Pre-Analysis Plan
Misperceptions of Career Incentives and Turnover: Evidence from Ethiopia

David Qihang Wu*

Maximiliano Lauletta†

Belay Mulat‡

1 Introduction

Numerous low- and middle-income countries are undergoing a process of industrialization, with governments actively pushing to transition away from subsistence-agriculture in favor of manufacturing. Part of this effort includes attracting foreign companies to start production in the country. Despite the fact that these manufacturing firms generally offer comparatively good formal job opportunities, turnover rates are high: over 30% of workers in large manufacturing firms quit within the first month of work (Blattman and Dercon, 2018). Given that industrialization is a relatively new process with most workers never having engaged in manufacturing before, it is possible for potential workers to

*Department of Economics, UC Berkeley, qihangwu@berkeley.edu.

†Department of Economics, UC Berkeley, maximiliano_lauletta@berkeley.edu.

‡Policies Studies Institute, belay.mulat@gmail.com

have misperceptions of the career incentives within manufacturing jobs, especially about long-run career trajectory.

In this paper, we study how misperceptions about career incentives in manufacturing can lead to high turnover rates in manufacturing firms. To this end, we conducted a field experiment in a major industrial Park in Ethiopia, which is the flagship industrialization project in the country, where dozens of foreign firms hire tens of thousands of workers in the textile industry (Hardy et al., 2022). We conducted a survey on 1,750 workers in the industrial park, in which we document significant misperceptions about career incentives: workers tend to be overly-optimistic about wages and the likelihood of being promoted into higher positions within a year, although there is substantial variation in priors.

We then conduct an information experiment in which we randomly select a subset of respondents and we provide them accurate information on wages and the likelihood of being promoted, both of which we calculate using a confidential survey conducted by the authorities of the industrial park. We then use administrative records linked to the survey IDs to track the turnover rates of workers in the treatment and control group. Our identification strategy relies on the random variation in posterior beliefs induced by the treatment.

Our paper contributes to several branches of the literature. The main contribution is to the literature that studies high turnover rates in manufacturing industries. The early literature focused on rich countries (Montgomery, 1989; Beckert, 2015; Farber, 1994, 1999), while more recent work has found high turnover rates in developing countries (Groh et al., 2016; Blattman and Dercon, 2018). These papers, however, only tangentially documented high turnover rates, without delving into the factors that drive this phenomenon. Our findings contribute to this literature by providing compelling evidence that misinformation about career incentives in manufacturing can drive the high turnover rates seen in industrial jobs in developing countries.

Our paper also contributes to the literature on frictions in job search. This literature has found that search frictions (Franklin, 2018; Abebe et al., 2018), matching frictions (Banerjee and Chiplunkar, 2022), and over-optimism (Spinnewijn, 2015; Banerjee and Chiplunkar, 2022) are significant factors behind low likelihoods of finding stable jobs. Our paper contributes to this literature by documenting how misinformation about career incentives can also constitute a significant friction in job search that then results in high turnover rates.

Finally, our paper contributes to the literature on behavioral job search. Research in this literature have documented that factors such as present focus (DellaVigna and Paserman, 2005) and the existence of reference points around past earnings (DellaVigna et al., 2017) can drive the job search effort among unemployed workers. Our paper contributes to this literature by documenting an additional behavioral factor behind job search efforts: misinformation about the career incentives within manufacturing jobs is a significant factor behind high turnover rates.

2 Research design and data

This study focuses on turnover in a major industrial park of Ethiopia. As a landmark industrialization project, the industrial park is a major labor recruiter of the local region, particularly in textile and garment industries. Every day, hundreds of job seekers come to the grading center of the industrial park, and wait in line in the hope for a job. The staff at the grading center collect requests from 22 large-scale firms every day, register most of the job seekers, and match them to firms based on a first-come-first-serve principle.

In the first round of baseline survey in May 2021, we sampled 547 workers in the grading center when they are waiting to be assigned a job (Round 1). To increase the sample size, in March 2022, we added another 1,203 new hires (Round 2). The study consists of the following steps:

1. Baseline survey: New hires take a baseline survey in the grading center to collect their initial perceptions of various job aspects and other variables (specified in Em-

pirical Strategy section). Additional grading test is conducted after the baseline to collect workers' cognitive skills, dexterity skills, and behavioral traits.

