Expected Benefits and Costs of Migration for Rural Youth: Experimental Evidence from India

Apurav Yash Bhatiya
Bhaskar Chakravorty
Clément Imbert
Roland Rathelot*

July 7, 2023

Abstract

We survey young job seekers in rural India to understand the determinants of enrollment in a government training program with guaranteed placement into urban jobs. Respondents are over-optimistic: they expect placement jobs to be closer to home and pay more than they do. We implement an RCT and provide them with objective information on the distribution of placement job locations or salaries. The intervention successfully corrects subjects' beliefs. Changes in beliefs affect their decision to enroll in the program. By revealed preferences, our estimates suggest that job seekers need to be paid double to work outside of their home state.

JEL Codes: D83, D84, R23, O15
Keywords: migration decisions, belief updating, urban labor markets, rural youth

---

*Bhatiya: University of Birmingham. Chakravorty: University of Warwick. Imbert: University of Warwick. Rathelot: Institut Polytechnique de Paris (ENSAE). We thank Wiji Arulamapalam, Travis Baseler, Catia Batista, Gaurav Chiplunkar, Mattia Fracchia, Sarah Frohnweiler, Matt Lowe, David McKenzie, Kate Orkin, Nina Roussille, Ashish Shenoy, and seminar participants at Birmingham, NOVAfrica Migration Workshop, DULBEA-ROA Workshop, ISI Delhi, Marseille, Newcastle, Bath, Oxford CSAE Conference 2023 and RES Annual Conference 2023 for useful comments at many stages of the projects. We also thank Kritarth Jha, Maximilian Lohnert, and Poonam Panda for their excellent research assistance. This project has been funded by EQUIP-ESRC. This project would not have been possible without the amazing collaboration of BRLPS and the Ministry of Rural Development. Bhatiya is affiliated with CAGE; Chakravorty with GLO; Imbert with BREAD, CEPR, EUDN, and JPAL; Rathelot with CREST, CEPR, and J-PAL. This trial was registered with the American Economic Association registry for randomized control trials, under the title "Mobilisation for Skill Training: Experimental Evidence from Bihar", and the trial number AEARCTR-0006003. The ethical approval was obtained from the Research Ethics Committee of the University of Warwick (HSSREC 36/17-18), and IFMR (IRB00007107; FWA00014616; IORG0005894).

1 Introduction

Large rural-urban wage gaps suggest that labor is misallocated across space in developing countries, with large negative effects on aggregate productivity (Gollin et al., 2014; Bryan and Morten, 2019; Tombe and Zhu, 2019). However, the origin of these gaps and the barriers to rural-urban migration are disputed. Low migration rates could stem from higher skill requirements in urban jobs (Lagakos and Waugh, 2013; Young, 2013), lack of credit or insurance (Bryan et al., 2014; Munshi and Rosenzweig, 2016), misinformation about urban opportunities (Baseler, 2022), and high monetary and non-monetary costs of migration (Imbert and Papp, 2020; Lagakos et al., 2023). It has proven challenging to disentangle and quantify the role of these different factors empirically as they operate jointly.

In this paper, we take advantage of a vocational training program, India's DDU-GKY (Deen Dayal Upadhyaya Grameen Kaushalya Yojana), which trains young rural workers and guarantees them placement in an urban job for free. The program thus alleviates most barriers to accessing urban jobs for rural workers, but information frictions, monetary and non-monetary costs of migration remain. We show that candidates to the program hold inaccurate beliefs about the jobs offered in the program in terms of location and salary. We experimentally correct these beliefs by providing factual information on placement jobs' location, salary, or both, and estimate the causal effect of beliefs on the decision to enroll into the program. By revealed preferences, the relative role of location and salary expectations in the decision to enroll provides us with an experimental estimate of migration costs.

Specifically, we surveyed 876 rural youth from Bihar (India) who attended 63 "mobilization" camps where prospective trainees learned about DDU-GKY from training provider and government representative. The survey suggests that the average candidate held overoptimistic expectations about placement opportunities: they expect 55% of jobs to be in their home state (the truth is 20%) and the average wage to be Rs. 9,800 (the truth is about Rs. 8,300). This may be due to self-selection of overoptimistic candidates into the camps, but "mobilizers" also had incentives to encourage over-optimistic beliefs in order to enroll more trainees. In any case, in our context, information frictions made rural young workers more willing to enroll in the program and migrate to urban areas.

We then provided information on the distribution of jobs provided by the program in the last year in terms of location (in/out of state) and salary (in 5 bins). Our intervention was successful in aligning beliefs with observed averages: posterior beliefs in the treated group were closer to the signal and significantly different from their own priors and from the posteriors of the control group. Our intervention corrected 62% of the initial bias on the job location and 90% of the bias on average salary. Belief updating was persistent: treated individuals held on to the updated beliefs up to four weeks after the intervention. We check that respondents adjusted their expectations about their own career if they completed the training, and not their expectations about outside options if they did not enroll.

Finally, we match the survey sample with administrative data on training enrollment and estimate the effect of salary and location expectations instrumented by the treatment assignment on the decision to enroll. We find that the decrease in salary expectations and in the perceived likelihood of finding a job in their home state makes the average treated candidates less likely to take part in the training program overall. This suggests that in our context misinformation about job opportunities increases willingness to migrate.

The relative effect of location and wage expectations on the decision to enroll also provides revealed preference estimates of migration costs. We estimate that rural job seekers require a salary that is twice as high to take up a job out of their home state. This echoes Tombe and Zhu (2019)'s findings that structurally estimated migration costs in China are twice as large across provinces than within (0.97 vs 0.45). Jobs out of state are located on average 10 times further away, which implies an elasticity of migration costs with respect to distance of 20%. Our estimates are similar in size to Bryan and Morten (2019)'s 15% estimate for Indonesia and an order of magnitude larger than their 2% estimate for the US. Our experimental estimates point to large costs of migration in the Indian context.

Our paper relates to four strands of the literature. First, we contribute to the literature that aims to understand the sources of rural-urban wage gaps in developing countries (Gollin et al., 2014; Bryan and Morten, 2019; Tombe and Zhu, 2019). The literature emphasizes the lack of skills among rural works (Young, 2013), financial constraints, and uninsured risk (Bryan et al., 2014; Munshi and Rosenzweig, 2016; Meghir et al., 2022). By contrast, we work in a context where skill mismatch, financial constraints, and risk are minimized by the offer of a free vocational training program with guaranteed placement. Instead, we focus on the role of rural job seekers' beliefs about urban jobs and their preferences about salary and location. The work closest to ours is Baseler (2022), who shows that rural workers in Kenya underestimate urban wages and that experimentally providing accurate information increases migration to the capital city.[1] In our setting, prospective candidates are on average over-optimistic about the urban placement jobs, so that accurate information reduces their willingness to join the program. We are the first to use experimental variation in beliefs about salary and location on the decision to enroll into the program to provide revealed preference estimates of migration costs. We find that rural young job seekers require to be paid double for a job located outside of their home state.

[1] In a similar vein, Frohnweiler et al. (2022) experimentally provide information on regional income differentials in Ghana and Uganda and find that it affects the destination choices but not the intention to migrate.

Qualitatively, our findings resonate with Kone et al. (2018)'s, who document substantial inter-state migration barriers in India, and with Imbert and Papp (2020) and Lagakos et al. (2023)'s findings of high non-monetary costs of seasonal migration in India and Bangladesh respectively. Quantitatively, our results are in line with the high migration costs estimated structurally by Bryan and Morten (2019) for Indonesia and Tombe and Zhu (2019) for China.

Second, the literature which is most closely related to ours in terms of design are recent lab-in-the-field and field experiments that study the determinants of international migration decisions (Shrestha, 2020; Bah and Batista, 2018; Batista and McKenzie, 2021; Bazzi et al., 2021; Bah et al., 2022).[2] On the one hand, Bah and Batista (2018) and Batista and McKenzie (2021) study the determinants of international migration intentions in a lab-in-the-field setting, with only reported migration intentions as an outcome. On the other, recent field experiments provide information on different aspects of the migration experience (e.g. intermediaries, mortality risk) and assess their effect on migration decisions without precisely identifying beliefs or preferences (Bazzi et al., 2021; Bah et al., 2022). One exception is Shrestha (2020)'s, who experimentally provides information on earnings abroad and on the probability of dying to potential international migrants in Nepal. He estimates the effect of beliefs about earnings and mortality risk on international migration decisions to compute the value of a statistical life. Like his, our design combines the advantage of a lab-in-the-field setting, by precise measurement of belief updating, with the advantage of a field experiment, enabling us to look at the real-world decision to be trained and placed in urban areas. We are the first to use this design to estimate migration costs.

Third, our paper adds to the literature that studies job search frictions and barriers to youth unemployment in developing countries (see, McKenzie (2017) for a review). Existing research highlights the importance of training (Alfonsi et al., 2020; Adhvaryu et al., 2023), of signaling one's skills (Carranza et al., 2020; Bassi and Nansamba, 2022), of search costs (Franklin, 2018; Abebe et al., 2021a,b), and information frictions (Hicks et al., 2011; Jensen, 2012). Recent contributions highlight the role of job seekers' often misplaced expectations about their labor market prospects to interpret the effect of experimental interventions aimed at improving their employment outcomes (Abebe et al., 2017; Banerjee and Sequeira, 2020; Alfonsi et al., 2022; Bandiera et al., 2023). We study a context in which most job search frictions are alleviated by the offer of a free training and placement program, which allows us to focus on the role of job seekers' beliefs and preferences. Our contribution is to directly manipulate beliefs through an information treatment and to estimate their effect on labor market decisions. We show that prospective candidates hold over-optimistic beliefs about

[2]The importance of beliefs in international migration has been emphasized since at least McKenzie et al. (2013), and recently by McKenzie and Yang (2022) in their recent literature survey.

the location and the pay of the placement job and that correcting these beliefs reduces enrollment in the program. Our results suggest that strong labor market preferences are important barriers for rural job seekers to access good (urban) jobs.[3]

Finally, there is a related and abundant literature on job search frictions in developed countries (Altmann et al., 2018; Belot et al., 2019, 2021; Kircher, 2022). Like the literature in developing countries, it includes structural work highlighting the importance of spatial frictions in job search (Van Ommeren and Fosgerau, 2009; Manning and Petrongolo, 2017; Marinescu and Rathelot, 2018; Schmutz and Sidibé, 2019). This strand of literature also includes lab-in-the-field experiments that estimate the value of non-monetary job amenities, such as commuting time (Mas and Pallais, 2017). Using an experimental design close to ours, Cullen and Perez-Truglia (2022) test the effect of information about pay inequality on employee motivation, and Jäger et al. (2022) the effect of information about outside options on job search intentions. To our knowledge, we are the first to use an experimental information treatment to quantify the value of job location in a real-world context.

2 Context and Experimental design

2.1 Context

India, like other developing countries, has large spatial differences in rural-urban wages. A cross-national comparison of internal migration by Bell et al. (2015) shows that India has among the lowest internal migration rate. In 2014, the Ministry of Rural Development (MoRD) launched the "Deen Dayan Upadhyaya Grameen Kaushal Yojana" (DDU-GKY) to tackle this challenge. DDU-GKY program is a residential training and placement program that targets unemployed rural youth aged 15-35 years from poor families and places them in jobs outside the home state (often in Delhi, Tamil Nadu and Kerala). The program focuses on rural youth, with mandatory coverage of females and socially disadvantaged groups. As compared to training-only programs, DDUGKY shifts the emphasis to training and placement, with a mandatory placement of at least 70% candidates. The government covers all the costs for the residential training, including accommodation and food, and provides financial support for trainees post-placement.[4]

[3] In the context of the DDU-GKY program, Banerjee and Chiplunkar (2022) inform placement officers about trainees' preferences regarding their placement jobs and finds that it leads to a better match and retention in the program. Chakravorty et al. (2023) find that informing trainees about placement jobs improves retention by inducing self-selection of trainees who are a better fit for the available jobs.

[4] DDU-GKY trainees receive a post-placement payment of Rs.1000 per month, for two months if placed in their home district, for three months if placed elsewhere within their home state, and for six months if placed outside of their home state.

