THE UNIVERSITY OF CHICAGO

ESSAYS ON ONLINE LABOR MARKETS IN LOW-INCOME COUNTRIES

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF ECONOMICS

BY
RUOXUAN WU

CHICAGO, ILLINOIS
AUGUST 2026

ORCID iD: https://orcid.org/0000-0003-3220-8458

To my family -

my grandparents, Fengying Zhao and Xiru Zhang,

and my parents, Wei Zhang and Guoqi Wu -

for your love across every distance.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

This dissertation would not have been possible without the support, guidance, and generosity of many people around me.

I am deeply grateful to my advisors, Christina Brown, Marianne Bertrand, and Michael Dinerstein, for their guidance and faith in me throughout this journey.

I have worked with Christina since my second year in the PhD program. She inspired my fieldwork in Pakistan, generously supported my time in the field, and introduced me to the research community there. Over the years, she spent countless hours discussing the smallest details of experimental design, going through comments from presentations, and pushing me to think more carefully and confidently about my work. She has been a true role model, and her way of doing research has profoundly shaped how I approach economics.

Marianne has always reminded me to focus on the bigger picture: what is truly important and interesting about my research. She continuously challenged me to think harder about my results and what they mean. Her curiosity, energy, and excitement for ideas made many difficult moments of research feel motivating again.

Michael has been equally thoughtful, patient, and caring. There were many moments when this project felt impossible to implement in the field, and Michael consistently encouraged me to keep going and helped me see a path forward. Coming from a more experimental perspective myself, I especially appreciate how much he pushed me to think more deeply about the underlying economic mechanisms and broader questions behind the work.

I am also grateful to be part of the development economics community at UChicago. I have benefited enormously from the generosity of faculty and fellow students, who have shaped both my research and my understanding of what it means to be a development economist. I also feel incredibly lucky for the friendships I found during graduate school. I learned so much from the people around me, and I will always cherish all the laughter, support, and shared struggles along the way.

I am especially indebted to the research community in Lahore, Pakistan, including my wonderful coauthors and friends at the Lahore School of Economics and the Centre for Economic Research in Pakistan. They welcomed me with extraordinary warmth and generosity, and shared their knowledge, time, and resources so openly. None of this work would have been possible without them.

Finally, I am endlessly grateful to my family for carrying me through every stage of this journey. I thank my parents for always making the effort to understand what I was going through during this process, even when it was difficult to explain. They constantly reminded me that there is much more to life than research and academic success. I owe far more to them than I could ever express in words. I am also deeply grateful to my partner, who has been beside me through this entire journey, from our pre-doc days at UChicago until now. Thank you for sharing every high and low with me, and for being my best friend, research partner, and life companion.

ABSTRACT

Remote-work technologies have fundamentally reshaped labor markets, creating new opportunities for workers in low- and middle-income countries (LMICs) to access high-quality jobs previously beyond their reach. This shift will profoundly change how people acquire skills, access jobs, and organize within the labor market over the coming decades. In this dissertation, I investigate frictions that impede workers' access to digital job opportunities and skill acquisition in the online labor market.

In the first chapter, joint with Hamna Ahmed and Zunia Tirmazee, we study entry barriers for workers in online labor markets. Hundreds of millions of workers in low- and middle-income countries seek digital jobs online but face entry barriers without an established reputation. Novice workers could offset this by lowering initial wages, yet few do. Our baseline survey points to two explanations: workers believe employers interpret low wages as low quality signals and are uncertain about their own abilities. We conduct two field experiments on a leading global freelancing platform to examine how these beliefs shape worker outcomes. In the demand-side experiment, we randomize wage offers by novice workers to 703 jobs and find that employers respond favorably to lower wage offers from novices, contrary to what workers believe. In the supply-side experiment with 481 novice workers, we randomly provide them with accurate information about employer responses and their performance. Correcting workers' beliefs increases their willingness to lower wages. Consistent with reputation models, effects are driven by high-ability novices with high returns to reputation once these frictions are removed. Simulations show

that without external intervention, worker learning about employer responses is slow and costly. Our findings highlight that information interventions can help workers in developing countries overcome reputation barriers and accelerate talent discovery in online labor markets.

In the second chapter, joint with Hamna Ahmed, Zunia Tirmazee, and Emma Zhang, we study how to incentivize digital skills investment among young women in low-income countries, a key prerequisite for participation in online labor markets. In particular, we study the role of intra-household payment and information targeting on the effectiveness of a conditional cash transfer (CCT) program using a randomized control trial. The program aims to boost completion of a digital skills training program among young females in urban Pakistan. Fixing the incentive size and daughters' knowledge about it, we cross randomize (1) the payment split between parents and daughters and (2) whether parents receive information about the daughters' incentive. We find that under asymmetric information about the CCT, incentivizing parents leads to a 103% increase in training completion compared to incentivizing daughters. When both parents and daughters know about the CCT, completion rates do not vary by the incentive split, consistent with the efficient collective household model. Our results suggest that in this parent-child context, incomplete information sharing is the main barrier to the optimal incentive targeting, instead of bargaining frictions on the future payment.

CHAPTER 1

BELIEFS, REPUTATION, AND BARRIERS TO ENTRY IN ONLINE LABOR MARKETS

1.1 Introduction

Inexperienced workers often struggle more than experienced ones to find jobs, particularly in low- and middle-income countries (LMICs).[1] A key challenge is their inability to credibly signal their skills to employers, particularly when their qualifications are difficult to evaluate or verify (Abebe et al., 2021; Bertrand et al., 2024). This problem is especially pronounced in the rapidly expanding online labor market, where hundreds of millions of workers from these countries seek digital job opportunities on global freelancing platforms (Datta et al., 2023). Although these platforms expand access to opportunities, information frictions about worker ability create substantial entry barriers for inexperienced workers without prior online job histories (Agrawal et al., 2016; Hilbert and Lu, 2020). Because employers cannot observe worker ability prior to hiring, they favor workers with established reputations, creating a "reputation trap" for inexperienced workers: they need jobs to build reputation but need reputation to secure jobs. Classic reputation models (Tirole, 1988; Stiglitz, 1989) suggest that inexperienced workers can overcome this problem through "introductory pricing": offering lower wages initially to attract employers and raising them

1. For example, in some regions, they are three times more likely to be unemployed than the general population (Alfonsi et al., 2020).

once reputation is established. Yet in practice, inexperienced workers rarely lower wages to build reputation. Why do they refrain from such reputation investment?

We study this question using detailed platform and survey data from LMIC freelancers, together with two randomized controlled trials on a leading global freelancing platform. We first document suggestive evidence that this pattern may reflect workers' beliefs about employer responses or their own abilities. Workers may hesitate to lower wages because they believe employers interpret low prices as signals of low quality.[2] Alternatively, workers who are uncertain about their own ability cannot predict their future performance or the reputation they will establish, making them reluctant to forgo initial earnings for future returns to reputation.

The first experiment tests whether employers truly interpret low wage offers from inexperienced freelancers as signals of low quality. We randomize wage offers by inexperienced (novice) and experienced (veteran) freelancers and compare employer responses. The second experiment examines how novices' beliefs shape wage offers. We randomly provide novices with information about employer responses and/or their own ability, and measure the effects on subsequent wage offers. Our results show that novices' inaccurate beliefs about employer responses and their own ability deter reputation investment and market entry, especially among high-ability workers. These frictions not only distort talent allocation in the online market, but are especially costly in LMICs, where limited local employment opportunities make access

2. This pattern has been well documented in product markets where consumers infer lower quality from lower prices when evaluating unfamiliar goods. See Rao and Monroe, 1989; Wathieu and Bertini, 2007; Erdem et al., 2008; Ashraf et al., 2013.

to online work particularly valuable for skilled workers.

We focus on data entry freelancers from LMICs on one of the largest global freelancing platforms. On the platform, employers post jobs with stated budgets, and freelancers apply by submitting wage bids. We define *undercutting* as freelancers bidding below the posted budget. When evaluating applicants, employers observe each freelancer's wage bid and reputation, represented by the number of completed jobs and job success rate on the platform. We combine comprehensive platform data on freelancers' profile characteristics, job histories, and applications with original survey and performance data collected from 481 novice freelancers across 37 LMICs hired for a standardized data entry task on the platform.

Our descriptive analysis reveals a puzzle. Reputation generates large returns to future earnings, and these returns are heterogeneous: public employer reviews increase hiring chances for high-performing novices but shut out low-performing ones (Pallais, 2014; Fazio et al., 2025). High-ability novices therefore stand to gain the most from securing initial jobs and building reputation early. This is especially important from an efficiency standpoint: novices in our data perform no worse than experienced workers with established reputations, suggesting the reputation trap may generate talent misallocation. Yet despite intense competition, wage bids vary little, even among novices with limited outside employment options.[3] Our original data points to two explanations. First, although novice performance varies substantially, work-

3. Our sample represents a substantial underemployed labor force with limited job opportunities. Nearly half report no offline work, and more than 80% have no completed jobs on this or any other freelancing platform. For these workers, overcoming initial entry barriers and building a reputation to secure digital jobs is especially valuable.

ers have limited knowledge of their abilities relative to others. Without knowing her type, a novice cannot predict whether she will perform well enough to recoup the initial earnings sacrifice, because a poor reputation may prevent her from securing future jobs. Second, novices believe that employers interpret low wage bids as signals of low quality, further deterring undercutting.

Motivated by these descriptive findings, our first experiment examines whether employers interpret low wage bids from novices as negative quality signals and thus whether undercutting is a viable entry strategy. To generous exogenous variation in wage bids, we create experimental profiles for novice and veteran freelancers with comparable characteristics, randomize their bids across 703 real jobs on the platform, and track employer responses.

The results are contrary to workers' baseline beliefs. Although novices are less likely than veterans to receive employer responses, low wage bids significantly increase employers' likelihood of viewing the application and contacting freelancers in both groups. Crucially, employers are more responsive to low bids from novices than from veterans: novice callback rates increase by 54 percent from a baseline of 3.41 percentage points, compared to a 16 percent increase from a baseline of 5.26 percentage points for veterans. These patterns are inconsistent with employers interpreting low wages as negative quality signals. Instead, employers appear to view lower bids by novices as a willingness to invest in reputation. One concern is that even if undercutting helps novices secure jobs, the future returns to reputation may not offset the initial earnings loss. Our back-of-the-envelope calculations suggest otherwise:

novices who consistently undercut can more than double their net earnings within the first year. Taken together, these findings suggest that undercutting can help novices overcome the reputation trap, raising the question of why so few workers adopt this strategy.

Our second experiment therefore examines how novices' beliefs about employer demand and their own ability affect wage bidding behavior. The experiment is informed by Tirole (1988), in which new entrants offer introductory pricing as an investment in reputation to overcome initial information frictions about their quality. In the original model, a separating equilibrium can emerge where high-ability novices offer low wages initially to build reputation for future transactions, while low-ability novices do not. Low wage offers by high-ability novices increase their chances of being hired and improve match quality. We show that the separating equilibrium breaks down when novices hold inaccurate beliefs about employer responses and their own ability. This framework yields testable predictions: correcting beliefs about demand encourages undercutting broadly, while correcting beliefs about own ability induces undercutting by high-ability novices.

We test these predictions in an experiment with 481 novice freelancers hired for the baseline data entry task, for whom we have measured their performance and beliefs about own ability and demand responses. Identifying the mechanisms using observational data alone is difficult because ability, beliefs, and wage bids are jointly determined. Our experiment generates exogenous shifts in beliefs by randomly providing novices with private information about their task performance relative to other novices and/or information about employer responses to wage bids. We then track their wage bids for a real job referred to them by us on the platform.[4] Leveraging baseline and endline surveys, we confirm that the treatments substantially shift novices' beliefs through the intended channels, without affecting perceptions about other aspects of the job application process.

We find that correcting novices' beliefs significantly increases their willingness to offer low wage bids, but through different channels. When novices learn that employers do not treat low bids as negative quality signals, the share who undercut nearly quadruples relative to the control group, from 5.5 percent to 20.5 percent, with no difference between high- and low-ability workers. By contrast, information about own performance increases the share of workers who undercut to 14.4 percent. This effect is concentrated among high-ability novices, especially those with more accurate baseline beliefs about employer demand. This pattern suggests that high-ability workers understand which strategies are effective but are uncertain about their own standing; receiving credible information about their ability induces them to invest in reputation through lower initial wages. When workers receive both types of information, we observe separation in wage bids: high-ability novices become five times more likely to undercut relative to the control group, while low-ability novices become only twice as likely. Moreover, we find no evidence that workers who undercut are more likely to renegotiate wages or produce lower-quality output. Taken together, the results show that misperceptions about employer demand and uncertainty about

4. We use this referral approach because job applications are not publicly observable on the platform. It allows us to observe application behavior while preserving workersâĂŹ incentives to appeal to prospective employers.

own ability both deter novices from investing in reputation. Information about employer responses encourages more novices to lower wage bids to secure initial jobs, while information about own ability primarily induces high-ability novices to invest in reputation through undercutting, improving talent allocation as in Tirole (1988).

The demand- and supply-side experiments show that inaccurate beliefs prevent novice workers from offering lower initial bids to overcome entry barriers. For policymakers, an important question is whether these belief frictions naturally dissipate with experience on the platform or whether external interventions are needed. Using estimates from both experiments, we simulate a Bayesian learning process to assess the time and monetary costs for perfectly rational workers to learn employer responses by experimenting with different bidding strategies. Even under systematic experimentation and perfect updating, learning is slow and costly, which help explain why these frictions may persist: a median worker with moderately inaccurate beliefs would require roughly 200 job applications and incur about $270 in application fees.

We also draw on recent evidence to argue that similar supply-side frictions exist among online freelancers in other service categories and in traditional offline labor markets in developing countries. More broadly, these findings help explain why skill training programs for inexperienced jobseekers alone often have limited effects on employment outcomes. While such programs improve workers' skills, novices still face barriers in credibly signaling those skills to employers and acquiring the initial reputation needed to secure future jobs. Our results instead point to the importance of external information provision (e.g., through mentorship programs) to facilitate market access and the growth of high-ability workers.

This paper bridges three fundamental issues that constrain young jobseekers in LMICs: workers' limited ability to signal their skills, their lack of information about the labor market and their own abilities, and the central role of initial reputation in securing jobs - particularly in online labor markets. We study these frictions in the increasingly important context of global freelancing platforms. Employers on these platforms rely heavily on workers' job histories to infer quality (Pallais, 2014; Agrawal et al., 2016; Barach and Horton, 2019), and alternative signals available to novice workers, such as micro-credentials or advertising, have limited effectiveness (Kässi and Lehdonvirta, 2022; Hannane, 2024). Building on classic reputation models (Tirole, 1988; Stiglitz, 1989), we provide causal evidence that novice workers can compensate for missing reputation by initially lowering wages. However, they under-adopt this strategy because of the belief frictions documented above. While prior work emphasizes upskilling freelancers in LMICs and finds limited effects on platform success (Baptista et al., 2023; Das et al., 2024; Fazio et al., 2025), our results show that when workers cannot credibly signal their skills, correcting belief frictions is essential for adopting strategies that overcome entry barriers.

Second, we contribute to growing body of work documenting jobseekers' inaccurate beliefs as a barrier to employment (see Caria et al., 2024 for a review). These beliefs are central because they shape job search strategies and, in our setting, determine whether workers are willing to sacrifice current earnings to build reputation. Prior work emphasizes two types of misbeliefs. The first concerns overly optimistic expectations about wages (Banerjee and Sequeira, 2023; Alfonsi et al., 2024; Kelley et al., 2024) or hiring probabilities (Chakravorty et al., 2024; Abebe et al., 2025; Bandiera et al., 2025), which can lead jobseekers to reject viable opportunities. We identify a new mechanism: misbeliefs about how employers interpret low wages as signals of worker quality. We experimentally show that even though novices are willing to work for less to prove themselves, pessimistic beliefs about employer responses deter them from adopting effective wage bidding strategies. The second concerns workers' imperfect information about their own abilities (Bassi and Nansamba, 2022; Carranza et al., 2022; Kiss et al., 2023). Correcting these misperceptions helps jobseekers target roles aligned with their strengths. We extend this insight by embedding it in a reputation model and show that performance feedback induces high-ability novices to lower wages to build reputation, generating separation in bidding behavior and potentially accelerating talent discovery on the platform.

Finally, we contribute causal evidence to the theoretical literature on prices as signals of quality, from both sides of a labor market. Theory predicts that market conditions determine whether sellers signal high quality through *high* prices (Wolinsky, 1983; Milgrom and Roberts, 1986; Judd and Riordan, 1994) or *low* prices (Tirole, 1988; Stiglitz, 1989). Empirical work has focused almost entirely on product markets. Empirical tests of these predictions in labor markets are scarce because worker types and beliefs are inherently unobservable. Roussille (2024) shows suggestive evidence that employers interpret higher wage requests as signals of higher ability on a US online labor platform. Closest to our work, Huang et al. (2025) examine how novice domestic-service workers in China use prices to signal quality, relying on initial worker characteristics as proxies for unobserved types. We advance this literature by collecting novel, directly measured data on workers' true abilities and their beliefs, and by experimentally varying both wage bids and information. This approach allows us to show that novice workers and employers hold systematically misaligned beliefs about how wages signal quality, and that these belief gaps are costly to correct in a market characterized by high search costs and intense competition.

The remaining sections are organized as follows. Section 1.2 provides background on the online freelancing platform and online labor markets more broadly. Section 1.3 describes the data collection and presents the motivating facts. Section 1.4 outlines the demand-side experimental design and present the results. Section 1.5 introduces the supply-side experimental design, motivated by a reputation model in the vein of Tirole (1988). Section 1.6 presents the results. Section 1.7 discusses policy implications for our findings and their generalizability to other labor market settings. Section 2.6 concludes.

1.2 Online Freelancing Markets

Online freelancing markets have become an important component of the global workforce. Recent estimates suggest that between 154 and 435 million people engage in online gig work, with the majority based in LMICs (Datta et al., 2023). These platforms connect workers with employers to perform and deliver remote, one-time tasks ranging from simple data entry to complex software development and graphic design. Workers are paid either per-project (fixed payment) or on an hourly basis.

For this study, we focus on fixed-payment, data entry jobs from one of the world's largest online freelancing platforms. This setting is well-suited to studying how novice workers from LMICs enter online labor markets for several reasons. First, data entry is highly prevalent: according to the World Bank, it represents the second most popular job category, accounting for 23% of total jobs after software development and technology (Datta et al., 2023). Second, these jobs are accessible to new entrants from LMICs, requiring less advanced skills and serving as a common entry point to the platform. Third, fixed-payment contracts dominate initial employment: 74% of data entry workers in our sample receive fixed-payment contracts in their first job. Together, these characteristics make this an ideal context for examining the early platform experiences of LMIC workers.

To understand how workers navigate this market, we outline the job matching process below. While we focus on a specific platform, this process shares key features with other major platforms. First, employers post a public job listing specifying the task description, required skill level, expected duration, and budget (either fixed payment or an hourly rate). Second, workers either apply directly to the posting or receive an invitation from the employer; for new workers, direct applications are the most common pathway. Third, to apply, workers submit a wage bid, a cover letter, and pay an application fee.[5] Wage bids can fall above or below the posted budget, provided they meet minimum requirements ($5 for fixed-payment projects, $3 per hour for

5. According to the platform, the application fee serves as a screening mechanism to reduce congestion by ensuring only serious applicants apply. While no public data exists on the fee amount, our primary data collection reveals fees between $1.20 and $1.65 per application. This fee requirement is unique to the studied platform, though workers on other platforms still incur time and effort costs when applying.

hourly projects). Bids are visible to employers, not to other applicants.

Once applications are submitted, employers review applicants through a dashboard displaying basic information: profile title, number of completed jobs and earnings, job success rates, self-reported skills, and wage bid. As shown in Appendix Figure A.1, the number of completed jobs, job success rates, and wage bid are the most salient features. Employers can also click on individual applications to review cover letters and full freelancer profiles, which typically include self-reported credentials and detailed job history. To move forward, employers contact workers through the platform's messaging system to either initiate contracts or request additional information. Although wage negotiation is possible at this stage, it rarely occurs (Barach and Horton, 2019), and workers are typically paid their proposed wage bid. Once a contract is signed, the job automatically appears on the worker's profile. Upon completion, both parties publicly rate each other on a one-to-five scale.

1.3 Data and Stylized Facts

In this section, we outline our main data sources, including platform data, worker performance, and baseline surveys, and document key stylized facts that motivate our experiments.

1.3.1 Platform Data Collection

We collect three sources of platform data on data entry freelancers to document competition and wage bidding patterns: profile characteristics, job history, and job applications. We begin by describing our sample and key variables before presenting the findings.

Basic Freelancer Profile We scraped listings of all data entry workers located in Bangladesh, India, and Pakistan in May 2024. These three countries account for over 50% of the global freelancer population (Online Labour Observatory, 2025). The sample comprises 10,818 freelancers. For each, we collected basic profile information visible to employers in the platform's search engine: number of completed jobs, total earnings, total hours worked, job success rate, preferred hourly rate, and agency association.

Job History To examine workers' career trajectories on the platform, we obtained detailed job histories for all 4,887 data entry workers in Pakistan, who represent over 45% of freelancers across the three countries. For each completed job, we recorded its duration, contract type, earnings, and employer rating.

Job Applications Application data is not publicly available on the platform. As part of the freelancer experiment, we collected application records from 2,332 workers outside our experimental sample to assess competition and wage bidding patterns. These records include wage bids, cover letters, and basic profile characteristics such as number of completed jobs and total earnings.

1.3.2 Stylized Facts from Platform Data

Using these data, we document three patterns that together motivates the central puzzle: why don't novice freelancers compete on prices to gain entry?

Fact 1. *Novice freelancers face significant entry barriers.*

We begin by examining how many freelancers successfully secure work on the platform. Among all data entry workers, 44% have never completed a job, and 62% have completed two or fewer. This aligns with broader evidence: International Labour Organization (2021) reports that 27-77% of freelancers across five major platforms have never completed a job, while Hilbert and Lu (2020) estimate that only 10% of registered freelancers in Latin America have ever secured platform work, with jobs highly concentrated among established workers. These patterns suggest that breaking into the market is a significant hurdle for new entrants.

Fact 2. *Freelancers' job prospects improve with the number of completed jobs.*

Next, we ask whether early jobs shape subsequent outcomes. Using job history data, we examine freelancers' growth trajectories. Appendix Figure A.2 plots wait time until the next job and log earnings from the next job against the number of cumulative jobs completed. Those with more completed jobs experience shorter gaps between jobs and faster earnings growth. While these patterns likely reflect selection on worker ability, they illustrate why securing initial jobs may be particularly valuable for high-ability novices. This is consistent with existing findings that initial reputations benefit high-performing novices and harm low-performing ones in securing future work (Pallais, 2014; Fazio et al., 2025).

Fact 3. *Competition is intense, yet wage bids vary little.*

Given the entry barriers and the returns to initial jobs, a natural question is whether novices compete on price. Application data reveals intense competition: each vacancy receives an average of 64 applications, 20% from experienced workers with at least three completed jobs. Despite this, there is little variation in wage bids. Appendix Figure A.3, 80% of applicants bid exactly at the posted budget, 15% bid below and 3% above. Among novices with fewer than three jobs, just 17% underbid.

This pattern is puzzling. If initial employment provides valuable reputation-building opportunities, rational novices should bid below experienced workers to increase their hiring probability. We next explore potential explanations for the limited variation in wage bids.

1.3.3 Measuring Worker Performance and Beliefs

The limited wage variation could be explained by two potential reasons.[6] First, in a competitive labor market where wages equal marginal products of labor, similar bids may reflect similar worker productivity. Second, workers may share beliefs about returns to wage undercutting that discourage bidding below posted budgets. Since

6. An alternative explanation is that workers avoid bidding down wages due to threats of social sanction. This is unlikely in our setting. Prior literature documenting such effects focuses on village economies where workers are in close networks. In online marketplaces, freelancers operate independently and their actions are not observed by other workers. We confirmed this in our qualitative survey with freelancers: few of them expressed concerns that bidding low wages would be deemed unfair to others.

neither novices' ability or beliefs can be observed from platform data, we collect primary data by hiring novices to complete a standardized task and baseline survey, then validating the performance measure through a follow-up task.

Main Sample Recruitment We posted a beginner-level data entry and online research job on the platform and hired freelancers over eight rounds between May and August 2024.[7] Eligible applicants had to meet three criteria: (1) based in LMICs, (2) specialized in data entry, and (3) completed no more than one job. To capture active job seekers, we selected eligible applicants on a first-come, first-served basis until all slots were filled.[8] The final sample comprised 481 novice freelancers.

Standardized Task Hired freelancers received detailed task instructions, which involved correcting existing entries against the original sources and adding new data from scanned PDF files into an Excel Spreadsheet within 24 hours.[9] To incentivize worker effort and ensure our performance measure reflected true ability, we informed workers that the top 3 performers would receive additional bonuses.

Worker performance on the data entry task was evaluated based on three objective dimensions: accuracy, speed, and following instructions.[10] Accuracy was measured

7. We conducted the hiring in batches because bringing all workers onboard simultaneously would be logistically challenging and delay performance evaluation across the sample.

8. To minimize attrition and reduce the risk of scams, we excluded freelancers who submitted unprofessional cover letters or proposed unreasonable wage bids.

9. Although workers were allowed to use optical character recognition (OCR) or AI tools, manual work was necessary since the scanned PDFs were of poor quality and the spreadsheet required a specific format.

