undertake the task of thinking about their long term plans. Stakes for the participants were also low with no costs to paying less attention to the information provided. Therefore, the study was not designed to provide a model for scaling information provision at a national level, but to examine in more detail mechanisms (updating of beliefs and intentions) critical to the success of information interventions. For a policy-maker looking to implement an information campaign to induce more optimal education decision making, these findings imply, all else constant, a limited potential for an information campaign to induce, on average, updating of beliefs regarding oneself, of a particular magnitude and in a given direction, despite accurate knowledge of information gaps in a particular population.

References

Arcidiacono, Peter, V Joseph Hotz, and Songman Kang. 2012. "Modeling college major choices using elicited measures of expectations and counterfactuals." *Journal of Econometrics*, 166(1): 3–16.

Avitabile, Ciro, and Rafael E De Hoyos Navarro. 2015. "The Heterogeneous effect of information on student performance: evidence from a randomized control trial in Mexico." *World Bank Policy Research Working Paper*, , (7422).

Bellemare, Marc F, Christopher B Barrett, and David R Just. 2013. "The welfare impacts of commodity price volatility: evidence from rural Ethiopia." *American Journal of Agricultural Economics*, 95(4): 877–899.

Bonilla, Leonardo, Nicolas L Bottan, and Andres Ham. 2016. "Information Policies and Higher Education Choices Experimental Evidence from Colombia."

Delavande, Adeline, and Basit Zafar. 2014. "University choice: the role of expected earnings, non-pecuniary outcomes, and financial constraints." *FRB of New York Staff Report*, , (683).

DellaVigna, Stefano, and Matthew Gentzkow. 2010. "Persuasion: Empirical Evidence." *Annual Review of Economics*, 2(1): 643–669.

Duflo, Esther, Rachel Glennerster, and Michael Kremer. 2007. "Using randomization in development economics research: A toolkit." *Handbook of development economics*, 4: 3895–3962.

Dupas, Pascaline. 2011. "Do teenagers respond to HIV risk information? Evidence from a field experiment in Kenya." *American Economic Journal: Applied Economics*, 3(1): 1–34.

Fryer Jr, Roland G. 2013. "Information and student achievement: evidence from a cellular phone experiment." National Bureau of Economic Research.

Gamboa, Luis, and Paul Rodríguez Lesmes. 2014. "Do Colombian students underestimate higher education returns?" UNIVERSIDAD DEL ROSARIO.

Hanna, Rema, Sendhil Mullainathan, and Joshua Schwartzstein. 2014. "Learning through noticing: Theory and evidence from a field experiment." *The Quarterly Journal of Economics*, 129(3): 1311–1353.

Harrison, Glenn W, and John A List. 2004. "Field experiments." *Journal of Economic literature*, 42(4): 1009–1055.

Hastings, Justine, Christopher A Neilson, and Seth D Zimmerman. 2015. "The effects of earnings disclosure on college enrollment decisions." National Bureau of Economic Research.

Jensen, Robert. 2010. "The (perceived) returns to education and the demand for schooling." *Quarterly Journal of Economics*, 125(2).

Kahneman, Daniel, and Amos Tversky. 1972. "Subjective probability: A judgment of representativeness." *Cognitive psychology*, 3(3): 430–454.

Kaufmann, Katja Maria. 2014. "Understanding the income gradient in college attendance in Mexico: The role of heterogeneity in expected returns." *Quantitative Economics*, 5(3): 583–630.

Kennedy, Peter E, et al. 1981. "Estimation with correctly interpreted dummy variables in semilogarithmic equations [the interpretation of dummy variables in semilogarithmic equations]." *American Economic Review*, 71(4).

Loyalka, Prashant, Chengfang Liu, Yingquan Song, Hongmei Yi, Xiaoting Huang, Jianguo Wei, Linxiu Zhang, Yaojiang Shi, James Chu, and Scott Rozelle. 2013. "Can information and counseling help students from poor rural areas go to high school? Evidence from China." *Journal of Comparative Economics*, 41(4): 1012–1025.

Maertens, Annemie. 2011. "Does education pay off? Subjective expectations on education in rural India." *Economic & Political Weekly*, 46(09): 58–63.

Manski, Charles F. 1993. "Adolescent econometricians: How do youth infer the returns to schooling?" In *Studies of supply and demand in higher education.* 43–60. University of Chicago Press.

Nguyen, Trang. 2008. "Information, role models and perceived returns to education: Experimental evidence from Madagascar." *Unpublished manuscript*, 6.

Oreopoulos, Philip, and Ryan Dunn. 2013. "Information and college access: Evidence from a randomized field experiment." *The Scandinavian Journal of Economics*, 115(1): 3–26.

Osman, Adam. 2014. "Occupational choice under credit and information constraints." *Available at SSRN 2449251.*

Pekkala Kerr, Sari, Tuomas Pekkarinen, Matti Sarvimäki, and Roope Uusitalo. 2015. "Post-Secondary Education and Information on Labor Market Prospects: A Randomized Field Experiment."

Shrestha, Maheshwor. 2016. "Get rich or die tryin': Perceived earnings, perceived mortality rate and the value of a statistical life of potential work-migrants from Nepal."

Sims, Christopher A. 2003. "Implications of rational inattention." *Journal of monetary Economics*, 50(3): 665–690.

Stinebrickner, Ralph, and Todd Stinebrickner. 2008. "The Effect of Credit Constraints on the College Drop-Out Decision: A Direct Approach Using a New Panel Study." *The American economic review*, 98(5): 2163–2184.

Wiswall, Matthew, and Basit Zafar. 2015*a*. "Determinants of college major choice: Identification using an information experiment." *The Review of Economic Studies*, 82(2): 791–824.

Wiswall, Matthew, and Basit Zafar. 2015*b*. "How Do College Students Respond to Public Information about Earnings?" *Journal of Human Capital*, 9(2): 117–169.

Zafar, Basit. 2013. "College major choice and the gender gap." *Journal of Human Resources*, 48(3): 545–595.