For the purpose of this study, we collaborated with the Bihar Rural Livelihood Promotion Society (BRLPS), who is in charge of DDU-GKY in the state of Bihar. We worked in "mobilization camps" organized by BRLPS in collaboration with the private partners in charge of training (called Project Implementing Agencies- PIAs). The camps usually begin with an introduction to the DDU-GKY program by a job resource person (JRP) from BRLPS. Next, "mobilizers" employed by the PIAs share information about their training center. From qualitative interviews, we learned that potential trainees were misinformed about DDU-GKY placement opportunities, i.e. that they overestimated the wages offered and underestimated how far the jobs were. We suspected that this misinformation could stem from mobilizers and JRP themselves, who have professional incentives to enroll the maximum number of candidates. The mismatch between trainees' expectations and placement opportunities contributes to high drop-out rates, a major concern for BRLPS.[5]

2.2 Intervention

We designed an information intervention to correct the labor market expectations of the potential trainees. At the end of the mobilization camp, we invited candidates to answer a few questions from the survey team. The survey measured candidates' priors on DDU-GKY jobs' location and salaries (see below). After these questions, respondents were randomly assigned to one of the four intervention arms (individual-level randomization). In the control group, the candidates watched a basic informational video about the DDU-GKY program, the training center, accommodation and food facilities, and classrooms. In the video, two past beneficiaries described their (positive) experience with DDUGKY. The control video did not provide any information on job location or wages offered. In the location treatment group, candidates watched the basic information video and one additional video that provided information on the distribution of placement job location for past DDU-GKY candidates. Similarly, in the salary treatment group, the second video showed the distribution of salaries of past placement jobs. In the salary treatment $\times$ location treatment group, the candidates watched all three videos.

Specifically, the intervention videos displayed 10 candidates who were allocated into two bins for the location treatment (inside state and outside state) and five bins for the salary treatment (less than Rs 6000 per month, Rs 6000 - 8000 per month, Rs 8000 - 10000 per month, Rs 10000 - Rs 12000 per month and more than Rs 12000 per month). Since the wages

[5]In Chakravorty et al. (2023), we document that 88% of enrolled candidates complete training, but only 45% join their placement job, and 33% are in their placement job after five months. We show that providing information about placement jobs to trainees has no effect on training completion or placement, but improves retention conditional on placement, which we interpret as evidence of improved self-selection into placement.

and job offers differ across male and female candidates (primarily due to different training sectors), the distributions were tailored to the gender of the candidate. [6] The distribution of wages and location for the placement job was obtained from a parallel project carried out the same year and in the same state (Chakravorty et al., 2023). Administrative data have incomplete information on the placement jobs of the candidates as PIAs don't focus on tracking candidates once they have left the training center. By contrast, surveys from Chakravorty et al. (2023) followed a sample of 2,488 DDU-GKY trainees from enrollment to five months after training completion with an attrition below 5%.

2.3 Data

Our research relies on primary data collected from three rounds of surveys. In addition, we used administrative data, which we matched with the survey data.

The baseline survey was administered to all participants in the mobilization camps after the trainees had received information from the JRP and/or the PIA mobilizer. It was a face-to-face interview with individual trainees. The baseline questionnaire first collected information about the probability of enrolling in the training and about their priors on the distribution of wages and location of DDUGKY jobs. Specifically, the survey asked 'After the training, if 10 people like you get a job. How many will get a job inside of Bihar, and how many will get a job outside of Bihar?" and 'After the training, if 10 people like you get a job. How many will get a job with a monthly salary of less than Rs 6000 / Rs 6000 - 8000 / Rs 8000 - 10000 / Rs 10000 - Rs 12000 / more than Rs 12000 per month'. To make it easier for respondents, we followed best practices from Delavande et al. (2011), and gave them ten marbles which we asked them to distribute into cups (one for each option). Then the survey provided information on the location and earnings distribution of DDUGKY jobs following the randomized treatment assignment and customized depending on the gender of the candidate. Finally, the survey measured posterior beliefs about wages and job location, following the same methodology as for the priors. In addition, it asked about the posterior probability to enroll in the training, expected earnings in a year if they completed the training, counterfactual earnings if they did not, and socio-economic characteristics.

The two follow-up surveys were conducted on the phone with the trainees one week and four weeks after the baseline survey for all respondents. Qualitative interviews with JRPs

[6]Appendix Figure A1 - A4 show snippets of the location and salary intervention videos for females and males, respectively. Appendix A provides a detailed transcript of each video.

and PIAs informed us that most candidates who want to enroll on the training program enroll within a week or 10 days of the mobilization camp. The objective of these surveys was to collect information about the posterior beliefs on wages and job location, expected and counterfactual earnings at the time when the candidates were making a decision to enroll in the program. The surveys also asked whether the candidate had visited the training center or enrolled in the program.

The administrative data comes from the management information system (MIS) of BRLPS and was compiled from the PIAs report to the state administration. It includes official information on candidate enrollment. We matched it to the survey dataset by mobile number, name and district of the candidate.

2.4 Summary statistics and balance tests

Our sample includes 876 candidates from 63 mobilization camps organized in Bihar.[7] The surveys were conducted between December 2019 and February 2020.[8] Information from the camp activity survey suggests that 74% of the camps were attended by the PIA mobilizer. All camps had the presence of a JRP. In 9.5% of the camps (6 out of 63), neither the JRP nor the mobilizer provided an introduction to the program. In 30% of the camps (19 out of 63), both the JRP and mobilizer spoke about the DDU-GKY program.

The summary statistics of our baseline variables are provided in Appendix Table A1. The average age of candidates in our sample was 20, and almost 58% were females. In terms of social category, 30% of the candidates came from the Scheduled Castes and Scheduled Tribes, and 55% were OBCs, which shows the pro-poor targeting of the DDU-GKY program. Both females and males say it would not be difficult for their family if they enroll in the training program and that there is almost 80% probability of enrolling in the program. This suggests JRPs target the candidates well: candidates who fulfill the program targeting and those who are eager to take part in the training were present in the mobilization camps. Balancing tests suggest that there were no issues with the randomization (Appendix Table A2). The attrition rate in both follow-up rounds is low (almost 6%) and similar across all treatment and control groups (Appendix Table A3).

Figure A5 shows the misperceptions in labor market beliefs. We measure misperceptions

[7] Our total survey sample was 880. However, in 4 camps there was only 1 candidate each. We exclude these camps from our analysis. Correia (2015) suggests that singleton observations together with mobilization camp fixed effects can overstate the statistical significance and lead to incorrect inference.

[8] The COVID-19 lockdowns were introduced in India towards the end of March 2020 and are unlikely to have affected the mobilization camps and the candidate's decision to enroll on the program.

by comparing the prior beliefs with the truth (signal). Less than 5% of the respondents' prior beliefs for the location fell within $\pm$ 5% of the signal. The majority of the respondents underestimated the number of candidates outside state, often by a large margin: the mean absolute error was 50%. On the average salary, the mean absolute error was 25%, only 12% of candidates' prior beliefs were within 5% of the signal, and a majority of the candidates overestimated the average salary. Appendix Figure A6 provides a density plot of the prior salary beliefs by gender. Females exhibited greater deviations from the truth than males: their priors were similar (Rs 9800 on average) but actual DDU-GKY placement jobs paid less for women (Rs 7600) than men (Rs 9000). One reason could be that females base their labor market expectations on the experience of migrants who are mostly male, while female-dominated sectors (e.g. garment factories) tend to offer lower pay.

3 Empirical Framework

3.1 Beliefs

Our empirical analysis follows our pre-registered pre-analysis plan.[9] We estimate the effect of our intervention on labor market beliefs regarding the location $j = l$ or the salary $j = s$ of DDU-GKY placement jobs for individual i present in mobilization camp c using the following specification:

$$Posterior_{ic}^{j} - Prior_{ic}^{j} = \gamma_j T_{ic}^{j} + X_{ic}'\alpha + \delta_c + \varepsilon_{ic} \tag{1}$$

$Prior_{ic}^{j}$ and $Posterior_{ic}^{j}$ denote the respondent i's prior and posterior distributions for DDU-GKY placement jobs' salary and location. Prior distributions are measured by the baseline survey before the intervention. Posterior distributions are measured either by the baseline survey after the intervention or by the two follow-up surveys. Location beliefs are measured as the number of trainees (out of 10) who get a job outside of Bihar. Salary beliefs are measured as the average expected salary, computed as the sum of the mean salary in each bin times the share of candidates (out of 10) assigned to each bin.[10] T_{ic}^{j} is an indicator variable equal to one if the candidate i received information about salary ($j = s$) or location ($j = l$). The coefficient of interest γ_j is the estimate of how treated individuals update their labor market beliefs on average as compared to those who did not receive the treatment j.

[9] American Economic Association registry for randomized control trials, under the title "Mobilisation for Skill Training: Experimental Evidence from Bihar", and the trial number AEARCTR-000600.

[10] We use Rs 5,000 as the mean salary in the "less than Rs 6,000" bin and Rs 13,000 as the mean salary in the "more than Rs 12000" bin. The appendix presents results using the median salary as a robustness check.

δ_c are mobilization camps fixed effects, and X_i denotes a vector of individual characteristics selected using a post-double-selection lasso (Belloni et al., 2014). Standard errors are clustered at the mobilization camp level.

To investigate further changes in beliefs, we carry out two additional pieces of analysis. First, the information intervention may change respondents' beliefs about the distribution of location and salaries of DDU-GKY placement jobs without affecting their expectations about what will happen to them personally if they enroll. Conversely, beyond the jobs offered by DDU-GKY, the intervention may change respondents' overall labor market outlook. To test this, we use survey questions that asked about respondents' expected earnings and location in a year in two scenarios: if they enrolled in DDU-GKY, and if they did not. We use specification 1 to check that the intervention did change respondents' expectations about their own future if they did the program and did not change their labor market expectations outside of the program. Second, providing information about job location may lead respondents to update their beliefs about salary, and vice versa: we test this by regressing changes in beliefs on both treatment dummies and their interaction.

3.2 Enrollment

Our goal is to understand how labor market expectations affect individuals' decisions to enroll in the training program. For this, we use a 2SLS estimation procedure similar to the one used by Cullen and Perez-Truglia (2022) and Jäger et al. (2022) in other contexts. Specifically, we instrument changes in beliefs in each dimension $j \in \{l, s\}$ $Posterior^j_{ic} - Prior^j_{ic}$ with the treatment indicators T^j_i, $Signal^j - Prior^j_{ic}$, which measures how far their priors were from the truth, and the interaction between them. The first stages write:

$$\begin{aligned} Posterior^j_{ic} - Prior^j_{ic} = {} & \beta^l_1 T^l_{ic} + \beta^l_2 (Signal^l - Prior^l_{ic}) + \beta^l_3 (Signal^l - Prior^l_{ic}) \times T^l_{ic} \\ & + \beta^s_1 T^s_{ic} + \beta^s_2 (Signal^s - Prior^s_{ic}) + \beta^s_3 (Signal^s - Prior^s_{ic}) \times T^s_{ic} \\ & + X'_{ic}\alpha + \delta_c + \varepsilon_{ic} \quad j \in \{l, s\} \end{aligned} \tag{2}$$

The coefficient β^j_1 captures any level shift in beliefs due to each treatment and β^j_3 captures any differential updating by individuals whose beliefs were further away from the signal.

The outcome is the difference between $I(Enrollment)^{Posterior}_{ic}$, a dummy variable for program enrollment, and $P(Enrollment)^{Prior}_{ic}$, the expected probability to enroll in the baseline survey prior to the intervention. The second stage of the estimation is:

$$I(Enrollment)_{ic}^{Posterior} - P(Enrollment)_{ic}^{Prior} = \beta_l(Posterior_{ic}^l - Prior_{ic}^l) + \beta_s(Posterior_{ic}^s - Prior_{ic}^s) + X_{ic}'\alpha + \delta_c + \varepsilon_{ic} \tag{3}$$

Both stages include mobilization camps fixed effects (δ_c), and a vector of individual characteristics (X_i) selected using a post-double-selection lasso (Belloni et al., 2014).

4 Results

4.1 Labor Market Beliefs

Figure 1a and 1b display graphically how the treatments changed labor market beliefs. Prior to the intervention, there is no difference in beliefs between the treatment and the control group. Respondents are over-optimistic: regarding location, they believe that a majority (55%) of placement jobs are in the state, when the truth (signal) is less than 20%, and regarding the salary, they believe that half of the jobs pay more than 10k, when the truth (signal) is less than 10%. After receiving the signal, the treatment group revise their expectations downward: they now attribute a 28% probability of having a job in the state, and a 30% probability of earning more than 10k. Beliefs do not change in the control.