10. To ensure these measures capture relevant dimensions of quality, we conducted extensive interviews with freelancers prior to the experiment and confirmed these three aspects were considered

as the percentage of correct entries, determined by comparing worker output against verified entries prepared by professional data entry specialists before the experiment. Speed measured the time between task release and final submission. Ability to follow instructions was scored from one to six, with each point reflecting adherence to a specific formatting requirement. To create comparable performance measures across workers, we ranked all freelancers within each round on each dimension and assigned quartile scores from 1 (lowest) to 4 (highest). The total performance score equals the sum of these three quartile scores, ranging from 3 to 12. Workers scoring above the median score within their round were classified as high-type, and others as low-type.[11]

Baseline Survey Following the standardized task, workers completed an online survey, which collected demographics and job search behaviors to complement the platform data. Since the limited wage variation could stem from either similar ability or similar beliefs, we measured workers' perceptions of both their own performance and the effectiveness of wage undercutting. To measure perceived performance, workers were incentivized to predict their performance quartile on each of the three evaluation dimensions, with cash bonuses for correct predictions. These predictions sum to a perceived performance score (ranging from 3 to 12) that can be compared to actual performance.

the most important performance indicators. In addition, this approach is similar to the performance measurement in Pallais (2014) and Kala and Lyons (2025).

11. When freelancers had the same score, we broke ties by sequentially comparing their accuracy, ability to follow instructions, and speed.

To measure perceived effectiveness of wage undercutting, we elicited worker beliefs about different approaches for novices to secure jobs, such as writing cover letters, paying the platform to boost their applications, and bidding below the job budget. Multiple strategies were included to mitigate concerns about experimenter demand effects. These beliefs can later be benchmarked against actual employer responses from the demand experiment to assess whether workers accurately understand the value of undercutting.

Follow-up Task To cross-validate our performance measure and compare novice versus veteran workers, we posted a second data entry job on the platform and extended 66 offers using stratified random sampling: 34 to in-sample and 32 to out-of-sample applicants. Among in-sample workers, we stratified by baseline performance and whether they bid below the posted budget to ensure a balanced sample for comparing subsequent performance across ability types and bidding strategies. Among out-of-sample applicants, we stratified by experience level, novice versus veteran, to allow for comparisons between new entrants and more established workers. Sixty-two applicants accepted the offer.

The task involved extracting data from financial reports into an Excel spreadsheet. Performance was measured using the same three dimensions as before, with one modification: because nearly half of workers followed instructions perfectly, we used a binary indicator for perfect compliance rather than quartiles. Therefore, the total performance score ranges from 3 to 10.

1.3.4 Stylized Facts from Performance and Belief Data

Our main sample comprises 481 novice freelancers from 37 LMICs. Table 1.1 reports the sample characteristics measured in the baseline survey. Freelancers are roughly evenly split by gender and average 29 years old. Three features are particularly notable. First, the sample represents a large idle, skilled workforce: nearly half (48%) are not currently employed offline, yet 83% hold at least a bachelor's degree. Second, workers struggle to break into the online market. Despite being registered for an average of 19 months, 87% have completed no jobs. Third, the entry barrier extends across platforms. While 44% have tried other freelancing platforms, only 18% have earned any money from them. Therefore, building reputation to access digital jobs is especially valuable for these novices, given their limited offline employment opportunities and persistent difficulty breaking into online markets.

Table 1.1: Freelancer Sample Characteristics

	Mean	Median	Min	Max
Female	0.52	1	0	1
Age	28.76	28	18	53
Bachelor Degree or Above	0.83	1	0	1
Employed Offline	0.52	1	0	1
Available as Full-Time Freelancer	0.36	0	0	1
Platform Tenure (Month)	19.65	12	0	278
Number of Applications/Month	9.80	7	1	73
Number of Completed Jobs	0.13	0	0	2
Present on Other Platforms	0.44	0	0	1
Has Earned from Other Platforms	0.18	0	0	1

Notes: This table reports sample characteristics for 481 freelancers from the baseline survey described in Section 1.3.3.

Next, we document patterns in novice freelancers' ability and beliefs to motivate potential explanations for the limited wage variation.

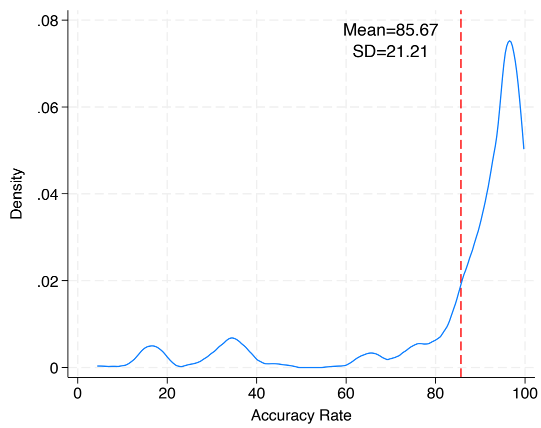
Fact 4. *Novices show substantial variation in task performance at baseline.*

We find heterogeneity in baseline task performance among novice workers along three dimensions: accuracy, speed, and ability to follow instructions. Figure 1.1 plots the distributions for each dimension and the total performance score, constructed by summing quartile rankings across the three measures. Accuracy and task completion time display wide dispersion, while instruction-following varies more moderately within the scale. The resulting total performance scores span the full range (3 to 12) and indicate substantial dispersion. This variation in performance contradicts the hypothesis that limited wage variation reflects similar productivity among novices.

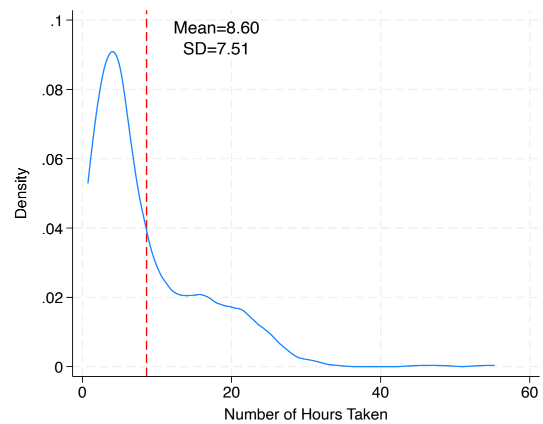
Given heterogeneity in novices' baseline performance, we compare novices and veterans in the follow-up task to evaluate a central efficiency question: are novices less capable than veterans, potentially justifying their difficulty securing jobs? Appendix Figure A.4 Panel A shows that novices outperform veterans on average. To ensure this pattern reflects genuine ability differences rather than noise, we examine whether performance among veterans increases with experience. Panel B confirms that among veterans, performance is positively correlated with completed jobs, as expected. Together, these patterns point to potential talent misallocation: capable novices face entry barriers despite outperforming many incumbents. That said, the sample is small ($n = 62$), so these findings should be interpreted as suggestive rather than conclusive.

Figure 1.1: Distribution of Task Performance Among Novices

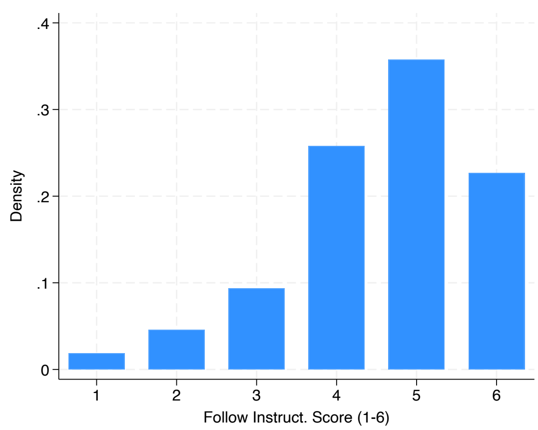
Panel A. Accuracy



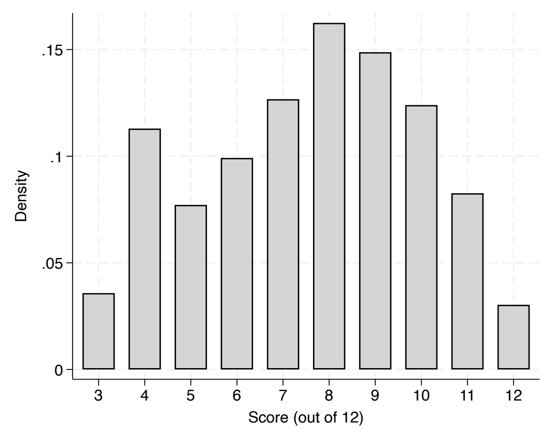
Panel B. Speed



Panel C. Follow Instruction



Panel D. Distribution of Performance Scores



Notes: This figure shows the distribution of the share of correct entries (accuracy), hours taken to complete the task (speed), the following instruction scores, and the total performance scores for the standardized data entry task.

Fact 5. *Novices have limited knowledge of their own performance relative to others.*

When workers lack accurate information about their ability, wage bids are shaped by perceived rather than actual performance. Figure 1.2 Panel A compares novices' true and perceived task performance and shows that most misperceive their performance quartile.[12] This suggests that limited wage variation is more likely driven by uncertainty about one's own ability rather than by similarity in true productivity.

A potential concern is that the gap between actual and perceived performance reflects measurement error rather than true misbeliefs. To address this, we examine the correlation between novices' performance in the two tasks and find a substantial positive relationship ($\rho = 0.53$), indicating that our performance measure captures a meaningful ability component rather than pure noise.

Fact 6. *Novices believe employers interpret low bids as signals of low quality.*

Lastly, we examine novices' belief about how employers respond to low wage offers. As shown in Figure 1.2 Panel B, 44% of freelancers agreed with the statement in the baseline survey: *"If a new freelancer offers a lower price, the client may think that their work is low quality."* This concern is consistent across worker types by baseline performance, suggesting that perceived employer responses may deter novices from underbidding to avoid sending a negative quality signal.

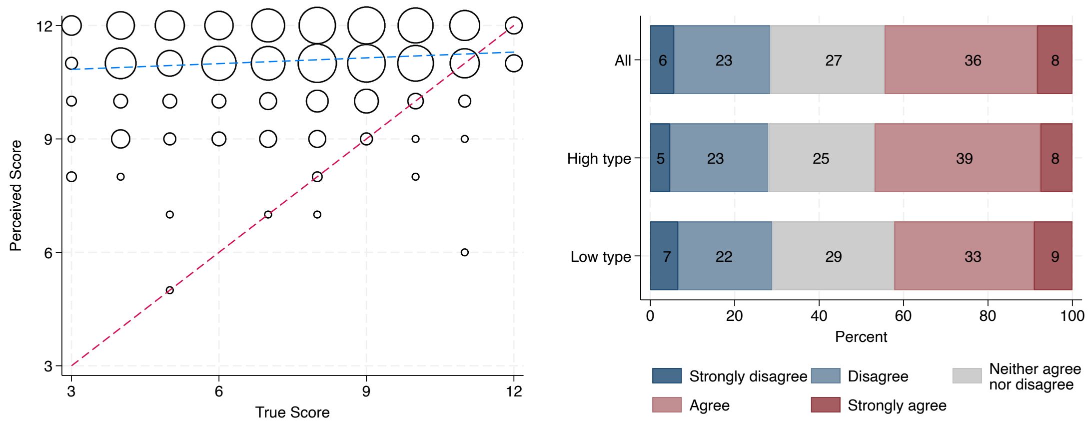
Taken together, these findings point to two mechanisms behind limited wage undercutting among novices: uncertainty about one's own ability and beliefs about how employers interpret low wages. These descriptive patterns motivate our experimental design, which causally tests these mechanisms on both the demand and supply sides.

12. Appendix Figure A.5 presents the comparison for each dimension and finds similar patterns.

Figure 1.2: Novices' Baseline Beliefs

Panel A. Actual versus Perceived Performance

Panel B. Low Bids = Low Quality to Employers



Notes: The figure reports novicesâĂŹ baseline beliefs about their own performance and the effectiveness of wage undercutting. Panel A plots true versus perceived standardized task scores; the red dashed line is the 45-degree line and the blue line shows the best linear fit. Marker size reflects the number of observations. Panel B shows responses to the statement: *âĂIJIf a new freelancer offers a lower price, the client may think their work is low quality.âĂİ* Bars correspond to the full sample, above-median (âĂIJhigh-typeâĂİ), and below-median (âĂIJlow-typeâĂİ) novices.

1.4 Demand Experiment

1.4.1 Experimental Design

Given novices' beliefs about employer responses to low wages, we first test whether employers in fact interpret low bids as signals of low quality and whether undercutting improves novices' hiring prospects. Answering this question with observational data is challenging for two reasons. First, employers' perceptions of worker quality cannot be directly observed. Second, hiring decisions depend on both wage bids and other application characteristics, making it difficult to isolate the causal effect of undercutting on employer responses.

To overcome these challenges, we leverage the following identification strategy. If employers treat low wages from novices as negative quality signals, then lowering a novice's bid reduces hiring costs but also lowers perceived quality, so the effect on demand is ambiguous. In contrast, for veterans with public reputations for quality, a lower bid primarily reduces cost and should increase demand. This yields a testable prediction: holding all other application attributes constant, lowering wages should increase employer demand less for novices than for veterans.

We test this hypothesis using an audit-style experiment on the platform, in which we create comparable novice and veteran profiles and compare employer responses to different wage bids from these profiles. The experimental estimates then allow us to quantify private returns to wage undercutting for novice workers through a back-of-the-envelope calculation.

Implementation

Freelancer Profile Creation To isolate the causal effect of wage bids on employer demand, we need to hold all profile attributes constant except the experience level. We therefore created four freelancer profiles located in Pakistan and specializing in data entry and virtual assistance.[13] Profile characteristics, including preferred hourly wages, self-reported skills, and language proficiency, were randomly assigned to reflect typical profiles in our descriptive data. Two profiles were randomly designated as novices with no job history. The remaining two were assigned veteran status, for

13. Pakistan has the worldâĂŹs third-largest online freelancer population and the highest share of data entry freelancers in our platform data.

which we built job histories between February and April 2025. Both veteran profiles were given identical numbers of completed jobs, job-specific ratings, and overall success rates to hold observable quality signals constant.

Job Listing and Application We compiled daily job listings from the platform between May and August 2025. Eligible jobs were data entry or virtual assistant positions with a fixed budget of at least $7 USD,[14] requiring entry- or intermediate-level skills. We submitted four applications per eligible job, each with a wage bid and a unique cover letter.[15] Within each experience level, one application was randomly assigned to undercut by bidding 80% of the posted job budget, while the other bid the full budget. We chose the 20% threshold because it corresponds to the most common range of wage undercutting observed in our job application data, as shown in Appendix Figure A.6. This design generated four application types per job: novice bidding 80% of the budget, novice bidding the full budget, veteran bidding 80% of the budget, and veteran bidding the full budget. Appendix Figure A.1 illustrates how these applications appeared on the employer dashboard.

We tracked two measures of employer responses using unique platform features. First, *employer view* captures when employers click on an application to review the cover letter and freelancer profile, which automatically generates a platform notification.

14. The platform sets a minimum price of $5 USD for fixed-budget projects. We applied a $7 USD threshold to ensure all submitted bids complied with this rule.

15. We construct a distinct cover letter template for each profile, which is used across all applications submitted by that profile. To maintain authenticity and reduce the risk of detection, we tailor each template to the specific job posting. In the main analysis, we control for profile and job fixed effects to isolate the effects of cover letters from the treatment effects.

Second, when employers contact applicants to request information or initiate an offer, the platform automatically sends an email notification to the applicant, which we record as *callback*. We declined all callback messages to avoid interfering with employers' hiring decisions and to maintain comparable job histories across experimental profiles.

Data

We submitted 2,972 applications to 743 job vacancies on the freelancing platform. In some cases, an application deviated from its pre-assigned wage or could not be submitted because the vacancy had already been closed or removed. We restricted the analysis sample to vacancies where all four applications were submitted with correct wages, yielding 2,812 applications to 703 vacancies.

Job and Employer Characteristics For each vacancy, we recorded basic job information, such as job title and budget, at the application time. Two weeks later, we tracked hiring outcomes and employer characteristics from the original job post, including the number of freelancers hired, employer location, tenure on the platform, number of past jobs posted, and employer ratings from previously hired freelancers. Appendix Table A.1 summarize these characteristics.

Application Characteristics and Outcomes For each application, we recorded the experience level (novice or veteran), wage bids, cover letter, submission time, employer view, and callback. Appendix Table A.2 summarizes the categories of callback messages.

1.4.2 Empirical Strategy

We estimate the causal effect of low wage bids on the likelihood of receiving employer responses by worker experience. The main regression specification is

$$Y_{ij} = \alpha + \beta_1 \text{Novice}_i + \beta_2 \text{LowBid}_{ij} + \beta_3 \text{Novice}_i \times \text{LowBid}_{ij} + \mu_j + \varepsilon_{ij} \quad (1.1)$$

where Y_{ij} is the response to the application from profile i to job j, $Novice_i$ is an indicator equal to one if profile i has zero completed jobs, $LowBid_i$ equals to one if the wage bid is 20% below the job budget, and μ_j is the job fixed effects to improve precision of our estimates. Standard errors are clustered at the job level. By comparing employer responses to wage bids across novice and veteran profiles, we map testable predictions to the signs of coefficients.

Prediction 1a. *If employers interpret lower wages as negative quality signals in the absence of other credible information, labor demand will be more responsive to low wage bids from veteran workers than from novices:* $\beta_2 > 0$*,* $\beta_3 < 0$*.*

Prediction 1b. *If low wages are not perceived as negative signals, undercutting will increase hiring chances for both novices and veterans, equally or more for novices:* $\beta_2 > 0$*,* $\beta_3 \geq 0$*.*

If Prediction 1a holds, low wage bids will have competing effects on demand through reduced hiring costs and lower perceived quality. When quality concerns dominate, low bids harm novices' chance of getting hired: $\beta_2 + \beta_3 < 0$. When hiring costs dominate, novices can improve their hiring chances by bidding low wages: $\beta_2 + \beta_3 >$

$0, \beta_3 < 0$.

1.4.3 Impact of Lowering Bids on Employer View and Callback

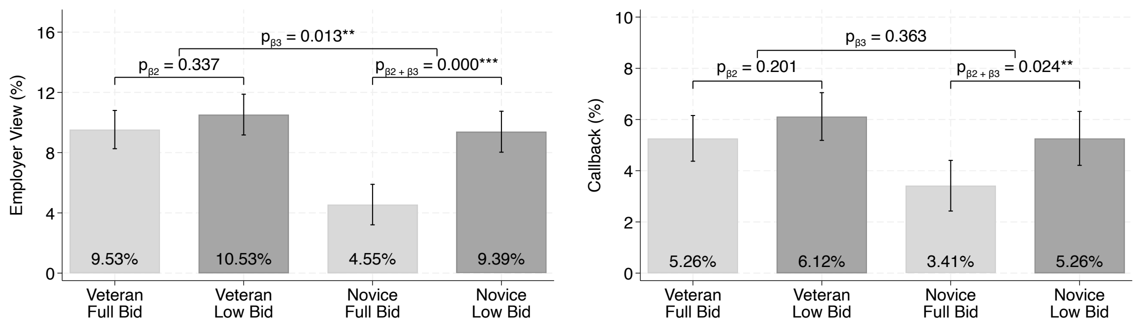
We find that low wage bids increase employer views and callbacks for both novices and veterans, with larger effects for novices. Figure 1.3 presents estimates from specification 1.1. Results are robust to alternative specifications and sample restrictions as shown in Appendix Table A.3.

Panel A shows effects on employer views. At the posted budget, veterans are more than twice as likely as novices to have their applications viewed. Undercutting to 80% of the budget more than doubles views for novices (from 4.55 to 9.39 percentage points), while increasing views only modestly for veterans (from 9.53 to 10.53 percentage points). Statistical tests on β_3 and $\beta_2 + \beta_3$ reject the hypothesis that employers interpret low bids from novices as negative quality signals ($p < 0.001$ and $p = 0.013$). Panel B reports similar patterns for callbacks. Low bids raise callback rates by 1.85 percentage points for novices and 0.86 percentage points for veterans, corresponding to increases of 54% and 16% relative to baseline. While the interaction is not statistically significant, undercutting significantly improves novices' callback probability as $\beta_2 + \beta_3$ is significantly different from zero ($p = 0.024$).

Appendix Figure A.7 presents the estimates as "demand curves", showing that employer views and callbacks decline with wage bids for novices and are more elastic than for veterans. These findings are consistent with Prediction 1b: employers do not treat low wage bids from novices as negative quality signals. Instead, lowering

Figure 1.3: Employer Responses to Low Bids

Panel A. Employer Views **Panel B.** Callback



Notes: The figures show results from regression 1.1 for employer views and callback. Each bar represents the sum of the control mean (veteran bidding full budget) and the relevant regression coefficients. We report the 95% confidence intervals based on the estimated standard errors of the linear combinations of the regression coefficients and p-values from Wald tests.

wages improves novices' hiring prospects, contrary to their baseline beliefs.

While employers are, on average, more likely to consider novice workers when they bid low wages, we observe meaningful heterogeneity across employer characteristics, as shown in Appendix Table A.4.

Two patterns stand out. First, less experienced employers exhibit more elastic demand for novices than more experienced ones. Measured by platform tenure and number of jobs posted, less experienced employers are substantially less likely to respond to novices who bid at the job budget than to veterans, but are much more responsive to those offering low wage bids. In contrast, experienced employers show a smaller differential response to undercutting. One explanation is that experienced employers may already have established relationships with freelancers, making them less sensitive to low bids from unknown applicants. By contrast, less experienced employers, who lack a stable pool of past collaborators, may rely more heavily on wage bids as a screening device when evaluating novice applicants. Second, higher-rated employers are significantly more likely to respond positively to low bids from novices than lower-rated employers. This pattern alleviates concerns about adverse selection, whereby undercutting might disproportionately attract exploitative or low-quality employers.

Taken together, these results suggest that lowering wage bids improves novices' hiring prospects, particularly with newer employers on the platform, without steering them toward lower-quality matches.

1.4.4 Returns to Lowering Wage Bids

Although lowering wage bids is an effective entry strategy, novices may still avoid undercutting if the future benefits from reputation building do not outweigh the immediate loss in earnings. To assess this concern, we conduct a back-of-the-envelope calculation comparing novices' job arrival and net earnings under two strategies: always bidding at the full job budget and consistently undercutting the budget by 20 percent. This exercise allows us to quantify the private returns to offering low wages.

We begin by modeling job arrival. Suppose a novice submits a applications, pays an application cost c_a per job, and receives an offer with probability p, which we approximate using callback rates from the demand experiment. Because jobs in this setting are usually short in duration, we assume novices are not capacity constrained and therefore always accept job offers. The expected number of jobs is then $n(a) = a \times p$.

Figure 1.4 Panel A plots this relationship between applications and accumulated jobs under the two bidding strategies. Undercutting increases p, thereby accelerating job arrival. For example, to secure the first four jobs, novices who undercut by 20 percent need to submit 76 applications whereas those bidding at the full budget require 118 applications. Given that novices in our baseline sample submit approximately 10 applications per month, undercutting shortens the time to four jobs by about four months.

Next, we compare novices' net earnings on the platform. Let the net return from the first job equal the wage bid minus the effort cost: $\pi_1 - c_e$. As workers accumulate experience, earnings grow at rate $\rho(n)$, reflecting the returns to reputation documented in the job history data. We calibrate these parameters using estimates from the demand experiment, administrative job histories, and the baseline survey; Appendix Table A.5 reports all parameter values and data sources. NovicesâĂŹ total net earnings as a function of applications can be written as:

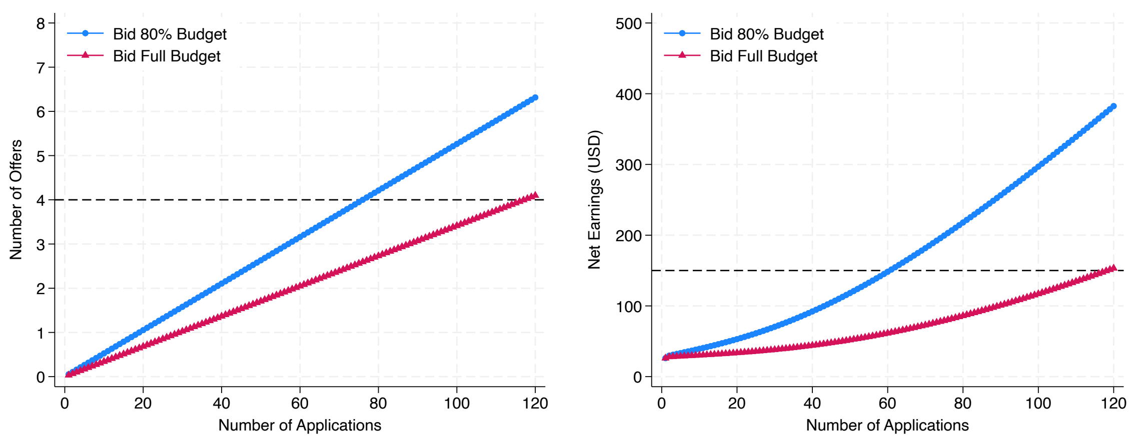
$$\Pi(a) = a \times (p \times (\pi_1 - c_e) \times \rho(n) - c_a)$$

Figure 1.4 Panel B plots net earnings under the two strategies. Undercutting dominates bidding at the full budget: novices earn more because they obtain jobs faster, experience earlier earnings growth, and incur lower application costs.