Figures & Tables

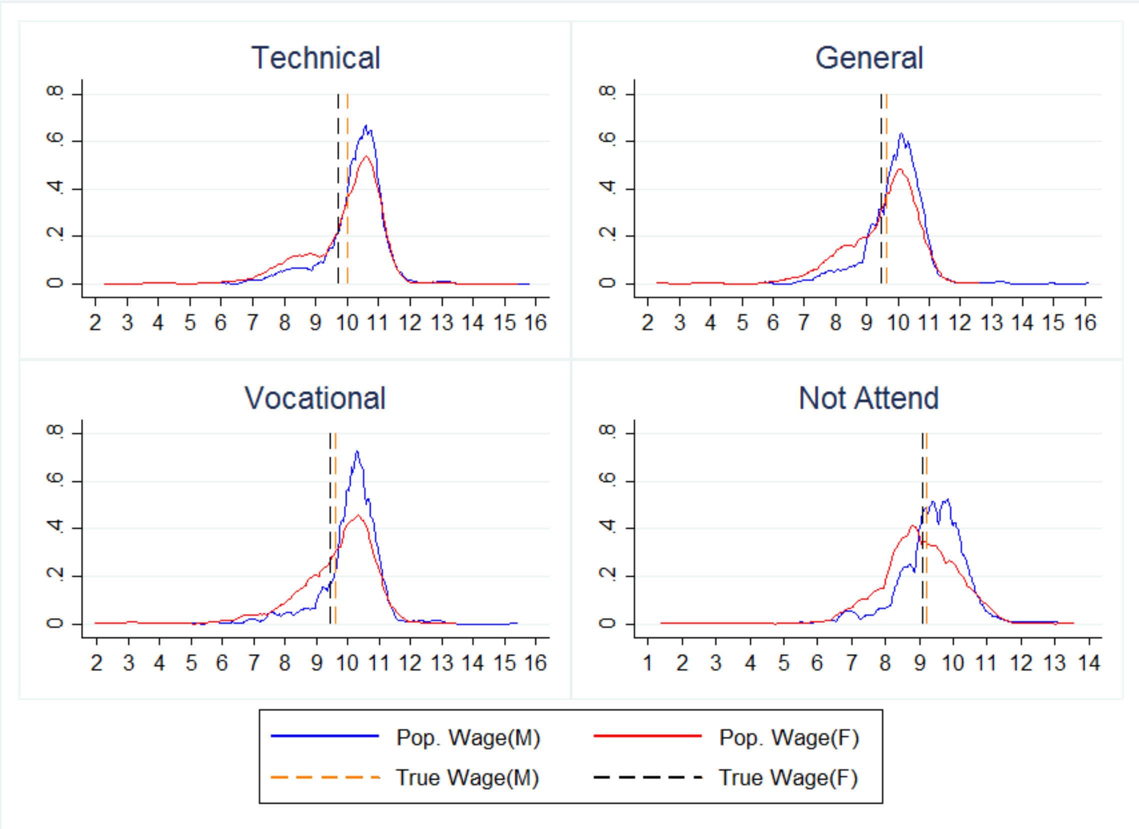


Figure 1: Log Population Wage Beliefs by Gender

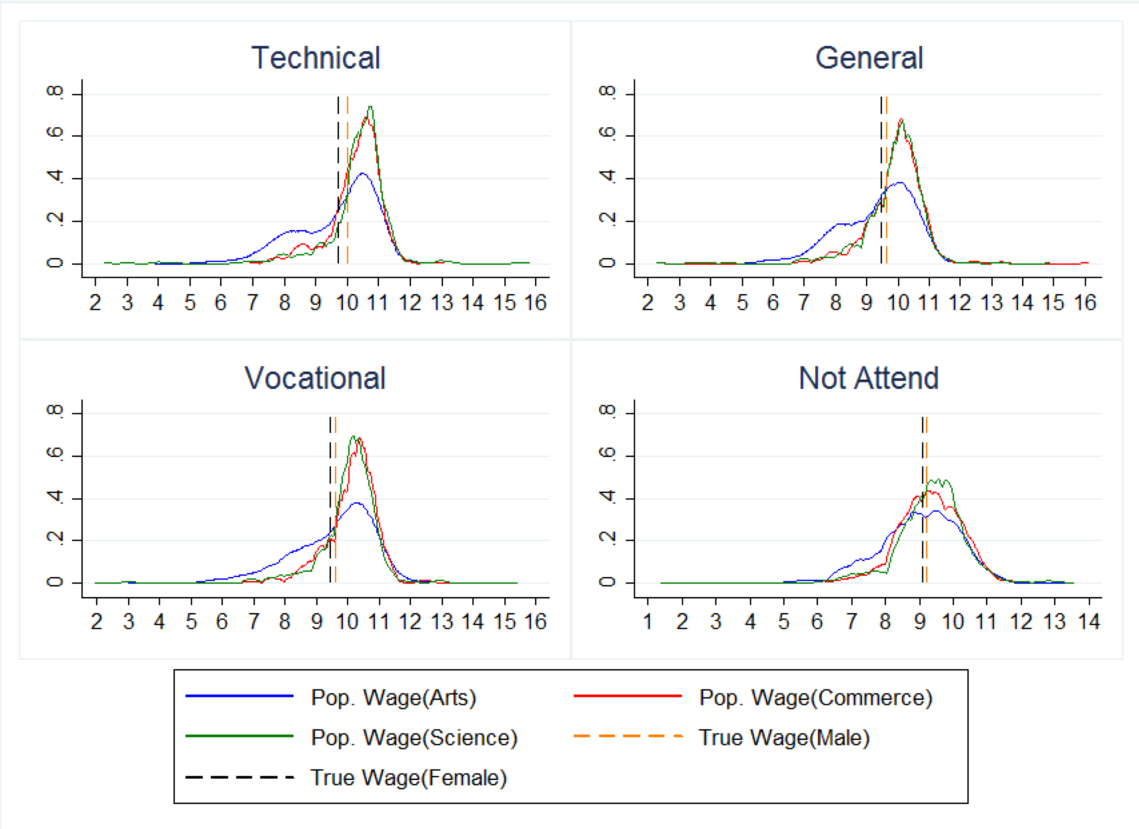


Figure 2: Log Population Wage Beliefs by Current Stream of Study

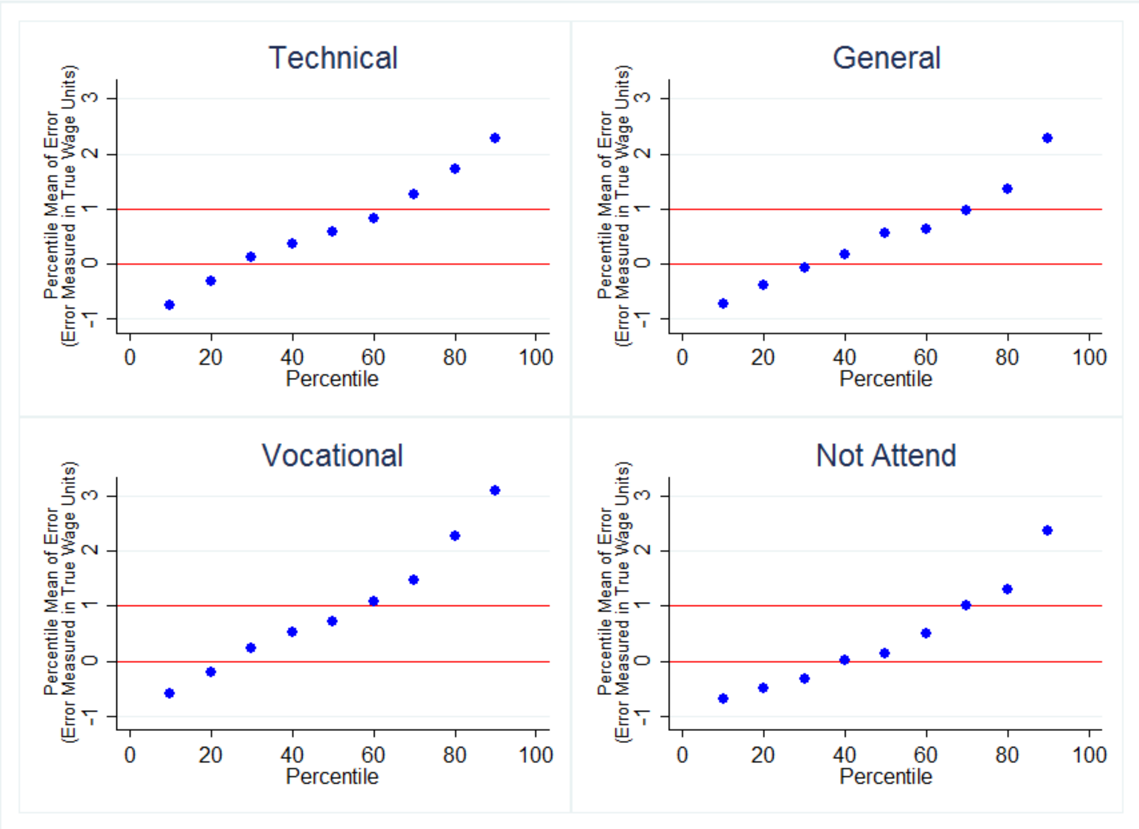


Figure 3: Scatter Plot of Baseline Population Errors Relative to True Wage (Zero-Error Line)