Table 1 presents the effects of the information intervention on changes in beliefs (posterior $-$ prior) about job location (Panel A) and average salary (Panel B), estimated using specification 1. To test whether the change in beliefs were persistent, we consider posterior beliefs at three different times: in the baseline survey just after the intervention (Column 1), one week after the intervention (Columns 2: Followup 1w), and four weeks after the intervention (Column 3: Followup 4w). The number of observations changes slightly across columns due to attrition. As Panel A in Table 1 shows, the control group who believed that 42% of placement jobs were outside the state at baseline barely updated their belief during the baseline survey (+5pp.) and in the following four weeks (+8pp.). By contrast, as Figure 1a showed, the treatment group updated their belief strongly upward (+25pp.) during the survey. One week after the survey, only half of this update remained (13%), but it had not decayed further four weeks later (12%). Table 1 Panel B turns to average salary expectations (in Rs 1,000). The control group at baseline believed that the placement jobs on average paid Rs. 9,873 and did not update their prior at all in the course of the following weeks. By contrast, the treatment group revised downwards their salary expectations by (Rs -1,463) during the survey, and again about half of that change was present a week later

(Rs -655), with almost no decay four weeks later (Rs -633). These results take the average salary as an outcome, but the information provided was about the whole salary distribution: in Appendix Table A4, we check that the treatment also shifted the median closer to the truth and reduced the variance of salary expectations.

The results so far focus on respondents' beliefs about the distribution of placement jobs, but it could be that the intervention did not change their expectations about what would happen to them personally if they enrolled in the program. We check this by using as outcome respondents' expectations about where they would be and how much they would earn if they completed the training program. As Table 2 Panel A and C show, respondents in the control group are even more optimistic about their own prospect a year after training than the average placement job: only 34% believe they will be out of state, and on average they expect to earn over Rs 13,000. Reassuringly, respondents in the treatment group become less optimistic, with an increase by 7pp. in the probability to be out of state, and a Rs 1,700 reduction in their expected wage. Over the course of the following four weeks, the effects strengthened for location (+10pp.) and weakened for salary (Rs -1,100).

Another important question is whether the information treatment changed their overall labor market outlook, including the jobs they could get outside of the training program. We investigate this using as outcome respondents' expectations about where they would be and how much they would earn if the did not complete the program As (Table 2 Panel B and D). Interestingly, respondents in the control group do not generally expect to migrate out of state (between 8 and 11% depending on the survey), and their salary expectations are low (between Rs 6,358 and 7,470 depending on the survey). These low salary expectations may in fact be accurate: in a companion project in the same context, we find that the salary of respondents who enroll but drop out of training is Rs 7,600 (Chakravorty et al., 2023). Reassuringly, the information treatment has no effect on respondents' expectations about their location or earnings if they did not complete the training.

Following our pre-analysis plan, we extend our analysis of belief updating in three more ways. First, we test whether information about salary changed beliefs about location and vice-versa: Appendix Table A5 suggests that there is no evidence of cross-treatment or interaction effects. Second, we consider belief updating separately by gender, caste and education level: as Appendix Figure A7 shows, all groups updated their beliefs on location, but the update on salary was stronger for female and less educated respondents. Third, we test whether the treatment group may have affected beliefs of the control group by telling them about the information they received. Appendix Table A6 shows that control candidates who (randomly) had a higher fraction of their peers (respondents from the same panchayat in the same mobilization camp) treated did update their beliefs on location, but the effect disappeared after four weeks, and there was no effect for beliefs on salary.

4.2 Enrollment and Migration Costs

We now examine whether the change in respondents' beliefs about the placement jobs and their own labor market prospects if they join the program actually influenced their decision to enroll. Table 3 presents the estimates of the first stage. The estimate of β_3^l in Columns 1 and 2 suggest that respondents who underestimated the probability of being placed out of state more also updated their beliefs more. Specifically, respondents whose prior were 10pp. below the signal (the average respondent is 35pp. below the signal) updated by 2.2pp. after one week and 2.4pp. after four weeks (both highly significant). Similarly, the estimates of β_3^s in Columns 3 and 4 suggest that the salary treatment had a stronger effect on the beliefs of respondents whose priors were further away from the truth: respondents who were Rs 1,000 below the signal (the average respondent was about Rs 2,000 below) updated their beliefs by Rs 137 (Rs 182 after four weeks).

Table 4 presents the estimates for the second stage, i.e. the effect of beliefs about placement jobs on the decision to enroll in the program.[11] We present results using as outcome either the difference between self-declared enrollment in the two follow-up surveys and intentions to enroll at baseline (Columns 1 and 2) or the difference between enrollment in the administrative data and intentions to enroll at baseline (Column 3). In the control group, intentions to enroll at baseline were high (79%), much higher than the self-reported measures of enrollment (19% after a week and 24% after four weeks), which were themselves higher than the enrollment rate confirmed in the administrative data (10%). These discrepancies were likely due to a combination of actual barriers to enrollment (e.g. availability of training, eligibility criteria, parental opposition, etc.) and experimenter demand effect: respondents likely anticipated that researchers expected them to enroll in the program.

Reassuringly, the estimated effects of beliefs on enrollment decisions are very similar across data sources, with more statistical precision for the more reliable administrative measure. The estimates in Table 4 Column 3 suggest that a 10pp decrease in the probability of being placed out of state would increase enrollment by 1.2pp (about 12%) and that a Rs 1,000 higher salary would increase enrollment by 2.4pp (about 24%).[12] Since the expected probability of migrating out of state increased by 12pp and the expected salary

[11]The Kleibergen-Paap F-stat for the joint significance of the two instruments and the Sanderson-Windmeijer partial F-stat for the instruments' joint significance in the two separate first-stage regressions all suggest that the instruments are strong.

[12]Following the pre-analysis plan, we estimate the effect of labor market beliefs on intentions to enroll in the program. Intentions to enroll declined over time, but at four weeks they were still much higher than even self-reported enrollment (60% as compared to 24%). We find no effect of beliefs on intentions (Table A7).

decreased by Rs 633, these estimates suggest that the location and the salary treatment reduced enrollment by about 1.4pp each.[13]

To understand how respondents trade off salary and location, we interpret the ratio of the two coefficients (salary/location). We perform 200 replicates of the mobilization camps with replacement and report the mean of this ratio and the 95% confidence interval at the bottom of the table. We find that the ratio is 2 and is close to the bootstrapped ratio (2.14 in Column 3). This suggests that candidates expect a salary that is twice as high for a job outside of their home state. By contrast, in the data we collected for a companion project (Chakravorty et al., 2023), DDU-GKY placement jobs out of state are only paid 3% higher than jobs in the state. The results suggest that migration costs are substantial in our context. They are quantitatively similar to Tombe and Zhu (2019)'s finding that inter-province migration costs in China are twice as large as within-province (0.97 vs 0.45). Using data from Chakravorty et al. (2023), we estimate that jobs out of state are located on average 10 times further away than jobs in the state. This implies an elasticity of migration costs to distance of 0.2, which is slightly higher than Bryan and Morten (2019)'s estimate for Indonesia (0.15) and an order of magnitude higher than their estimate for the US (0.02).

Following our pre-analysis plan, we explore heterogeneity by gender, caste, and education levels. We find evidence of higher migration costs for SC & ST than for higher castes (ratio of -3.9 and -1.3 respectively), and for less educated than more educated (ratios of -1.9 and 0.8 respectively), but the 95% confidence intervals overlap (Appendix Table A9).

5 Conclusion

This paper studies the role of information friction and location preferences in rural-urban migration in India. We conducted a field experiment with prospective candidates for a training program that guarantees disadvantaged rural job seekers free placement in a formal job in urban areas. We find that candidates were over-optimistic: they expected placement jobs to be closer to home and pay more than they did. We experimentally informed them about the probability that jobs were outside the state, about the wage distribution, or both. We show that the intervention makes job seekers more pessimistic and changes their decision to enroll. Revealed preference estimates of the trade-off between location and salary imply that rural job seekers require double pay to work outside of their home state. This suggests that large rural-urban wage gaps are needed to compensate for the disutility of rural-urban migration in India.

[13] We deviate from the pre-analysis plan and estimate the effect of receiving any information treatment on enrollment: the estimate is large (-3.9pp or -40%) but insignificant (Table A8).

References

Abebe, G., A. S. Caria, M. Fafchamps, P. Falco, S. Franklin, and S. Quinn (2021a): "Anonymity or distance? Job search and labour market exclusion in a growing African city," *The Review of Economic Studies*, 88, 1279–1310.

Abebe, G., A. S. Caria, and E. Ortiz-Ospina (2021b): "The selection of talent: Experimental and structural evidence from ethiopia," *American Economic Review*, 111, 1757–1806.

Abebe, G. T., S. Caria, M. Fafchamps, P. Falco, S. Franklin, S. Quinn, and F. Shilpi (2017): "Job fairs: Matching firms and workers in a field experiment in Ethiopia," *World Bank Policy Research Working Paper*.

Adhvaryu, A., N. Kala, and A. Nyshadham (2023): "Returns to on-the-job soft skills training," *Journal of Political Economy (forthcoming)*.

Alfonsi, L., O. Bandiera, V. Bassi, R. Burgess, I. Rasul, M. Sulaiman, and A. Vitali (2020): "Tackling youth unemployment: Evidence from a labor market experiment in Uganda," *Econometrica*, 88, 2369–2414.

Alfonsi, L., M. Namubiru, and S. Spaziani (2022): "Gender Gaps: Back and Here to Stay? Evidence from Skilled Ugandan Workers during COVID-19," .

Altmann, S., A. Falk, S. Jäger, and F. Zimmermann (2018): "Learning about job search: A field experiment with job seekers in Germany," *Journal of Public Economics*, 164, 33–49.

Bah, T. L. and C. Batista (2018): "Understanding willingness to migrate illegally: Evidence from a lab in the field experiment," *Nova School of Business and Economics Working Paper*.

Bah, T. L., C. Batista, F. Gubert, and D. McKenzie (2022): "Can information and alternatives to irregular migration reduce "backway" migration from the Gambia?" NOVAFRICA Working Paper Series wp2203.

Bandiera, O., V. Bassi, R. Burgess, I. Rasul, M. Sulaiman, and A. Vitali (2023): "The Search for Good Jobs: Evidence from a Six Year Field Experiment in Uganda," Manuscript.

Banerjee, A. and G. Chiplunkar (2022): "How important are matching frictions in the labor market? experimental & non-experimental evidence from a large Indian firm," *Working Paper*.

Banerjee, A. V. and S. Sequeira (2020): "Spatial mismatches and imperfect information in the job search," *CEPR Discussion Paper No. DP14414*.

Baseler, T. (2022): "Hidden income and the perceived returns to migration," *American Economic Journal: Applied Economics (forthcoming)*.

Bassi, V. and A. Nansamba (2022): "Screening and signalling non-cognitive skills: experimental evidence from Uganda," *The Economic Journal*, 132, 471–511.

Batista, C. and D. J. McKenzie (2021): "Testing classic theories of migration in the lab," *CEPR Discussion Paper No. DP16469*.

Bazzi, S., L. Cameron, S. G. Schaner, and F. Witoelar (2021): "Information, intermediaries, and international migration," Tech. rep., National Bureau of Economic Research.

Bell, M., E. Charles-Edwards, P. Ueffing, J. Stillwell, M. Kupiszewski, and D. Kupiszewska (2015): "Internal migration and development: Comparing migration intensities around the world," *Population and Development Review*, 41, 33–58.

Belloni, A., V. Chernozhukov, and C. Hansen (2014): "Inference on treatment effects after selection among high-dimensional controls," *The Review of Economic Studies*, 81, 608–650.

Belot, M., P. Kircher, and P. Muller (2019): "Providing advice to jobseekers at low cost: An experimental study on online advice," *The Review of Economic Studies*, 86, 1411–1447.

——— (2021): "Eliciting time preferences when income and consumption vary: Theory, validation & application to job search," *Tinbergen Institute Discussion Paper 2021-013/V*.

Bryan, G., S. Chowdhury, and A. M. Mobarak (2014): "Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh," *Econometrica*, 82, 1671–1748.

Bryan, G. and M. Morten (2019): "The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia," *Journal of Political Economy*, 127, 2229–2268.

Carranza, E., R. Garlick, K. Orkin, and N. Rankin (2020): "Job Search and Hiring With Two-Sided Limited Information About Workseekers' Skills," *Economic Research Initiatives at Duke (ERID) Working Paper*.

Chakravorty, B., W. Arulampalam, A. Y. Bhatiya, C. Imbert, and R. Rathelot (2023): "Can Information about Jobs Improve the Effectiveness of Vocational Training? Experimental Evidence from India," *SSRN*.

Correia, S. (2015): "Singletons, cluster-robust standard errors and fixed effects: A bad mix," *Technical Note, Duke University*, 7.