The magnitudes are economically meaningful. To earn the first $150 USD, novices need 61 applications when consistently undercutting, compared to 119 applications when bidding at the full budgetâĂŤa difference equivalent to six months of search. Over one year (120 applications), workers who never undercut earn about $153 USD, whereas those who always undercut earn $382 USD, nearly 2.5 times as much. The $229 USD gap corresponds to roughly 20 percent of annual income in low-income countries.

Figure 1.4: Expected Number of Jobs and Net Earnings

Panel A. Expected Number of Jobs **Panel B.** Expected Net Earnings



Notes: This figure presents back-of-the-envelope calculations comparing two bidding strategies for novice freelancers: bidding at 80 percent of the job budget and bidding at the full budget. Panel A shows the expected number of jobs as a function of applications. Panel B shows cumulative net earnings (USD), incorporating wage bids, effort costs, earnings growth from reputation, and application costs. Parameter values and data sources are reported in Appendix Table A.5. Dashed lines mark four jobs (Panel A) and $150 (Panel B) for reference..

These estimates are illustrative and may not fully capture actual returns for several reasons. First, we only measure effects of undercutting by 20 percent, while actual undercutting amounts vary widely (Appendix Figure A.6). Second, we assume one-to-one conversion from callbacks to job offers, which likely provides an upper bound on platform earnings since we do not directly observe conversion rates. Third, we compare pure strategies rather than the mixed strategies workers might employ in practice. Workers may optimize by undercutting selectively based on job characteristics or market conditions.

Finally, our calculation captures private returns and abstracts from general equilibrium effects. A natural concern is that widespread undercutting practices by novices could trigger a race to the bottom in wages. However, this need not be the case. First, Cui (2025) documents market expansion in a similar online setting, where low initial bids by novices substantially increased total matches on the platform. Second, we find evidence of slack demand: only 54 percent of jobs in our demand experiment were filled, as shown in Appendix Table A.1. This suggests scope for price competition to move the extensive margin by increasing the likelihood of successful job matches, rather than merely reallocating existing matches across workers.

Despite these limitations, the results demonstrate that wage undercutting offers substantial economic gains for workers in developing countries seeking to overcome entry barriers in online platforms. This raises a key question: given these returns, why don't novice workers offer lower wages to break into the market?

1.5 Freelancer Experiment: Design

The freelancer experiment examines why few novice workers offer low wages to overcome entry barriers. The baseline survey identifies two potential frictions: novices' limited information about their own abilities and inaccurate belief that employers interpret low wages as low quality signals. We test how correcting these frictions affects novices' bidding decisions.

1.5.1 *Conceptual Framework*

Our experiment is motivated by a standard reputation model (Tirole, 1988) with supply-side belief frictions. We outline the main intuition here and present the full model in A.2. In this labor market, employers live for one period, while workers live for two: they enter as novices and, if hired, become veterans with publicly observable reviews. Employers post jobs with a stated budget, and workers apply by submitting wage bids at or below this budget.[16] Novices differ in innate ability (high or low), which is known to themselves but not to employers ex ante. Employers can infer veterans' ability from reviews. For novices, employers form beliefs based on observed wage bids and may incur a cost to screen applicants. Workers choose wage bids to maximize expected lifetime earnings, and employers hire the applicant who generates the highest expected return.

In the baseline model, high-ability novices offer lower wages initially to compensate employers for hiring them and build reputation for future jobs, while low-ability novices maximize current earnings since they won't be hired again once their type is revealed. This strategy can be sustained when the future returns to reputation exceed the short run cost of lower wages for high-ability novices. This condition is likely to hold in our setting, given the high returns to reputation and relatively

16. This is empirically motivated: only 3% of applicants in our job application data bid above the budget.

low outside options. In equilibrium, low wage bids serve as credible signals of high ability, allowing employers to infer quality without incurring screening costs. This yields a separating equilibrium in which only high-ability novices undercut, are more likely to be hired, and deliver higher-quality output.[17]

Two supply-side belief frictions can cause the market to deviate from the separating equilibrium. First, if novices are uncertain about their own ability, wage bids reflect expected rather than actual ability, weakening the signaling value of low bids. Employers must then rely more on costly screening, making novices more expensive to hire. As a result, fewer high-ability novices are hired, and match quality worsens. Second, if novices believe that employers interpret low wages as signals of low quality and place greater weight on these concerns in hiring, they pool at the posted budget to avoid being perceived as low quality. This eliminates the informational content of wage bids. Employers must then incur additional screening costs to distinguish among novices and, with identical bids, cannot update their beliefs about the relationship between wages and quality. Both frictions impede high-ability novices from signaling their quality, slowing talent discovery and limiting their ability to overcome entry barriers.

17. The baseline model rests on assumptions consistent with recent empirical evidence. First, the assumption that employers live for only one period rules out incentives on the demand side to invest in training. This is realistic in our setting: the average hourly contract length among novices in our data is 6.4 hours, and most freelancers work with unique employers (Pallais, 2014). Second, we assume that low-ability workers do not improve their productivity through accumulating platform experience because once their abilities are revealed, they will not acquire future jobs on the platform. Fazio et al. (2025) find suggestive evidence that low ratings on initial jobs reduce workers' chances of securing further work. Thus, online platforms provide limited scope for skill accumulation after a bad initial signal. As a result, only high-type novices have strong incentives to offer low wages in their first period in order to establish a reputation for future employment.

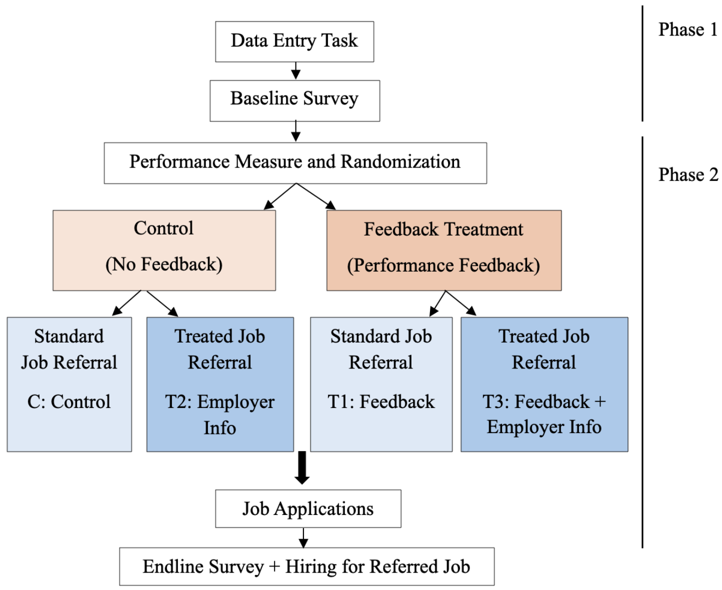
Our experimental design shuts down each of the frictions by exogeneously providing novices with information about their own ability and/or about employer responses to low wage bids. The framework yields three predictions. First, in the absence of misperceptions about employer responses, providing novices with information about their own ability induces separation in wage bids, driven by high-ability novices. Second, correcting beliefs about employer responses increases undercutting among novices. Third, when both frictions are alleviated, the increase in undercutting should again be concentrated among high-ability novices. Next, we describe the experimental design.

1.5.2 Timeline and Randomization

Our experimental design consists of two phases. In the first phase, we hired novice freelancers for a standardized data entry task to measure their baseline performance and beliefs. In the second phase, we randomly provided these workers with information about their performance and employer responses, and tracked their subsequent job applications as main outcomes. Figure 1.5 presents the experimental design.

Section 1.3.3 describes the first phase in detail, including the sample recruitment, the standardized task and performance evaluation, and the baseline survey. Following the performance evaluation, we cross-randomized two treatments at the individual level: the *feedback treatment* and the *employer information treatment*. The randomization was stratified by whether workers performed above the median (high-type) or below the median (low-type).

Figure 1.5: Freelancer Experiment Timeline



Notes: This figure shows the freelancer experimental design. This design was repeated for 8 rounds and each round took place over 2 weeks.

Feedback Treatment To address workers' limited knowledge about their performance, we randomly assigned half of the sample to receive private feedback about their baseline performance. Workers were told that we evaluated their work based on accuracy, speed, and following instructions relative to other novice freelancers hired for this job. For each dimension, they received a score from 1 to 4 stars, representing their quartile rank. Below is an example of the feedback.

> *Hi! We reviewed your data entry work and compared it with other new freelancers we hired. We looked at three things: accuracy of the entry, the time taken to complete the task, and how well you followed instructions. Your results:*

Accuracy: ★★ - below average, but not in the bottom 25%

Speed: ★ - bottom 25%

Following instructions: ★★★ - above average, but not in the top 25%

Please note that this will not affect our review for you. We ONLY share this feedback with you and hope this helps you grow and succeed on this platform.

A few days after the feedback treatment, we posted a data entry job from another employer accounts on the platform. Appendix Figure A.8 shows that the job budgets were comparable to typical data entry jobs as measured in the demand experiment and well above the $5 minimum for fixed-payment projects. We then shared the job link with all workers in the experimental sample through our original account.[18] The standard referral message was as follows:

Hi! A former colleague of ours is hiring for a data entry job and asked us to help spread the word widely. They mentioned they are open to working with new freelancers. If you're interested, feel free to apply directly! Since we are not involved in the hiring process, we won't be able to make individual recommendations.

Employer Information Treatment To address workers' belief that employers treat low wage bids as negative signals, we cross-randomized half the sample to receive additional information about the employer. The treated referral message was

18. Within each round, we posted one job from two different employer accounts to avoid overcrowding one application pool. Workers were randomly assigned to one job opportunity.

as follows:

> *Hi! A former colleague of us is hiring for a data entry job and asked us to help spread the word widely. They mentioned they are open to working with new freelancers and* ***won't judge quality of new freelancers based on their proposed prices.*** *If you're interested, feel free to apply directly! Since we are not involved in the hiring process, we won't be able to make individual recommendations.*

This referral design helps establish the credibility of the employer information for novices while preserving their incentives to attract employers unfamiliar with their abilities. One concern is that the treated message might eliminate workers' incentives to signal *high* quality by bidding low wages. However, baseline survey data suggest this is unlikely: fewer than 4 percent of workers believed that employers would interpret low wage bids from novices as high-quality signals. The neutral tone of the message also minimizes experimenter demand effects or suspicion about the job, which we later confirm using responses from the endline survey.

After the treatment, we tracked applications received within two days, as most freelancers responded within 24 hours. Application data, including wage bids and cover letters, were scraped from the employer dashboard and include both in-sample and out-of-sample applicants. Our main outcomes are whether novices in the experimental sample submit an application and the wage bids they propose.

Endline Survey We administered the endline survey to in-sample novices after the two-day application window. The survey measures workers' perceptions of the referred job and re-elicits their beliefs about their performance relative to other novices. These data allow us to assess treatment validity, that is, whether the intervention shifted novices' beliefs as intended. For those who did not apply, we also asked about their reasons for not submitting an application. After completing the survey, we closed the initial contract and left a public but uninformative review for all in-sample novices.

Follow-up Task We randomly offered the follow-up job to a small subset of applicants. Section 1.3.3 describes the task content and hiring procedure in detail. We then tracked offer acceptance and worker performance to examine treatment effects on these downstream outcomes.

1.5.3 Experiment Validity

We now report several checks that support the validity of the experimental implementation. First, we verify that baseline freelancer characteristics are balanced across treatment and control groups. These characteristics were collected through the baseline survey, including demographics, work status, platform experience, and performance measures. We supplement these with public profile characteristics scraped from the platform, including number of jobs completed, preferred hourly wages, location, English proficiency, professional certificates, and other background details. Appendix Table A.6 reports summary statistics and compares freelancers along 20 baseline and profile characteristics. Of the 80 tests, two are statistically significant at the 5% level and three are statistically significant at 10% level, as expected under random chance.

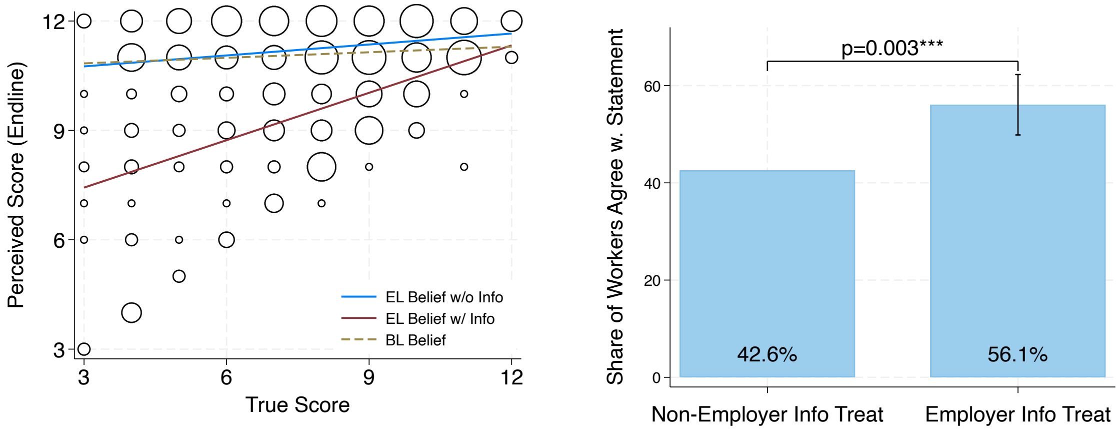
Second, attrition in the endline survey is minimal. Only 1.2 percent of the experimental sample did not complete the endline survey because their accounts were blocked by the platform and we were not able to reach them.

Finally, we provide evidence of treatment validity by showing that workers correctly interpreted the information provided. Figure 1.6 shows the treatments worked as intended. Panel A compares novices' perceived performance scores at the endline with their actual scores by treatment status. Workers who did not receive feedback show little change in their assessments: their endline beliefs closely align with the baseline ones, which indicates little updating. In contrast, workers who received private performance feedback revise their beliefs substantially, with perceived scores moving closer to the true performance level represented by the 45-degree line. In the treated groups, this updating occurs across the performance distribution rather than being driven by high- or low-type novices alone. Among novices with lower baseline performance, belief revisions are larger but the updated beliefs are still away from the truth. Among higher-performing novices, belief adjustments are smaller but their updated beliefs align more closely with their true performance. We also show in Appendix Figure A.9 that similar patterns emerge when examining the three dimensions of performance separately. Overall, these results confirm that the feedback treatment successfully shifts workers' beliefs about their performance relative to other novices.

Figure 1.6: Treatment Effects on Worker Beliefs

Panel A. Beliefs about Own Performance

Panel B. Employers Don't Infer Quality from Bids



Notes: This figure reports the treatment effects on worker endline beliefs about own performance and employer responses. Panel A plots workers' actual performance scores against their perceived scores measured at the endline. The blue line represents the best linear fit for workers without performance feedback and the red line for workers with performance feedback. The dotted line shows the best linear fit of baseline perceived performance against true performance. Panel B reports the share of workers who "agree" or "strongly agree" with the statement: *"The employer (of the referred job) will not judge my quality based on my wage bids."* We report the 95% confidence intervals and p-values from the Wald test in Panel B.

Panel B examines whether the employer information treatment alleviated novices' concerns that employers interpret low wage bids as signals of poor quality. In the endline survey, workers were asked to indicate whether they agree with several statements about the referred job. One statement reads: *"The employer (of the referred job) will not judge my quality based on my wage bids."* We find that in the treated groups, the proportion of workers agreeing with the statement rose by 13.5 percentage points, representing a 32 percent increase from the non-treated mean. This difference is statistically significant ($p = 0.003$), indicating that the employer information treatment meaningfully shifted workers' beliefs about how employers interpret wage bids. Appendix Figure A.10 further reports novices' baseline and endline beliefs about employer responses to low wage bids and shows that the treatment increases the share of workers holding the correct belief, while baseline beliefs remain similar across treatment groups. This provides additional evidence that the feedback treatment shifted beliefs as intended rather than reflecting pre-existing differences between treated and non-treated workers.

Importantly, the treatment does not appear to affect workers' perceptions of the credibility or competitiveness of the job itself. Appendix Figure A.11 shows that 83 percent of novices view the referred job posting as trustworthy and 85 percent believe competition for the job is strong, with no significant differences across treatment groups.

1.6 Freelancer Experiment: Results

In this section, we outline the empirical strategy and present the experimental results on job applications, wage bids, offer acceptance, and performance to test the predictions from the conceptual framework. We then examine the role of baseline beliefs to shed light on the underlying mechanisms and evaluate alternative hypotheses.

1.6.1 Empirical Strategy

We estimate the following specification to measure the treatment effects on novice freelancersâĂŹ likelihood of wage undercutting:

$$\text{Undercut}_i = \alpha + \beta_1 \text{Feedback}_i + \beta_2 \text{EmployerInfo}_i + \beta_3 (\text{Feedback+EmployerInfo})_i \tag{1.2}$$
$$+ \eta_r + \varepsilon_i$$

where Undercut$i$ is an indicator equal to one if novice i submits a wage bid below the posted budget. Feedback$_i$, EmployerInfo$_i$, and (Feedback+EmployerInfo)$_i$ are treatment indicators. η_r denotes experiment-round fixed effects, and ε_i is an error term; we report Huber-White robust standard errors.The signs of the coefficients β's capture the average treatment effects on novice i's decision to undercut and correspond to the following testable predictions:

Prediction 2a. *If novice workers hold inaccurate beliefs about employer responses to low wage bids and have limited information about their own ability, all three treatments will affect their likelihood of wage undercutting: $\beta_m \neq 0$ for all $m \in 1, 2, 3$.*

Prediction 2b. *If novice workers believe that employers interpret lower wage bids as negative signals of quality, alleviating this concern (EmployerInfo) will increase their likelihood of wage undercutting: $\beta_2 > 0$.*

The signs of β_1 and β_3 depend on novices' prior beliefs about their own ability. In the conceptual framework, novices whose beliefs adjust downward are less likely to undercut as a form of reputation investment, whereas those whose beliefs adjust upward are more likely to do so. Therefore, the signs are ambiguous *ex ante*.

To capture differential treatment effects by novice ability, we include interactions of treatment dummies with novices' baseline performance to the specification:

$$\begin{aligned}\text{Undercut}_i =& \alpha + \theta_1 \text{Feedback}_i + \theta_2 \text{EmployerInfo}_i + \theta_3 (\text{Feedback+EmployerInfo})_i \\ & + \delta \text{High}_i + \gamma_1 \text{Feedback}_i \times \text{High}_i + \gamma_2 \text{EmployerInfo}_i \times \text{High}_i \\ & + \gamma_3 (\text{Feedback+EmployerInfo})_i \times \text{High}_i + \eta_r + \varepsilon_i\end{aligned} \tag{1.3}$$

where $High_i$ indicates worker i's baseline performance was above the median. In this specification, δ captures differences in the likelihood to undercut between high- and low-type workers in the control group, θ's capture the average treatment effects for low-type workers, and γ's capture the differential treatment effects for high-type workers. We can map signs of these coefficients to the following testable predictions on wage bids by novice ability.

Prediction 3a. *If novice workers have limited information about their own ability, wage bids will not vary by worker ability:* $\delta = 0, \gamma_2 = 0$.

Prediction 3b. *If novice workers hold accurate beliefs about employer responses to low wage bids, providing them information about their ability (Feedback) will induce separation in wage bids, with high-type workers more likely to undercut:* $\gamma_1 > 0$.

Prediction 3c. *If novice workers hold inaccurate beliefs about employer responses to low wage bids, providing them information about both their ability and employer response (Feedback + EmployerInfo) will induce high-type workers to undercut:* $\gamma_3 > 0$.

1.6.2 Treatment Effects on Worker Outcomes

Job Application and Wage Bids

We find high compliance among novice freelancers in applying for the follow-up job across all four treatment groups. Figure 1.7, Panel A shows that over 80% of freelancers applied, with no evidence of differential attrition across treatment groups. Endline survey evidence is consistent with this pattern: among non-applicants, the most commonly cited reasons for not applying were a perceived low likelihood of selection due to competition (28%) and insufficient funds to cover the application fee (25%). This suggests that the treatments do not affect participation decisions and that subsequent results are not driven by selection into the applicant pool.

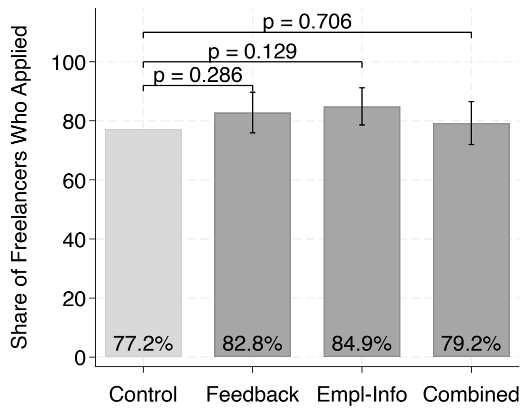
We next examine the main outcome: wage bids by novices. Figure 1.7, Panel B shows the share of freelancers bidding below the job budget across treatment groups.[19] Consistent with our descriptive evidence, undercutting is rare in the control group: only 5.5 percent of freelancers bid below the budget.[20] All three treatments significantly shifted the share of novices who undercut, consistent with Prediction 2a.

19. Only one worker in our sample proposed a wage above the job budget.

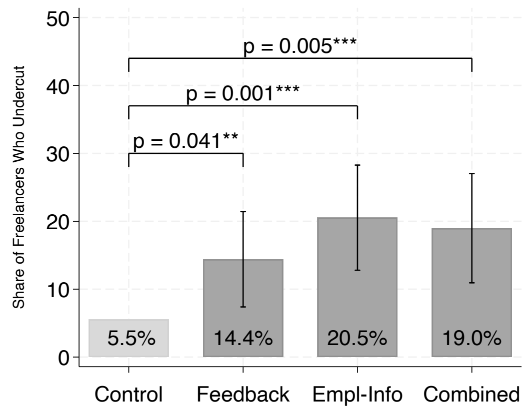
20. This rate is lower than in the descriptive data, likely because experimental participants received a job referral, which may have provided a perceived advantage over other applicants.

Figure 1.7: Treatment Effects on Job Application and Wage Bids

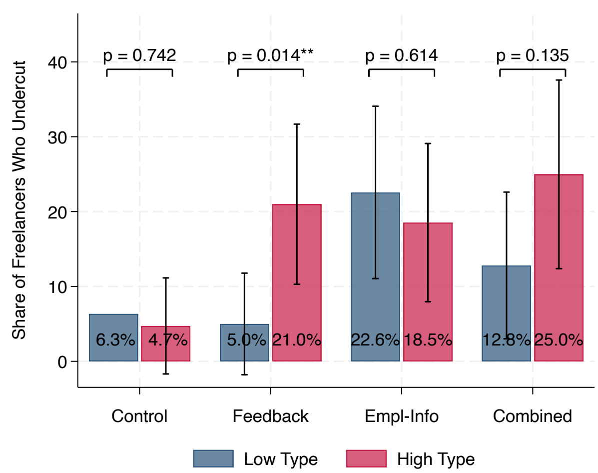
Panel A. Application Decisions



Panel B. Bids Below Job Budget



Panel C. Bids Below Job Budget by Worker Type



Notes: The figures show treatment effects on job application and wage bids. Panel A shows the treatment effects on whether workers apply to the referred job and Panel B shows effects on whether workers bid below the job budget conditional on applying based on regression 1.2. Panel C show heterogeneous treatment effects on wage undercutting by worker baseline performance (above or below median) from regression 1.3. Each bar represents the sum of the control group mean and the relevant regression coefficients. We report the 95% confidence intervals based on the estimated standard errors of the linear combination of regression coefficients and p-values from Wald tests. $p < 0.10^{*}, p < 0.05^{**}, p < 0.01^{***}$.

Relative to the control group, freelancers are 8.9 percentage points more likely to bid below the budget in the *Feedback* group ($p = 0.041$), 15 percentage points more in the *Employer-Info* group ($p = 0.001$), and 13.5 percentage points more in the *Feedback+Employer-Info* group ($p = 0.005$). These effects indicate that inaccurate beliefs about employer responses and own ability at baseline discourage novices from bidding low, and that correcting these beliefs substantially influences wage-setting behavior. The results also support Prediction 2b: providing information about employer responses increases novices' willingness to undercut. Since the effects of the *Feedback* and *Feedback+Employer-Info* treatments are ambiguous ex-ante, we explore heterogeneity in baseline beliefs to explain the observed patterns in Section 1.6.3.

The standard reputation model predicts that high- and low-ability novices differ in their propensity to adopt "introductory pricing". Do they respond differently to the treatments? Figure 1.7, Panel C shows the main results separately for novices who perform above (high-type) or below (low-type) the median in the baseline task. In the absence of performance feedback, the two types are equally likely to bid at the job budget in both the control group ($p = 0.742$) and the *Employer-Info* group ($p = 0.614$). Once performance feedback is provided, treatment effects are driven by high-type workers. High-type novices are four times more likely to bid below the budget than low-type novices in the *Feedback* group ($p = 0.014$) and twice as likely in the *Feedback+Employer-Info* group ($p = 0.135$). This pattern is consistent with the model, in which high-type novices have greater returns to securing initial jobs to build reputation and are therefore more likely to undercut initially.

Taken together, these results identify supply-side belief frictions that deter novices from using low wages as an entry strategy. In particular, uncertainty about own ability prevents the separating equilibrium in wage bids predicted by standard reputation models. Providing information about employer responses encourages both types to undercut, leading to pooling in wage bids. In contrast, providing information about own ability induces high-type novices to undercut. This separation improves match quality: high-type workers are more likely to be hired, deliver higher-quality output, and build reputations, thereby accelerating talent discovery in the market.