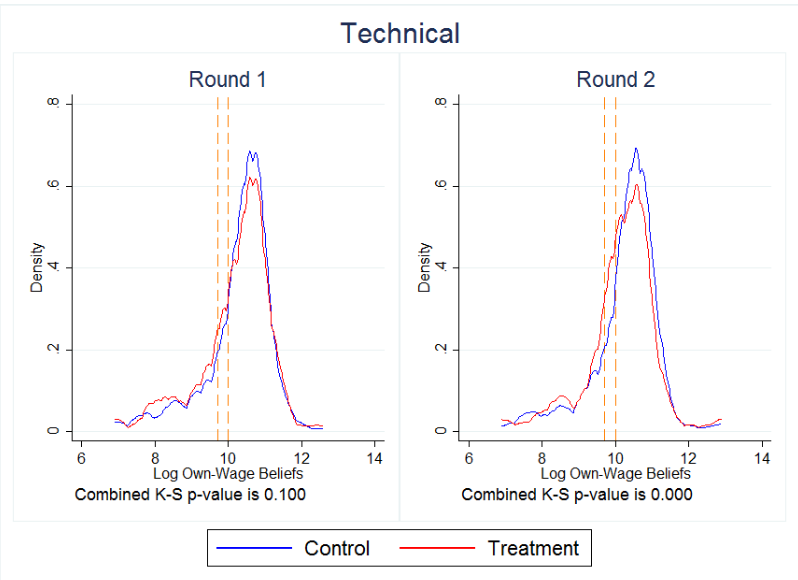


Figure 4: Pre and Post Distributions of Own Wage Beliefs for Technical Track; Range 1-99 Percentile

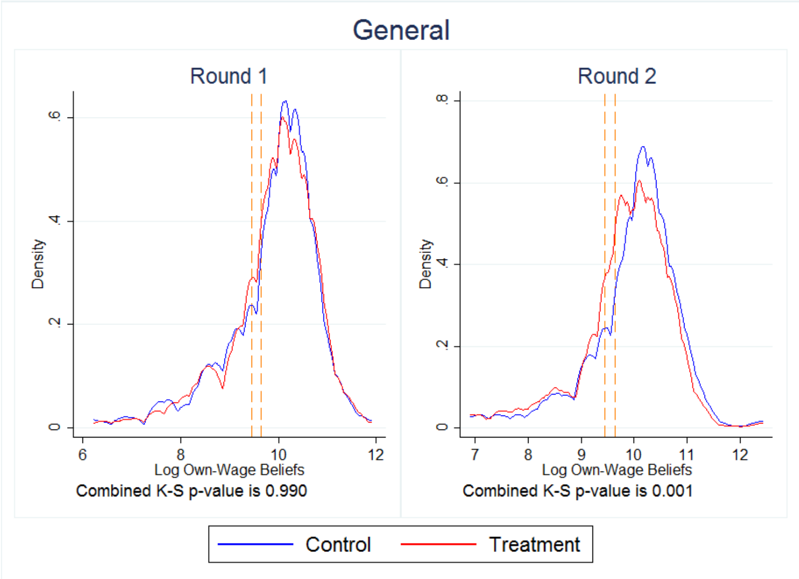


Figure 5: Pre and Post Distributions of Own Wage Beliefs for General Track; Range 1-99 Percentile

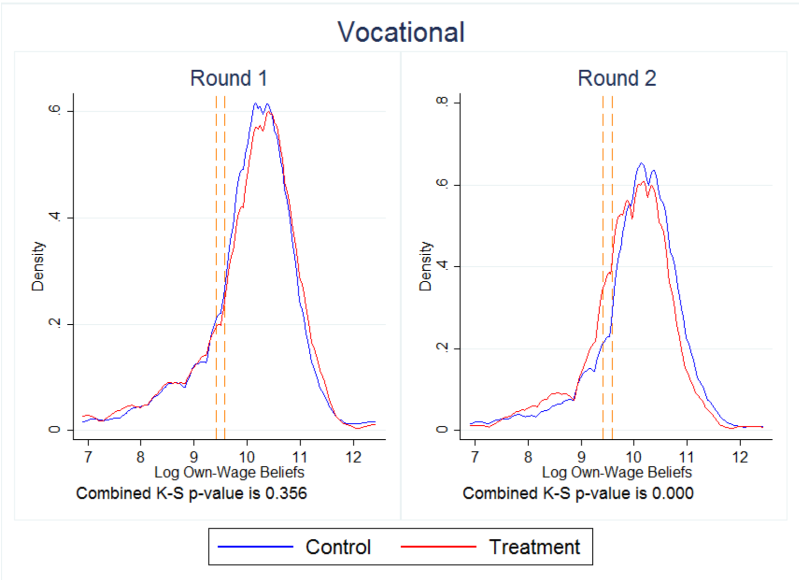


Figure 6: Pre and Post Distributions of Own Wage Beliefs for Vocational Track; Range 1-99 Percentile

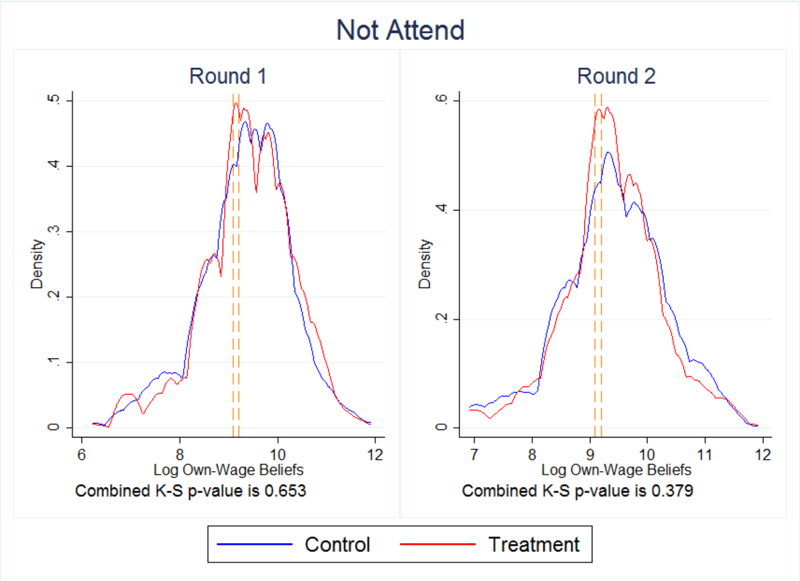


Figure 7: Pre and Post Distributions of Own Wage Beliefs for Non-Attendance Track; Range 1-99 Percentile

Table 1: % of Students Who Overestimate Earnings at Baseline

Full Sample:			
	Overall	Males	Females
Technical	71.41%	72.24%	70.43%
General	64.13%	66.06%	61.86%
Vocational	75.34%	83.39%	65.86%
Not Attend	60.79%	70.55%	49.29%
Arts:			
	Overall	Males	Females
Technical	55.49%	53.22%	57.34%
General	49.52%	49.79%	49.30%
Vocational	61.08%	71.24%	52.80%
Not Attend	53.18%	63.09%	45.10%
Commerce:			
	Overall	Males	Females
Technical	77.19%	73.19%	81.20%
General	73.35%	74.47%	72.22%
Vocational	82.52%	87.23%	77.78%
Not Attend	62.47%	72.34%	52.56%
Science:			
	Overall	Males	Females
Technical	81.72%	84.03%	77.09%
General	70.34%	71.15%	68.72%
Vocational	82.84%	88.80%	70.95%
Not Attend	66.79%	74.23%	51.96%