Cullen, Z. and R. Perez-Truglia (2022): "How much does your boss make? The effects of salary comparisons," *Journal of Political Economy*, 130, 766–822.

Delavande, A., X. Giné, and D. McKenzie (2011): "Eliciting probabilistic expectations with visual aids in developing countries: how sensitive are answers to variations in elicitation design?" *Journal of Applied Econometrics*, 26, 479–497.

Franklin, S. (2018): "Location, search costs and youth unemployment: experimental evidence from transport subsidies," *The Economic Journal*, 128, 2353–2379.

Frohnweiler, S., B. Beber, and C. Ebert (2022): "Information Frictions, Belief Updating and Internal Migration: Evidence from Ghana and Uganda," *IZA Discussion Paper*.

Gollin, D., D. Lagakos, and M. E. Waugh (2014): "The agricultural productivity gap," *The Quarterly Journal of Economics*, 129, 939–993.

Hicks, J. H., M. Kremer, I. Mbiti, and E. Miguel (2011): "Vocational education voucher delivery and labor market returns: A randomized evaluation among Kenyan youth," *Report for Spanish Impact Evaluation Fund (SIEF) Phase II*.

Imbert, C. and J. Papp (2020): "Costs and benefits of rural-urban migration: Evidence from India," *Journal of Development Economics*, 146, 102473.

Jäger, S., C. Roth, N. Roussille, and B. Schoefer (2022): "Worker beliefs about outside options," *The Quarterly Journal of Economics (forthcoming)*.

Jensen, R. (2012): "Do labor market opportunities affect young women's work and family decisions? Experimental evidence from India," *The Quarterly Journal of Economics*, 127, 753–792.

Kircher, P. (2022): "Job search in the 21st Century," *Journal of the European Economic Association*, 20, 2317–2352.

Kone, Z. L., M. Y. Liu, A. Mattoo, C. Ozden, and S. Sharma (2018): "Internal borders and migration in India," *Journal of Economic Geography*, 18, 729–759.

Lagakos, D., A. M. Mobarak, and M. E. Waugh (2023): "The welfare effects of encouraging rural-urban migration," *Econometrica*, 91, 803–837.

Lagakos, D. and M. E. Waugh (2013): "Selection, agriculture, and cross-country productivity differences," *American Economic Review*, 103, 948–80.

Manning, A. and B. Petrongolo (2017): "How Local Are Labor Markets? Evidence from a Spatial Job Search Model," *American Economic Review*, 107, 2877–2907.

Marinescu, I. and R. Rathelot (2018): "Mismatch unemployment and the geography of job search," *American Economic Journal: Macroeconomics*, 10, 42–70.

Mas, A. and A. Pallais (2017): "Valuing Alternative Work Arrangements," *American Economic Review*, 107, 3722–3759.

McKenzie, D. (2017): "How effective are active labor market policies in developing countries? a critical review of recent evidence," *The World Bank Research Observer*, 32, 127–154.

McKenzie, D., J. Gibson, and S. Stillman (2013): "A land of milk and honey with streets paved with gold: Do emigrants have over-optimistic expectations about incomes abroad?" *Journal of Development Economics*, 102, 116–127.

McKenzie, D. and D. Yang (2022): "Field and Natural Experiments in Migration," *World Bank, Washington, DC*.

Meghir, C., A. M. Mobarak, C. Mommaerts, and M. Morten (2022): "Migration and informal insurance: evidence from a randomized controlled trial and a structural model," *The Review of Economic Studies*, 89, 452–480.

Munshi, K. and M. Rosenzweig (2016): "Networks and misallocation: Insurance, migration, and the rural-urban wage gap," *American Economic Review*, 106, 46–98.

Schmutz, B. and M. Sidibé (2019): "Frictional Labour Mobility," *Review of Economic Studies*, 86, 1779–1826.

Shrestha, M. (2020): "Get Rich or Die Tryin': Perceived Earnings, Perceived Mortality Rates, and Migration Decisions of Potential Work Migrants from Nepal," *The World Bank Economic Review*, 34, 1–27.

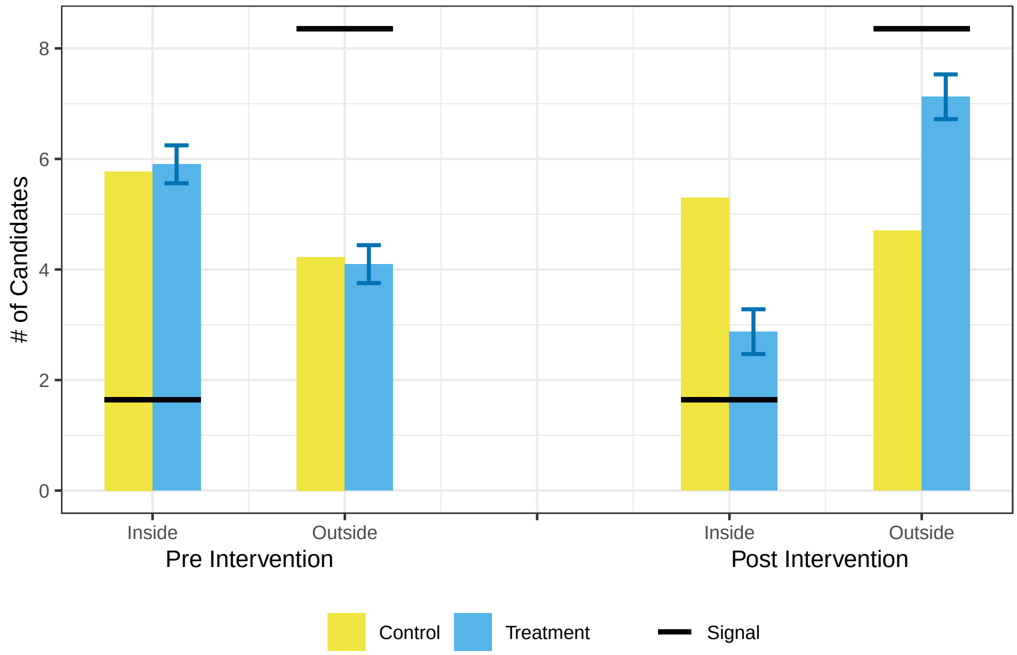
Tombe, T. and X. Zhu (2019): "Trade, Migration, and Productivity: A Quantitative Analysis of China," *American Economic Review*, 109, 1843–72.

Van Ommeren, J. and M. Fosgerau (2009): "Workers' marginal costs of commuting," *Journal of Urban Economics*, 65, 38–47.

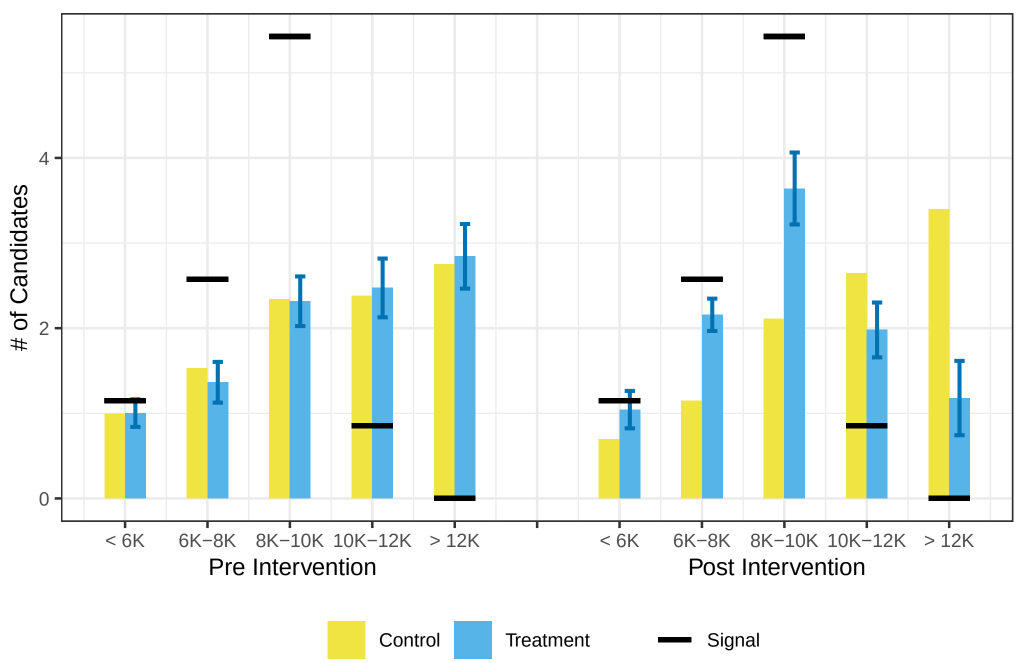
Young, A. (2013): "Inequality, the urban-rural gap, and migration," *The Quarterly Journal of Economics*, 128, 1727–1785.

Tables and Figures

Figure 1: Effect of Treatment on Labor Market Beliefs



(a) Location



(b) Salary

Notes: Panel A shows the distribution of the number of candidates inside and outside the state from the baseline survey pre- and post-intervention. Panel B shows the distribution of number of candidates from the baseline survey pre- and post-intervention in 5 salary bins: below Rs 6K, 6K-8K, 8K-10K, 10K-12K and above Rs 12K. The error bars show the 95% CI on the coefficient of an indicator variable for the information treatment. "Signal" is the information provided by the treatment videos.

Table 1: Effect of Treatment on Labor Market Beliefs

	Posterior − Prior		
	Baseline Posterior (1)	Followup 1w (2)	Followup 4w (3)
Panel A: Location (Candidates Outside State)			
Location Treatment	2.495***	1.276***	1.228***
	(0.223)	(0.221)	(0.254)
Mean DV [Control]	0.474	0.890	0.812
Prior [Control]	4.227	4.291	4.215
Panel B: Salary (Earnings Distribution Mean)			
Salary Treatment	-1.463***	-0.655***	-0.633***
	(0.125)	(0.132)	(0.129)
Mean DV [Control]	0.506	0.001	0.117
Prior [Control]	9.873	9.856	9.886
# of Camps	63	62	63
Camp FE	Yes	Yes	Yes
Observations	876	823	826

Notes: This table presents the effect of the location treatment (Panel A) and the salary treatment (Panel B) on how the respondents update their labor market beliefs (Posterior - Prior). Column 1 measures the outcomes after the intervention during the baseline survey. Columns 2 and 3 measure the outcomes at the follow-up surveys one week and four weeks after the intervention, respectively. The outcome variables in Panel A measure the number of candidates (out of 10) who will get a job outside state. The outcome variables in Panel B measure earnings distribution mean calculated using the number of candidates in each bin. All outcomes in Panel B are scaled by 1000. Standard errors are clustered at the camp level. The control variables chosen by a post-double-selection lasso procedure were a dummy for females and having the RSBY document (Panel A) and a dummy variable for having the NREGA job card (Panel B). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Effect of Treatment on Own Career Expectations (1 year later)

	Posterior		
	(1) Baseline Posterior	(2) Followup 1w	(3) Followup 4w
Panel A: Respondent Outside of State if Completes Training			
Location Treatment	0.072**	0.096***	0.097***
	(0.028)	(0.032)	(0.032)
Mean DV [Control]	0.337	0.396	0.376
Panel B: Respondent Outside of State if Does Not Complete Training			
Location Treatment	-0.010	-0.016	0.013
	(0.018)	(0.020)	(0.019)
Mean DV [Control]	0.113	0.093	0.077
Panel C: Respondent Salary if Completes Training			
Salary Treatment	-1.700***	-1.049***	-1.094***
	(0.299)	(0.316)	(0.299)
Mean DV [Control]	13.173	13.959	13.442
Panel D: Respondent Salary if Does Not Complete Training			
Salary Treatment	-0.440	-0.133	0.093
	(0.357)	(0.431)	(0.420)
Mean DV [Control]	6.358	7.470	7.128
# of Camps	63	62	63
Camp FE	Yes	Yes	Yes
Observations	876	823	825

Notes: This table presents the estimation of own career expectations (location and salary) one year later on the treatment status. Column 1 measures the outcomes after the intervention during the baseline survey. Columns 2 and 3 measure outcomes at the follow-up surveys one week and four weeks after the intervention, respectively. The dependent variable in Panels A and B indicates whether the respondent expects to be outside of Bihar one year later with training (expectation; Panel A) and without training (counterfactual; Panel B). The dependent variables in Panels C and D measure the average monthly salary one year later with training (expectation; Panel C) and without training (counterfactual; Panel D). All outcomes in Panels C and D are scaled by 1000. Standard errors are clustered at the camp level. The control variables were chosen by a post-double-selection lasso procedure. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Heterogeneity by Signal (First Stage Regressions)