Other Application Features and Job Performance

We have shown that providing information about own ability and employer response increases novices' propensity to bid below the posted budget, particularly among the high-ability workers. A potential concern is that these effects are driven by changes in other aspects of the application. We examine two alternative explanations: effort in cover letter writing and the timing of wage bargaining. First, since wage bids and cover letters are the two main choices applicants make in the application, treatments may shift effort from bidding to cover letter writing. Using cover letters from all 390 in-sample applicants for the follow-up job, we proxy application effort by cover letter length. Appendix Figure A.12 shows no treatment effects on cover letter length, suggesting that changes in application effort do not explain the observed patterns in wage bids.

Second, since we observe only the wage bids submitted at the application stage, novices may strategically bid low to attract attention and renegotiate after receiving an offer. To assess this, we randomly offered the job to 34 in-sample novices and tracked offer acceptance, as described in Section 1.3.3. We find no evidence of such strategic bargaining. All but one freelancer accepted the offer at their proposed wage and completed the task; the remaining freelancer declined due to time constraints.

A final concern is that novices who undercut may exert less effort, leading to lower-quality output. We test this using worker performance on the follow-up task. Appendix Figure A.13 Panel A shows that freelancers who bid lower wages performed, on average, better than those who did not undercut, although the difference is not statistically significant. Panel B reports task performance by treatment group and similarly shows no statistically significant differences. These patterns are consistent with novices using low bids to secure initial jobs while maintaining effort to build their reputation. Given the small sample size ($n = 33$), however, these findings should be interpreted as suggestive rather than causal.

1.6.3 Mechanisms

What Drives the Heterogeneity?

The heterogeneous treatment effects by novice ability are consistent with predictions from the standard reputation model in Tirole (1988), but raise a question: if most workers believe they are high performers at baseline, why does feedback induce high types to undercut when they presumably already know their quality? To answer this, we examine heterogeneity in the accuracy and confidence of novices' beliefs about their own performance. We measure accuracy as the absolute difference between perceived and actual baseline performance, and confidence as workers' self-reported confidence in the accuracy of their self-assessment, measured in the baseline and endline surveys.[21] Appendix Table A.7 reports the results.

Panel A, columns 1 to 4 show that the treatment effects for high-type novices are concentrated among those with more accurate priors in the feedback-only treatment relative to the control group. This pattern partly reflects a mechanical relationship: since most workers perceive themselves as high performers, true high types tend to have more accurate priors. Columns 5 and 6 split the sample by whether confidence in the accuracy of own performance assessment increases from baseline to endline. The effects are concentrated among workers whose confidence increases. This suggests that feedback primarily validates high types' prior beliefs rather than shifting them.

An alternative explanation is that high-type novices respond more because they better understand the feedback. We test this using Figure 1.6, Panel A, which compares changes in workers' beliefs about their own performance from baseline to endline. While high-performing novices' beliefs move closer to the truth in the feedback group, low-performing novices exhibit larger updates from baseline. This indicates that low-type novices also process the feedback and adjust their beliefs about their own performance, which is inconsistent with differential comprehension as the primary mechanism.

21. In both baseline and endline surveys, novices were asked to evaluate their own performance on the baseline task and indicate how confident they are in their guess, on a scale of one to five.

Is There Complementarity Between Treatments?

Since the interventions target different sources of misbeliefs, a natural question is why the combined treatment exhibits limited complementarity. We use baseline data on workers' beliefs about bidding strategies to address this. Appendix Table A.8 reports the results. Columns 1 and 2 show that the treatment effects for high-ability novices in the feedback group are concentrated among those who initially hold more accurate beliefs that employers do not interpret low prices as signals of low quality. Columns 3 to 6 further show that these workers are more likely to believe that undercutting is effective for high types.

These patterns yield two insights. First, high-ability workers understand which strategies are effective but remain uncertain about their own ability. The feedback treatment therefore reinforces their self-assessment and induces them to adopt the undercutting strategy they already believe is effective. Second, because these workers hold correct priors about employer responses, the second treatment adds little, leading to limited complementarity.

In summary, the interventions address belief frictions through different channels. Correcting workers' misperception about employer responses to low wage bids encourage those with inaccurate beliefs about the demand side to offer low wages to secure jobs. Resolving workers' uncertainty about own abilities strengthens high-ability workers' incentives to invest in reputation through low wage offers. This improves the allocation of talent by encouraging more efficient sorting across novice types.

1.7 Discussion

1.7.1 How Fast Can Novices Learn about Demand?

Our experimental results show that novice freelancers often avoid lowering their initial wage bids because they hold incorrect beliefs that employers interpret low prices as signals of low quality. This raises an important policy question: can novices eventually overcome these belief frictions through trial and error on the platform, or is external intervention needed to correct these misperceptions? To shed light on this question, we simulate an ideal learning environment in which a fully rational novice experiments systematically with wage bids and updates her beliefs about demand according to BayesâĂŹ rule. This exercise provides a conservative benchmark for how long and how costly the learning process can be in the absence of external intervention.[22]

We model a novice freelancer conducting an A/B test by alternating between two bidding strategies when submitting job applications: bidding at the full budget or undercutting by 20 percent. Each application yields a binary outcome indicating whether the novice receives a callback. The novice starts with a 50/50 prior over two possible states of the world:

Model 1 ("Low wage bids $\neq$ low quality"): Employers do not penalize low bids. Undercutting by 20 percent increases callback rates from 3.41% to 5.26% for novices.

22. This simulation focuses on learning about employer responses to wage bids. Novices face a similar learning challenge in assessing their own abilities, as they need to secure jobs and receive employer feedback to update beliefs about their performance. Therefore, learning about one's ability would impose similar (if not more) time and financial costs as learning about demand.

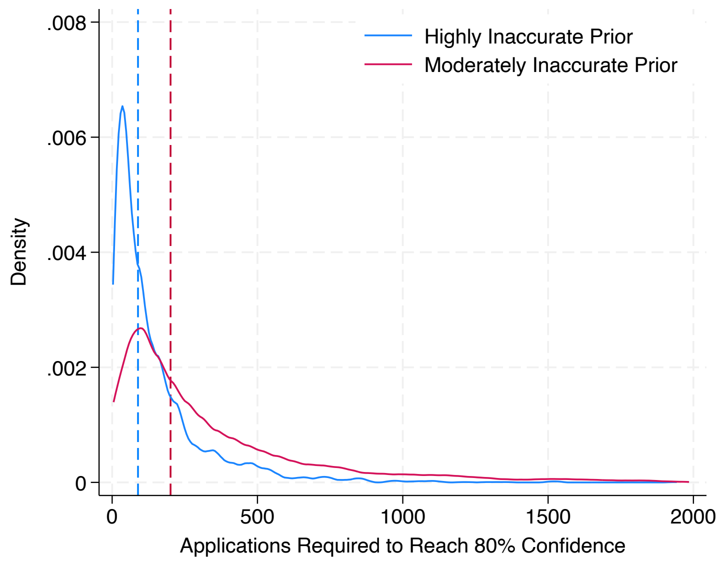
Model 2 ("Low wage bids = low quality"): Employers respond less to novices who submit lower bids. The callback rate at the full budget is 3.41%. We consider two degrees of misbelief about undercutting. Under moderate misbeliefs, novices believe the callback rate declines by 20 percent, from 3.41% to 2.73%. Under strong misbeliefs, they believe the callback rate declines by 50 percent, to 1.71%.

After each application, the novice updates beliefs about the two models using Bayes' rule based on whether a callback is received. Learning is complete once the posterior probability assigned to the true model reaches 80%. We simulate 3,000 novice workers, each submitting up to 2,000 applications, and record the number of applications required to reach this threshold.

The simulation reveals that even under systematic experimentation and perfect Bayesian updating, learning is slow. Figure 1.8 plots the distribution of applications required to learn under the two types of inaccurate priors. Over 99% novices reach 80% certainty of the true model within 2,000 applications. The median worker with moderately inaccurate beliefs needs 201 applications to reach 80% confidence, while with strongly inaccurate beliefs needs 89 applications. Learning is faster under stronger misbeliefs because each successful application is more informative about the true model and leads to belief update. Nevertheless, learning remains slow overall because callbacks are rare and most applications provide little information. As a result, even perfectly rational workers cannot quickly infer how employers respond to wage undercutting.

The learning process translates into substantial real-world costs. Novice freelancers

Figure 1.8: Distribution of Applications Required to Learn True Demand



Notes: The figure plots the simulated distribution of applications required for novice freelancers with inaccurate prior beliefs to learn the true demand model (model 1). Results are based on 3,000 simulated workers who alternate between high and low bids and update their beliefs using Bayes' rule. The dotted line indicates the median number of applications needed to reach 80% confidence in the true model.

in our baseline survey submit on average 10 applications per month at an average cost of $1.35 USD per application (excluding workers' effort costs). A novice with moderately inaccurate beliefs would need 20 months of experimenting at the cost of $271 USD, which is equivalent to 24% of the average annual income in low-income countries according to the World Bank.

These costs help rationalize why belief frictions may persist in practice and point to the importance of external interventions that reduce the cost of learning, rather than relying on workers to discover these patterns on their own. Even under ideal learning conditions where workers experiment systematically and update beliefs perfectly, the time and financial costs of experimentation are substantial. In reality, workers may experiment less frequently or update beliefs imperfectly, making learning even slower. As a result, many novices may never experiment enough to discover that lowering wage bids improves hiring chances, preventing them from adopting effective bidding strategies to overcome entry barriers in online labor markets. To facilitate worker learning and market entry, existing government initiatives, such as the Ajira Digital Program in Kenya and the Mastering the World of Online Freelancing projects in Jordan and Lebanon, incorporate mentorship and business strategy guidance (Datta et al., 2023). Similar components could be integrated into digital training programs to help workers form accurate expectations about job search in global markets.

1.7.2 Generalization to Other Settings

Next, we discuss how our findings translate to other labor market settings, especially in developing countries.

Online Labor Markets

Our analysis focuses on data entry services on one of the largest online freelancing platforms, so a natural question is how much the entry barriers and frictions we document extend to other service categories and platforms. As discussed in Section 1.3.2, similar entry barriers arise broadly: inexperienced workers struggle to get hired because employers are uncertain about their ability.

One potential difference is that workers in higher-skill categories, such as graphic design or programming, may have access to alternative signals of ability, such as portfolios or certifications, rather than relying on lower initial wages. However, these signals may be less informative than they appear. Using profile data from 3,461 graphic designers on the same platform, we find that 96% list a portfolio and 40% report certifications. When such signals are nearly universal, they provide limited differentiation across workers. This pattern is consistent with Kässi and Lehdonvirta (2022), who find that micro-credentials have little effect on job outcomes for new entrants.

There is also evidence that introductory pricing is used in higher-skill categories. Cui (2025) shows that freelancers specializing in PHP adopt low initial prices to build reputation on Freelancer.com. At the same time, these investments remain below the social optimum, leading to inefficient hiring of novice workers. This suggests that the underuse of introductory pricing is not unique to low-skill services.

A remaining question is whether the belief frictions we document are present in these settings. There is limited direct evidence, as our study is the first to elicit freelancers' beliefs about their own ability and employer responses. That said, similar frictions have been documented in offline labor markets, and there are reasons to expect them to persist online. In particular, novices from developing countries may have limited information about their standing in the global talent pool or about preferences of foreign employers. Platform design may further slow learning: on many platforms, workers cannot observe competing bids, making it difficult to infer demand. In contrast, platforms that reveal bid ranges may facilitate faster learning and stronger price competition.

Offline Labor Markets in Developing Countries

While this paper studies online marketplaces, our results are relevant in offline labor markets. These markets share three features. First, both settings resemble spot markets, characterized by high unemployment, high search frictions, and short job spells (Breza and Kaur, 2025). Wages in these markets exhibit limited dispersion relative to underlying productivity (Breza et al., 2018; Breza et al., 2019; Cefala et al., 2024; K, 2025). Second, inexperienced workers face substantial barriers to entry. Because educational credentials often provide weak signals of ability, novices struggle to credibly convey their productivity to employers (Bertrand et al., 2024). As a result, employers rely heavily on observable work experience as a proxy for quality.[23] Third, workers often have limited information about both their own abilities and employer preferences (Carranza et al., 2022; Kiss et al., 2023; Caria et al., 2024). Together, these features suggest that supply-side belief frictions may contribute to compressed wages and inefficient allocation, particularly for high-ability novices.

At the same time, three features of offline labor markets may limit the direct applicability of our results. First, the scope for reputation building may be more limited: job spells are typically short, and past performance is often difficult to verify across employers (Donovan et al., 2023). This weakens workers' incentives to accept lower

23. For example, evidence from Burundi shows that employers prefer candidates with prior experience in low-skill jobs over otherwise identical inexperienced applicants (Ndayikeza, 2025).

initial wages as an investment to reveal their true ability. However, when workers remain with the same employer, returns to experience can be substantial (Breza and Kaur, 2025) because worker performance information becomes available over time. This mirrors the reputation-building process on online platforms, where performance history is available to all potential employers. In both settings, high-ability workers have incentives to invest in reputation, improving match quality. Second, subsistence constraints may restrict workers' ability to undercut wages. Nevertheless, recent evidence shows that wage undercutting does occur even among low-wage manual laborers in rural settings (Breza et al., 2019; K, 2025), indicating that workers engage in strategic wage-setting despite financial constraints. Third, wage undercutting may be discouraged by social norms in rural labor markets, even in the absence of belief frictions (Breza et al., 2019). As a result, the mechanisms we document are likely to be most relevant in urban labor markets and sectors with higher returns to experience, such as services and manufacturing, where both reputation and strategic pricing play a larger role.

1.8 Conclusion

This paper studies why novice workers from LMICs under-adopt introductory pricing to overcome entry barriers in online labor markets. In the demand-side experiment, we show that employers respond favorably to low wage bids from novices and do not interpret them as signals of low ability. Back-of-the-envelope calculations further indicate that undercutting is a profitable entry strategy.

Despite these returns, few novices adopt it. Survey evidence points to two frictions: incorrect beliefs about how employers interpret low wages and limited information about their own ability. In a supply-side experiment, we show that providing workers with information about employer responses and their relative performance increases novices' likelihood of proposing low initial wages. The effect is driven by high-ability novices after receiving performance feedback, consistent with standard reputation models. This separation in wage bids improves talent allocation by increasing high-ability novices' chances of being hired and producing high-quality output.

We then examine whether novices can correct their misbeliefs through on-platform search and application. Simulations show that even under Bayesian updating, learning is slow and costly, as most applications receive no response and thus offer little scope for learning. This highlights the value of external information provision, such as mentorship or guidance on bidding strategies, in helping novices overcome entry barriers. While we do not observe how employers use wage bids or other profile signals to infer worker quality, better information on employer decision-making would help high-ability novices optimize their application strategies and improve entry into the market.

Nevertheless, the implications of these findings can be extended beyond online labor markets to broader questions of information provision and market efficiency. Our results suggest that policymakers can improve talent discovery and worker outcomes by providing skill assessments for new entrants and sharing information about employer decision-making processes. More broadly, the findings indicate that policies addressing information frictions on the supply side, rather than the demand side, may be more effective for improving job matching in markets characterized by high entry barriers and low-information environments.

CHAPTER 2

INTRA-HOUSEHOLD INCENTIVE DESIGN: AN EXPERIMENT ON PARENT-CHILD DECISION MAKING DYNAMICS IN PAKISTAN

2.1 Introduction

Conditional cash transfers (CCTs) are a popular policy instrument for promoting human capital investment, particularly in low-income countries. The vast majority of CCTs target only parents or children, although these investment decisions often involve both parties. When preferences are not aligned, the identity of the transfer recipient could lead to different outcomes due to household bargaining frictions. For example, the recipient may withhold information from the non-recipient or fail to bargain on the transfer. Can policymakers enhance the effectiveness of CCTs by designing incentives to address household frictions between parents and adolescent children?

We study this question in the context of a digital skills training program for adolescent girls in Lahore, Pakistan. The government of Pakistan heavily subsidizes supplemental digital skills training programs for the youth, as 91 percent of young adults in Pakistan lack basic digital skills, such as copying and pasting text (World Skills Clock 2023) and only 42 female youth possess digital skills for every 100 males in Punjab (UNICEF 2023). While many programs offer sizable cash rewards to trainees upon completion, training providers still report low female participation and completion rates.[1]

We partner with the Punjab Skills Development Fund (PSDF) to offer a short-term digital skills training program with cash incentives upon completion to female students aged 16 and over. PSDF is Pakistan's largest skills development fund and an organization explicitly tasked to boost digital skills for young women by the national government.[2] Our experimental sample consists of 489 households across 8 girls-only, government high schools and colleges in Lahore. As girls in our sample had low computer literacy at baseline, the program offered a selection of beginner-level online trainings on Coursera, ranging from bookkeeping to digital marketing. Upon program completion, girls received a globally recognized Coursera certificate, and their household received 3,000 PKR, equivalent to 7.5 percent of the monthly household income in our sample.[3]

To test information and bargaining frictions within the household, we fixed the total household reward at 3,000 PKR and cross varied (1) the split of incentive paid to parents (0 PKR, 1500 PKR, or 3000 PKR) and (2) whether parents learned about the total incentive size (partial or full information). This 3-by-2 design results in six experiment groups, where the control group is the status-quo case when parents

1. For example, PSDF offers an incentive of 5,000-6,000 PKR upon program completion for the standard Coursera trainings. Administrative data from PSDF shows that females account for 20.2 percent among those signed up for the program and only 8.9 percent of them complete the course.

2. In April 2023, the Pakistan Senate passed a resolution to equip women with digital skills and to fund digital empowerment programs for women through the Punjab Skills Development Fund (PSDF) (Pakistan Today 2023).

3. At the time of the study, 3,000 PKR was approximately equivalent to 10.5 USD.

receive neither the incentive nor information about the incentive. When girls signed up for the program at school, they were randomly assigned to one of the six arms and received full information on the total household incentive and the incentive split. They were also made aware of what information their parents would receive from the training provider. We then contacted parents to request consent for their daughters' participation in this external training program and informed them about the incentive offer. After the training provider received the parental permission, students enrolled and attended the courses through computer labs on campus.

We track differences by incentive contract across two main outcomes: parental permission for participation and program completion. For girls who received parental consent, we attribute any non-completion to their own decisions of dropping out. As girls were asked to complete the program in their school's computer laboratory during a free-period, parents could not monitor daughters' action or withdraw their consent without notifying the research team. We also find no evidence of parents discouraging daughters' participation or forcing them to drop out at the endline.[4]

We find that when parents are not informed about the total incentive, allocating the entire incentive to parents instead of the daughter increases completion by 16.45 percentage points (pp), which represents a 103% increase from the control group's mean of 15.94 percent. This result violates the income-pooling hypothesis, allowing us to reject the unitary model of decision making between parents and daughters. Once we impose information symmetry on the total incentive, differences in completion be-

4. See Section 2.3 for more discussion on separability of actions.

tween payment allocation groups shrink substantially, ranging between 0.78pp and 3.09pp.[5] Taken together, our experiment highlights two findings. First, parents and adolescent daughters make skill investment decisions jointly, even when trainings are conducted entirely at school, but there is incomplete information sharing. Second, once policymakers impose information symmetry, we find no evidence of inefficient bargaining, as differences across payment allocation are both small in magnitude and not statistically significant.

This project contributes to multiple strands of literature. Our main contribution is to provide evidence of incomplete information sharing in parent-daughter decision making and assess its implications for the effectiveness of different incentive designs. Specifically, we test whether the identity of the incentive recipient matters under information symmetry about the incentive. Most studies on conditional cash transfers cannot address this question because the incentive targets either parents or children. A few papers experimentally test the relative effectiveness of incentivizing parents versus children and find no significant differences (Baird et al. 2011, Berry 2015, Levitt et al. 2016, de Walque and Valente 2023). However, their designs do not target possible information withholding between household members. In our setting, information frictions may arise when daughters want to pursue skills training but strategically disclose partial information about the returns to their action. This friction is particularly relevant in setting where women have limited agency and

5. Compared to other papers that rejects efficient household decision making, we find smaller differences in outcomes under information symmetry. See Appendix Table B.8 for the range of inefficiencies and citations.

need approval from other household members for their actions.[6] Our results suggest that payment targeting has little impact on completion rates when both parents and daughters have perfect information on the payment amount.

Our work also adds new empirical evidence of efficient collective bargaining to the literature on household decision making. Many empirical studies on spousal decision making reject the unitary household model (see Donni and Molina (2018) for a summary) and some (Ashraf 2009, Lowe and McKelway 2024) find evidence against efficient bargaining under perfect information. Yet, little is known about how parents and adolescent children make decisions. To the best of our knowledge, we are the first to experimentally evaluate both the unitary and the efficient collective household models in the parent-child context. Previous studies (Bursztyn and Coffman 2012, Bergman 2021, de Walque and Valente 2023) only observe household-level outcomes that require joint inputs from parents and children, such as school attendance and test scores. Our design allows us to decompose program completion into actions by each party, thereby testing predictions from the efficient collective model. Moreover, the majority of empirical papers (see for example Duflo (2003), Duflo and Udry (2004), De Mel et al. (2009)) use cash transfers as exogenous shocks to one household member's income and identify its impact on consumption as evidence for household bargaining. We analyze how household members bargain on program completion in order to obtain the conditional transfer.[7]

6. This is commonly observed in spousal decisions about women's consumption, fertility, and employment (Ashraf 2009, Castilla and Walker 2013, Ashraf et al. 2014, Lowe and McKelway 2024).

7. See Section 2.5 for a full discussion on differences in bargaining environments.

Finally, we contribute to the growing literature on barriers to female labor force participation in developing countries via human capital accumulation. Past studies (Dean and Jayachandran 2019; Bursztyn et al. 2020; Heath and Tan 2020; McKelway 2021; Lowe and McKelway 2024) emphasize on intra-household decision making regarding women's labor force participation. In our sample, the majority of girls only wanted to work in skilled, white-collar jobs. However, among those who did not receive parental consent, nearly half reported not comfortable with cutting and pasting text. We document how inefficiency in household bargaining can cause under-investment in young women's essential income-generating skills, which may further exacerbate their ability to work.

The rest of this paper is organized as follows. Section 2.2 generates testable hypotheses on how different incentive contracts affect program completion under the unitary and the efficient collective household models. Sections 2.3 and 2.4 describe our experiment and results, mapped to the testable hypotheses. Section 2.5 discusses the implications and caveats of our results on household decision making between parents and children, and Section 2.6 concludes.

2.2 Theoretical Framework

We build on canonical models of the unitary household by Samuelson (1956) and the collective household with efficient bargaining by Browning and Chiappori (1998) to generate predictions on the relationship between different incentive contracts and program completion.

Suppose there is a skills training program that offers an incentive s for successful completion. Daughters are targeted for the program, meaning they have full information about the incentive. Daughters cannot hide their decision to participate,[8] however, they may choose to hide the incentive payment to fund private consumption. Parents must provide consent for their daughters to participate in the program and cannot perfectly monitor their child's progress within the program.

Let s be the total incentive in currency that is paid to one household. s^p is the amount that is paid to the parent and $(s - s^p)$ is paid to the daughter. Daughters have full information on s and $(s - s^p)$, while parents may not know about s and therefore $(s - s^p)$.

We denote $\theta \in [0, 1]$ as the proportion of daughters' incentive $(s - s^p)$ that is revealed to the parents. The total incentive s is fixed, but s^p and θ may vary at the household level and represent differences in incentive targeting to parents and the extent of information asymmetry within the household.

2.2.1 Unitary household

The unitary model Samuelson (1956) posits that households behave as a single individual with one payoff maximization problem. There is effectively one utility function

8. In alternate settings, daughters may need to travel outside the house or use the shared home computer to attend the training sessions. Even though parental consent might not be explicitly required by the training provider in these cases, daughters will still need to obtain parents' approval for their participation.

u^{hh} and one cost function c^{hh}.[9] The household decision function is characterized as follows.

$$I_{\text{completes}}\{u^{hh} + s > c^{hh}\}$$

Program completion is efficient when benefits outweigh costs at the household level. Otherwise, households will choose not complete training.

Hypothesis 1: Under the unitary model, completion does not vary with s^p or θ.

2.2.2 Collective household

In the collective household, parents and children have different preference and cost parameters (u^p, c^p) and (u^d, c^d) and they make decisions by solving separate payoff maximization problems. The preference parameters u^p and u^d are assumed stable, encompassing all beliefs about the benefits of completing the training. Similarly, the cost parameters c^p and c^d are stable, incorporating all monetary, social, and effort costs of completing the training.

We use the methodology described in Browning and Chiappori (1998) to evaluate effi-

9. There are two prominent theories that link individual preference parameters to household behavior consistent with the unitary model: Samuelson (1956) and Becker (1974, 1991). Samuelson (1956) explains a collective household may still behave like a unitary one when the household welfare function already incorporates each member's preferences. In our context, this maps to a setting where bargaining is unconditionally cooperative (see Subsection 2.2.2). Becker's Rotten Kid Theorem (1974, 1991) finds the household may behave as a single unit when an altruistic 'patriarch' uses transfers to motivate the behavior of other members. The Rotten Kid Theorem hinges on a patriarch's ability to monitor children's behavior. As parents cannot perfectly monitor girls' attendance, this model cannot be applied to our setting.

cient[10] collective households which (1) makes no assumptions on the specific method of household bargaining so long as outcomes are pareto-optimal and (2) assumes that household members divide income to fund private consumption. Browning et al. (2006) clarifies the sequence of steps required to correctly analyze household behavior under the efficient collective framework. We proceed as recommended.