2. Information treatment: We randomly select half of the days to be treated clusters. On the treated days, 2/3 of the new hires are provided with the benchmark information of promotion and salary of upper-level positions. The benchmark information is constructed by a current worker survey conducted by the government also validated by our own firm surveys. In Round 2 baseline, right after the information treatment, we elicit workers' updated perceptions on the job aspects regarding career incentives. For most specifications, we will use Round 2 sample as the main sample, and use Round 1 sample for robustness.

   The benchmark information in the intervention is provided by the Ethiopian Investment Commission (EIC). The EIC surveyed 1,378 workers from October 2020 to February 2021. We calculated the proportion of workers who started as entry-level workers but were promoted to upper-level positions (quality check, team leader, supervisor) 1 year after and within 2 years, and the average salary of upper-level positions. We also conducted interviews with 10 major employers in the industrial park in November 2021 and collected the same information; employers reported very similar numbers as we calculated. For Round 2 baseline, we updated the benchmark information using the follow-up data from EIC. Given the lack of administrative data on workers' salary, this is the best available data source to calculate promotion likelihood and salary after promotion.

3. End-line survey: We follow up each worker at least one year after the baseline survey. We plan to conduct the end-line survey in March 2023 (for Round 1 sample, it will be two years after the baseline survey). We want to wait at least 1 year to allow workers to fully update their perceptions of long-run career incentives and potentially be promoted. We will collect their updated perceptions, employment, income,

and their career plans in the future.

4. Administrative data collection: We will first obtain turnover data from Ethiopian Investment Commission who has the access to the turnover records for each worker. We will then visit each firm in the sample and collect personnel records on sampled workers. Personnel records will include current base salary, current total salary, current position, productivity measure, and number of absent days in the last 30 days.

3 Effect of Misperceptions on Turnover

3.1 Measuring Misperceptions

Building on Blattman and Dercon (2018), we select quantifiable job aspects, confirm with firms the importance of these job aspects, classify them into three categories, and compare them to objective benchmark information obtained from firm surveys and the current worker surveys conducted by the government:

1. Amenities: These are the work requirements or benefits that affect workers' utility and apply to every worker from all career levels regardless of effort or performance. These quantifiable amenities include: Days of work per week; Hours of work per day; Overtime hours per week; Minutes of break per day; Percentage of major firms providing free or subsidized transportation; Percentage of major firms providing free or subsidized food.

2. Bonus and incentives: These are the benefits that depend on workers' effort or performance, but do not depend on workers' career levels. These include: Percentage of major firms providing attendance bonus; Salary increase in entry-level if workers stay for 6 months; Relative performance pay of top 10^{th}-percentile workers.

3. Career ladder: These are the benefits that depend on workers' career levels, including (1) Entry-level career information: Average total salary of entry-level workers; Percentage of workers who start from entry level; (2) Upper-level career information: Percentage of workers promoted from entry to upper level after one year; Average total salary of upper-level workers.

We measure misperceptions as the difference between perceptions and the truth divided by the truth. For example, if a worker has a 15% baseline misperception of entry-level salary, we interpret it as "this worker's baseline perception of entry-level salary is 15% more than the truth", and so on. For each aspect, we measure baseline misperception and end-line misperception. For the four aspects in the category "Career ladder", we also measure the updated misperceptions after the treatment.