	Posterior − Prior			
	Location		Salary	
	(1) Followup 1w	(2) Followup 4w	(3) Followup 1w	(4) Followup 4w
Location Treatment	0.439*	0.261	0.137	0.398**
	(0.251)	(0.324)	(0.147)	(0.171)
Location (Signal − Prior)	0.653***	0.595***	0.083***	0.108***
	(0.060)	(0.052)	(0.025)	(0.027)
Location (Signal − Prior) × Treatment Location	0.215***	0.236***	-0.044	-0.080**
	(0.051)	(0.063)	(0.032)	(0.039)
Salary Treatment	0.133	0.145	-0.396***	-0.258*
	(0.289)	(0.223)	(0.118)	(0.147)
Salary (Signal − Prior)	0.027	0.113	0.655***	0.636***
	(0.107)	(0.079)	(0.044)	(0.048)
Salary (Signal − Prior) × Salary Treatment	0.211*	0.110	0.137**	0.182***
	(0.124)	(0.104)	(0.055)	(0.065)
Mean DV [Control]	0.890	0.812	0.001	0.117
Prior [Control]	4.291	4.215	9.856	9.886
# of Camps	62	63	62	63
Camp FE	Yes	Yes	Yes	Yes
Observations	823	826	823	826

Notes: This table presents the first stage regressions as described in Section 3. The dependent variables in Columns 1 and 2 measure the number of candidates (out of 10) who will get a job outside state. The dependent variables in Columns 3 and 4 measure earnings distribution mean calculated using the number of candidates in each bin. All outcomes in Columns 3 and 4 are scaled by 1000. Followup 1w and Folloupw 4w indicate the follow-up surveys one week and four weeks after the intervention respectively. Standard errors are clustered at the camp level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Effect of Beliefs on Training Enrollment

	I(Enrollment) − P(Enrollment Prior)		
	(1) Followup 1w	(2) Followup 4w	(3) Admin
Location (Posterior − Prior)	-0.007	-0.010	-0.012**
	(0.007)	(0.006)	(0.006)
Salary (Posterior − Prior)	0.021**	0.024**	0.024***
	(0.010)	(0.011)	(0.008)
Mean DV [Control]	-0.602	-0.544	-0.684
P(Enrollment Prior) [Control]	0.787	0.787	0.787
Enrollment [Control]	0.187	0.238	0.103
KP F Stat	67.59	62.22	62.22
F Stat (Salary)	130.7	145.3	145.3
F Stat (Location)	70.54	112.37	112.37
Bootstrapped Ratio Mean	-2.81	-2.55	-2.14
Bootstrapped Ratio 95% CI	[-21.76, 10.99]	[-19.17, 17.40]	[-5.77, -0.50]
# of Camps	62	63	63
Camp FE	Yes	Yes	Yes
Observations	823	826	826

Notes: This table presents the 2SLS estimates of the respondents' updated beliefs on their updated probability to enroll in the training program. The updated beliefs are instrumented using the treatment status as described in Section 3. Columns 1 and 2 measure outcomes at the follow-up surveys one week and four weeks after the intervention respectively. Column 3 measures the enrollment outcome from the administrative dataset. The KP F stat is the Kleibergen-Paap F-stat for the joint significance of the two instruments in the first-stage regression. The F-stat (Salary) and F-stat (Location) are the Sanderson-Windmeijer partial F-stat for the instruments' joint significance in the two separate first-stage regressions. Standard errors are clustered at the camp level. The control variables chosen by a post-double-selection lasso procedure were a dummy for having the NREGA job card. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Online Appendix

Appendix Tables and Figures

Table A1: Summary Statistics

Variable	Mean	SD	Minimum	Maximum	Signal
Panel A: Socio-Demographics (Full Sample)					
Female	0.575	0.495	0	1	
Age	20.42	3.366	17	35	
I(Education $\geq$ Higher Secondary)	0.578	0.494	0	1	
Religion: Hindu	0.929	0.257	0	1	
Social Category: SC or ST	0.303	0.460	0	1	
Social Category: OBC	0.556	0.497	0	1	
Social Category: General	0.122	0.328	0	1	
Social Category: Prefer No Answer	0.0194	0.138	0	1	
Number of Observations			876		
Panel B: Prior Labor Market Beliefs (Females)					
Location (Candidates Outside State)	3.762	2.774	0	10	9
Salary (monthly average - Rs)	9836	1679	5000	13000	7600
... Less than Rs 6000 per month	1.050	1.542	0	10	2
... Rs 6000 - Rs 8000 per month	1.597	1.814	0	10	3
... Rs 8000 - Rs 10,000 per month	2.179	2.308	0	10	5
... Rs 10,000 - Rs 12,000 per month	2.472	2.528	0	10	0
... More than Rs 12,000 per month	2.702	2.967	0	10	0
Difficulty to family during training [0-10]	3.014	3.642	0	10	
P(Enrollment)	0.802	0.277	0	1	
Panel C: Prior Labor Market Beliefs (Males)					
Location (Candidates Outside State)	5.215	2.557	0	10	7
Salary (monthly average - Rs)	9824	1708	5000	13000	9000
... Less than Rs 6000 per month	1.097	1.622	0	10	0
... Rs 6000 - Rs 8000 per month	1.527	1.632	0	10	2
... Rs 8000 - Rs 10,000 per month	2.202	1.798	0	10	6
... Rs 10,000 - Rs 12,000 per month	2.511	2.219	0	10	2
... More than Rs 12,000 per month	2.664	2.825	0	10	0
Difficulty to family during training [0-10]	3.618	3.439	0	10	
P(Enrollment)	0.767	0.286	0	1	

Notes: This table presents summary statistics on socio-demographic characteristics (Panel A) for the full sample and prior labor market beliefs for females (Panel B) and males (Panel C). The prior labor market beliefs are presented separately by gender due to differences in signal by gender.

Table A2: Balance Statistics

Variable	Control Mean	Treatment Mean			p-value		
		Salary	Location	Salary $\times$ Location			
	(1)	(2)	(3)	(4)	(2) vs (1)	(3) vs (1)	(4) vs (1)
Panel A: Socio-Demographic Variables							
Female	0.580	0.550	0.547	0.624	0.527	0.494	0.361
Age	20.31	20.47	20.42	20.47	0.624	0.749	0.626
I(Education $\geq$ Higher Secondary)	0.591	0.558	0.561	0.603	0.504	0.536	0.805
Religion: Hindu	0.927	0.939	0.924	0.926	0.634	0.884	0.946
Religion: Muslim	0.0466	0.0519	0.0314	0.0437	0.789	0.448	0.882
Religion: Prefer No Answer	0.0259	0.00866	0.0448	0.0306	0.278	0.238	0.770
Social Category: SC or ST	0.290	0.303	0.296	0.319	0.774	0.898	0.525
Social Category: OBC	0.591	0.550	0.534	0.555	0.400	0.244	0.458
Social Category: General	0.0933	0.121	0.157	0.114	0.382	0.0482	0.526
Social Category: Prefer No Answer	0.0259	0.0260	0.0135	0.0131	0.996	0.359	0.343
Panel B: Prior Labor Market Beliefs							
Location (Candidates Outside State)	4.249	4.641	4.345	4.258	0.148	0.724	0.974
Salary (monthly average - Rs)	9860	9791	9684	9989	0.677	0.290	0.436
... Less than Rs 6000 per month	1	1.087	1.139	1.044	0.574	0.370	0.777
... Rs 6000 - Rs 8000 per month	1.539	1.649	1.749	1.332	0.514	0.218	0.222
... Rs 8000 - Rs 10,000 per month	2.352	2.100	2.148	2.179	0.219	0.324	0.400
... Rs 10,000 - Rs 12,000 per month	2.378	2.550	2.480	2.528	0.464	0.667	0.523
... More than Rs 12,000 per month	2.731	2.615	2.484	2.917	0.683	0.389	0.512
Difficulty to family during training [0-10]	3.352	2.861	3.552	3.341	0.158	0.570	0.973
Difficulty to family 1 year outside State [0-10]	4.021	3.450	3.583	3.812	0.120	0.237	0.571
P(Enrollment)	0.786	0.786	0.792	0.783	0.996	0.831	0.925
Number of Observations				880			

Notes: To check that our randomization achieved a balance between treatment and control at baseline, we estimate the following model for each control variable X_i^l and test the null of no difference between the treatment groups and control group ($\beta_s = 0$, $\beta_l = 0$ and $\beta_{sl} = 0$).

$$X_i = \beta_s T_i^s + \beta_l T_i^l + \beta_{sl} T_i^s \times T_i^l + \varepsilon_i$$

Table A3: Attrition

	Attrition	
	(1) Followup 1w	(2) Followup 4w
Location Treatment	-0.001	-0.012
	(0.025)	(0.028)
Salary Treatment	-0.011	-0.026
	(0.024)	(0.025)
Location Treatment × Salary Treatment	0.006	0.008
	(0.032)	(0.033)
Mean DV [Control]	0.062	0.067
# of Camps	63	63
Camp FE	Yes	Yes
Observations	876	876

Notes: Standard errors are clustered at the camp level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A4: Effect of Treatment on Salary Distribution: Mean, Median and Variance

	Posterior − Prior		
	(1) Baseline Posterior	(2) Followup 1w	(3) Followup 4w
Panel A: Mean of Salary Distribution			
Salary Treatment	-1.463***	-0.655***	-0.633***
	(0.125)	(0.132)	(0.129)
Mean DV [Control]	0.506	0.001	0.117
Prior [Control]	9.873	9.856	9.886
Panel B: Median of Salary Distribution			
Salary Treatment	-1.643***	-0.740***	-0.655***
	(0.159)	(0.170)	(0.169)
Mean DV [Control]	0.753	0.187	0.155
Prior [Control]	9.567	9.560	9.608
Panel C: Variance of Salary Distribution			
Salary Treatment	-1.221***	-0.464	-0.787*
	(0.312)	(0.408)	(0.453)
Mean DV [Control]	0.289	0.899	1.235
Prior [Control]	5.670	5.752	5.657
# of Camps	63	62	63
Camp FE	Yes	Yes	Yes
Observations	876	823	826

Notes: This table presents the effect of the salary treatment on the changes in the mean (Panel A), median (Panel B) and the variance (Panel C) of the salary distribution. Column 1 measures the outcomes after the intervention during the baseline survey. Columns 2 and 3 measure the outcomes at the followup survey one week and four weeks after the intervention respectively. All outcomes in Panel A are scaled by 1000. Standard errors are clustered at the camp level. Control variables were selected using a post double-selection lasso. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A5: Effect of Treatment on Labor Market Beliefs

	Posterior − Prior		
	Baseline Posterior	Followup 1w	Followup 4w
	(1)	(2)	(3)
Panel A: Location (Candidates Outside State)			
Location Treatment	2.433***	1.342***	1.235***
	(0.316)	(0.291)	(0.327)
Salary Treatment	-0.402	-0.240	-0.245
	(0.284)	(0.309)	(0.272)
Location Treatment ×	0.095	-0.142	-0.030
Salary Treatment	(0.436)	(0.408)	(0.427)
Mean DV [Control]	0.474	0.890	0.812
Prior [Control]	4.227	4.291	4.215
Panel B: Salary (Earnings Distribution Mean)			
Location Treatment	-0.059	0.055	0.080
	(0.153)	(0.143)	(0.189)
Salary Treatment	-1.380***	-0.442**	-0.509***
	(0.153)	(0.181)	(0.174)
Location Treatment ×	-0.168	-0.417*	-0.241
Salary Treatment	(0.223)	(0.234)	(0.253)
Mean DV [Control]	0.506	0.001	0.117
Prior [Control]	9.873	9.856	9.886
# of Camps	63	62	63
Camp FE	Yes	Yes	Yes
Observations	876	823	826

Notes: This table presents the effect of the location treatment (Panel A) and the salary treatment (Panel B) on how the respondents update their labor market beliefs (Posterior - Prior). Column 1 measures the outcomes after the intervention during the baseline survey. Columns 2 and 3 measure the outcomes at the follow-up surveys one week and four weeks after the intervention, respectively. The outcome variables in Panel A measure the number of candidates (out of 10) who will get a job outside state. The outcome variables in Panel B measure earnings distribution mean calculated using the number of candidates in each bin. All outcomes in Panel B are scaled by 1000. Standard errors are clustered at the camp level. The control variables chosen by a post-double-selection lasso procedure were a dummy for females and having the RSBY document (Panel A) and a dummy variable for having the NREGA job card (Panel B). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A6: Spillover Effects of Treatment