First, we describe the one-period problem where the incentive payment is made after the completion of a skills training program. Then, we characterize the separate payoff maximization problems for parents and children. The maximization problems will depend on parents and children's ability to divide the incentive payment. We generate bounds on this sharing rule from parents' and children's problems which will make program completion incentive compatible for both parties. Finally, we analyze how the bounds change with differences in payment targeting and information symmetry to generate testable predictions of observable behavior.

Negotiation on a Future Payment

In order for households to receive s, parents and daughters must both remain interested in the course. Parents cannot force their daughters to join or complete the training, and daughters cannot enroll in the training without permission from their

10. Efficient collective households are often also referred to as households where bargaining is cooperative.

parents.[11,12] To reach an agreement, they must negotiate a sharing rule ($\lambda \in [0, 1]$) to divide the payment. We characterize efficient household bargaining as the ability for parents and daughters to agree on the sharing rule so that benefits are greater than costs for both parties. The parameter of interest is the feasible range for λ such that both parties remain interested in completing the course. As we will derive below, the feasible range of λ depends on the exogenous targeting of incentives and information. The larger the feasible range of λ, the more likely households are to complete the program and receive incentive s under efficient bargaining. If parents and daughters do not agree on the sharing rule, households will not complete the program and receive nothing. Outcomes are pareto optimal, because neither party can be made better off by investing in the training at the expense of the other.

Decision Processes

We assume the timeline of household bargaining is as follows.

(i) Girls receive their skills training offer and full information on s and s^p. If they want to participate, they can choose proportion $\theta \in [0, 1]$ of the payment $(s - s^p)$ to reveal to parents and retain $(1 - \theta)(s - s^p)$ as the private return.

(ii) Parents are informed about their daughter's interest in the program, and of s^p

11. In the experimental setting, parents only learn about this program after daughters agree to participate. Since the program is held during school time, parents cannot force their daughters to attend the sessions. They can only nudge their daughters. As mentioned earlier, parents also cannot monitor progress very easily, as daughters are completing the trainings in school.

12. We require written parental consent for participation as PSDF previously attempted to offer digital skills trainings in schools without parental permission, and they received backlash from parents and school administrators.

by the training provider. Parents also learn about $\theta(s - s^p)$ from daughters.

(iii) Parents and daughters negotiate on the public return $\theta(s - s^p) + s^p$ known to both parties.[13] Parents can keep proportion $\lambda \in [0, 1]$ of the public return by the sharing rule.

(iv) Post-completion, parents keep their share of the public return, $\lambda(\theta(s-s^p)+s^p)$, and daughters keep their share of the public return plus their private return, $(1 - \lambda)(\theta(s - s^p) + s^p) + (1 - \theta)(s - s^p)$.

Equations 2.1 and 2.2 are the two respective decision functions of the parents and daughters. We use these functions to generate predictions on observable actions between parents, daughters, and the household by deriving bounds on the sharing rule, λ.

$$I^p_{\text{consents}}\{u^p + \lambda(\theta(s - s^p) + s^p) > c^p\} \tag{2.1}$$

$$I^d_{\text{completes}}\{u^d + (1 - \theta)(s - s^p) + (1 - \lambda)(\theta(s - s^p) + s^p) > c^d\} \tag{2.2}$$

Parents: Rearrange Equation 2.1 to isolate λ in the parents' decision function.

$$I^p_{\text{consents}} \left\{ \begin{array}{ll} \lambda > \frac{c^p - u^p}{\theta s + s^p(1-\theta)} & \text{if } \theta(s - s^p) + s^p > 0 \text{ Case 1} \\ u^p > c^p & \text{if } \theta s + s^p(1 - \theta) = 0 \text{ Case 2} \end{array} \right\}$$

13. The parents portion of the incentive is negotiated on, because daughters in this context are not working outside of the household. The majority of consumption by daughters is indirectly funded via income to parents.

We ignore Case 2, because payments do not enter the parents' decision function and parents always give consent. Case 1 generates a lower bound on payment division which will make consent to daughters' participation incentive compatible to parents.

$$\lambda > \frac{c^p - u^p}{\theta s + s^p(1-\theta)} = \lambda_{lb} \tag{2.3}$$

Hypothesis 2: Under the efficient collective model, when parents have full information on the household incentive, parental consent does not vary with s^p.

Proof: Plug $\theta = 1$ into Equation 2.3. This result does not depend on s^p.

$$\lambda_{lb}|_{\theta=1} = \frac{c_p - u^p}{s}$$

Daughters: Rearrange Equation 2.2 to isolate λ within the daughter's decision function.

$$I^d_{\text{completes}} \begin{cases} \lambda < \frac{u^d - c^d + s}{\theta(s-s^p)+s^p} & \text{if } \theta(s-s^p)+s^p > 0 \text{ Case 1} \\ u^d + s > c^d & \text{if } \theta(s-s^p)+s^p = 0 \text{ Case 2} \end{cases}$$

Again, we ignore Case 2, because the daughter's decision does not depend on information or payment targeting and she will complete the program only if she captures the entire incentive. Hence, we derive an upper bound on payment division which will make program completion incentive compatible for the daughters.

$$\lambda < \frac{u^d - c^d + s}{\theta(s - s^p) + s^p} = \lambda_{ub} \tag{2.4}$$

Hypothesis 3: Under the efficient collective model, when parents have full information, the girls' choice to complete or drop out does not vary with s^p.

Proof: Plug $\theta = 1$ into Equation 2.4. This result does not depend on s^p.

$$\lambda_{ub}|_{\theta=1} = \frac{u^d - c^d + s}{s}$$

Remark 1: Under the collective household framework, we do not observe the unconditional decision of the girls to drop out of the program. All daughter's decisions are observed, conditional on parental consent. To test Hypothesis 3 on the data, we rely on the assumption that under information symmetry, Hypothesis 2 holds and parental consent does not vary with s^p. Hence, once information symmetry is imposed, the girl's drop-out decision conditional on parental consent is the same as the drop-out decision unconditional on parental consent.

Households: We combine Equations 2.3 and 2.4 to generate predictions on household completion.

$$\lambda_{ub} - \lambda_{lb} = \frac{u^d + u^p + s - (c^d + c^p)}{\theta(s - s^p) + s^p} \tag{2.5}$$

Hypothesis 4: Under the efficient collective household, when parents have full information, completion does not vary with s^p.

Proof: Plug $\theta = 1$ into Equation 2.5. This result does not depend on s^p.

$$\lambda_{ub} - \lambda_{lb}|_{\theta=1} = \frac{u^d + u^p + s - (c^d + c^p)}{s}$$

Remark 2: Hypothesis 4 comes directly from Hypotheses 2 and 3. We report this because it allows us to confirm that household-level results are robust to misspecification of the bargaining timeline.[14]

Remark 3: Under full information, both the unitary model and the collective model predict that household completion does not vary with s^p. *Hence, if we are unable to reject efficient bargaining in the full information setting, we must rely on differences in payment targeting under information asymmetry to properly evaluate the unitary model.*[15] In particular, we test the income pooling property under the unitary model.

Remark 4: The collective model cannot predict ex-ante how completion will change with incentive allocation under information asymmetry because the partial derivative of Equation 2.5 with respect to s^p depends on unobserved preference and cost parameters $\{u^d, c^d\}$ and $\{u^p, c^p\}$.

Table 2.1 summarize all testable hypotheses generated by the two models of the household. We can reject one framework if any of its hypotheses are rejected with a reasonable statistical test.[16]

14. Our full sample of results on household bargaining is on the set of households in which daughter and parents both complete a survey. The timeline stated in Section 2.22.2.22.2.2 requires that bargaining happens after parents are contacted by the training provider. However, if bargaining happens before, it may be possible that we have selective attrition into the parents survey. See B.2 for further discussion. We do not find much evidence of selective attrition in our sample.

15. See Samuelson (1956), Becker (1974, 1991) or footnote 6 for more examples of situations in which a collective household may behave in ways that are indistinguishable from the unitary household.

16. See Appendix Table B.5 for the list of every reasonable statistical test.

Table 2.1: Testable Hypotheses on s^p and θ

A. Unitary household
H1: Household completion does not vary with s^p or θ.
B. Efficient collective household
H2: Under full information, parental consent does not vary with s^p.
H3: Under full information, the girl's decision to complete or drop-out does not vary with s^p.
H4: Under full information, household completion does not vary with s^p.

Notes: This table presents testable hypotheses regarding the impact of payment targeting and information symmetry on the main outcomes, under the unitary and efficient collective models respectively.

2.3 Experimental Design

Through PSDF, we delivered the first supplemental digital skills training program to 8 women's government schools in Lahore. Given the low digital literacy in our sample, PSDF offered 6 beginner-level, Coursera courses for students to choose from and an incentive of 3,000 PKR to households upon course completion within 6 weeks.[17]

2.3.1 Description of the Experiment

We map our experiment directly to the model in Section 2.2 by exogenously varying the split of the incentive to parents, s^p, and information on the full incentive amount to parents, θ. We select $s^p \in \{0, 1500, 3000\} = \{0\%, 50\%, 100\%\}$ and cross-randomize whether parents are informed about their share as well as the full amount,

17. The courses are: (1) Better Business Writing in English, (2) Fundamentals of Graphic Design, (3) Foundations of Digital Marketing & E-commerce, (4) Intuit Bookkeeping, (5) Work Smarter with Microsoft Excel, and (6) How to Create a Website. These courses take on average a total of 15 hours over 4 weeks to complete and are designed to help students acquire basic digital skills for more advanced courses. These courses differ from the Coursera programs offered by PSDF, which last 4-6 months and are intended to make trainees job-ready after finishing the course. Regular trainees would receive 5,000-6,000 PKR for completing these programs.

allowing $\theta \in$ {Partial, Full}. This results in 6 contract groups as shown in Figure 2.1, along with the information given to parents.

Figure 2.1: Targeting Incentive and Information to Parents

Increasing Parental Info of Incentive

	0% reward to parents	50% reward to parents	100% reward to parents
Full info	*Program Info* + *0 PKR to Parents* + *3000 PKR to Daughter*	*Program Info* + *1500 PKR to Parents* + *1500 PKR to Daughter*	*Program Info* + *3000 PKR to Parents* + *0 PKR to Daughter*
Partial info	*Program Info*	*Program Info* + *1500 PKR to Parents*	*Program Info* + *3000 PKR to Parents*

Increasing Share of Incentive Assigned to Parents

Notes: Incentive and information targeting to parents by treatment arms. Daughters always have complete information about the program, the incentive, the allocation rule, and the information shared by the research team with their parents.

All daughters have full information on the total household incentive and the incentive split. To ensure full information transparency on girls' side when they made the decision to join the program, we also told the girls exactly what information their parents would receive.[18] We refer to the group where parents are not paid and not given full information on the household incentive as the control group ($\{s^p, \theta\} = \{0, \text{Partial}\}$). This is to mimic the status-quo in terms of program targeting, where

18. For example, girls assigned to {50%,Partial} were told: *"You will receive 1,500 rupees in cash, and your parents/legal guardian will receive 1,500 rupees via mobile payment at the end of the program. Your parents/legal guardian will not be informed of your share of the payment."* The girls assigned to {50%,Full} were told: *"You will receive 1,500 rupees in cash, and your parents/legal guardian will receive 1,500 rupees via mobile payment at the end of the program. We will inform your parents/legal guardian that you were selected to receive 1,500 rupees."*

girls are targeted to receive the incentive for task-completion and have total control over information sharing with parents.

2.3.2 Experiment Timeline

In December 2022, we held information sessions with PSDF at the selected high schools and colleges. To impose some effort on the girls to sign up, we asked attendees who were interested in enrolling in the program to complete a paper form, where they were quizzed on the information session and had to write about why they wanted to participate in the trainings. We received complete forms from 1,442 students following the information sessions.

Given the program capacity of 300 completed certificates, we screened candidates based on their responses to the attention check and open-ended questions and 792 students passed.[19] These students were randomized into contract groups at the individual level[20] and invited to the baseline survey. The randomization was stratified on school ID and a dummy variable for self-reported computer use above average.

The baseline survey took place in schools and students were notified of their contract group only after completing the baseline survey. Enumerators informed them that the training provider would need to contact their parents for permission because the

19. We told students to fill in the sign-up form as completely as possible and that their response would be used for screening purposes. The screening criteria includes: having complete personal identifiable information (e.g. full name, student number), personal and parent's contact information, consistent program choices, correct answers to all attention check questions, and answered at least one open-ended question about motivation for joining the program.

20. If siblings from the same household signed up for the program, only one sibling was randomly selected to complete baseline.

program was delivered by an external organization and offered monetary incentives for completion. To ensure the separate targeting of incentive between daughters and parents, we informed girls that upon program completion they would receive their portion, $s - s^p$, in cash and that parents would receive s^p via mobile money transfer. Of the 489 girls who completed the baseline survey, only one girl reported that she was no longer interested in the program upon hearing her contract offer.

Following the baseline survey, we attempted to phone the parents of 488 girls who were still interested in the program. We provided parents the relevant information on the course and the incentive payment should their daughter complete the training program.[21] A text message of the program offer was sent to parents after the call.[22] Parents could give consent to their daughters' participation only by replying to the text message within two days, otherwise this offer would expire. We adopted this approach to give some time for negotiation within the household and to avoid experimenter demand bias caused by enumerators requesting consent over the phone. We successfully surveyed and texted 416 parents. Of the parents surveyed, 313 texted

21. For example, parents from households assigned to {50%, Partial} were told: *"We used a lottery to determine the reward for [student's name]. Congratulations, you as the legal guardian will receive 1,500 rupees via mobile payment if [student's name] is selected and successfully completes this free program. The reward will be disbursed after the program finishes."* Parents in households assigned to {50%, Full} were told: *"We used a lottery to determine the reward for [student's name]. Congratulations, your household has won 3,000 rupees which we will give 1,500 rupees to [student's name] in cash and send the remaining 1,500 rupees to you via mobile payment if [student's name] is selected and completes the program successfully. The reward will be disbursed after the program finishes."* If parents had any questions about the course or about PSDF, they were directed to a hot-line that was set up by PSDF.

22. The text message to parents is in Urdu and reads as follows: *"[student's name] has been given priority for a spot to participate in the xxx training program. [Description of the program offer] Text 1 to accept the offer. Text 0 to decline. Please respond to us within 2 days, otherwise this offer will expire."*

us back within 2 days granting permission.[23]

Students who received parental permission were then contacted during school to enroll into the PSDF Coursera platform. They were requested to complete the digital skills training courses in the computer lab of their school, during the government mandated daily free-period.[24] Once enrolled in the program, students needed to watch the online tutorials and receive satisfactory marks on all of the required assignments in six weeks. At the end of the program, PSDF invited all 119 girls who completed the training to a certificate ceremony and disbursed the incentive to girls via cash and sent to parents via mobile money.

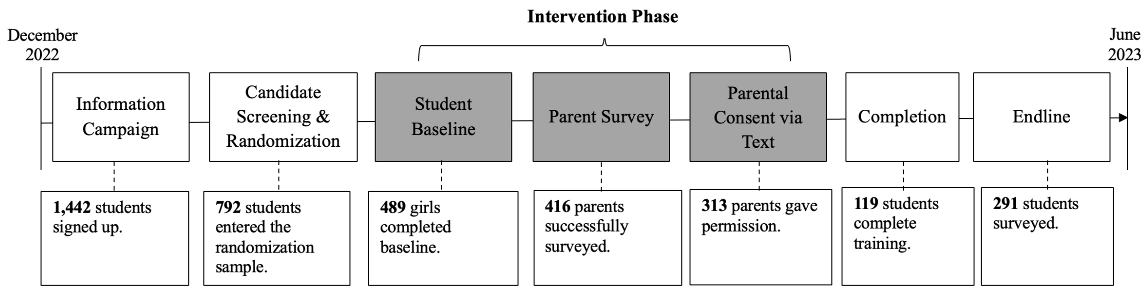
After PSDF closed the Coursera platform for all students, we invited the 489 girls who filled out the baseline survey to complete the endline survey. The endline survey was designed to help us understand why girls dropped out of the program and to measure digital skills through a self-assessment test and a short quiz. We surveyed 291 students in person and via phone.

In Figure 2.2, we summarize the timeline and document the sample size in each part of the research.

23. We conducted a back-check survey on 65% of households who completed the parents survey but did not text us back. We find evidence that parents who did not text us back seemed to not be interested. See Appendix B.2.2 for further discussion.

24. Anecdotal evidence suggests that students usually take the free-period for leisure. In the baseline survey, most students also anticipated that the program would take up their leisure time.

Figure 2.2: Experiment timeline



Notes: This figure reports the timeline of the fieldwork and the sample size at each stage.

2.3.3 Sample Characteristics and Main Outcomes

We present baseline characteristics for all 489 households and show balance across treatment arms in Table 2.2. Our full sample has an average household size of 6 people with a monthly household income of 41,620 PKR, which is about the average in urban Lahore in 2020.[25] We find that 97.8 percent of girls are interest in participating in the labor force after graduation and 73.4 percent think only skilled, white collar jobs are appropriate for women.

Our main outcomes of interest are program completion, parental consent, and girls' drop-out. The first two outcomes are collected from the PSDF Coursera platform and the text messages. Conditional on parental consent, we attribute any decision to drop-out of the program as daughter's choice. One concern is that parents might withdraw their consent after the program started, so we cannot decompose household

25. According to Pakistan Social and Living Standards Measurement Survey 2019-2020, the average monthly household income in urban Lahore is 45,900 thousand PKR. However, Pakistan has experienced high inflation rates from 9.7 percent in 2020 to 19.9 percent in 2022 (World Bank (2022)), so we expect the average monthly household income in 2022 to be higher than the 2020 statistics.

Table 2.2: Summary Statistics and Balance Check

	Full	Partial Info			Full Info			
	Sample	0%	50%	100%	0%	50%	100%	P-Value
N	489	69	85	90	82	69	94	
				Panel A: Personal Characteristics				
High School Student	0.225	0.275	0.212	0.278	0.171	0.217	0.202	0.531
	(0.418)	(0.450)	(0.411)	(0.450)	(0.379)	(0.415)	(0.404)	
Plans to Work	0.978	0.971	0.988	0.978	0.976	0.971	0.979	0.965
	(0.148)	(0.169)	(0.108)	(0.148)	(0.155)	(0.169)	(0.145)	
Only Skilled Jobs Appropriate for Women	0.734	0.783	0.718	0.656	0.756	0.812	0.713	0.265
	(0.442)	(0.415)	(0.453)	(0.478)	(0.432)	(0.394)	(0.455)	
Own a Personal Smart Phone	0.108	0.116	0.059	0.078	0.195	0.087	0.117	0.143
	(0.311)	(0.323)	(0.237)	(0.269)	(0.399)	(0.284)	(0.323)	
				Panel B: Household Characteristics				
Household Size (in Members)	6.427	6.304	6.600	6.544	6.244	6.565	6.309	0.679
	(1.812)	(1.428)	(2.178)	(1.831)	(1.495)	(2.193)	(1.633)	
Monthly Income (PKR in Thousands)	41.620	42.392	41.582	39.379	42.656	40.250	43.334	0.657
	(17.735)	(17.723)	(17.353)	(17.520)	(20.501)	(16.668)	(16.594)	
Father Works	0.890	0.797	0.906	0.922	0.939	0.841	0.904	0.111
	(0.314)	(0.405)	(0.294)	(0.269)	(0.241)	(0.369)	(0.296)	
Mother Works	0.135	0.188	0.106	0.122	0.159	0.130	0.117	0.742
	(0.342)	(0.394)	(0.310)	(0.329)	(0.367)	(0.339)	(0.323)	
Father with HS Degree or Above	0.681	0.754	0.671	0.633	0.646	0.652	0.734	0.439
	(0.467)	(0.434)	(0.473)	(0.485)	(0.481)	(0.480)	(0.444)	
Mother with HS Degree or Above	0.556	0.551	0.506	0.578	0.549	0.580	0.574	0.932
	(0.497)	(0.501)	(0.503)	(0.497)	(0.501)	(0.497)	(0.497)	
Has Computer at Home	0.689	0.594	0.765	0.722	0.683	0.681	0.670	0.324
	(0.463)	(0.495)	(0.427)	(0.450)	(0.468)	(0.469)	(0.473)	

Notes: This table reports baseline summary statistics and balance of personal and household characteristics, and p-values from a joint test of equality of means. Standard deviations reported in parentheses. Summary statistics on household monthly income are from the parent survey (sample size = 416). We replace missing values of household monthly income with the average household income at the respondent's school.

completion outcome into separate actions by parents and daughters.[26] To address this concern, we leverage a key feature of our setting: the girls can only take the trainings on campus during school hours. Once parents granted consent, they could not perfectly monitor daughters' participation or withdraw their permission without

26. There are very few settings in which one-time parental consent is credible. If trainings were held outside of school hours or if they were held at home, parents might implicitly withdraw consent by refusing to drive daughters to the training or by restricting their daughter's access to the computer at home.

notifying the research team or pulling their daughters out of school. Appendix Figure B.1 documents the reasons why girls who were given consent dropped out of the program. No student reported being discouraged by a parent or other member in the household.

To evaluate the unitary model of the household and the overall impacts of the payment and information targeting on completion, we use household-level completion outcomes on the full sample of 489 households that have a student baseline survey. To evaluate whether household bargaining is efficient under information symmetry, we restrict our sample to the 416 households that have a completed parent survey because we need to separate parent's refusal to consent from girl's decision to drop out. We find no evidence of selective attrition into the parent survey.[27]

2.4 Estimation and Results

The specification for all outcomes is:

$$
\begin{aligned}
y_i =& \beta_0 + \beta_1\{50\%, \text{Partial}\}_i + \beta_2\{100\%, \text{Partial}\}_i + \beta_3\{0\%, \text{Full}\}_i + \beta_4\{50\%, \text{Full}\}_i \\
&+ \beta_5\{100\%, \text{Full}\}_i + \boldsymbol{\gamma X_i} + \boldsymbol{\delta S_i} + \varepsilon_i
\end{aligned} \tag{2.6}
$$

where i denotes a household, y_i is the outcome of interest, and the control group $\{0\%, \text{Partial}\}$ is omitted. The vector $\boldsymbol{X_i}$ contains 10 baseline survey variables and

27. See B.2 for a detailed discussion on non-responses in the parent survey.

$\boldsymbol{S_i}$ strata fixed effects. We report bootstrapped confidence intervals and p-values. Parameters of interest (β's) can be directly mapped to hypotheses presented in Table 2.1. Table 2.3 lists the statistical tests and model predictions.[28]

Table 2.3: Model Predictions and Statistical Tests

		Treatment Groups	
Household Models	Hypothesis	Partial Info	Full Info
Unitary	H1	$\beta_1 = \beta_2 = 0$	$\beta_3 = \beta_4 = \beta_5$
	Income Pooling	$\beta_2 = 0$	
Efficient Collective	H2, H3, H4		$\beta_3 = \beta_4 = \beta_5$

Notes: Remark 3 from Section 2.2 discusses the income pooling test of the unitary model.

2.4.1 Effects on Program Completion

Figure 2.3 shows household completion rates across treatment groups for all households that had a girl's baseline survey.[29] In the control group, 15.94 percent of students completed the program. We find that treating parents with either payments or information about payments pushed completion rates up by 7.32pp to 16.45pp. The completion rate is the highest at 32.39 percent when parents are assigned the entire payment under partial information.

28. All possible statistical tests of the models are presented in in Appendix Table B.5. Appendix Tables B.6 and B.7 report all relevant regressions, bootstrapped p-values, and Fisher's exact p-values robust to relevant sample restrictions and the exclusion of covariates.

29. We also report all household level results in Appendix Table B.6 restricted to the sample where we were able to successfully complete the parents survey and deliver information on s^p and s. The full sample and the parent survey sample represent intention to treat estimates and treatment on the treated estimates, respectively. For the sake of evaluating the unitary model of the household, estimates from the full sample of students who completed baseline is the most robust and policy relevant.

We highlight two distinct patterns. First, we observe very little differences in completion rates for households where parents were given full information. As suggested by Remark 3 from Section 2.2 and Table 2.3, we cannot distinguish the unitary household from an efficient collective one as both models yield the same prediction. Hence, to evaluate the unitary model, we need to restrict our analysis to the partial information treatment.

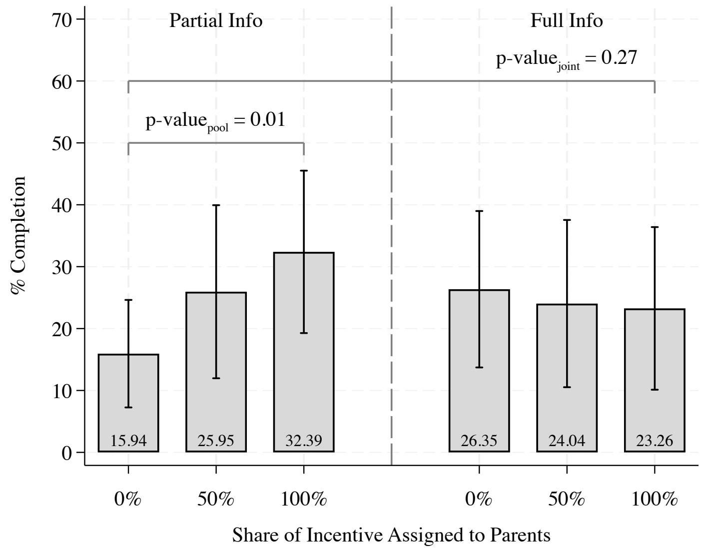
Second, in the partial information group, completion increases with the share of incentive targeted to parents. This pattern contradicts Hypothesis 1, as completion should not vary with s^p under the unitary model. We conduct a classic test of income pooling, which is to compare household completion where 100% of payment is allocated to either the student or the parents.[30] When parents are paid instead of daughters, we find that completion rates increase by 103% and is statistically significant at the 1% level.