Table 2: Impact of the Information Treatment on Own Wage Beliefs for Full Sample

	(1)	(2)	(3)	(4)
Panel A:	Log Own Wage (Technical)	Log Own Wage (General)	Log Own Wage (Vocational)	Log Own Wage (Not Attend)
post x treatment	-0.00742	-0.0969	-0.140**	-0.0679
	(0.0564)	(0.0592)	(0.0578)	(0.0649)
Observations	2,961	2,961	2,961	2,955
Panel B:	Log Own Wage (Tech./NA)	Log Own Wage (Gen./NA)	Log Own Wage (Voc./NA)	
post x treatment	0.0606	-0.0300	-0.0743	-
	(0.0638)	(0.0633)	(0.0632)	-
Observations	2,955	2,955	2,955	-
School FE	YES	YES	YES	YES
Stream FE	YES	YES	YES	YES

Notes: Standard errors clustered at the survey session level. Number of clusters: 106.
*** $p<0.01$ ** $p<0.05$ * $p<0.1$. All main effects are included.
Results are robust to addition of baseline controls.

Table 3: Impact of the Information Treatment on Own Wage Beliefs by Baseline Error

	(1) Log Own Wage (Technical)	(2) Log Own Wage (General)	(3) Log Own Wage (Vocational)	(4) Log Own Wage (Not Attend)
Under-estimators:				
Panel A:				
post x treatment	0.0803	-0.0238	0.0499	0.0535
	(0.130)	(0.114)	(0.134)	(0.0992)
Observations	847	1,058	730	1,148
	Log Own Wage (Tech./NA)	Log Own Wage (Gen./NA)	Log Own Wage (Voc./NA)	
Panel B:				
post x treatment	0.146	0.0801	-0.0505	-
	(0.112)	(0.102)	(0.126)	-
Observations	1,169	1,338	897	-
Over-estimators:				
	Log Own Wage (Technical)	Log Own Wage (General)	Log Own Wage (Vocational)	Log Own Wage (Not Attend)
Panel C:				
post x treatment	-0.0434	-0.131**	-0.188***	-0.118
	(0.0565)	(0.0638)	(0.0586)	(0.0742)
Observations	2,114	1,903	2,231	1,807
	Log Own Wage (Tech./NA)	Log Own Wage (Gen./NA)	Log Own Wage (Voc./NA)	
Panel D:				
post x treatment	0.00196	-0.120	-0.0830	-
	(0.0758)	(0.0726)	(0.0713)	-
Observations	1,786	1,617	2,058	-
School FE	YES	YES	YES	YES
Stream FE	YES	YES	YES	YES

Notes: Standard errors clustered at the survey session level. Number of clusters: 102 (under)- 106 (over). *** p<0.01 ** p<0.05 * p<0.1. All main effects are included. Results are robust to addition of baseline controls.

Table 4: Impact of the Information Treatment on Own Wage Beliefs by Stream

	(1)	(2)	(3)	(4)
	Log Own Wage (Technical)	Log Own Wage (General)	Log Own Wage (Vocational)	Log Own Wage (Not Attend)
Arts:				
Panel A:				
post x treatment	0.0756	-0.140	-0.147	-0.124
	(0.130)	(0.122)	(0.136)	(0.116)
Observations	1,010	1,010	1,010	1,010
	Log Own Wage (Tech./NA)	Log Own Wage (Gen./NA)	Log Own Wage (Voc./NA)	
Panel B:				
post x treatment	0.200	-0.0159	-0.0223	-
	(0.138)	(0.125)	(0.139)	-
Observations	1,010	1,010	1,010	-
Commerce:				
	Log Own Wage (Technical)	Log Own Wage (General)	Log Own Wage (Vocational)	Log Own Wage (Not Attend)
Panel C:				
post x treatment	-0.0455	-0.182*	-0.163**	0.169
	(0.0805)	(0.102)	(0.0809)	(0.102)
Observations	917	917	917	912
	Log Own Wage (Tech./NA)	Log Own Wage (Gen./NA)	Log Own Wage (Voc./NA)	
Panel D:				
post x treatment	-0.215**	-0.349***	-0.330***	-
	(0.103)	(0.119)	(0.0904)	-
Observations	912	912	912	-
Science:				
	Log Own Wage (Technical)	Log Own Wage (General)	Log Own Wage (Vocational)	Log Own Wage (Not Attend)
Panel E:				
post x treatment	-0.0534	0.0232	-0.108	-0.218**
	(0.0725)	(0.0831)	(0.0730)	(0.110)
Observations	1,034	1,034	1,034	1,033
	Log Own Wage (Tech./NA)	Log Own Wage (Gen./NA)	Log Own Wage (Voc./NA)	
Panel F:				
post x treatment	0.167*	0.238***	0.102	-
	(0.0984)	(0.0835)	(0.0903)	-
Observations	1,033	1,033	1,033	-
School FE	YES	YES	YES	

Notes: Standard errors clustered at the survey session level. Number of clusters: 90-92 (range). *** p<0.01 ** p<0.05 * p<0.1. All main effects are included. Results are robust to addition of baseline controls

Table 5: Differential Impact of Information on Own Wage Beliefs for Females
By Current Stream of Study

Dependent Variables:	(1) Log Own Wage (Tech./NA)	(2) Log Own Wage (Gen./NA)	(3) Log Own Wage (Voc./NA)
Panel A; Arts/Humanities:			
post x treatment	0.0505	-0.183	-0.205
	(0.181)	(0.137)	(0.155)
post x treatment x female	0.274	0.315	0.332
	(0.267)	(0.252)	(0.245)
Observations	1,010	1,010	1,010
Panel B; Commerce:			
post x treatment	-0.0922	-0.159	-0.208*
	(0.125)	(0.144)	(0.116)
post x treatment x female	-0.248	-0.376*	-0.235
	(0.205)	(0.218)	(0.187)
Observations	912	912	912
Panel C; Science:			
post x treatment	0.0272	0.110	-0.0899
	(0.105)	(0.0778)	(0.0861)
post x treatment x female	0.377**	0.369**	0.529***
	(0.178)	(0.166)	(0.170)
Observations	1,033	1,033	1,033
School FE	YES	YES	YES

Notes: Standard errors clustered at the survey session level. Number of clusters: 90. *** $p<0.01$ ** $p<0.05$ * $p<0.1$. All main effects and two-way interactions included. Results are robust to addition of baseline controls

Table 6: Differential Impact of Information on Own Wage Beliefs by Gender
For Over-estimators in Commerce & Science

	(1)	(2)	(3)
Dependent variable:	Log Own Wage (Tech./NA)	Log Own Wage (Gen./NA)	Log Own Wage (Voc./NA)
Commerce-Males:			
post x treatment	-0.315**	-0.325*	-0.302***
	(0.150)	(0.169)	(0.108)
Observations	270	272	348
Commerce-Females:			
post x treatment	-0.361**	-0.655***	-0.455***
	(0.148)	(0.164)	(0.134)
Observations	329	314	323
Science-Males:			
post x treatment	0.0952	0.0105	-0.153*
	(0.132)	(0.0999)	(0.0840)
Observations	493	409	575
Science-Females:			
post x treatment	0.393*	0.363*	0.517**
	(0.204)	(0.208)	(0.193)
Observations	235	203	235
School FE	YES	YES	YES

Notes: Standard errors clustered at the survey session level. Number of clusters: 44-74 (range). *** p<0.01 ** p<0.05 * p<0.1. All main effects are included. Results are robust to addition of baseline controls.