	Posterior − Baseline Posterior			
	Followup 1w		Followup 4w	
	(1)	(2)	(3)	(4)
Panel A: Location (Candidates Outside State)				
Location Treatment	-1.128***	-2.336***	-1.376***	-1.972***
	(0.230)	(0.643)	(0.219)	(0.669)
Share Treated		-1.803***		0.117
		(0.671)		(0.594)
Location Treatment × Share Treated		2.622**		0.853
		(1.071)		(0.980)
Mean DV [Control]	0.407	0.407	0.464	0.464
Baseline Posterior [Control]	4.701	4.701	4.701	4.701
Panel B: Salary (Earnings Distribution Mean)				
Salary Treatment	0.755***	-1.003**	0.793***	-0.665*
	(0.106)	(0.404)	(0.122)	(0.380)
Share Treated		-0.307		-0.178
		(0.455)		(0.415)
Salary Treatment × Share Treated		0.656		0.122
		(0.698)		(0.675)
Mean DV [Control]	-0.520	-0.520	-0.385	-0.385
Baseline Posterior [Control]	10.379	10.379	10.379	10.379
Camp FE	Yes	Yes	Yes	Yes
Observations	823	823	826	826

Notes: This table presents the effect of the treatment on labor market beliefs for location (Panel A) and salary (Panel B). The share of treated respondents is defined as the share of treatment within a peer group (defined as mobilization camp × panchayat). The outcome variables in Panel A measure the number of candidates (out of 10) who will get a job outside state. The outcome variables in Panel B measure earnings distribution mean calculated using the number of candidates in each bin. All outcomes in Panel B are scaled by 1000. Standard errors are clustered at the camp level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A7: Effect of Beliefs on Training Enrollment Intentions

	Probability to Enroll (Posterior − Prior)		
	(1) Baseline Posterior	(2) Followup 1w	(3) Followup 4w
Location (Posterior − Prior)	0.002	-0.003	-0.009
	(0.002)	(0.006)	(0.007)
Salary (Posterior − Prior)	0.006	0.010	0.013
	(0.005)	(0.008)	(0.010)
Mean DV [Control]	0.049	-0.123	-0.175
P(Enrollment) [Control]	0.787	0.787	0.787
KP F Stat	99.24	67.59	62.22
F Stat (Salary)	123.7	130.7	145.3
F Stat (Location)	102.1	70.54	112.37
# of Camps	63	62	63
Camp FE	Yes	Yes	Yes
Observations	876	823	826

Notes: This table presents the 2SLS estimates of the respondents' updated beliefs on their updated probability to enroll the training program. The updated beliefs are instrumented using the treatment status as described in the Section 3. Column 1 measures the outcomes after the intervention during the baseline survey. Columns 2 and 3 measure outcomes at the follow-up surveys one week and four weeks after the intervention respectively. The KP F stat is the Kleibergen-Paap F-stat for the joint significance of the two instruments in the first-stage regression. The F-stat (Salary) and F-stat (Location) are the Sanderson-Windmeijer partial F-stat for the instruments' joint significance in the two separate first-stage regressions. Standard errors are clustered at the camp level. Control variables selected using post double-selection lasso. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A8: Effect of Treatment on Training Enrollment

	I(Enrollment) $-$ P(Enrollment Prior)		
	(1)	(2)	(3)
	Followup 1w	Followup 4w	Admin
Panel A: Any Treatment			
Any Treatment	-0.045	-0.062*	-0.039
	(0.037)	(0.036)	(0.030)
Panel B: All Treatment Arms			
Location Treatment	-0.049	-0.027	-0.046
	(0.046)	(0.047)	(0.039)
Salary Treatment	-0.051	-0.097**	-0.035
	(0.045)	(0.038)	(0.032)
Location Treatment ×	0.063	0.067	0.046
Salary Treatment	(0.059)	(0.061)	(0.043)
Mean DV [Control]	-0.602	-0.544	-0.684
Prior [Control]	0.787	0.787	0.787
Admission [Control]	0.187	0.238	0.103
# of Camps	62	63	63
Camp FE	Yes	Yes	Yes
Observations	823	826	826

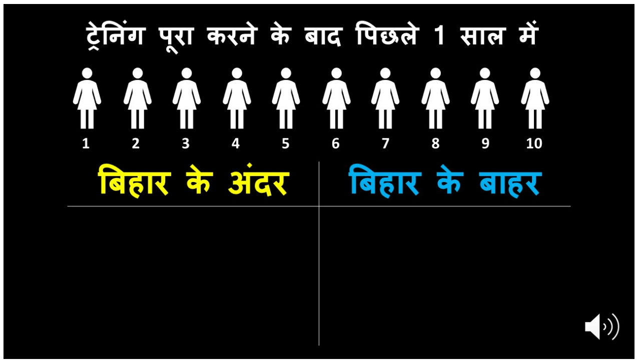
Notes: This table presents the reduced form estimates of the treatment on the updated probability to enroll in the training program. In Panel A, Any Treatment is a dummy variable that takes a value of one for respondents receiving either treatment and zero for the control group. In Panel B, we consider all treatment arms. Columns 1 and 2 measure outcomes at the follow-up surveys one week and four weeks after the intervention respectively. Column 3 measures the enrollment outcome from the administrative dataset. Standard errors are clustered at the camp level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A9: Effect of Beliefs on Training Enrollment
(Bootstrapped across sub-samples)

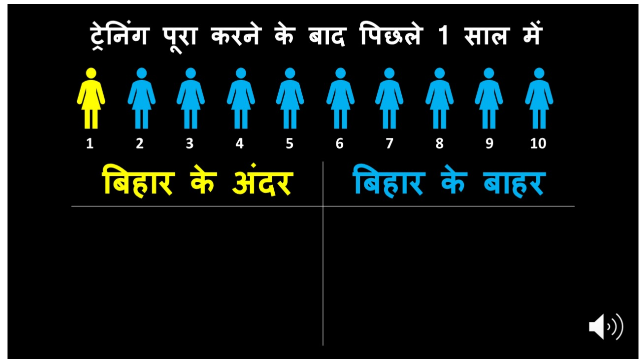
	I(Enrollment) − P(Enrollment Prior)						
	(1) Full Sample	(2) Females	(3) Males	(4) General & OBC	(5) SC & ST	(6) High Educ	(7) Low Educ
Location (Posterior − Prior)	-0.012**	-0.004	-0.010	-0.013	-0.010	0.005	-0.028***
	(0.006)	(0.007)	(0.011)	(0.007)	(0.012)	(0.007)	(0.009)
Salary (Posterior − Prior)	0.024***	0.022*	0.015	0.015	0.050**	0.009	0.050***
	(0.009)	(0.013)	(0.015)	(0.010)	(0.020)	(0.009)	(0.016)
Mean DV [Control]	-0.684	-0.733	-0.614	-0.706	-0.623	-0.678	-0.692
Prior [Control]	0.787	0.813	0.750	0.804	0.748	0.800	0.768
Enrollment [Control]	0.103	0.080	0.135	0.097	0.125	0.121	0.076
Bootstrapped Ratio Mean	-2.14	0.157	-0.841	-1.267	-3.896	0.780	-1.855
Bootstrapped Ratio 95% CI	[-5.77, -0.50]	[-24.07, 85.24]	[-17.70, 9.08]	[-9.38, 1.95]	[-103.62, 50.54]	[-20.51, 13.21]	[-4.58, -0.60]
# of Camps (Clusters)	63	52	51	58	45	58	53
Camp FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	826	473	345	555	244	483	334

Notes: The total number of replications for each column is 200. Bootstrapped standard errors are clustered at the camp level. Control variables selected using post double-selection lasso. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

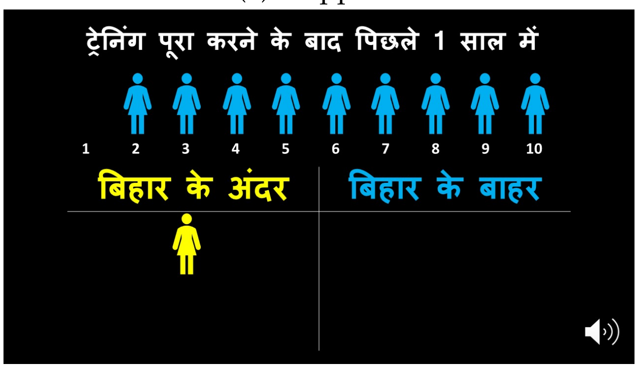
Figure A1: Location Intervention Video Snippets (Female)



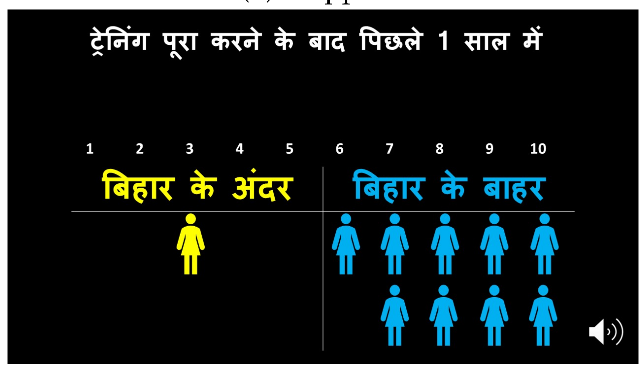
(a) Snippet 1



(b) Snippet 2

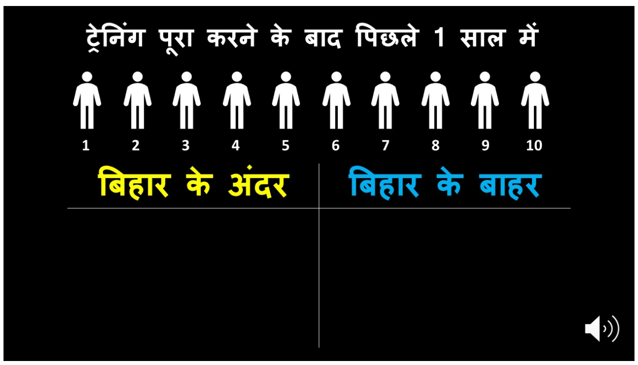


(c) Snippet 3

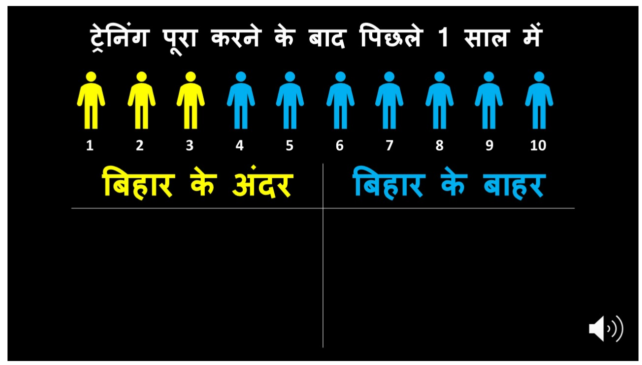


(d) Snippet 4

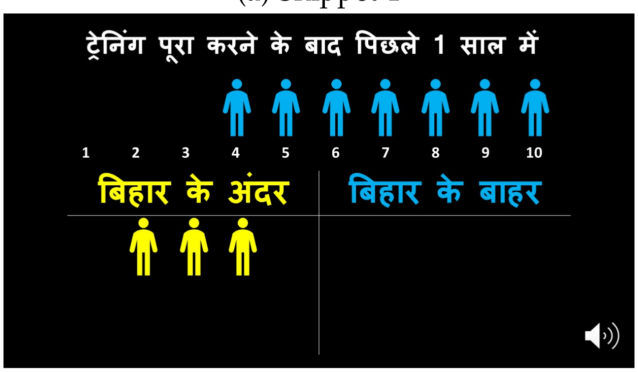
Figure A2: Location Intervention Video Snippets (Male)