2.4.2 Effects on Parental Consent and Girl's Drop-out

We now restrict our analysis to households that have both a daughter's and parent's baseline survey. This allows us to evaluate whether changes in household completion were driven by parent's refusing consent or daughters dropping out. To evaluate the efficient collective model, we perform the statistical test as shown in Table 2.3 in the full information group.

30. Appendix Table B.6 reports p-values for two other statistical tests that would contradict H1 under the unitary household.

Figure 2.3: Program Completion Rates *(Full Sample)*



Notes: We report completion rates across the treatments in the full analysis sample ($N = 489$), derived from a regression of program completion on treatment indicators with strata fixed effects and controls for whether the student is in high school, in the last year of school, household size, working status of the father and mother, high school graduation status of the father and mother, if there is a computer at home, if the student has their own cell phone, and if the student plans to work. Bootstrapped 95% confidence intervals are reported, along with a bootstrapped p-value from a joint test of null effects for all treatments and a bootstrapped p-value of a two-sided test on income pooling ($\beta_2 = 0$). Bootstrap results are generated from 10,000 draws.

Panel A of Figure 2.4 shows that within this sample, completion rates are remarkably similar under full information. Panels B and C break down household completion by evaluating parental consent and girl's drop-out. Bootstrapped p-values for joint tests of equality on payment targeting under full information are 0.82, 0.88, and 0.96 for H2, H3, and H4 of Table 2.1. While our confidence intervals are large due to smaller sample sizes, differences in magnitude are small under full information, while they are large under partial information. Hence, our results seem consistent with the efficient collective household model, and we are unable to reject pareto-efficient bargaining on payments under full information.

To address potential concerns with the small sample, we also report Fisher's exact p-values in Appendix Table B.7 and find them essentially identical to their bootstrapped counterparts. Appendix Figure B.5 plots the full distributions of bootstrapped group means under full information and the distributions mostly overlap. In addition, we compare our estimates with those in other papers that reject the efficient collective household model as shown in Appendix Table B.8.[31] Differences in our point estimates across payment targeting under full information, adjusted from percentage points to percentages, range from 6% to 10%, which is on the smaller end.

2.5 Discussion of results

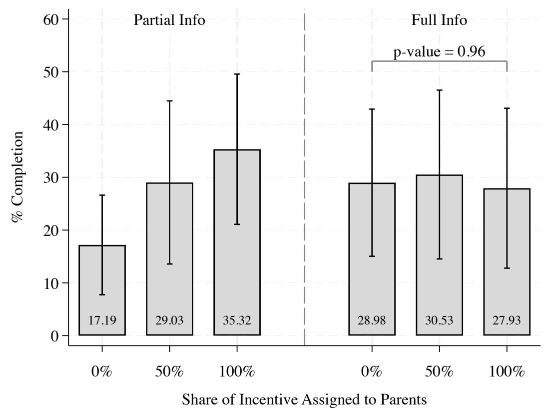
2.5.1 Why Does Completion Increase with s^p Under Partial Information?

Recalling Remark 4 of Section 2.2, a standard collective model with no assumptions on preferences and cost structures $\{u^d, c^d\}$ and $\{u^p, c^p\}$ cannot generate predictions on completion rates across s^p under partial information. Yet, experimental results show that under partial information, completion rates increase in s^p. We discuss two explanations for why completion rates increase when parents are allocated larger

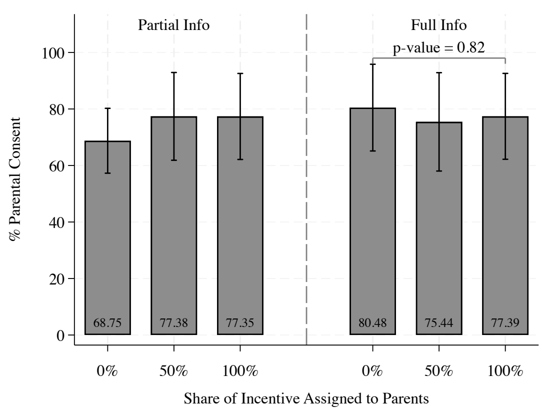
31. See Munro (2018) for a review of intra-household experiments and sample sizes.

Figure 2.4: Household Bargaining Outcomes

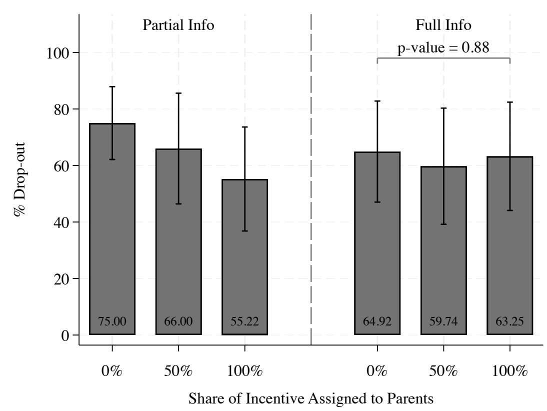
Panel A. Program Completion *(Parent Survey Sample)*



Panel B. Parental Consent *(Parent Survey Sample)*



Panel C. Girls' Drop-out *(Parental Consent Sample)*



Notes: Program completion, parental consent, and girls' drop-out across the six treatments, derived from a regression of the outcomes on treatment indicators with strata fixed effects and controls for whether the student is in high school, in the last year of school, household size, working status of the father and mother, high school graduation status of the father and mother, if there is a computer at home, if the student has their own cell phone, and if the student plans to work. Panels A and B report completion and parental consent among households that answered the parent survey (N = 416). Panel C reports girls' drop-out conditional on parental consent (N = 313). Bootstrapped 95% confidence intervals are reported, along with bootstrapped p-values from the joint test on equality of treatment effects under full information. Bootstrap results are generated from 10,000 draws.

shares of the incentive.

Explanation 1: Parents are able to influence completion through reminders.

Post-program focus groups with treated households reveal some parental nudging to encourage daughters' program completion through reminders and sharing updates with relatives.

We incorporate parental nudging to the collective household model and derive its predictions in Appendix B.5.1. In the adjusted model, it is costless for parents to nudge daughters. When a greater payment amount is assigned to parents, they exert a stronger influence on daughters' actions to ensure the household can attain the reward. Daughters gain indirect utility by complying to what parents want but lose own monetary incentive to complete the program. In this case, our experimental results imply that the positive effect of parental nudging dominates the negative effect of daughters losing incentive for completion.

Explanation 2: Payment to parents reveals information.

It is possible that the act of paying parents acts as information on the true value of s. In this case we rewrite, $\theta(s^p) \in [0,1]$ as a function that is increasing in s^p. Under this explanation, the support of $\lambda|_{\theta \in [0,1]}$ will converge to that of $\lambda|_{\theta=1}$ as we increase payment to parents.

This explanation is also consistent with our experimental results, as we cannot detect a statistically significant difference in completion between $\{100\%, \text{Partial}\}$ and all treatment groups $\{s^p, \text{Full}\}$.

2.5.2 Implications for Incentive Design

We reject the unitary model in the context of parents and daughters, but find that once information symmetry on payment is imposed, differences in household completion rate from payment targeting are minimal. Our results have important implications for incentive design around skills training programs for young adults still living in the household. Policymakers often trade-off training completion with externalities of the payment itself. These externalities include costs of delivering payments or bargaining frictions. Our findings suggest that if information symmetry can be imposed, this trade-off may not exist. That is, policymakers can simply design payment targeting that minimizes externalities of the payments while obtaining similar completion rates. Otherwise, when information symmetry cannot be easily imposed, our results show that household completion is the highest when parents are paid the entire incentive.

We highlight two discussion points around the full information group. First, we note that parents lack the ability to monitor the girls once they begin the training program. Several studies (Bursztyn and Coffman (2012), Bergman (2021), de Walque and Valente (2023)) find evidence that randomizing the ability for parents to monitor the actions of their children has impacts on school attendance and test scores. It is possible that had we introduced monitoring, completion would have increased with payment targeting to parents even under full information.

Second, we note that negotiation outcomes are conditional on the ability to split the incentive according to the sharing rule. If one party does not believe that the other will comply, then targeting more money to this party will incentivize consent or discourage drop-out even under full information. We derive the alternative model with beliefs on transfer frictions in Appendix B.5.2 and reject its prediction using results from the full information group.

2.5.3 Can We Generalize Our Results to Different Bargaining Settings?

Our results can inform incentive design under intra-household decision making when bargaining happens before a single, future payment that is made conditional on an observed action (ex-ante bargaining). While this setup generalizes to many other incentivized programs, we discuss main considerations for incentive design in other bargaining settings.

We start by considering the one-time payment that comes before an observed household action. This type of scenario is common when a policymaker is trying to incentivize household spending using CCT. Household members negotiate how to spend the money after receiving the CCT (ex-post bargaining). The key distinction between the ex-ante and the ex-post bargaining environments is the friction on ability to split the incentives. Frictions in the ex-ante bargaining case depend entirely on beliefs around the other party's compliance to the agreed sharing rule. In the ex-post case, frictions may also arise in actual ability to transfer money between agents. This friction on transferring money between agents is often studied in the spouse context where agents have separate savings methods or bank accounts. This friction may be less relevant in our context, as girls in our sample receive all money for private consumption from family.

We end by considering the multi-period bargaining context. This scenario is most analogous to working outside of the household (Lowe and McKelway (2024)), or longer incentivized educational programs where payments are made conditional on attendance every period (Baird et al. (2011), Bursztyn and Coffman (2012), de Walque and Valente (2023)). Beliefs around the other party's compliance to the sharing rule should reflect the true friction as agents update their priors by the start of every period. Hence, the actual ability to transfer incentives or income becomes highly relevant for outcomes such as labor supply and program take-up. Results in this multi-period scenario likely converge to the ex-post bargaining outcomes.

2.5.4 Program Evaluation

We document self-reported digital literacy by completion status in Appendix Figure B.2, B.3 and Table B.1. Our results cannot be interpreted as causal effects of online training programs as we did not recruit large enough of a sample to hold a pure control group[32], yet there are some striking patterns. Basic digital skills within our sample are very low, as Appendix Table B.1 shows that among girls who did not enroll in the program, 47.67 percent are not confident cutting and pasting text. Likewise, 73.42

32. Due to the lack of power, we are not able to estimate the program effect on the treated using the first stage of treatment assignment on program completion.

percent of girls in our baseline survey felt that it was only appropriate for women like them to work in skilled, white collar jobs. Such jobs are guaranteed to require basic knowledge of computer skills, suggesting that young women lack essential skills to pursue their desired career. We show some improvements in self-reported confidence with essential computer tasks, providing hope that girls who participated are more likely to continue acquiring basic digital skills on their own, without need for official trainings.

2.6 Conclusion

In many contexts, educational decisions for teenagers are made jointly with parents. Our study sheds light on how policymakers can leverage the dynamics of intra-household decision-making to effectively incentivize young women's skills investment.

Our findings reveal that targeting parents with payments for program completion or information about payments can improve program take-up and completion. Hence, we reject the unitary household model between parents and young adult daughters. Moreover, we document that inefficiencies in household negotiation on incentive payments come through information asymmetry and not through payment targeting. When young women face household opposition for investing in their income-generating skills, our results have several implications for the optimal incentive design. If the policymaker can impose information symmetry on payment, payment targeting does not impact completion. If the program designer cannot impose information symmetry, payments should be targeted towards parents to maximize household completion.

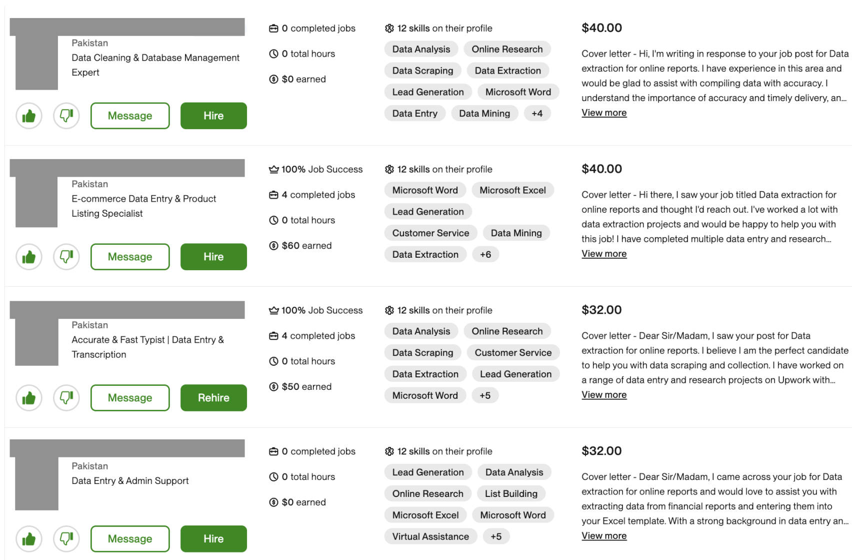
We think that studying household bargaining between parents and young adults and identifying other frictions to decision making, especially in the context of longer-term human capital investments and labor force participation, is a promising area for future research. Future work might also explore how family dynamics affect bargaining between different household members (e.g., parents and sons, the role of extended family members). Finally, identifying a scalable solution that can successfully boost basic digital literacy skills for middle-income teenage girls in Pakistan is an important area of research, as these girls are some of the least likely to participate in the labor force due to social stigma around unskilled work coupled with basic skills gaps (Jayachandran (2021)).

APPENDIX A

BELIEFS, REPUTATION, AND BARRIERS TO ENTRY IN ONLINE LABOR MARKETS

A.1 Figures and Tables

Figure A.1: Freelancer Profiles for Audit Study

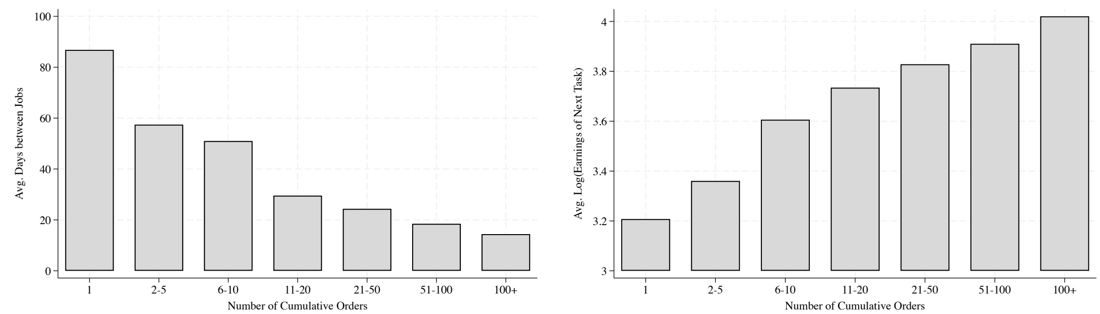


Notes: This is a screenshot taken from the employer's dashboard on the online freelancing platform. The original job post has a fixed budget of 40 USD. The applications are sent by four freelancer profiles in the audit study. The top and the bottom applications are sent by inexperienced freelancers and the middle ones by experienced freelancers.

Figure A.2: Hiring Prospects and Completed Jobs

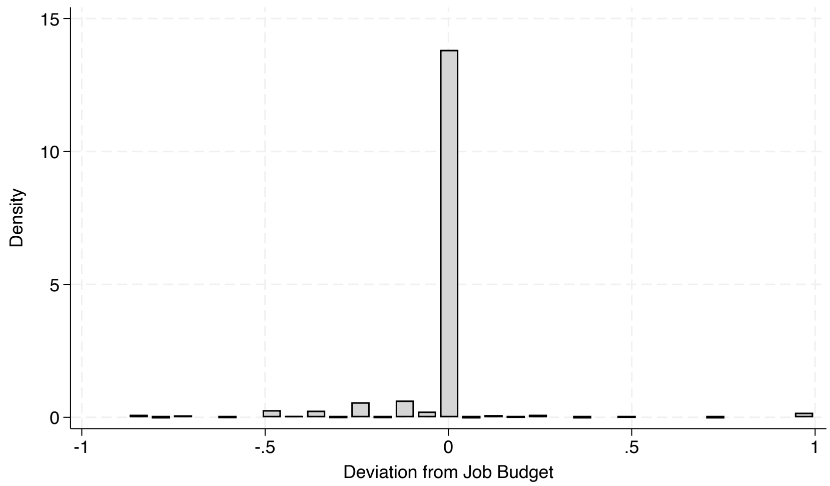
Panel A. Gap between Jobs (Days)

Panel B. Log Earnings of Future Jobs



Notes: The figures plot the average number of days between jobs (Panel A) and the log of earnings from the subsequent job (Panel B) against the number of completed jobs. The job-level data was collected for all active Pakistani workers specialized in data entry on the studied platform by June 2024.

Figure A.3: Distribution of Deviation in Proposed Wages

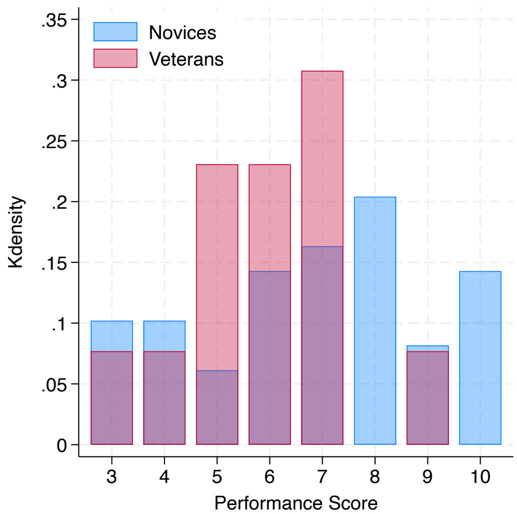


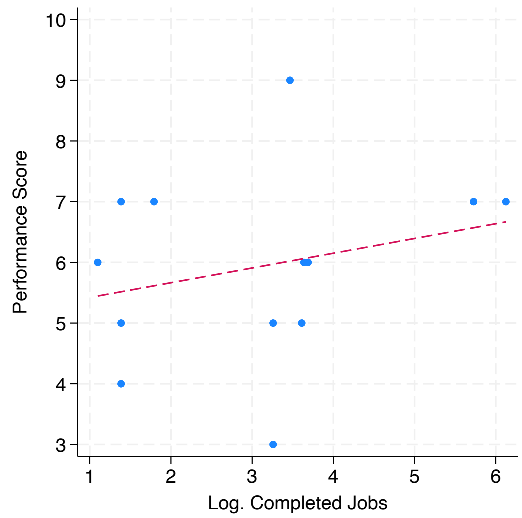
Notes: This figure shows the distribution of deviation in proposed wages from the job budget, measured in the application data.

Figure A.4: Follow-up Task Performance

Panel A. Novices versus Veterans

Panel B. Veteran Performance by Experience

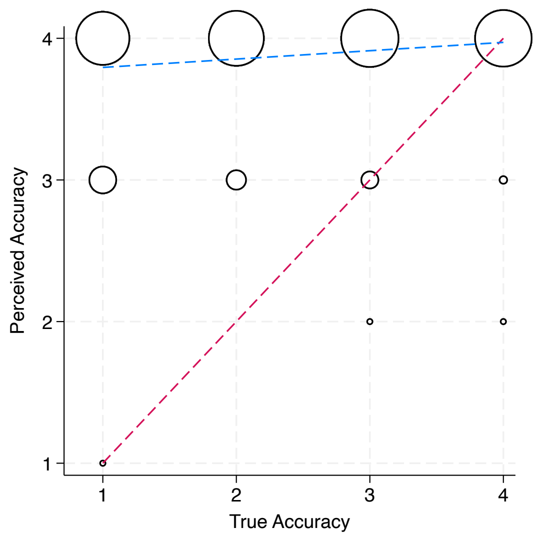




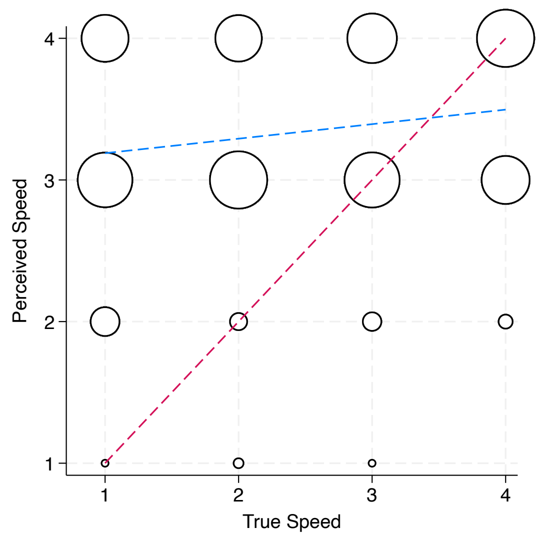
Notes: This figure describes follow-up task performance. Panel A plots the distribution of follow-up performance scores for novices ($n = 49$) and veterans ($n = 13$). Performance score ranges from 3 to 10, computed as the sum of quartile ranking in accuracy and speed, and a binary variable in whether workers followed task instructions perfectly. Panel B plots the relationship between log completed jobs and performance scores among veteran workers.

Figure A.5: Actual versus Perceived Performance by Each Dimension

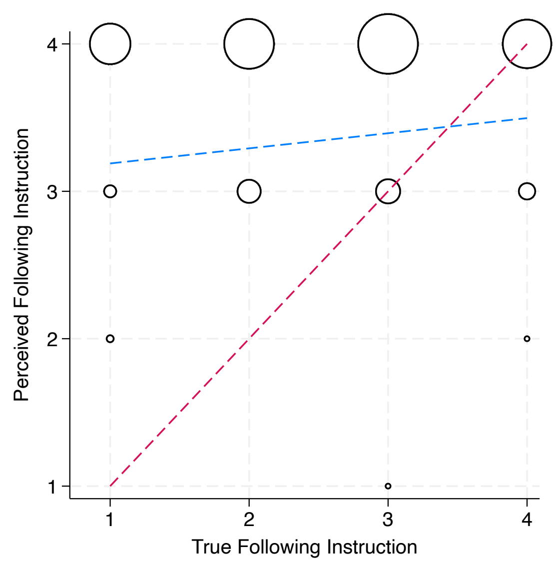
Panel A. Accuracy



Panel B. Speed

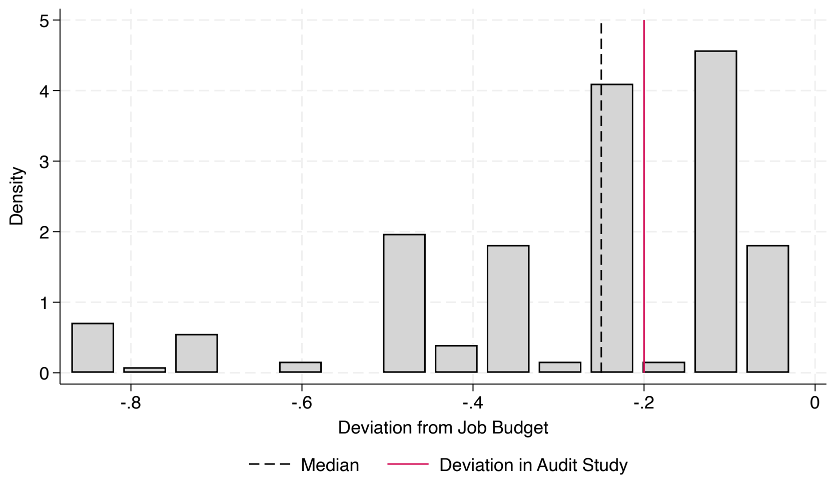


Panel C. Follow Instruction



Notes: This figure shows the actual and perceived scores for accuracy, speed, and following instructions, with the true scores on the x-axis and the perceived scores on the y-axis. The red dotted line is the 45-degree line and the blue line shows the best linear fit. The marker size is relative to the number of observations.

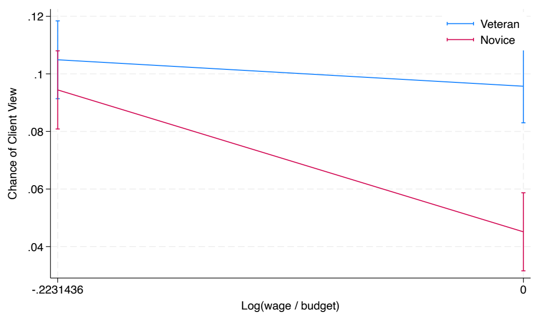
Figure A.6: Distribution of Wage Bids Below Job Budget



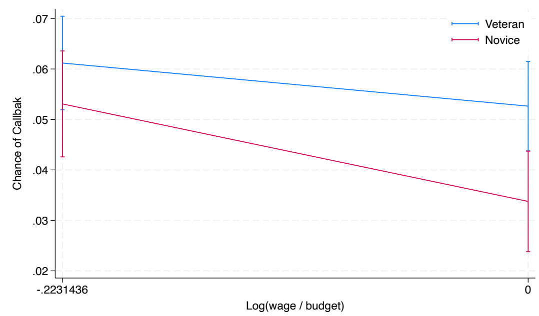
Notes: This figure shows the distribution of wage bids below the job budget, measured in the application data. The dotted black line shows the median value of the deviation and the solid red line represents wage undercutting by 20%, which is what we proposed in the audit study.