Table 7: Revision in Own-Wage Beliefs as a Continuous Function of Baseline Error

	(1) Arts	(2) Commerce	(3) Science	(4) Arts	(5) Commerce	(6) Science
Dependent variable:	$OwnWage^{T}_{ijt=2} - OwnWage^{T}_{ijt=1}$					
Attendance Tracks:						
treatment x error	-0.00509	-0.0206*	-0.00523	-0.00548	-0.0214**	-0.00814
	(0.00945)	(0.0107)	(0.0105)	(0.00954)	(0.00916)	(0.00991)
Observations	1,473	1,344	1,494	1,473	1,344	1,494
Non Attendance Track:						
treatment x error	-0.00723	-0.00834	-0.0268***	-0.00600	-0.0160	-0.0289***
	(0.0119)	(0.0129)	(0.00997)	(0.0121)	(0.0123)	(0.00925)
Observations	491	448	498	491	448	498
School FE	YES	YES	YES	YES	YES	YES
Add. Baseline Controls	NO	NO	NO	YES	YES	YES

Notes: Standard errors clustered at the survey session level. Number of clusters: 88-90 (range)
*** p<0.01 ** p<0.05 * p<0.1. All main effects are included.

Table 8A: Impact of the Information Treatment on Enrollment for Full Sample

Dependent Variable:	Enrollment Log-Odds
post x treatment	0.0612
	(0.126)
post x treatment x general	0.0415
	(0.112)
post x treatment x vocational	0.0425
	(0.101)
Observations (3 tracks x round)	8,880
School FE	YES
Stream FE	YES

Notes: Standard errors clustered at the survey session level. Number of clusters: 106. *** p<0.01 ** p<0.05 * p<0.1. All main effects and two-way interactions are included. Results are robust to addition of baseline controls.

Table 8B: Marginal Effects on the Predicted Probability of Enrollment
(Parameter estimates from Table 8A)

	Technical	General	Vocational	Not Attend
	-0.904%	0.856%	0.452%	-0.404%

Table 9A: Impact of the Information Treatment on Enrollment
By Current Stream of Study

	Arts/Humanities	**Commerce**	**Science**
Dependent variable:	Enrollment Log Odds	Enrollment Log Odds	Enrollment Log Odds
post x treatment	0.102	-0.255	0.317**
	(0.212)	(0.210)	(0.157)
post x treatment x general	0.0617	0.162	-0.0913
	(0.191)	(0.179)	(0.163)
post x treatment x vocational	0.0416	0.283*	-0.179
	(0.184)	(0.169)	(0.144)
Observations	3,030	2,751	3,099
School FE	YES	YES	YES

Notes: Standard errors clustered at the survey session level. Number of clusters: 90. *** $p<0.01$ ** $p<0.05$ * $p<0.1$. All main effects and two-way interactions are included. Results are robust to addition of baseline controls.

Table 9B: Marginal Effects of the Treatment on the Predicted Probability of Enrollment
(Parameter estimates from Table 9A)

	Technical	General	Vocational	Not Attend
Arts	-1.232%	1.636%	0.451%	-0.855%
Commerce	-4.705%	1.456%	2.772%	0.477%
Science	3.456%	-0.505%	-2.283%	-0.668%

Table 10: Impact of the Information Treatment on Enrollment
By Current Stream of Study; Effects by Gender

	Arts/Humanities	**Commerce**	**Science**
Dependent Variables:	Enrollment Log Odds	Enrollment Log Odds	Enrollment Log Odds
Panel A; Males:			
post x treatment	0.292	0.0811	0.174
	(0.287)	(0.259)	(0.193)
post x treatment x general	0.264	-0.100	-0.0386
	(0.234)	(0.191)	(0.175)
post x treatment x vocational	-0.0121	0.120	-0.203
	(0.230)	(0.213)	(0.161)
Observations	1,374	1,377	2,070
Panel B; Females:			
post x treatment	-0.0566	-0.611*	0.589**
	(0.301)	(0.322)	(0.283)
post x treatment x general	-0.0857	0.404	-0.184
	(0.281)	(0.289)	(0.310)
post x treatment x vocational	0.0823	0.454*	-0.129
	(0.264)	(0.250)	(0.270)
Observations	1,656	1,374	1,029

Notes: Standard errors clustered at the survey session level. Number of clusters: 49-77 (range). *** p<0.01 ** p<0.05 * p<0.1. All main effects and two-way interactions are included. Results are robust to addition of baseline controls.

Table 11: Impact of the Treatment on Borrowing Probability

	(1) Full Sample	(2) Arts	(3) Commerce	(4) Science
treatment	0.0665*	-0.0360	0.0571	0.147***
	(0.0353)	(0.0573)	(0.0555)	(0.0399)
mean dep. control group	0.5612	0.5447	0.5267	0.6088
Observations	1,437	491	448	498
School FE	YES	YES	YES	YES

Notes: Standard errors clustered at the survey session level. Number of clusters: 88-106 (range). *** p<0.01, ** p<0.05, * p<0.1.
Results are robust to addition of baseline controls.

Appendix 1 Survey Details

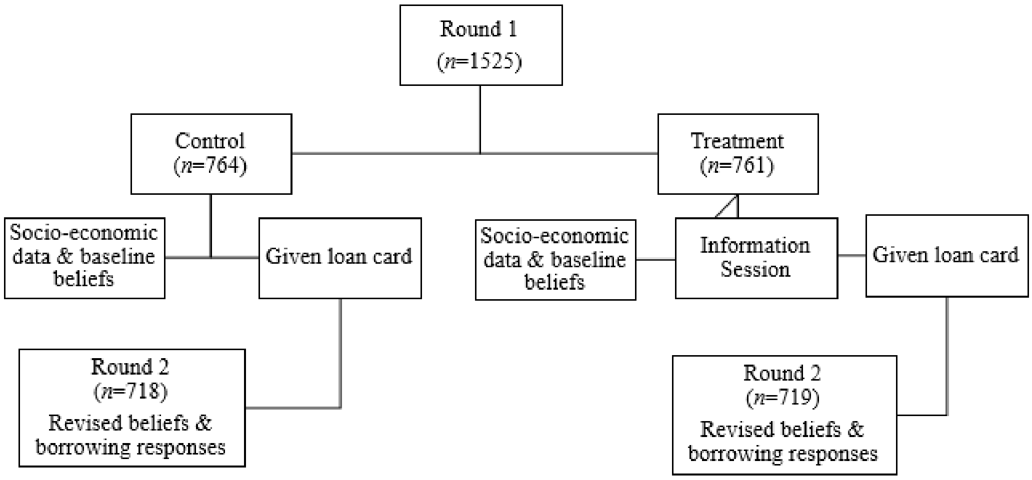


Figure A1: Survey Structure & Experimental Design

English Translation of Information Script

" Your contribution to the first round of the survey is now over. However, before you leave we would like to talk to you about a few more things. Now onwards, you do not need to fill anything else on the tablet, only pay attention towards the screen.

You have with you the printout of the graph and table that you are seeing on the screen (**attached at the end of the script**). You are encouraged to take this home with you.

With the help of the information on the screen, we would like to provide you with some findings from recent survey data collected by the Government of India. Every few years, the Government surveys a sample of people from all Indian states and asks them about what occupation they are currently engaged in and their weekly earnings in that occupation. This survey allows us to see what the average earnings are for different individuals without having to guess or just go by what we hear from a few people. The information presented uses data on roughly 40,000 people. The data was collected between 2009 & 2012.