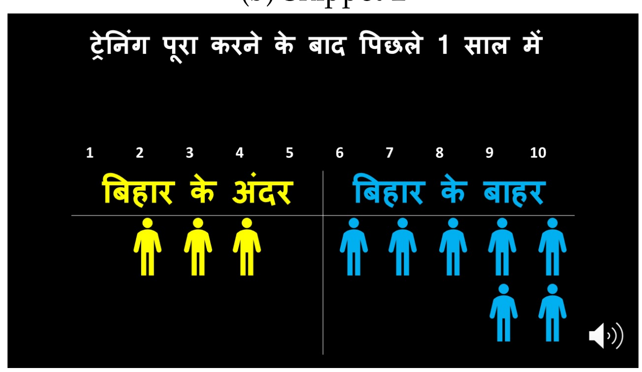
(a) Snippet 1



(b) Snippet 2

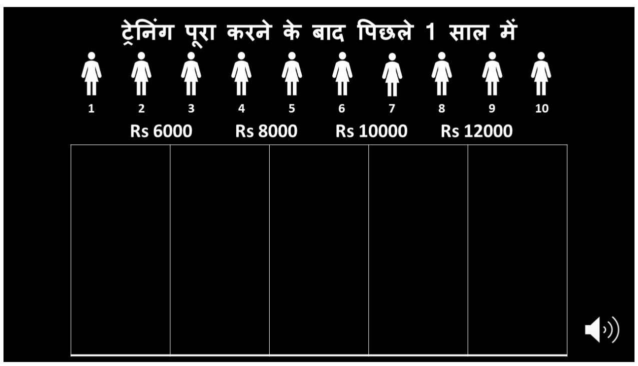


(c) Snippet 3

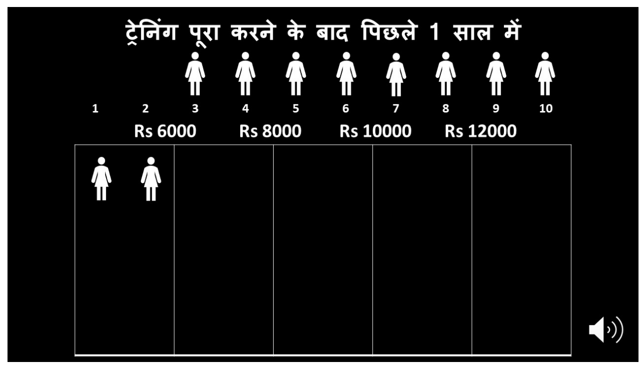


(d) Snippet 4

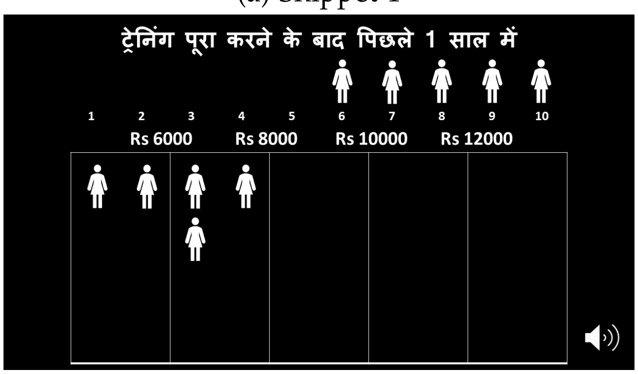
Figure A3: Salary Intervention Video Snippets (Female)



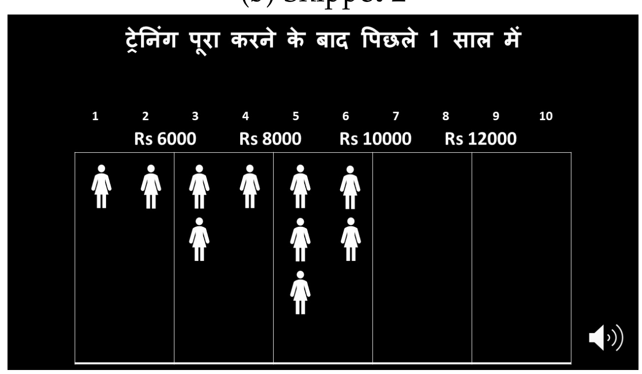
(a) Snippet 1



(b) Snippet 2

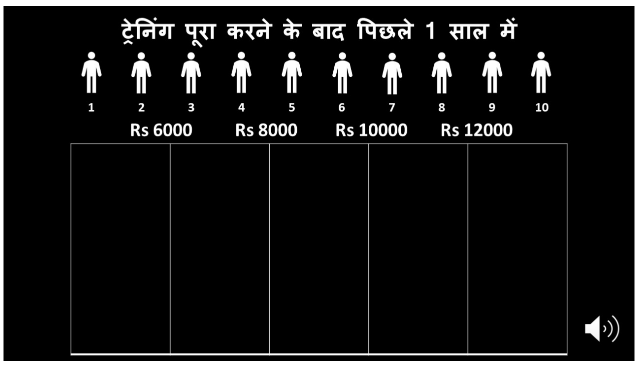


(c) Snippet 3

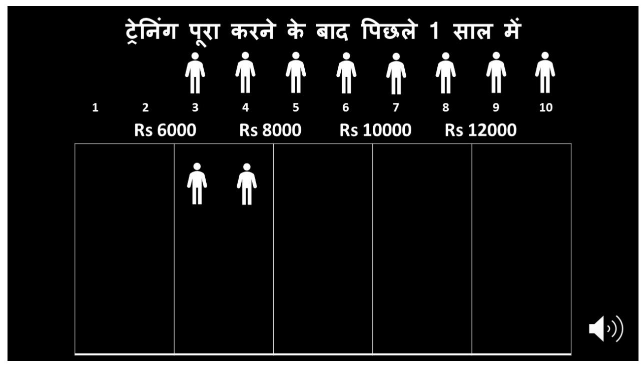


(d) Snippet 4

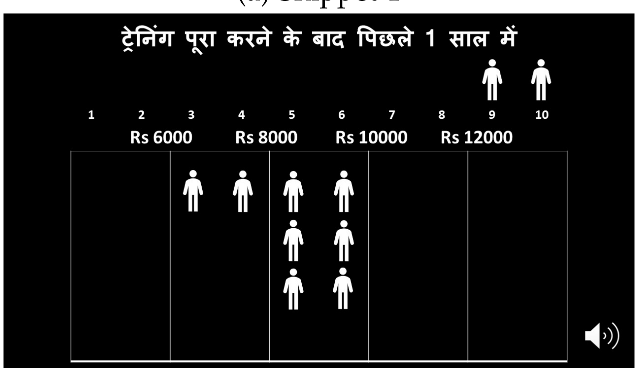
Figure A4: Salary Intervention Video Snippets (Male)