Figure A.7: Demand Response

Panel A. Employer View

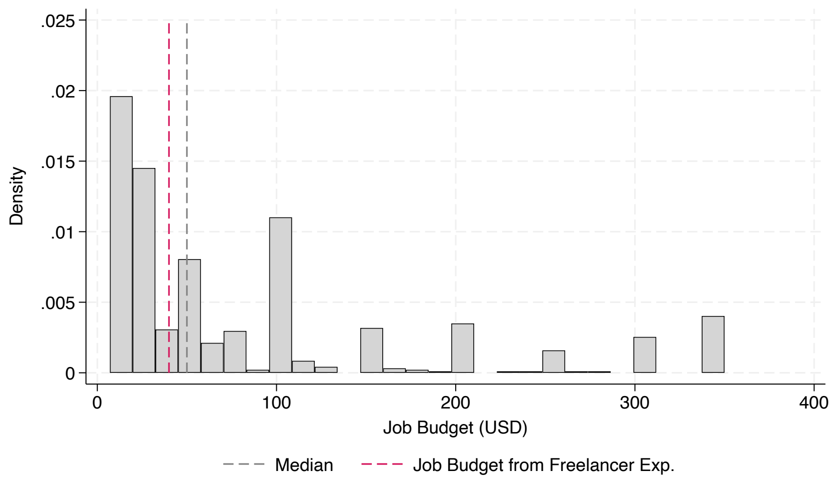


Panel B. Callback



Notes: This figure plots mean employer views (Panel A) and callbacks (Panel B) by wage bid and worker experience. The horizontal axis shows log(wage/budget), where 0 corresponds to bidding at the posted budget and âĹŠ0.223 corresponds to bidding at 80% of the budget. Vertical bars denote 95% confidence intervals.

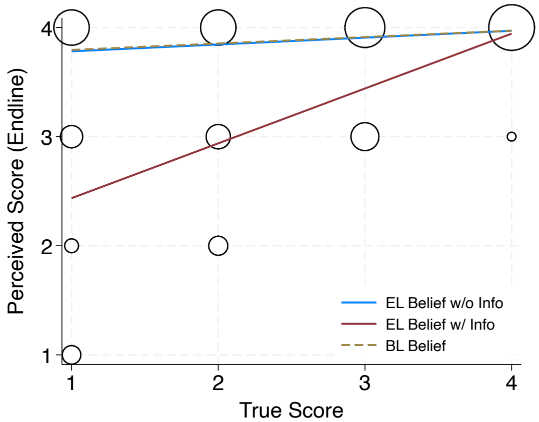
Figure A.8: Distribution of Budget for Data Entry Jobs



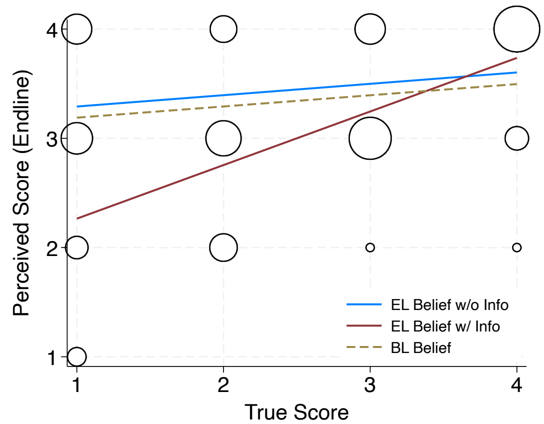
Notes: This figure shows the distribution of job budget from 703 data entry jobs in the demand experiment. The red dotted line shows where the job budget for the referred job falls in the freelancer experiment.

Figure A.9: Treatment Effects on Worker Beliefs by Performance Dimension

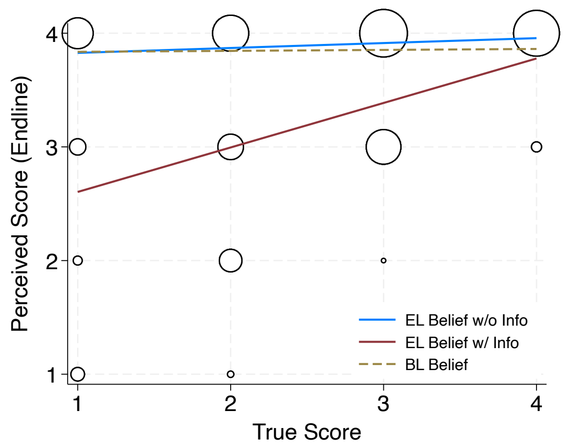
Panel A. Accuracy



Panel B. Speed

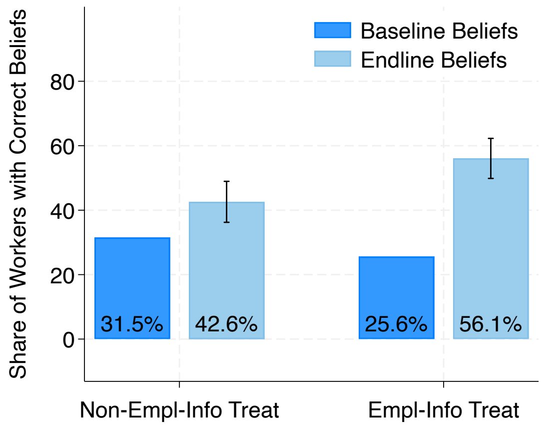


Panel C. Follow Instruction



Notes: This figure plots workers' actual score for each performance dimension (accuracy, speed, and following instruction) against their perceived score measured at the endline. The blue line represents the best linear fit for workers without performance feedback and the red line for workers with performance feedback. The dotted line shows the best linear fit of baseline perceived performance against true performance.

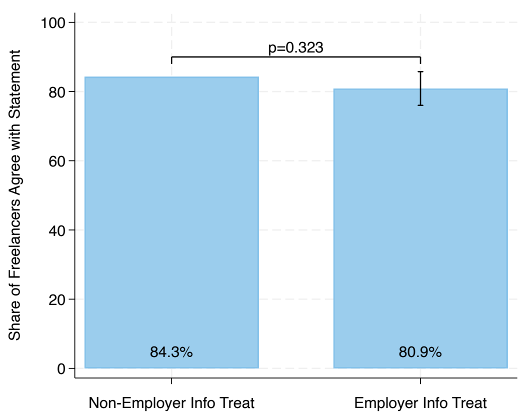
Figure A.10: Baseline and Endline Beliefs by Treatment



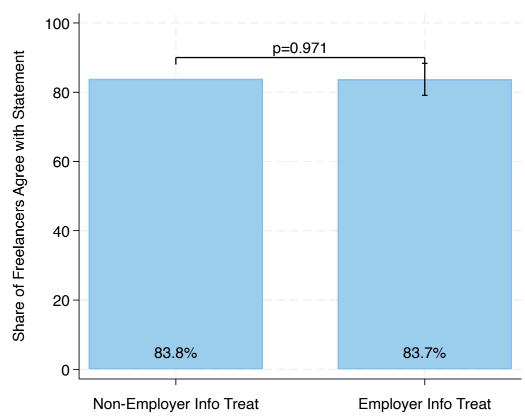
Notes: This figure plots the share of workers who "disagree" or "strongly disagree" with the statement in the baseline survey: *"If a new freelancer offers a lower price, the client may think that their work is low quality."*; and those who "agree" or "strongly agree" with the statement in the endline survey: *"The employer (of the referred job) will not judge my quality based on my wage bids."* We report the 95% confidence intervals and p-values from the Wald test.

Figure A.11: Treatment Effect on Worker Perceptions of Referred Job

Panel A. This job opportunity seems trustworthy.

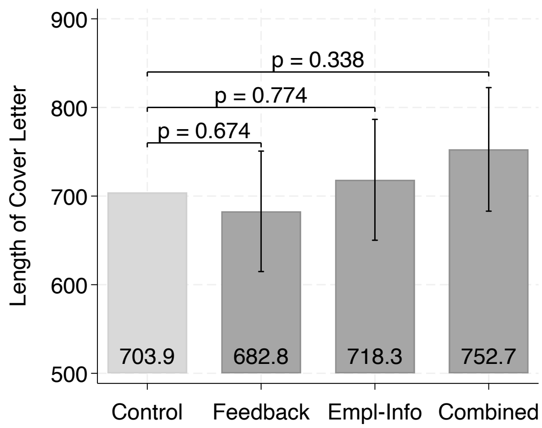


Panel B. Competition may be strong.



Notes: This figure reports the treatment effects on worker beliefs about the trustworthiness and competition for the referred job. We report the 95% confidence intervals in both panels and p-values from the Wald test.

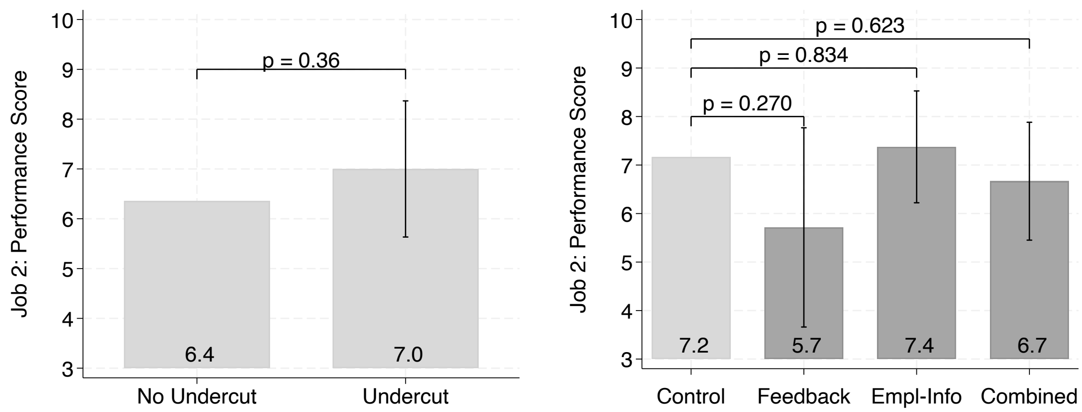
Figure A.12: Treatment Effect on Cover Letter Length



Notes: This figure reports the effects of treatments on the cover letter length by in-sample novices ($N = 390$) who applied to the follow-up task. We report the 95% confidence intervals and p-values from the Wald test in both panels.

Figure A.13: Treatment Effect on Follow-up Task Performance

Panel A. Performance by Whether Undercut **Panel B.** Performance by Treatment



Notes: This figure reports the effects of undercutting and treatments on performance (score between 3 and 10) by in-sample novices who completed the follow-up task. We report the 95% confidence intervals and p-values from the Wald test in both panels.

Table A.1: Job and Employer Characteristics

	Median	Mean	Min	Max	N
Job Characteristics					
Job Budget (USD)	50	112	7	8,000	743
Any Freelancer Hired	1.00	0.54	0.00	1.00	743
No. Freelancer Hired	1.00	0.73	0.00	14.00	743
Employer Characteristics					
Employer Tenure (Month)	24	40	0	281	720
Employer from HICs	1.00	0.85	0.00	1.00	720
No. Jobs Posted	21	139	1	3,468	720
Employer Rating (Out of 5)	4.99	4.86	1.75	5.00	565

Notes: This table reports characteristics of jobs and employers in the analysis sample of the employer experiment. The first panel shows job characteristics and hiring outcomes indicated on the platform, at least two weeks after the initial job post. The second panel shows employer characteristics. *Employer Tenure* is measured by the gap between when the employer first joined the platform and when the job was posted. *Employer from HICs* indicates whether the employer is located in a high-income country, using World Bank country classifications by income level for 2024-2025. *No. Jobs Posted* refers to the total number of jobs the employer has posted since joining the platform. *Employer Rating* is a measure of reputation (from one to five) generated by the platform based on rating by freelancers previously hired by the employer. New employers without any hiring history do not have a rating.

Table A.2: Type of Callback Messages

Type	Count	Share (%)
Request for additional info	32	22.53
Task instructions	30	21.13
Request for work samples	29	20.42
Confirm interest	21	14.79
Interview request	16	11.27
Other	13	9.15
Job offer	1	0.70
N	142	

Notes: This table categorizes callback messages received during the employer experiment. Shares are reported as percentages of the total number of callbacks (N = 142). "Request for additional information" refers to employers asking applicants to provide details such as availability or relevant experience. "Task instructions" refers to employers sharing details about the assignment. "Request for work samples" includes requests for past work samples or for applicants to complete a sample task. "Confirm interest" refers to employers checking whether the freelancer remains interested in the position. "Interview request" refers to invitations to interview. "Other" includes messages that do not fall into the preceding categories, such as a greeting. "Job offer" refers to direct offers of employment for the posted position.

Table A.3: Treatment Effects on Employer Responses

	Client Viewed				Contacted			
	Analysis Sample			Full Sample	Analysis Sample			Full Sample
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Novice (β_1)	-0.050	-0.043	-0.061	-0.041	-0.018	-0.025	-0.021	-0.025
	(0.011***)	(0.013***)	(0.015***)	(0.012***)	(0.008**)	(0.009***)	(0.010**)	(0.009***)
Undercut (β_2)	0.010	0.009	0.009	0.010	0.009	0.008	0.008	0.008
	(0.010)	(0.010)	(0.009)	(0.010)	(0.007)	(0.007)	(0.006)	(0.006)
Novice $\times$ Undercut (β_3)	0.038	0.039	0.051	0.035	0.010	0.010	0.014	0.010
	(0.015**)	(0.015**)	(0.017***)	(0.015**)	(0.011)	(0.011)	(0.012)	(0.010)
Constant	0.095	0.082	0.114	0.080	0.053	0.050	0.057	0.049
Test: $\beta_2 + \beta_3 = 0$	0.000	0.000	0.000	0.000	0.024	0.026	0.026	0.019
Linear Prob.	$\times$	$\times$		$\times$	$\times$	$\times$		$\times$
Logit			$\times$				$\times$	
Job FE	$\times$	$\times$		$\times$	$\times$	$\times$		$\times$
Profile FE		$\times$		$\times$		$\times$		$\times$
No. Applications	2812	2812	2812	2972	2812	2812	2812	2972

Notes: This table reports treatment effects on employer views and callbacks by worker experience and wage undercutting. Columns (1) and (5) present the baseline specification corresponding to Figure 1.3. Columns (2) and (6) include freelancer profile fixed effects. Columns (3) and (7) report logit estimates. Columns (4) and (8) extend the original analysis to the full sample with profile and job fixed effects. The omitted group is veteran applicants without wage undercutting. We report p-values from the statistical test for $\beta_2 + \beta_3 = 0$ from all specifications. Standard errors are clustered at the job level. $p < 0.10^{*}, p < 0.05^{**}, p < 0.01^{***}$.

Table A.4: Heterogeneous Treatment Effects on Employer Responses

	Job Budget		Employer Tenure		HIC Employer		No. Jobs Posted		Employer Rating	
	< Median (1)	≥ Median (2)	< Median (3)	≥ Median (4)	= 0 (5)	= 1 (6)	< Median (7)	≥ Median (8)	< Median (9)	≥ Median (10)
Panel A: Employer View										
Novice (β_1)	-0.057	-0.043	-0.074	-0.023	-0.029	-0.052	-0.083	-0.014	-0.030	-0.073
	(0.016***)	(0.016***)	(0.018***)	(0.013*)	(0.036)	(0.012***)	(0.018***)	(0.014)	(0.018)	(0.017***)
Undercut (β_2)	0.006	0.013	0.003	0.014	0.000	0.010	0.006	0.011	0.015	-0.015
	(0.014)	(0.016)	(0.016)	(0.014)	(0.024)	(0.012)	(0.016)	(0.013)	(0.017)	(0.015)
Novice × Undercut (β_3)	0.039	0.038	0.057	0.023	0.049	0.038	0.074	0.006	0.023	0.055
	(0.020*)	(0.023)	(0.024**)	(0.020)	(0.043)	(0.017**)	(0.024***)	(0.020)	(0.026)	(0.024**)
Omitted Mean	0.099	0.092	0.116	0.075	0.137	0.088	0.122	0.069	0.098	0.099
No. Applications	1328	1484	1344	1392	408	2328	1344	1392	1064	1092
Panel B: Employer Contact										
Novice (β_1)	-0.021	-0.016	-0.030	-0.003	0.000	-0.019	-0.030	-0.003	-0.011	-0.029
	(0.012*)	(0.011)	(0.013**)	(0.010)	(0.024)	(0.009**)	(0.012**)	(0.011)	(0.016)	(0.011**)
Undercut (β_2)	0.018	0.000	0.006	0.014	0.000	0.012	0.009	0.011	0.011	0.015
	(0.011)	(0.008)	(0.011)	(0.008*)	(0.014)	(0.007)	(0.009)	(0.010)	(0.012)	(0.010)
Novice × Undercut (β_3)	-0.003	0.022	0.018	-0.003	0.029	0.003	0.021	-0.006	-0.004	0.004
	(0.014)	(0.016)	(0.017)	(0.014)	(0.030)	(0.012)	(0.015)	(0.016)	(0.022)	(0.012)
Omitted Mean	0.057	0.049	0.065	0.037	0.088	0.045	0.057	0.046	0.071	0.044
No. Applications	1328	1484	1344	1392	408	2328	1344	1392	1064	1092

Notes: This table examines the heterogeneity in employer's response to wage undercutting by novice and veteran freelancers. For *Job Budget, Employer Tenure, No. Jobs Posted, and Employer Rating*, we divide the analysis sample by the median and for *HIC Employer*, the analysis sample is split by whether the employer is located in a high-income country, using World Bank country classifications by income level for 2024-2025. Panel A reports the heterogeneous treatment effects on whether employers viewed the application and Panel B on whether employers reached out to the applicant. All regressions control for job fixed effects. The omitted group is applications by veteran workers without wage undercutting. Standard errors are clustered at the job level. $p < 0.10^{*}, p < 0.05^{**}, p < 0.01^{***}$.

Table A.5: Data Sources for Back-of-the-Envelope Calculation

	Parameter Values		Source
	Bid 80%	Bid Full	
Contact Probability (p)	5.26%	3.41%	Demand Experiment
Initial Job Earnings (π_1)	50	40	Demand Experiment
Application Fee (c_a)	1.35	1.35	Demand Experiment
Earning Growth by Completed Jobs	-	-	Job History Data
Application Sent per Month	10	10	Freelancer Survey
Job Effort Cost ($c_e = w * h$)	12.5	12.5	Freelancer Survey
Hourly reservation wage (w)	5	5	Freelancer Survey
Hours taken per task (h)	2.5	2.5	Freelancer Survey

Notes: This table reports the key parameters used in the back-of-the-envelope calculation of the returns to wage undercutting, along with their data sources. The initial job earnings reflect the original job budgets posted by employers before any wage discounting. These values differ across bidding strategies because, in the demand experiment, jobs that responded to lower bids had a higher median budget than those that responded to full-price bids. Earnings growth is computed as a function of the number of completed jobs using the job history data, so no single value is reported here. Application behavior and effort costs are drawn from the baseline freelancer survey. The job effort cost equals the product of novices' self-reported reservation wage and the hours required to complete their first platform job.

Table A.6: Balance Test for Freelancer Experiment

	All	Control	Info		Signal		Info+Signal		p-val
N	481	118	117		126		120		
	Mean (SD)	Mean (SD)	Mean (SD)	β_1 (p-val)	Mean (SD)	β_2 (p-val)	Mean (SD)	β_3 (p-val)	
Panel A: Personal Characteristics									
Female	0.52	0.56	0.55	-0.01	0.47	-0.09	0.50	-0.06	0.45
	(0.50)	(0.50)	(0.50)	(0.86)	(0.50)	(0.16)	(0.50)	(0.36)	
Age	28.76	29.16	28.55	-0.62	28.96	-0.19	28.35	-0.81	0.72
	(5.98)	(6.67)	(5.87)	(0.45)	(5.65)	(0.81)	(5.74)	(0.32)	
Bachelor Degree or Above	0.83	0.79	0.85	0.06	0.83	0.04	0.85	0.06	0.59
	(0.38)	(0.41)	(0.36)	(0.24)	(0.38)	(0.46)	(0.36)	(0.21)	
Employed Offline	0.52	0.52	0.58	0.06	0.50	-0.02	0.48	-0.03	0.44
	(0.50)	(0.50)	(0.50)	(0.32)	(0.50)	(0.80)	(0.50)	(0.61)	
Full-Time Freelancer	0.36	0.35	0.35	0.00	0.33	-0.02	0.40	0.05	0.67
	(0.48)	(0.48)	(0.48)	(0.95)	(0.47)	(0.71)	(0.49)	(0.40)	
Platform Tenure (Month)	19.65	22.25	17.50	-4.75	19.00	-3.28	19.87	-2.36	0.59
	(25.70)	(34.46)	(19.90)	(0.20)	(24.62)	(0.39)	(21.51)	(0.53)	
Present on Other Platforms	0.44	0.47	0.40	-0.07	0.48	0.00	0.42	-0.06	0.52
	(0.50)	(0.50)	(0.49)	(0.24)	(0.50)	(0.98)	(0.50)	(0.38)	
Has Earned from Other Platforms	0.18	0.20	0.20	-0.01	0.19	-0.01	0.14	-0.06	0.56
	(0.39)	(0.40)	(0.40)	(0.89)	(0.39)	(0.80)	(0.35)	(0.22)	
High Type	0.50	0.47	0.53	0.06	0.49	0.02	0.49	0.02	0.86
	(0.50)	(0.50)	(0.50)	(0.40)	(0.50)	(0.79)	(0.50)	(0.79)	
Panel B: Profile Characteristics									
Number of Completed Jobs	0.13	0.13	0.14	0.01	0.09	-0.04	0.18	0.06	0.18
	(0.35)	(0.33)	(0.35)	(0.80)	(0.28)	(0.32)	(0.41)	(0.24)	
Hourly Rate (USD)	7.91	7.71	7.89	0.19	7.93	0.22	8.10	0.39	0.95
	(5.06)	(4.59)	(5.15)	(0.77)	(4.99)	(0.72)	(5.53)	(0.56)	
Low Income Country	0.09	0.09	0.09	-0.00	0.10	0.00	0.07	-0.03	0.80
	(0.28)	(0.29)	(0.29)	(0.99)	(0.29)	(0.97)	(0.25)	(0.44)	
Lower-Middle Income Country	0.78	0.74	0.75	0.02	0.78	0.04	0.84	0.10	0.17
	(0.42)	(0.44)	(0.43)	(0.79)	(0.42)	(0.45)	(0.37)	(0.04**)	
Upper-Middle Income Country	0.13	0.16	0.15	-0.01	0.13	-0.03	0.09	-0.07	0.32
	(0.34)	(0.37)	(0.36)	(0.89)	(0.33)	(0.45)	(0.29)	(0.11)	
Fluent in English	0.97	0.95	0.95	0.00	0.99	0.04	0.97	0.03	0.07*
	(0.18)	(0.22)	(0.22)	(0.99)	(0.09)	(0.05*)	(0.16)	(0.30)	
Portfolio	0.65	0.61	0.69	0.08	0.70	0.09	0.61	-0.00	0.28
	(0.48)	(0.49)	(0.46)	(0.20)	(0.46)	(0.16)	(0.49)	(0.97)	
Certificate	0.43	0.41	0.41	-0.00	0.50	0.08	0.37	-0.04	0.27
	(0.49)	(0.49)	(0.49)	(0.96)	(0.50)	(0.20)	(0.49)	(0.52)	
Verified Certificate	0.08	0.04	0.06	0.02	0.14	0.09	0.09	0.05	0.06*
	(0.28)	(0.20)	(0.24)	(0.55)	(0.35)	(0.01**)	(0.29)	(0.13)	
Number of Skills	11.12	11.48	10.74	-0.73	11.28	-0.20	10.97	-0.51	0.43
	(3.72)	(3.65)	(3.74)	(0.13)	(3.66)	(0.66)	(3.83)	(0.29)	
Employment History	0.95	0.97	0.97	0.01	0.95	-0.01	0.92	-0.05	0.23
	(0.21)	(0.18)	(0.16)	(0.69)	(0.22)	(0.58)	(0.28)	(0.11)	

Notes: This table summarizes personal (panel A) and profile (panel B) characteristics of all participants in the freelancer experiment. Personal characteristics are measured from the baseline survey and the data entry task. Profile characteristics are public information collected from freelancer profiles on the platform. We show means and standard deviations within treatment arms as well as coefficients and p-values on the treatment indicators from the regression: $Y_i = \alpha + \beta_1 \text{Info}_i + \beta_2 \text{Signal}_i + \beta_3 (\text{Info+Signal})_i + \eta_r + \varepsilon_i$. The last column reports a joint test of mean values across treatment groups. $p < 0.10^{*}, p < 0.05^{**}, p < 0.01^{***}$.