Now, we would like to draw your attention to the graph on the screen. This graph shows the average monthly earnings of four groups of people. The graph on the left side is for men and the graph on the right side is for women. According to the graph on the left, for men who have completed a technical degree, their average monthly earnings are around 22,070 rupees. Similarly, if we talk about men who have completed a general degree, their average monthly earnings are 15,280 rupees. Those who have obtained a diploma or completed certificate course, their average monthly earnings are around 14,500 rupees and for men not studying after intermediate, their average monthly earnings are around 9,970 rupees.

Similarly, for women who have completed a technical degree, their average monthly earnings are around 16,450 rupees. If we talk about women who have completed a general degree, their average monthly earnings are 12,750 rupees. Those women who have completed a diploma or certificate course, their average monthly earnings are around 12,200 rupees and for women not studying after intermediate, their average monthly earnings are around 8,900 rupees.

It is important to keep in mind that average earnings do not imply that every individual in that group earns the average amount. Some people earn more than the average and some people earn less. For this reason, for every higher education group, we will now try to explain to you what the lower & higher amounts earned by people in that group are.

Now we would like to draw your attention towards the table in the slide. If we look at the data of men and women who have obtained a technical degree, we see that 25 percent of men earn approximately 11,780 rupees or less and 25 percent of women earn approximately 6,200 rupees or less. In the category of men who earn a technical degree, 95 percent of people, in a month, earn 51,400 rupees or less and 95 percent of women who have earned a technical degree, earn approximately 42,800 rupees or less.

This means that 51,400 rupees is the 95th percentile of men who have technical degrees and 42,800 rupees is the 95th percentile of women who earn technical degrees. This also means that very few men in this group earn more than 51,400 rupees per month and very few women earn more than 42,800 rupees per month.

*Before proceeding, ask all students if they understand the meaning of "percentile" and if they have any questions. (**Enumerators were encouraged to have a discussion around the concept of a percentile.**)*

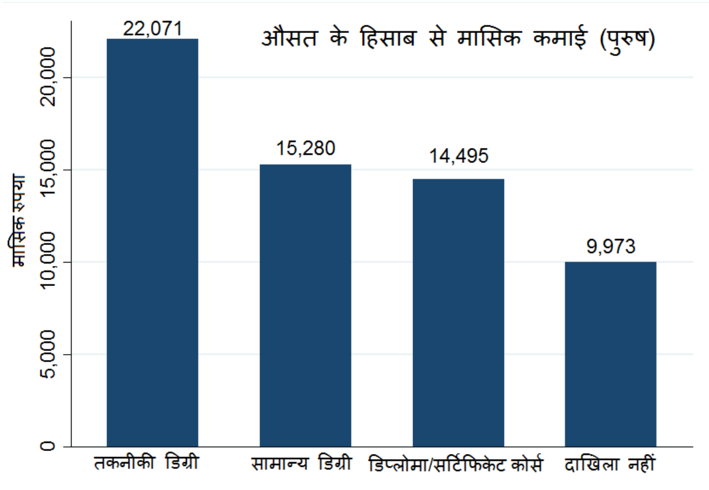
Let's talk more about the data of men and women who have obtained a general degree. The 25th percentile for men in this group is 7,500 rupees and for women it is 4,500 rupees. The 95th percentile for men in this group is 36,000 rupees and for women it is around 32,120 rupees. As discussed before, this means that very few men in the group earn more than 36,000 rupees per month and very few women in this group earn more than 32,120 rupees per month.

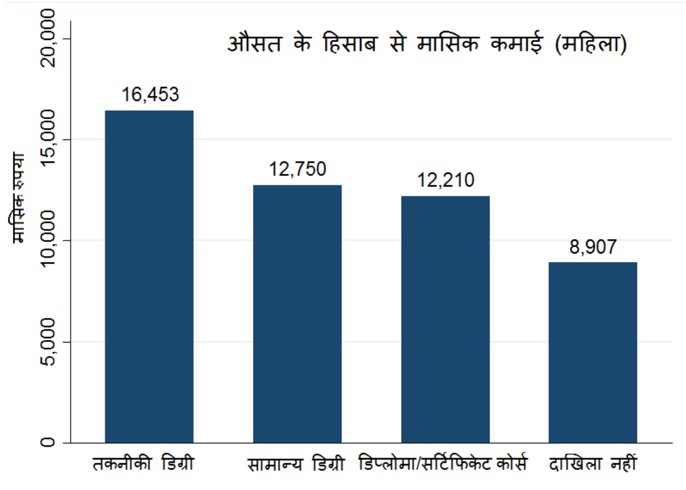
Now, let's talk about the data of men and women who have obtain a vocational degree. The 25th percentile for men in this group is around 6,430 rupees and for women it is 4,500 rupees. The 95th percentile for men in this group is 36,000 rupees and for women it is around 30,000 rupees. As discussed before, this means that very few men in the group earn more than 36,000 rupees per month and very few women in this group earn more than 30,000 rupees per month.

Finally, let's talk about the data of men and women who do not study beyond intermediate. The 25th percentile for men in this group is around 4,290 rupees and for women it is around 2,680 rupees. The 95th percentile for men in this group is around 26,000 rupees and for women it is 24,990 rupees. As discussed before, this means that very few men in the group earn more than 26,000 rupees per month and very few women in this group earn more than 24,990 rupees per month. Again, the information that we have discussed with you today, is also given in the printout that is with you. Please look at it again at home. "

(Information Sheet on Next Page. Students in the treatment group took this sheet home and the same sheet was projected on screen during the information discussion.)

सरकारी डेटा से कुछ जानकारी





	तकनीकी डिग्री		सामान्य डिग्री		डिप्लोमा / सर्टिफिकेट कोर्स		दाखिला नहीं	
	पुरुष	महिला	पुरुष	महिला	पुरुष	महिला	पुरुष	महिला
25 प्रतिशतक[1]	11,786	6,206	7,500	4,500	6,429	4,500	4,286	2,679
95 प्रतिशतक[2]	51,429	42,857	36,000	32,117	36,000	30,000	26,001	24,990

[1] इसका मतलब है की 25 प्रतिशत पुरुष / महिला हर माह दिए गये रकम या इससे कम कमाते हैं।

[2] इसका मतलब है की 95 प्रतिशत पुरुष / महिला हर माह दिए गये रकम या इससे कम कमाते हैं।

Figure A2: Image of Information Sheet Accompanying Script

English Translation of Loan Script

“We want to place in front of you, one other hypothetical situation, for you to think at home about. Not all students are able to continue their studies after intermediate. Often, their household income is not enough for them to continue their studies or for them to enroll in a higher education program of their choice. In some such situations, bank or non-bank institutions are able to offer higher education loans, at fair interest rates, which you have to repay after completing your higher education.

Some of you must be wondering what is meant by the term “interest rate”. Suppose you take a loan of 100 rupees on which there is a 10% interest rate. When repaying this loan, you have to return 110 rupees. This additional 10 rupees that you pay is your interest. Fair interest rate means a rate that is neither too low or neither too high.

Please read out the questions on the loan card given to the students.

Loan for Higher Education

Q1. Suppose that someone (bank or non-bank institute) offers you a loan to enroll in a higher education course of your choice. This loan is available at a fair interest rate and you have to repay the loan only after you complete your higher education. Do you think you would want to accept such a loan?

- **Yes or No**

Q2. If yes, then how much would you like to borrow on a yearly basis? Remember, you would have to repay the loan after completing your higher education.