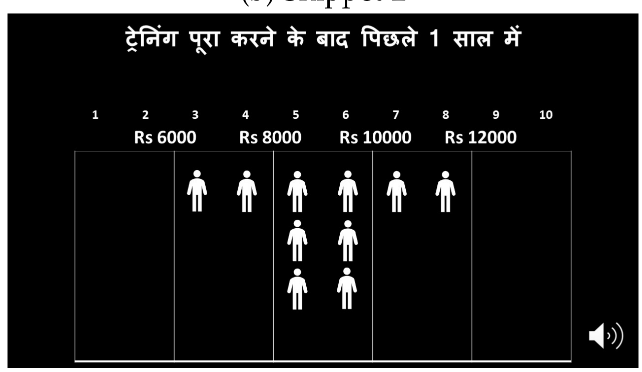
(a) Snippet 1



(b) Snippet 2

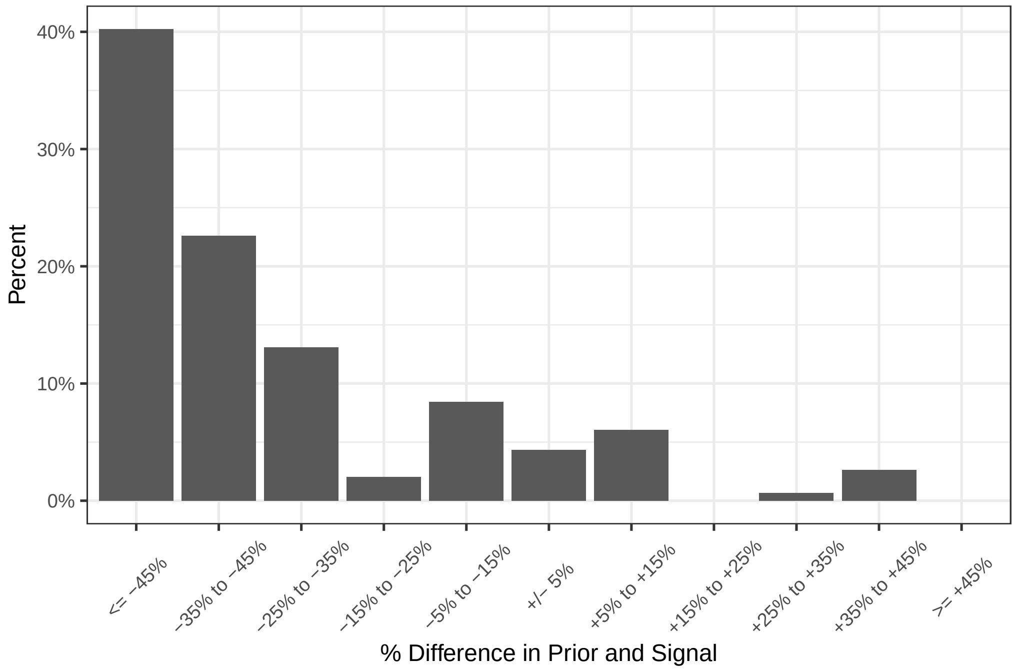


(c) Snippet 3

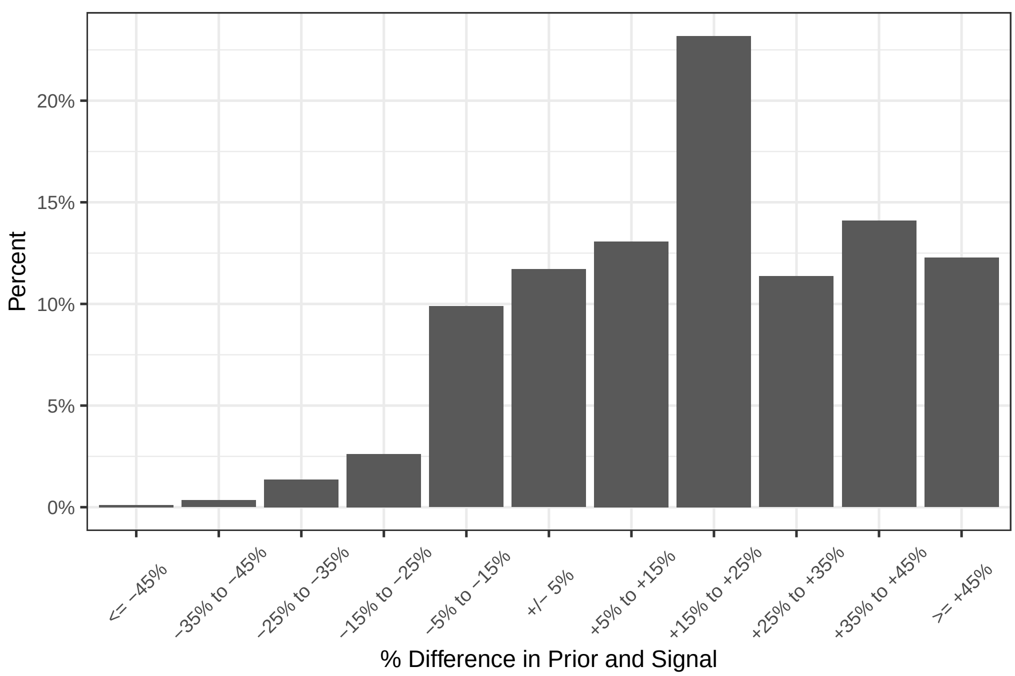


(d) Snippet 4

Figure A5: Misperceptions in Prior Beliefs

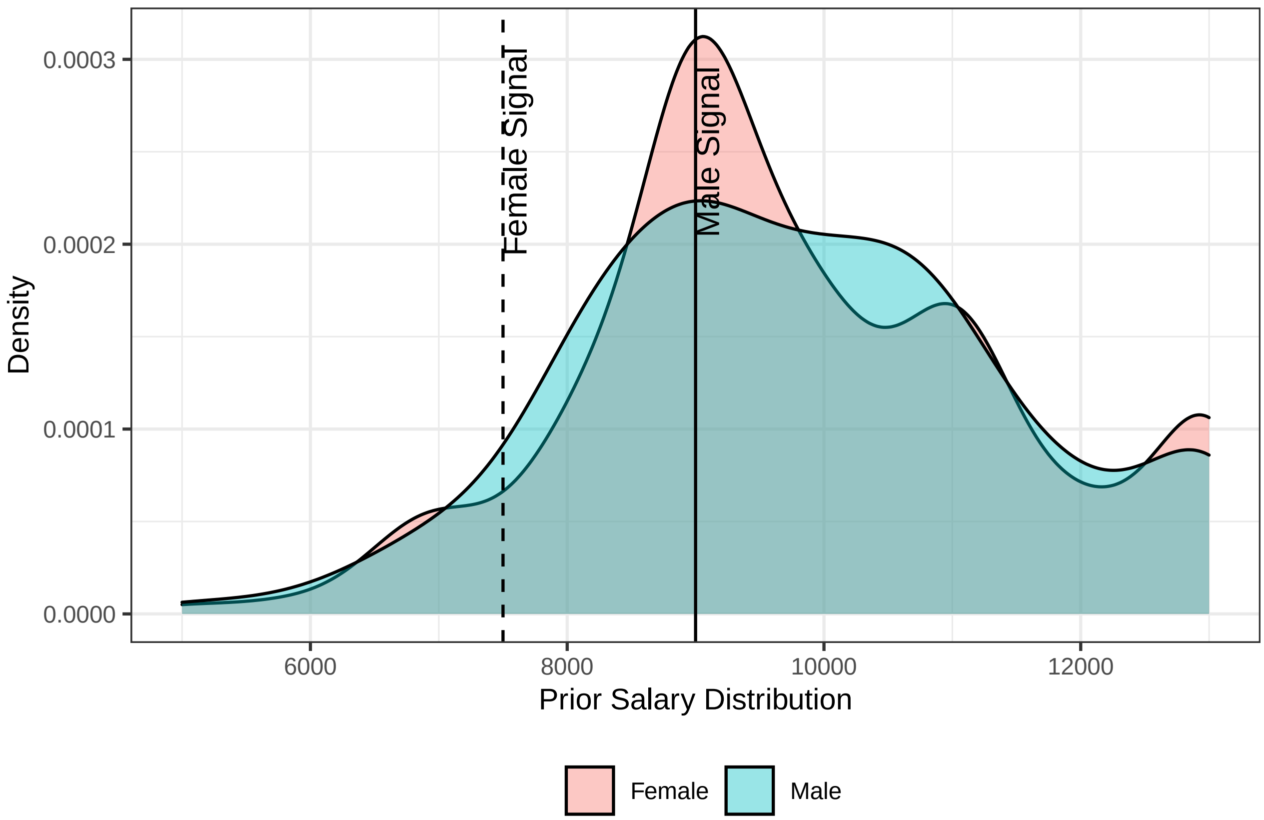


(a) Location (Candidates Outside State)



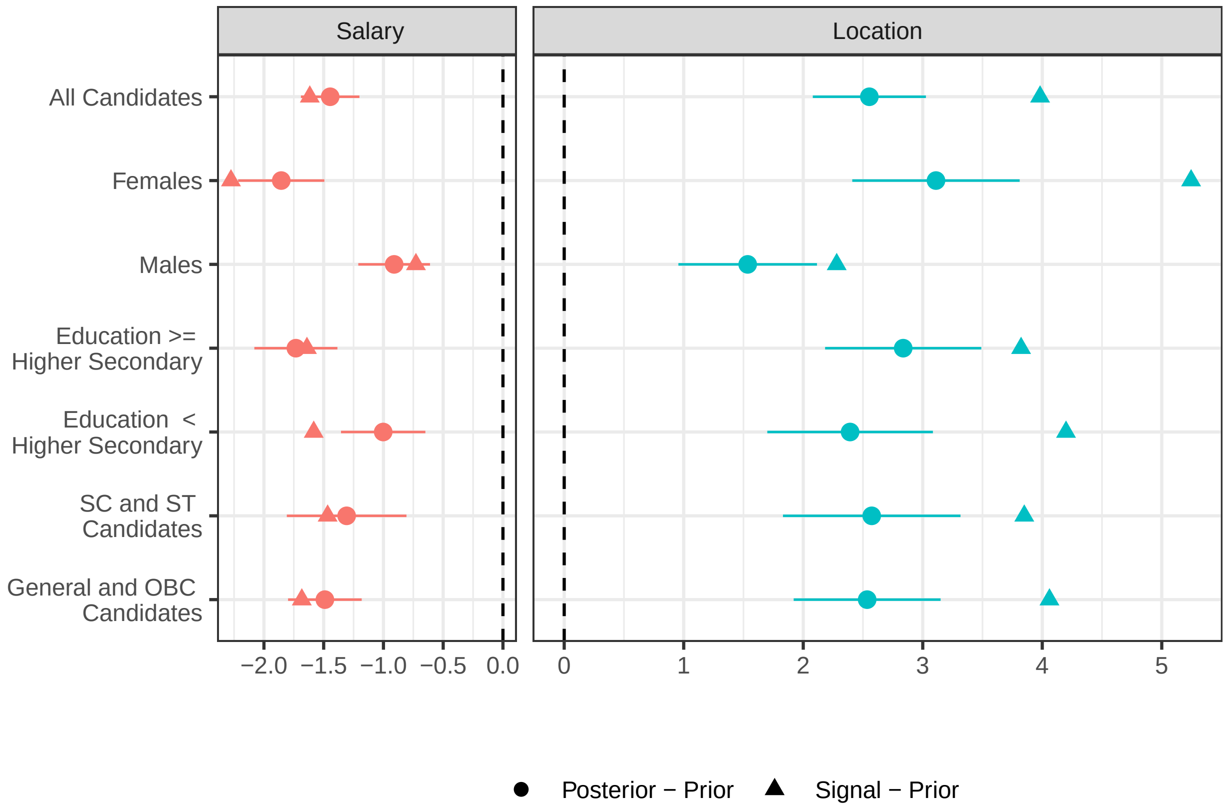
(b) Salary (Earnings Distribution Mean)

Figure A6: Prior Salary Distribution by Gender



Notes: The figure shows prior salary distribution by gender. The vertical line shows the truth/signal by gender. The actual salary for males rests at the 38th percentile of the prior salary distribution. For the females, the signal intersects the prior distribution at the 8th percentile.

Figure A7: Heterogeneity in Labor Market Beliefs



Notes: The figure shows heterogeneity in treatment effect for the salary and location intervention for sub-samples by gender, education levels, and social category of candidates. The circles and error bars show the point estimate and 95% CI on the indicator variable for the salary treatment (red color) and the location treatment (green color) regressed on the outcome variable: posterior - prior. The triangle shows the average gap between the signal and the prior. Posterior/Prior for salary is the earnings distribution mean calculated using the number of candidates in each bin. Posterior/Prior for location is the number of candidates outside state. The negative x-axis is scaled by 1000.

A Video Transcripts

A.1 Introduction Video

Voiceover: In the households of the village where there was not much enthusiasm so far, today there is hope. The young men in the rural areas, and especially the young women of the villages, who had never imagined their future outside the threshold of their houses, are today dreaming big and giving wings to their dreams because of their skills and self-confidence.

Now they are getting jobs in the organized work sector of big and metro cities.

Now happiness and smile never leaves their faces.

For lakhs of 15-35 years old rural youth, Indian Government has initiated this Rural Skill Development program.

To bring the youth from rural areas to the best training institutes and companies, this program is run on a public private partnership model.

Youth from rural areas of this country are brought and given free of cost training. Arrangements are also made for their free of cost boarding and lodging.

During the training, candidates are given books and uniform as well in the DDUGKY program.

DDUGKY program has opened lakhs of such opportunities for young men and women across this country, so that it has enabled them to write their own future with their own hands."

Female candidate: "I come from a poor family. Our family works on the farms and I have studied while working on the farms myself. My parents have taught me with great difficulty. I got to know about this free of cost training, DDU-GKY. I enquired where to get the form for this training and where is this happening. They called me that we have to leave for ranchi. . . . The facilities are good here. We had to live in hostel, the food was good.. three months we got training there. It was good, we used to have fun and play, everything was there. It feels good when we get our salaries. If we are independent people will give us importance and talk with respect.."

Male candidate: "I have my mother and father at home. We are 7 siblings, 3 brothers and 4 sisters. Before this I used to work as a daily laborer. I did not study much. I have passed my matriculation, that too with much difficulty, while working. I worked as a labor worker in a construction site where they make buildings. I worked as a helper for the masons. About DDU-GKY, they told this was a good course and they will teach us computers.."

Voiceover: "Their progressing steps towards their own brighter present are also making a stronger and developed future for India. This will turn this nation into a place of skilled individuals."

"My skill is my identity."

A.2 Intervention Video: Salary (Male)

In this video we will tell you about the monthly salary distribution of the male candidates after their training completion, in the last one year under the DDU-GKY skill development program.

Through our survey we have come to know that, after completing the training in the last one year, if 10 candidates like you got jobs, then nobody got a job for a monthly salary below Rs 6000. After completing the training in the last one year, if 10 candidates like you got jobs, then 2 male candidates got a job for monthly salary ranging between Rs 6000 to Rs 8000. Through our survey we have come to know that, after completing the training in the last one year, if 10 candidates like you got jobs, then 6 male candidates got a job for monthly salary ranging between Rs 8000 to Rs 10000. Through our survey we have come to know that, after completing the training in the last one year, if 10 candidates like you got jobs, then 2 male candidates got a job for monthly salary ranging between Rs 10000 to Rs 12000. After completing the training in the last one year, if 10 candidates like you got jobs, then nobody got a job for a monthly salary above Rs 12000.

Through this video we learn that after completing the training in the last one year, if 10 candidates like you got jobs, then nobody got a job for a monthly salary below Rs 6000, 2 male candidates got a job for monthly salary ranging between Rs 6000 to Rs 8000, 6 male candidates got a job for monthly salary ranging between Rs 8000 to Rs 10000, 2 male candidates got a job for monthly salary ranging between Rs 10000 to Rs 12000 and nobody got a job for a monthly salary above Rs 12000.

Thank you for paying attention to this.

A.3 Intervention Video: Salary (Female)

In this video we will tell you about the monthly salary distribution of the female candidates after their training completion, in the last one year under the DDU-GKY skill development program.

Through our survey we have come to know that, after completing the training in the last one year, if 10 female candidates like you got jobs, then 2 female candidates got a job for a monthly salary below Rs 6000. Through our survey we have come to know that, after completing the training in the last one year, if 10 female candidates like you got jobs, then 3 female candidates got a job for monthly salary ranging between Rs 6000 to Rs 8000.

Through our survey we have come to know that, after completing the training in the last one year, if 10 female candidates like you got jobs, then 5 female candidates got a job for monthly salary ranging between Rs 8000 to Rs 10000. Through our survey we have come to know that, after completing the training in the last one year, if 10 female candidates like you got jobs, then nobody got a job for monthly salary ranging between Rs 10000 to Rs 12000. After completing the training in the last one year, if 10 female candidates like you got jobs, then nobody got a job for a monthly salary above Rs 12000.

Through this video we learn that after completing this training in the last one year, if 10 female candidates like you got jobs, then 2 female candidates got a job for a monthly salary below Rs 6000, 3 female candidates got a job for monthly salary ranging between Rs 6000 to Rs 8000, 5 female candidates got a job for monthly salary ranging between Rs 8000 to Rs 10000, nobody got a job for monthly salary ranging between Rs 10000 to Rs 12000 and nobody got a job for a monthly salary above Rs 12000.

Thank you for paying attention to this.

A.4 Intervention Video: Location (Male)

In this video we will tell you about the job location of the male candidates after their training completion, in the last one year under the DDU-GKY skill development program.

Male candidates who got placed inside Bihar are in yellow color and male candidates who were placed outside Bihar are in blue color.

Through our survey we have come to know that, after completing the training in the last one year, if 10 male candidates like you got jobs, out of them 3 male candidates got a job inside Bihar and through our survey we have come to know that, after completing the training in the last one year, if 10 male candidates like you got jobs, out of them 7 male candidates got a job outside Bihar.

Through this video we learn that after completing the training in the last one year, if 10 male candidates like you got jobs, out of them 3 male candidates got a job inside Bihar and 7 male candidates got a job outside Bihar.

Thank you for paying attention to this.

A.5 Intervention Video: Location (Female)

In this video we will tell you about the job location of the female candidates after their training completion, in the last one year under the DDU-GKY skill development program.

Female candidates who got placed inside Bihar are in yellow color and female candidates who were placed outside Bihar are in blue color.

Through our survey we have come to know that, after completing the training in the last one year, if 10 female candidates like you got jobs, out of them 1 female candidate got a job inside Bihar and through our survey we have come to know that, after completing the training in the last one year, if 10 female candidates like you got jobs, out of them 9 female candidates got a job outside Bihar.

Through this video we learn that after completing the training in the last one year, if 10 female candidates like you got jobs, out of them 1 female candidate got a job inside Bihar and 9 female candidates got a job outside Bihar.

Thank you for paying attention to this.