3.2 Causal effect of misperceptions of career incentives on turnover

3.2.1 Efficacy of the information treatment

We first estimate the first-stage effect of information treatment on updated misperceptions on career incentives using a Bayesian update specification. Let $P_i^{x,0}$ be worker i's prior belief of job aspect x, $P_i^{x,1}$ the posterior belief immediately after the information provision, $P_i^{x,2}$ the posterior belief in the follow-up survey, $P_i^{x,s}$ the signal provided by the survey team. Bayesian learning implies that, after the signal is provided (information treatment), the mean of the posterior belief should be a weighted average between the signal and the mean of the prior belief; the weight α, ranging from 0 and 1, is determined by the variance of the prior and the variance of the signal. This prediction can be summarized as follows:

$$\log(P_i^{x,t}) - \log(P_i^{x,0}) = \alpha_t\big(\log(P_i^{x,s}) - \log(P_i^{x,0})\big), \text{ where } t \in \{1, 2\}$$

To empirically test the first-stage effect of information treatment on belief update, we use the following specification:

$$\log(P_i^{x,t}) - \log(P_i^{x,0}) = \tau + \alpha_t T_{c(i)} \cdot \big(\log(P_i^{x,s}) - \log(P_i^{x,0})\big) \tag{1}$$
$$+\beta_t\big(\log(P_i^{x,s}) - \log(P_i^{x,0})\big) + \epsilon_i, \text{ where } t \in \{1,2\}$$

α_1 and α_2 are the parameters of interest. In particular, α_1 shows the weight by which treated workers immediately update their perceptions when presented the benchmark information compared to control workers, α_2 the weight by which treated workers retain the perceptions over one year compared to control workers. Other parameters are not the main focus but provide some useful information. In particular, β_1 captures the spurious reversion towards the signal among control worker, β_2 shows the weight by which control workers update their perceptions without being presented benchmark information. β_2 can be useful to infer the extent to which workers update their perceptions by being exposed to other information provided by firms or among their social network.

Given that all firms provide information of entry-level salary and amenities on the first day of work but almost zero firm provides information of career progression, our predictions are as follows:

- $x \in \{$Amenities, bonus and incentives, entry-level career information$\}$: $\alpha_1 = \alpha_2 = 0$
- $x \in \{$Upper-level career information$\}$: $\alpha_1 > \alpha_2 > 0$

3.2.2 Effect of misperceptions on turnover: Instrumental variable approach

To causally identify the effect of misperceptions of career incentive on turnover, a simple comparison between treated and control group may not be ideal because the treatment effect depends on the prior belief in two ways. First, treated workers with higher prior belief than benchmark may react in a different direction than those with lower prior belief than benchmark (direction). Second, the first-stage effect is higher among those whose prior belief is more biased, and a 0-1 treatment variable may not be the most efficient instrument to capture the variation (magnitude).

To tackle this issue, we follow Cullen and Perez-Truglia (2022) and adopt the following instrumental variable approach. Specifically,

$$Y_i^t = \pi + \delta \log(P_i^{x,t}) + \eta \log(P_i^{x,0}) + A_i \phi + u_i \tag{2}$$

$$\log(P_i^{x,t}) = \kappa + \gamma_t T_{c(i)} \cdot \left(\log(P_i^{x,s}) - \log(P_i^{x,0}) \right) + \zeta \log(P_i^{x,0}) + A_i \psi + v_i \tag{3}$$

Equation 3 corresponds to the first stage of the IV regression, a variation of the Bayesian update model-derived equation 1. Equation 2 is the second stage of the IV regression. In particular, the main parameter of interest is δ, interpreted as the magnitude change in outcome Y_i^t caused by a 1 percentage change in perception $P_i^{x,t}$, *i.e.* worker i's perception on job aspect x at the end of the baseline. The reason we use $\log(P_i^{x,t})$ as the main independent variable instead of the bias measure $\log(P_i^{x,t}) - \log(P_i^{x,0})$ is to keep a flexible functional form in the estimation. In the main specification, we will use $P_i^{x,1}$ as the main independent variable, that is, the immediate updated perception of x (promotion likelihood or salary after promotion) at the end of baseline in Round 2. In the robustness check, we will use $P_i^{x,2}$, the updated perception in the follow-up survey, to verify the results.

The vector of additional control variables A_i is included to reduce the variance of the error term, generating higher precision in the estimates. It will include the following categories:

- Demographics: Age, marital status, whether the worker speaks Sidamagna (the main local language), whether the worker speaks Amharic (the main work language), religion, whether worker's family is from Hawassa
- Education and experience: Whether the worker has any college education, whether the worker has any work experience
- Cognitive ability: See Section 3.3.