____________________ rupees

Figure A3: Loan Card

We will ask you about your response to this question when we meet tomorrow. In the meantime, please think about this at home and if possible discuss this with your mother/father or other family members. We are very eager to learn what you think about this, so please do not forget to attend tomorrow’s survey session in room [**room number**] at [**time**].”

Appendix 2 Balance of Baseline Variables & Correlates of Stream of Study

Table A1: Balance of Baseline Variables

	(1) Control	(2) Treatment	(3) p-value
Age	17.24	17.30	0.19
% Male	0.53	0.55	0.45
% Scheduled Tribe	0.33	0.34	0.70
% Hindu	0.65	0.63	0.39
Asset Index	7.52	7.82	0.10
HH Facility Index	2.62	2.63	0.87
% Own Land	0.74	0.69	0.04
Board Exam Score	61.16	61.17	0.98
% Grades Repeated	0.14	0.15	0.69
% Father in Contact	0.91	0.92	0.48
% Father High School	0.18	0.20	0.22
% Father Family Business	0.11	0.14	0.11
% Father Salaried Job	0.21	0.21	0.88
% Mother High School	0.08	0.09	0.91
% Mother Housewife	0.60	0.61	0.51
Average Older Sibling Edu.	5.27	5.20	0.45
Enroll Probability (Tech)	36.96	35.27	0.24
Enroll Probability (Gen)	32.89	32.05	0.47
Enroll Probability (Voc)	22.50	24.16	0.15
Enroll Probability (NA)	7.65	8.53	0.23
% Arts Stream	0.34	0.34	0.84
% Commerce Stream	0.31	0.31	0.99
% Science Stream	0.35	0.35	0.86

Columns (1) and (2) show sample means
Column (3) shows p-values of OLS regressions on a treatment group dummy.

Table A2: Balance of Baseline Variables by Stream

	Arts			Commerce			Science		
	(1) Control	(2) Treatment	(3) p-value	(4) Control	(5) Treatment	(6) p-value	(7) Control	(8) Treatment	(9) p-value
Age	17.41	17.35	0.60	17.20	17.42	0.02	17.09	17.16	0.43
% Male	0.48	0.42	0.19	0.48	0.53	0.29	0.63	0.70	0.11
% Scheduled Tribe	0.42	0.47	0.21	0.34	0.26	0.07	0.22	0.27	0.25
% Hindu	0.56	0.47	0.05	0.63	0.65	0.66	0.75	0.75	0.96
Asset Index	6.40	6.37	0.94	7.75	8.34	0.09	8.43	8.76	0.24
HH Facility Index	2.14	2.04	0.52	2.64	2.87	0.19	3.06	2.99	0.66
% Own Land	0.74	0.73	0.74	0.66	0.58	0.07	0.79	0.75	0.20
Board Exam Score	55.74	56.09	0.69	60.77	60.41	0.72	66.82	66.70	0.91
% Grades Repeated	0.18	0.19	0.74	0.14	0.15	0.88	0.10	0.11	0.80
% Father in Contact	0.87	0.90	0.24	0.91	0.93	0.40	0.95	0.93	0.30
% Father High School	0.13	0.18	0.22	0.18	0.18	0.34	0.23	0.22	0.89
% Father Family Business	0.09	0.10	0.79	0.12	0.18	0.09	0.13	0.15	0.47
% Father Salaried Job	0.16	0.16	0.99	0.23	0.21	0.52	0.24	0.27	0.43
% Mother High School	0.04	0.05	0.62	0.05	0.09	0.11	0.16	0.12	0.22
% Mother Housewife	0.52	0.51	0.93	0.63	0.67	0.37	0.65	0.66	0.76
Average Older Sibling Edu.	4.88	4.75	0.49	5.23	5.08	0.43	5.75	5.80	0.77
Enroll Probability (Tech)	29.75	26.30	0.11	33.95	33.56	0.98	46.67	45.31	0.53
Enroll Probability (Gen)	33.85	34.46	0.77	39.72	37.70	0.34	25.93	24.83	0.58
Enroll Probability (Voc)	25.78	26.96	0.58	19.70	20.26	0.79	21.76	24.88	0.10
Enroll Probability (NA)	10.62	12.28	0.31	6.63	8.48	0.12	5.64	4.98	0.47

Notes: p-values are from OLS regressions on a treatment group dummy.

Table A3: Balance of Baseline Variables using the F-test Approach

Dependent variable:	treatment group dummy			
	All	Arts	Commerce	Science
Age	0.0137	-0.0328	0.0669*	0.0291
	(0.020)	(0.035)	(0.036)	(0.038)
% Male	-0.0237	-0.0757	0.0861	0.0171
	(0.044)	(0.075)	(0.081)	(0.091)
% Scheduled Tribe	0.0827*	0.0978	-0.0455	0.148*
	(0.050)	(0.087)	(0.091)	(0.088)
% Hindu	0.00925	-0.0519	-0.0128	0.0775
	(0.044)	(0.080)	(0.079)	(0.081)
Asset Index	0.0162**	0.0016	0.00948	0.0238*
	(0.008)	(0.016)	(0.014)	(0.013)
HH Facility Index	-0.0187	-0.00943	0.00968	-0.0274
	(0.015)	(0.028)	(0.027)	(0.027)
% Own Land	-0.0491	-0.000787	-0.0735	-0.00928
	(0.044)	(0.090)	(0.076)	(0.076)
Board Exam Score	-0.00034	0.00302	0.00113	-0.00315
	(0.002)	(0.004)	(0.004)	(0.003)
% Grades Repeated	-0.0602	0.0429	-0.151	-0.117
	(0.053)	(0.086)	(0.098)	(0.101)
% Father High School	0.077	0.13	0.0777	0.0303
	(0.047)	(0.101)	(0.084)	(0.073)
% Father Family Business	0.123**	0.0304	0.0186	0.268***
	(0.062)	(0.134)	(0.108)	(0.100)
% Father Salaried Job	0.0458	-0.0718	0.0796	0.0881
	(0.052)	(0.111)	(0.097)	(0.080)
% Mother High School	-0.0191	0.151	0.0155	-0.109
	(0.068)	(0.182)	(0.138)	(0.089)
% Mother Housewife	0.0153	-0.101	0.0712	0.0371
	(0.041)	(0.077)	(0.078)	(0.068)
Average Older Sibling Edu.	-0.0173	0.000116	-0.0443*	-0.0206
	(0.012)	(0.021)	(0.023)	(0.023)
Enroll Probability (Tech)	-0.0026	-0.00338	-0.00363	0.00433
	(0.001)	(0.002)	(0.003)	(0.003)
Enroll Probability (Gen)	-0.00226	-0.00426*	-0.00334	0.00561
	(0.001)	(0.002)	(0.003)	(0.004)
Enroll Probability (Voc)	-0.00178	-0.00283	-0.0034	0.00522
	(0.002)	(0.002)	(0.003)	(0.003)
School Fixed Effects	Yes	Yes	Yes	Yes
F-test that all coef. are 0	1.340	0.910	1.160	1.116
p-value of F-test	0.153	0.564	0.298	0.295