- Dexterity skills: See Section 3.3.
- Intrinsic motivation: See Section 3.3.

We expect that workers will lower posterior belief of relevant job aspects x (promotion likelihood and salary after promotion), are more likely to drop out before signing the contract, but less likely to drop out after signing the contract:

- $Y =$ Quit before signing the contract: $\delta < 0$
- $Y =$ Quit after signing the contract: $\delta > 0$

3.2.3 Exclusion restrictions and alternative mechanisms

From the main results above, we expect to establish two facts: (i) Treated workers update perceptions of promotion likelihood and salary after promotion; (ii) Workers who updated perceptions of career incentives present different turnover behaviors. The main exclusion restriction assumption to establish the causation is $E\big[\big(\log(P_i^{x,s}) - \log(P_i^{x,0})\big) \cdot T_{c(i)} \cdot u_i\big] = 0$. In other words, for all levels of prior belief of job aspect x, the clustered treatment is not correlated with unobserved factors captured in the error term ϵ_i'.

One potential violation of this assumption may happen when the treatment provides general information different than the career trajectory of the assigned firm. Treated workers may update the perceptions of the average career trajectory, but when they are assigned to a firm with higher (or lower) promotion likelihood or salary after promotion, they may be less (or more) likely to quit before signing the contract and rejoin the complex in the hope for a better draw. Mathematically, $E[T_{c(i)} \cdot u_i] \neq 0$.

We plan to test this alternative interpretation as follows. First, we measure directly whether worker i ever rejoins the complex. If there is no treatment effect on workers' rejoining the complex, it is unlikely that treated workers strategically game the firm allocation system with better information. Second, we include firm fixed effects and cluster standard errors within firm (workers who never join any firm will be considered as one group). If the main results hold, the treatment effect is unlikely to be explained by within-firm correlation.

Another potential violation of this assumption is that worker's characteristics correlated with prior belief may affect the retention of information. For example, suppose workers with higher cognitive ability are more likely to have overly high prior of promotion likelihood; meanwhile, they are also more likely to retain information when treated. This is the case when $E\big[\big(\log(P_i^{x,s}) - \log(P_i^{x,0})\big) \cdot T_{c(i)} \cdot u_i\big] > 0$, leading to overestimation of parameters of interest.

We plan to test this alternative interpretation as follows. First, we look at what observed characteristics predict higher retention of information in Equation 1. Then, we include these characteristics interacting treatment status in the control vector A_i of Equation 2 and 3. If main results still hold, then the alternative mechanism above is unlikely, although it is still possible that we omit some other unobserved characteristics correlated both with prior beliefs and with information retention.

3.2.4 Spillover

People gossip. Our main specification looks at the effect of clustered treatment and confines information diffusion largely within cohorts, but it is likely that workers spread information outside cohorts through their social network, which may affect the turnover among control workers and lead to biased estimate.

We plan to test and control for across-cohort spillover by exploiting the baseline information of workers' social network. For each worker, we elicit the number of family and friends who have worked in the complex before. We then present 5 random names from the treatment group from the last 14 days and ask how many of them they personally know. We will include the two characteristics interacting with treatment status in the control vector A_i of Equation 2 and 3.

In addition, if we assume the across-cohort spillover is weaker than within-cohort spillover, we can estimate within-cohort spillover as an upper-bound of across-cohort spillover. To estimate within-cohort spillover, we will run Equation 2 and 3 but only on control workers. The estimate of δ should be interpreted as the comparison between control in control clusters and treated cluster, the latter exposed to potential within-cluster spillover.

3.3 Heterogeneity and sorting of workers

So far, we estimate the effect of misperceptions on average turnover before and after signing the contract. A more subtle potential effect is on sorting. For instance, workers of higher skills with over-pessimistic prior belief may be less likely to quit after treatment because they stand a higher chance of being promoted; workers of higher skills with over-optimistic prior belief, however, may also tend to stay because they may want to exert more effort to get the promotion. There is potentially a trade-off between exerting more effort to get the higher salary after promotion and leaving complex for a higher-paid job, and the trade-off varies for workers of different types. In this subsection, we will first look at the heterogeneous treatment effect on workers' turnover. In the next subsection, we will look at the treatment effect on workers' effort specifically.

The econometric specification of heterogeneity analysis is the following:

$$Y_i^t = \pi' + \delta_1 \log(P_i^{x,t}) + \delta_2 \log(P_i^{x,t}) \cdot B_i + \eta' \log(P_i^{x,0}) + A_i\phi' + u_i' \tag{4}$$

$$\begin{pmatrix} \log(P_i^{x,t}) \\ \log(P_i^{x,t}) \cdot B_i \end{pmatrix} = \kappa' + \mathbf{\Gamma} \begin{pmatrix} T_{c(i)} \cdot \big(\log(P_i^{x,s}) - \log(P_i^{x,0})\big) \\ T_{c(i)} \cdot \big(\log(P_i^{x,s}) - \log(P_i^{x,0})\big) \cdot B_i \end{pmatrix} + \zeta' \log(P_i^{x,0}) + A_i\psi' + v_i' \tag{5}$$

We require two assumptions for exclusion restriction: $E\big[\big(\log(P_i^{x,s}) - \log(P_i^{x,0})\big) \cdot T_{c(i)} \cdot u_i'\big] = 0$, $E\big[\big(\log(P_i^{x,s}) - \log(P_i^{x,0})\big) \cdot T_{c(i)} \cdot B_i \cdot u_i'\big] = 0$. The first one is the same as discussed before. The second assumption may be violated if workers' unobserved characteristics affect the treatment implementation given prior beliefs and heterogeneity of interest B_i.

We focus on the following heterogeneity: (i) Cognitive ability, (ii) Dexterity skills specific to garment industry, (iii) intrinsic motivation. Specifically, we construct three indices with the following data:

1. Cognitive ability: We first conduct a 12-question Raven test on each worker and compute a Raven score from the test. We then conduct a short memory test to measure the extent to which they remember a number sequence. In addition, we ask two simple questions to test their knowledge of current affairs (the year when Prime Minister Abiy Ahmed won Nobel Peace Prize; number of regions in Ethiopia). We will extract the principal component from these measures as a cognitive index, and group workers into high-cognitive type (cognitive index above median) and low-cognitive type (cognitive index below median).

2. Dexterity skills specific to garment industry: We conduct two simple games to measure workers' dexterity relevant to sewing and coordination. The first game requires workers to thread three needles within a minute. The second game requires workers to take 10 pin balls from a box, put each pin ball through a tube and drop it in a different box. Both games were inspired from the grading center of the industrial park who used to conduct grading test on new workers. We will extract the principal component from the two measures as a dexterity index, and group workers in high-dexterity type (dexterity index above median) and low-dexterity type (dexterity index below median).

3. Intrinsic motivation: We collect additional survey questions on why workers want to work here and their career plans. We will use the following questions to extract a intrinsic motivation index: (i) whether workers apply for the job because the job is interesting/good for skill development; (ii) how long workers plan to stay in the company; (iii) whether they plan to start their own business. We will group workers

into high-intrinsic-motivation type (intrinsic motivation index above median) and low-intrinsic-motivation type (intrinsic motivation index below median).

3.4 Internal margin: Effect of misperceptions on effort

Misperceptions may affect remaining workers' effort and thus productivity. Suppose a worker is over-optimistic of promotion likelihood and thinks he stands a high chance getting promotion with low effort. After the information treatment, he updates the real promotion likelihood, but instead of quitting, he decides to stay but exert more effort to get the promotion in the future. This provides more insights in the "internal" margin in addition to sorting: among the workers who stay, it is likely that misperceptions of career progression may affect their effort relevant to getting promotion.

To causally identify the effect of misperceptioins on effort, one major challenge is that workers who stay are already selected. A simple comparison between treated and control workers' effort measured from firms' personnel records is not causal because treated workers who stay in the firm may be systematically different, as it might be suggested in the previous subsection. To tackle this challenge, we will use Equations 2 and 3 to estimate average treatment effect on the following set of outcomes that do not select on workers who stay:

- Effort: Whether worker stays in the firm and has zero absent days in the last month.
- Productivity: We will utilize the productivity measure from the firm personnel record. We generate a relative productivity measure as follows: For each worker, we subtract their productivity measure from the average productivity measure of the same team in the same firm and divide it by the average. The dependent variable in the regression is whether worker stays in the firm and has above-average productivity.