Table A4: Correlates of Students' Current Stream of Study

	Mean(Commerce)-Mean(Arts)		Mean(Science)-Mean(Arts)		Mean(Science)-Mean(Commerce)	
	(1)	(2)	(3)	(4)	(5)	(6)
	Diff. in Mean	p-value	Diff. in Mean	p-value	Diff. in Mean	p-value
Age	-0.068	0.325	-0.257	0.000	-0.188	0.002
% Male	0.052	0.102	0.217	0.000	0.165	0.000
% Scheduled Tribe	-0.149	0.000	-0.201	0.000	-0.052	0.064
% Hindu	0.117	0.000	0.235	0.000	0.118	0.000
Asset Index	1.657	0.000	2.212	0.000	0.554	0.013
HH Facility Index	0.662	0.000	0.934	0.000	0.271	0.019
% Own Land	-0.112	0.000	0.036	0.171	0.148	0.000
Board Exam Score	4.682	0.000	10.850	0.000	6.168	0.000
% Grades Repeated	-0.042	0.079	-0.081	0.000	-0.039	0.066
% Father in Contact	0.035	0.063	0.050	0.004	0.015	0.344
% Father High School	0.024	0.339	0.069	0.007	0.045	0.091
% Father Family Business	0.057	0.013	0.049	0.023	-0.009	0.724
% Father Salaried Job	0.065	0.019	0.097	0.000	0.032	0.267
% Mother High School	0.032	0.033	0.098	0.000	0.066	0.001
% Mother Housewife	0.129	0.000	0.136	0.000	0.007	0.813
Average Older Sibling Edu.	0.341	0.007	0.950	0.000	0.609	0.000
Enroll Probability (Tech)	5.718	0.000	17.950	0.000	12.232	0.000
Enroll Probability (Gen)	4.558	0.004	-8.775	0.000	-13.333	0.000
Enroll Probability (Voc)	-6.388	0.000	-3.038	0.036	3.350	0.013
Enroll Probability (NA)	-3.889	0.000	-6.137	0.000	-2.248	0.003

Table A5: Baseline Relationship between Enrollment Intentions & Own Wage Beliefs

Dependent variable:	Probability of Enrollment							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All	Arts	Commerce	Science	All	Arts	Commerce	Science
prob. enjoy coursework	0.444***	0.317***	0.450***	0.548***				
	(0.0244)	(0.0395)	(0.0490)	(0.0377)				
graduation prob.	0.142***	0.217***	0.179***	0.0258				
	(0.0262)	(0.0419)	(0.0491)	(0.0438)				
prob. parental approval	0.256***	0.218***	0.261***	0.262***	0.417***	0.411***	0.426***	0.393***
	(0.0241)	(0.0421)	(0.0448)	(0.0381)	(0.0127)	(0.0228)	(0.0240)	(0.0202)
employment prob.	0.123***	0.0506	0.0878*	0.227***	0.181***	0.111***	0.164***	0.264***
	(0.0284)	(0.0470)	(0.0526)	(0.0453)	(0.0209)	(0.0340)	(0.0390)	(0.0353)
log own wage	2.320***	0.0836	2.711**	4.204***	2.033***	0.145	2.659***	3.843***
	(0.733)	(1.102)	(1.268)	(1.491)	(0.525)	(0.813)	(0.931)	(1.045)
Constant	-49.51***	-15.17	-54.98***	-77.15***	-26.98***	-2.594	-33.63***	-49.88***
	(7.130)	(10.60)	(12.52)	(14.64)	(4.748)	(7.314)	(8.405)	(9.553)
Observations (student x track)	4,572	1,557	1,407	1,608	6,092	2,076	1,873	2,143
R-squared	0.472	0.382	0.473	0.572	0.344	0.249	0.352	0.431
Non-Attendance Track	NO	NO	NO	NO	YES	YES	YES	YES
Attendance Tracks	YES	YES	YES	YES	YES	YES	YES	YES
Student FE	YES	YES	YES	YES	YES	YES	YES	YES

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix 3 Further Results

Table A6: OLS & Quantile Regressions of Baseline Error on Stream of Study

Dependent variable:	$error^T$					
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	q10	q25	q50	q75	q95
Technical Track:						
Arts	-5.438***	-0.585***	-18.71***	-1.620***	-0.193**	-0.516***
	(0.640)	(0.172)	(0.116)	(0.287)	(0.0980)	(0.169)
Commerce	-0.885	-0.0543	0	-0.280	-0.0219	-0.138
	(0.539)	(0.190)	(5.310)	(0.209)	(0.115)	(0.134)
Mean for Science	6.843***	-9.793***	8.676***	10.49***	10.95***	11.89***
	(0.354)	(0.175)	(0.0937)	(0.141)	(0.0743)	(0.116)
General Track:						
Arts	-4.296***	-0.438***	-3.387***	-16.20**	-0.292***	-0.183**
	(0.715)	(0.132)	(0.210)	(7.384)	(0.0991)	(0.0800)
Commerce	0.532	0.148	0	0	0	0
	(0.605)	(0.173)	(6.464)	(0.165)	(0.108)	(0.146)
Mean for Science	4.621***	-9.586***	-6.328***	9.875***	10.58***	11.40***
	(0.430)	(0.136)	(0.220)	(0.120)	(0.0524)	(0.0731)
Vocational Track:						
Arts	-4.253***	-0.894***	-18.10***	-0.843***	-0.218	0.135
	(0.628)	(0.196)	(0.214)	(0.213)	(0.143)	(0.145)
Commerce	0.104	0	0.479	0.192*	0.0858	0.135
	(0.526)	(0.201)	(0.319)	(0.100)	(0.0915)	(0.131)
Mean for Science	6.868***	-9.104***	8.627***	10.15***	10.84***	11.52***
	(0.352)	(0.200)	(0.211)	(0.101)	(0.0821)	(0.0956)
Non-Attendance Track:						
Arts	-2.421***	-0.342***	-0.537***	-4.318*	0	-0.0761
	(0.662)	(0.0664)	(0.193)	(2.491)	(0.0825)	(0.178)
Commerce	-0.682	0	-0.296	-0.618	0.101	-0.0513
	(0.625)	(0.111)	(0.262)	(0.760)	(0.0818)	(0.166)
Mean for Science	2.976***	-9.301***	-8.668***	8.307***	9.906***	11.29***
	(0.456)	(0.0742)	(0.197)	(0.660)	(0.0509)	(0.137)
Observations	1,524	1,524	1,524	1,524	1,524	1,524

Notes: Robust Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Reference category is Science.

Table A7: Probability of Non-Attendance by Stream

	Unconstrained Choice Set		Constrained Choice Set	
	(1)	(2)	(3)	(4)
Dependent Variable:	NA prob.	NA prob.	NA prob.	NA prob.
Arts	5.992***	4.574***	6.551**	0.887
	(0.968)	(1.043)	(2.999)	(3.347)
Commerce	2.131***	1.372*	9.214***	6.549**
	(0.766)	(0.802)	(3.095)	(3.184)
School FE	YES	YES	YES	YES
Baseline Controls	NO	YES	NO	YES
Observations	1,524	1,524	1,524	1,524

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Reference category in Science.

Table A8: Effect of the Treatment on Own-Wage Belief Updating
by Baseline Enrollment Probability

	Unconstrained Choice Set		Constrained Choice Set	
	(1)	(2)	(3)	(4)
Dependent variable:	$\mid OwnWage^T_{ijt=2} - OwnWage^T_{ijt=1} \mid$			
treatment#baseline enroll. prob.	0.000616	0.000616	-0.000885**	-0.000885**
	(0.000667)	(0.000671)	(0.000434)	(0.000437)
School FE	YES	YES	YES	YES
Track FE	YES	YES	YES	YES
Baseline Controls	NO	YES	NO	YES
Observations	3,784	3,784	3,784	3,784

Notes: Clustered Standard Errors in Parenthesis. *** p<0.01, ** p<0.05, * p<0.1.
All main effects are included in the model. Sample restricted to Commerce & Science.