3.5 Long-run effect of misperceptions on workers

Eventually, misperceptions of career incentives of the industrial park may affect workers' long-run economic outcomes. Information provision absent, it may take a long time for workers to realize their misperceptions of promotion and salary after being promoted. By then, workers who have over-optimistic prior may have wasted too much time on the job and lost other better opportunities; workers who have over-pessimistic prior may have quit too early and missed the chance of being promoted. We plan to collect survey outcomes one year after the baseline on workers' employment, income and expenditure, health, and aspirations to understand the full picture of the effect of misperceptions of career incentives. We will use Equations 2 and 3 to estimate average treatment effects.

The following is a list of outcomes in this section. For each category above, we will first report sharpened q-value to address multiple inferences, and then generate a principal component index and run the same specification on the index.

- Employment outcomes: (1) Whether the worker is currently employed; (2) Whether the worker is currently employed in a formal job; (3) Whether the worker signs a permanent contract on the job; (4) Salary of the current job if employed; (5) Whether the worker works in garment industry.

- Income and expenditure: (1) Total income received from work in the last month; (2) Total income received from family and friends in the last month; (3) Total expenditure on food in the last month; (4) Total expenditure on leisure goods in the last month (cosmetics, hair salon, etc).

- Health: (1) Whether the worker skips a meal in the last week; (2) Whether the worker falls ill in the last month.

- Aspiration: (1) Whether the worker plans to marry within 5 years if single; (2) Whether the worker plans to stay in the current job; (3) Whether the worker plans to move to other cities for better job opportunities; (4) Whether the worker plans to start their own business within 5 years.

References

Girum Tefera Abebe, Stefano Caria, Marcel Fafchamps, Paolo Falco, Simon Franklin, and Simon Quinn. Anonymity or distance? job search and labour market exclusion in a growing african city. 2018.

Abhijit Banerjee and Gaurav Chiplunkar. How important are matching frictions in the labour market? experimental & non-experimental evidence from a large indian firm. *Working paper*, 2022.

Sven Beckert. *Empire of cotton: A global history*. Vintage, 2015.

Christopher Blattman and Stefan Dercon. The impacts of industrial and entrepreneurial work on income and health: Experimental evidence from ethiopia. *American Economic Journal: Applied Economics*, 10(3):1–38, 2018.

Stefano DellaVigna and M Daniele Paserman. Job search and impatience. *Journal of Labor Economics*, 23(3):527–588, 2005.

Stefano DellaVigna, Attila Lindner, Balázs Reizer, and Johannes F Schmieder. Reference-dependent job search: Evidence from hungary. *The Quarterly Journal of Economics*, 132 (4):1969–2018, 2017.

Henry S Farber. The analysis of interfirm worker mobility. *Journal of Labor Economics*, 12 (4):554–593, 1994.

Henry S Farber. Mobility and stability: The dynamics of job change in labor markets. *Handbook of labor economics*, 3:2439–2483, 1999.

Simon Franklin. Location, search costs and youth unemployment: experimental evidence from transport subsidies. *The Economic Journal*, 128(614):2353–2379, 2018.

Matthew Groh, Nandini Krishnan, David McKenzie, and Tara Vishwanath. The impact of soft skills training on female youth employment: evidence from a randomized experiment in jordan. *IZA Journal of Labor & Development*, 5(1):9, 2016.

Morgan Hardy, Gisella Kagy, Christian Meyer, Eyoual Tamrat, and Marc Witte. The impact of firm downsizing on workers: Evidence from ethiopia's ready-made garment industry. *Working paper*, 2022.

David Montgomery. *The fall of the house of labor: the workplace, the state, and American labor activism, 1865-1925*. Cambridge University Press, 1989.

Johannes Spinnewijn. Unemployed but optimistic: Optimal insurance design with biased beliefs. *Journal of the European Economic Association*, 13(1):130–167, 2015.