Table A.7: Heterogeneity by Beliefs of Ability

Dependent Var:	Abs(Actual-Perceive Performance Score)				Confidence in Perceived Score	
Wage Undercutting	1st Quartile (1)	2nd Quartile (2)	3rd Quartile (3)	4th Quartile (4)	Decreased (5)	Increased (6)
Panel A: Feedback Only versus Control						
Treated	-0.500	-0.000	-0.004	0.016	-0.012	-0.056
	(0.368)	(0.000)	(0.092)	(0.064)	(0.062)	(0.057)
High-Type	-0.444	0.050	-0.067	0.000	-0.010	-0.056
	(0.373)	(0.051)	(0.067)	(0.000)	(0.063)	(0.057)
Treated × High-Type	0.686	0.108	0.226	0.000	0.158	0.286
	(0.382*)	(0.101)	(0.172)	(0.000)	(0.097)	(0.135**)
Omitted Mean	0.500	0.000	0.067	0.037	0.069	0.056
No. Obs	51	45	46	46	142	46
Panel B: Employer Info Only versus Control						
Treated	0.167	0.250	0.161	0.180	0.175	0.069
	(0.100)	(0.159)	(0.115)	(0.095*)	(0.081**)	(0.135)
High-Type	-0.444	0.050	-0.067	0.000	-0.010	-0.056
	(0.370)	(0.051)	(0.067)	(0.000)	(0.063)	(0.056)
Treated × High-Type	0.000	-0.133	-0.161	0.000	-0.048	0.112
	(0.000)	(0.185)	(0.115)	(0.000)	(0.109)	(0.181)
Omitted Mean	0.500	-0.000	0.067	0.037	0.069	0.056
No. Obs	47	55	46	50	151	47

Notes: This table examines the heterogeneity in treatment effects by workers' accuracy of own ability assessment and confidence in accuracy of their guess. Panel A reports the heterogeneous treatment effects for feedback-only and control groups and Panel B on employer info-only and control groups. All regressions control for round fixed effects. The omitted group is low-type workers in the control group. Standard errors are clustered at the job level. $p < 0.10^{*}, p < 0.05^{**}, p < 0.01^{***}$.

Table A.8: Heterogeneity by Beliefs of Employer Response

Dependent Var: Wage Undercutting	Employer Think Low Price=Low Quality		High-Type Should Undercut		Undercut Works Better for High-Type	
	Disagree (1)	Agree (2)	Disagree (3)	Agree (4)	Disagree (5)	Agree (6)
Panel A: Feedback Only versus Control						
Treated	-0.040	0.131	0.062	-0.069	0.007	-0.049
	(0.081)	(0.105)	(0.073)	(0.094)	(0.076)	(0.111)
High-Type	0.051	-0.091	0.050	-0.069	0.018	-0.071
	(0.088)	(0.080)	(0.068)	(0.094)	(0.081)	(0.100)
Treated × High-Type	0.262	-0.051	-0.041	0.373	0.102	0.309
	(0.116**)	(0.131)	(0.096)	(0.131***)	(0.108)	(0.148**)
Omitted Mean	0.040	0.091	0.000	0.111	0.034	0.111
No. Obs	110	78	84	104	109	79
Panel B: Employer Info Only versus Control						
Treated	0.240	0.088	0.217	0.122	0.201	0.099
	(0.099**)	(0.094)	(0.088**)	(0.098)	(0.086**)	(0.111)
High-Type	0.051	-0.091	0.050	-0.069	0.018	-0.071
	(0.102)	(0.099)	(0.091)	(0.104)	(0.101)	(0.105)
Treated × High-Type	-0.170	0.130	-0.172	0.079	-0.078	0.075
	(0.138)	(0.136)	(0.126)	(0.140)	(0.129)	(0.159)
Omitted Mean	0.040	0.091	0.000	0.111	0.034	0.111
No. Obs	103	95	84	114	122	76

Notes: This table examines the heterogeneity in treatment effects by workers' baseline beliefs about employer response. Panel A reports the heterogeneous treatment effects for feedback-only and control groups and Panel B on employer info-only and control groups. All regressions control for round fixed effects. The omitted group is low-type workers in the control group. Standard errors are clustered at the job level. $p < 0.10^{*}, p < 0.05^{**}, p < 0.01^{***}$.

A.2 Reputation Model with Belief Frictions

A.2.1 Setup

Consider a labor market with a unit mass of workers and employers. Workers, indexed by i, live for two periods, while employers, indexed by j, live for one period. Workers hired in the first period become *veterans* in the second period, while those in their first period or not hired previously are referred to as *novices*. At the end of each period, all employers and second-period workers exit the market and are replaced by a new generation.

At the beginning of the period, each employer posts an identical job on the online platform with an exogenous budget $\bar{w}_j$. The budget $\bar{w}_j$ represents the maximum amount the employer is willing to pay for an online freelancer, based on the cost of hiring a local worker offline.

Workers differ in their innate ability $\theta_i \in \{H, L\}$ to execute the task. We assume that workers know their own type upon entering the market. Workers decide whether to apply to a job by submitting a wage bid $w_{ij} \leq \bar{w}_j$ or not applying at all. [1] If hired, the worker is paid w_{ij}; otherwise, the worker receives the outside option $c_\theta = \{c_H, c_L\}$.

Employers derive value $v_\theta = \{v_H, v_L\}$ from a worker's output but cannot observe the worker type prior to hiring. Instead, they observe proposed wages for all applicants

1. This is empirically supported: fewer than 3% freelancers in our application data proposed wages higher than the budget.

and public reviews r_i for veteran workers. For veterans, employers infer their type from their review history. For novices, employers form beliefs about their type based on their proposed wages w_{ij}: a novice is high type with probability $\gamma(w_{ij})$ and low type with probability $1 - \gamma(w_{ij})$. If all novices propose the same wage, or if wages are uninformative about worker type, employers may screen candidates through interview at costs s_j and obtain additional quality cues λ_{ij}. The employer's objective is to maximize the expected return from hiring:

$$\max_i \{E\left[v_i | r_i, w_{ij}, \lambda_{ij} \times 1[\text{screening}]\right] - w_{ij} - s_j \times 1[\text{screening}], \bar{v}_j\}$$

where $\bar{v}_j$ is the employer's outside option. Employers hire from the platform only if expected returns exceed $\bar{v}_j$.

Once a task is completed, employers observe the output and leave a public rating on the platform that reveals the worker's type. Following a bad review, low-type veterans will not receive repeated offers in the next period. High-type veterans, in contrast, will be hired again at $\bar{w}_j \in \{c_H, v_H\}$. The worker's problem is therefore to set wages in the novice period to maximize expected lifetime earnings on the platform:

$$\max_{w_{ij}} (w_{ij} - c_i) + \beta(\bar{w}_j - c_i) \times 1[\theta_i = H]$$

The timeline of events within each period is as follows:

(i) A new generation of workers and employers joins the platform. Each employer posts a job with budget $\bar{w}_j$.

(ii) Workers observe $\bar{w}_j$ and decide whether to apply by proposing a wage w_{ij}.

(iii) Employers observe the proposed wage w_{ij} and public reviews r_i if they exist, and possibly incur s_j to screen inexperienced candidates. They make at most one offer to the candidate that generates the highest return.

(iv) Hired workers produce the output and their type is revealed.

(v) All employers and workers in their second period exit the market.

A.2.2 Market Equilibrium

We examine labor market outcomes under the three scenarios: (i) the benchmark case as in Tirole (1988), (ii) novices having incomplete information about own ability, (iii) novices believing that employers interpret low wage offers as signals for low quality for novices. The experimental treatments are designed to shift employers and workers across these scenarios.

Proposition 1. *In the benchmark case, a perfect Bayesian equilibrium exists when returns to reputation for high-types exceed the difference in outside option between types: $\beta(\bar{w}_j - c_H) \geq c_H - c_L$. High-type novices propose wages below the job budget at $w_{Hj} = c_L < v_H$ in period 1, while low-types propose $\bar{w}$. Employers interpret low wages as credible quality signals and hire all high-type novices and veterans in* *equilibrium.*

Proof. Even though employers cannot observe the type of novice workers prior to hiring, high-type novices may distinguish themselves from low-type peers by proposing lower wages in period 1 as an investment in their reputation to generate repeated hiring on the platform. That is, high-type novices offer wages w_{Hj} to maximize their lifetime earnings:

$$\max_{w_{Hj}} \Pi_H(w_{Hj}) = (w_{Hj} - c_H) + \beta(\bar{w}_j - c_H)$$

The optimization is subject to the following constraints:

$$\begin{cases} \Pi_L(w^*_{Hj}) = (w^*_{Hj} - c_L) \leq 0 & \text{Incentive Compatibility (IC)} \\ \Pi_H(w^*_{Hj}) = (w^*_{Hj} - c_H) + \beta(\bar{w}_j - c_H) \geq 0 & \text{Participation Constraint (IR)} \end{cases}$$

High-type novices set wages equal to the low-type's outside option to ensure that at w^*_{Hj}, low-type novices will incur a loss in lifetime earnings, but not themselves. Together, these two constraints imply that wage separation by types can be sustained if for high-type novice workers, returns to reputation for high-type novices exceeds the cost differential between types: $\beta(\bar{w}_j - c_H) \geq c_H - c_L$.

Consistent with this wage-setting strategy, employers can infer high quality from low wages proposed by novices such that $\gamma'(w_{ij}) < 0$. No costly screening is required. Hence, in equilibrium, employers hire all high-type novices at $w^*_{Hj} = c_L$ and veterans.

Proposition 2. *When workers have limited information on own type, there is no complete separation in proposed wages by unobserved type. Employers rely on costly screening to distinguish among novice workers. In equilibrium, novices proposing* $w_{ij} < \lambda_{ij} v_H + (1 - \lambda_{ij}) v_L - s_j$ *and high-type veterans are hired. Compared to the benchmark case, fewer novices are employed and average output quality is lower.*

Proof. Suppose neither workers nor employers know worker type upon entering the market. Workers believe that they belong to the high type with probability μ_i and the low type with probability $1 - \mu_i$. Once hired, they learn their true type by observing public reviews.

Under imperfect information about their own types, novice workers propose wages w_{ij} to maximize their expected lifetime earnings given their beliefs μ_i:

$$\max_{w_{ij}} \Pi(w_{ij}, \mu_i) = w_{ij} - (\mu_i c_H + (1 - \mu_i) c_L) + \beta \cdot \mu_i (\bar{w}_j - c_H)$$

In this case, proposing lower wages no longer signals true high ability but rather reflects novice workers' belief about being the high-type. Consequently, employers no longer view proposed wages as credible signals of novices' ability and therefore incur costly screening to discern types through the additional signal λ_{ij}. Novice workers may still be hired if they can compensate for the employer's screening cost by offering lower wages. That is

$$\lambda_{ij} v_H + (1 - \lambda_{ij}) v_L - w_{ij} - s_j \geq 0$$

However, due to the additional screening costs, employers hire fewer novices in equilibrium and average output quality is lower than the benchmark case.

Proposition 3. *When employers interpret low wage offers by novices as signals of low quality, labor demand is more responsive to undercutting by veterans than by novices. If quality concerns dominate cost considerations in hiring decisions, novice workers pool at the job budget* $\bar{w}_j < \lambda_{ij} v_H + (1 - \lambda_{ij}) v_L - s_j$ *and high-type veterans with* $\bar{w}_j \in \{c_H, v_H\}$ *are hired. Compared to the benchmark case, hiring novices becomes more costly and fewer are employed.*

Proof. Now suppose employers interpret higher proposed wages as signals of higher quality, even in the absence of other credible signals: $\gamma(w_{ij})$ is increasing in w_{ij}.

Under this condition, demand for novices is less responsive to proposed wages than demand for veterans. If workers offer lower wages, this reduces hiring costs for both novices and veterans, but also reduces perceived returns for novices only. Whether proposing lower wages improves novices' hiring chances depends on the relative strength of the cost and perceived quality channels.

When the cost channel dominates, high-type novice workers can still increase their hiring probability by proposing lower wages as in the standard case. Once employers hire novice workers who undercut and discover their type, they can update their beliefs about $\gamma(w_{ij})$ and thus, employers' perception will be corrected over time.

When quality concerns dominate, novice workers avoid undercutting due to concerns about negative signals and propose wages equal to the job budget, $w_{ij} = \bar{w}_j$. Without variation in proposed wages, employers need to incur screening cost s_j to obtain additional quality cues λ_{ij}. Since novice workers with unknown type charge the same as veteran workers in this case, employers will always prefer the experienced ones. Nevertheless, novices may be hired if the expected returns can compensate for employers' screening costs. That is

$$\lambda_{ij} v_H + (1 - \lambda_{ij}) v_L - \bar{w}_j - s_j \geq 0$$

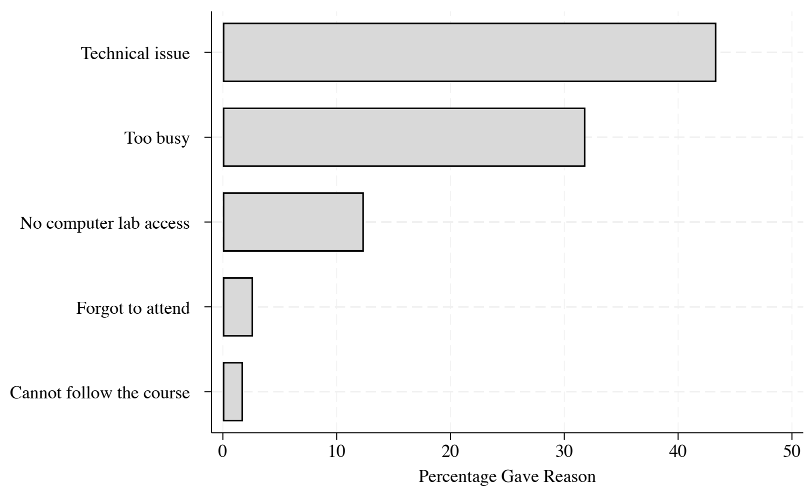
As employers do not observe proposed wages below $\bar{w}_j$, they cannot update their beliefs about $\gamma(w_{ij})$ over time. In equilibrium, employers hire fewer novices and average output quality is lower than the benchmark case.

APPENDIX B

INTRA-HOUSEHOLD INCENTIVE DESIGN: AN EXPERIMENT ON PARENT-CHILD DECISION MAKING DYNAMICS IN PAKISTAN

B.1 Endline Results

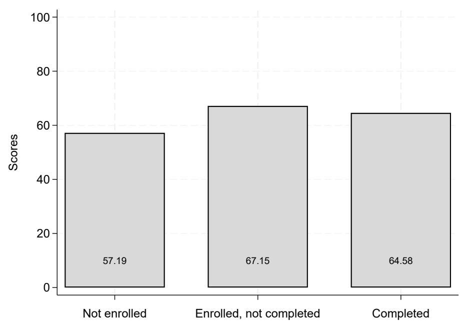
Figure B.1: Reasons for Drop-out



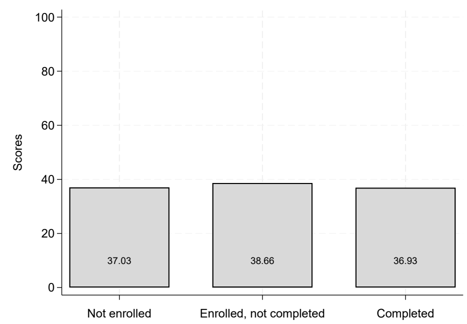
Notes: This figure shows the top reasons for program drop-out given by 113 students during the endline survey. *Technical issue* means that students encountered some technical difficulties when enrolling in the online course or submitting their assignment on the Coursera webpage. *Too busy* means that students were pre-occupied with other responsibilities. *No computer lab access* means that students couldn't use computers at school to attend the course. *Forgot to enroll* means that students forgot to attend the enrollment session or the training sessions at school. *Cannot follow course* means that students found the course difficult to follow.

Figure B.2: Endline Computer Scores

Panel A. Self-Assessment Test

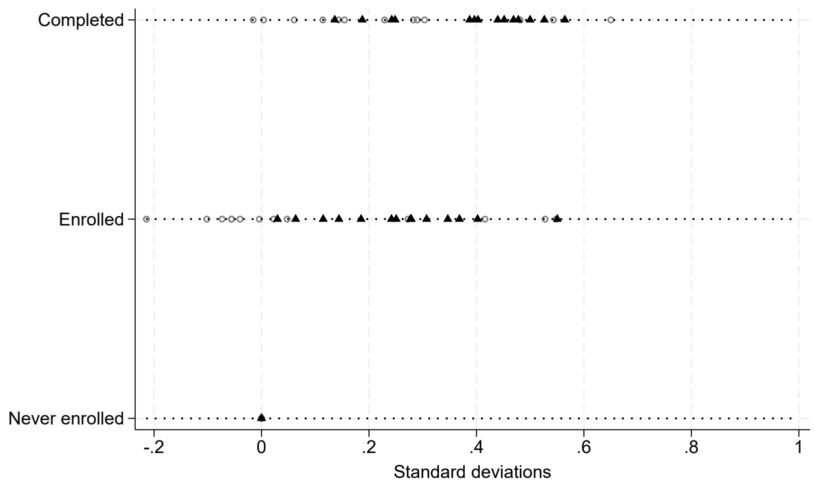


Panel B. Quiz



Notes: This figure shows students' computer skills measured at the endline. We conducted a self-assessment test for 171 students who participated in the in-person endline. Panel A reports average scores from the self-assessment test by program status. *Not enrolled* refers to students from the baseline sample who did not have the parental consent and those who had the consent but did not enroll (N = 86). *Enrolled, not completed* refers to students who started the course but did not complete after 6 weeks (N = 26). *Completed* refers to students who completed the course in time (N = 59). Panel B reports average scores from a short quiz conducted for all 291 students who participated in the endline in person or over phone. Among 291 students, 164 never enrolled in the program, 43 enrolled but did not complete, and 84 completed the program.

Figure B.3: Essential vs non-essential digital skills at endline



Notes: This figure shows differences in self-reported digital literacy skills by 171 students during the endline survey. Responses are converted to standard deviations, within each treatment arm, relative to the students who never enrolled in the training. Triangles represent essential digital skills that are absolutely required for any type of white collar work. Circles represent intermediate or more niche digital skills that may not be necessary for every type of white collar job. Students report their level of confidence from 1 to 5. 1 is "I have never done this before". 2 is "I could do this with help". 3 is "I can do this alone, but might make some mistakes". 4 is "I can do this alone with confidence". 5 is "I can teach others to do this task". From Appendix Table B.1, essential skills are categorized as "Cut and paste text within a document", "Change font size, style, and color", "Create bulleted or numbered lists", "Create a new spreadsheet", "Use math functions such as sum and mean", "Open and print an email attachment", "Use cc and bcc to manage email recipients", "Use Calendar to assist in time management", "Use a search engine like google", "Save a file and locate that file", "Search and find a missing file", "Create new folders", correctly choosing a browser from the list, and correctly identifying a computer folder. All other tasks from the computer test and Appendix Table B.1 are deemed intermediate digital skills.

Table B.1: Self-Reported ability to complete computer tasks

	Not comfortable with this task = 1		
	Not Enrolled	Enrolled, Not Completed	Completed
	(1)	(2)	(3)
Word Processing			
Cut and paste text within a document	47.67	30.77	22.03
Change font size, style, and color	46.51	34.62	22.03
Create bulleted or numbered lists	55.81	42.31	37.29
Create a hyperlink	83.72	76.92	69.49
Excel Spreadsheets			
Create a new spreadsheet	59.30	53.85	42.37
Use math functions such as sum and mean	53.49	38.46	35.59
Create a graph and adjust the properties	60.47	65.38	52.54
Email			
Open and print an email attachment	54.65	46.15	37.29
Use cc and bcc to manage email recipients	87.21	76.92	79.66
Use Calendar to assist in time management	55.81	50.00	50.85
Internet			
Use a search engine like google	36.05	30.77	28.81
Create a blog	76.74	61.54	67.80
Create a website	82.56	76.92	81.36
Insert an audio file, image, video, and podcast onto a website	69.77	53.85	49.15
File Management			
Save a file and locate that file	55.81	38.46	32.20
Search and find a missing file	52.33	34.62	35.59
Create new folders	38.37	30.77	20.34
Media Files			
Create and edit an audio recording	36.05	38.46	32.20
Create and edit a video recording	37.21	38.46	28.81
Upload and download a video from a website	29.07	46.15	20.34
Observations	86	26	59

Notes: This table reports people self-reported ability to complete various computer tasks. The self-reported measures were collected through the in-person endline survey of 171 students. The original measure was on a scale of 1 to 5. 1 means I have never done this before, 2 I could do this with help, 3 I can do this alone, but might make some mistakes, 4 I can do this alone with confidence, and 5 I can teach others to do this task. A dummy variable was generated which would take the value of 1 if the original scale was less than 4. A lower number is interpreted as having more digital skills.

Table B.2: Attrition by program choice

Program Choice	Baseline	Enrolled	Completed	In-person endline
	(1)	(2)	(3)	(4)
Better business writing in English	97	39	28	32
Fundamentals of graphic design	175	66	36	62
Foundations of digital marketing & e-commerce	85	38	22	28
Intuit bookkeeping	7	2	2	2
Work smarter with Microsoft excel	90	34	24	32
How to create a website	35	17	7	15
Total	489	196	119	171

Notes: This table reports the number of students at each stage of the study by program choice.

B.2 Non-reponses from parents

B.2.1 Attrition on the parents survey

One of the main objectives of our experiment design was to look for evidence against efficient bargaining between parents and daughters. Tests against efficiency depend on both differences in parental consent and girls' drop-out. To properly measure both of these outcomes, we can only consider households where we have both a completed parent's survey, as we are only able to request parental permission from parents who complete the parent survey.

While it is possible to interpret all parents who did not complete the parents survey as refusing consent, it is impossible to interpret the daughters drop-out decision under parent survey attrition. Girls cannot drop out if their parents have never been texted. We proceed by showing that there is little evidence of selective attrition on the parents survey.

Argument 1: We find very little attrition as a whole, meaning any differences across group may be subject to finite sample bias.

First, we have high completion rates on the parents survey. Figure 2.2 shows that 416 parents completed the parent survey, of the 488 girls who completed baseline and were still interested in the training after hearing about their contract. That is a completion rate of 85% on an un-incentivized phone survey.

We should be extremely hesitant in interpreting differential attrition by contract group due to small sample bias as there are only 72 households who did not complete the parents survey and there are 6 contract groups. Appendix Table B.3 shows that of the number of households who did not complete the parents survey per group ranges from 5 to 20. Only 20 households have parents answer the phone and explicitly refuse the survey. Refusals by contract group range from 2 to 7 households. We do not observe any pattern of refusals by contract groups.

Table B.3: Reason for non-response on parents survey

	Partial information			Full information		
	0%	50%	100%	0%	50%	100%
	(1)	(2)	(3)	(4)	(5)	(6)
Refused	2	5	2	2	2	7
No answer	2	6	7	5	10	11
Other	1	1	2	3	3	2
Total	5	12	11	10	15	20

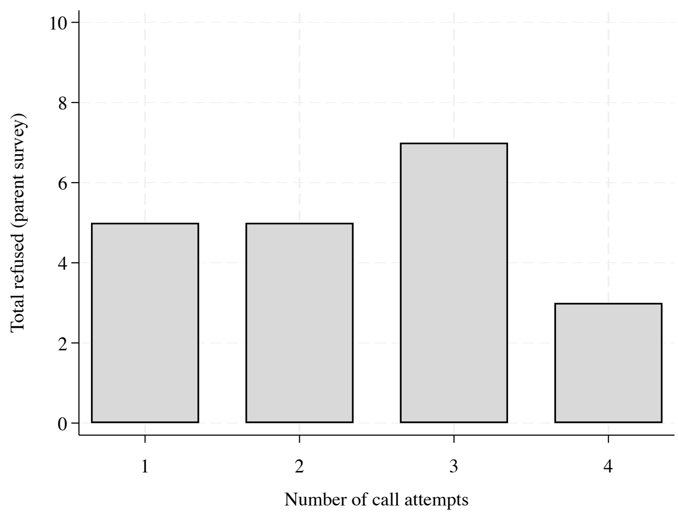
Notes: This table shows parents' reasons for non-response on the phone survey by information treatment (partial or full) and payment targeted to parents (0%, 50%, or 100% of total incentive). The total sample size is 72.

Argument 2: Despite the small sample, we find no interpretable patterns in the data suggesting selective attrition.

We do not have parent surveys on 52 households because (1) phone numbers were disconnected or incorrect or (2) no one answered the phone after 5 attempts. While it is possible that girls gave us incorrect phone numbers, we find it unlikely, as it is common within Pakistan to use prepaid SIM cards which expire, meaning that parent phone numbers are subject to change.

We may be worried that households who did not answer the phone after 5 attempts effectively refused the survey. This would be the case if refusal was correlated with number of attempts. While we may have small-sample bias within refusals, Appendix Figure B.4 shows no clear pattern between refusals and number of call attempts.

Figure B.4: No relationship between probability of refusal and number of attempts



Notes: This figure shows correlation between the number of call attempts and the number of refusals ($N = 20$).

B.2.2 Non-response on permission texts

Our results on parental consent for daughters to participate in the skills training takes the entire set of parents who were texted. This is all 416 parents. We consider only responses of parents who texted back saying “Yes" as having given consent. We count parents who texted back saying “No" or parents who did not text us back within 2 days as having not given permission.

We feel strongly that parents who did not text us back were not interested in the program. First, all parents were texted on the same phone number where they willingly completed an un-incentivized phone survey with an enumerator. They were explicitly told to expect a text from us and given instructions on the requirements to respond within 2 days. Parents who did not receive a text could have tried calling back the enumerators to communicate with the research team that they did not receive a text.

Second, we randomly conducted a back-check survey of households who completed the parents survey and did not text us back. This was to confirm that our enumerators were actually completing the calls, and our messages were not being automatically filtered. Appendix Table B.2 B.4 shows that of the 67 households audited, only 1 parent did not recall receiving a call. Of the parents audited, 64% recall receiving our text message. Of the households who did not receive our text, we see no difference by contract group.

Table B.4: Reason for non-response on permission texts

	Partial information			Full information		
	0%	50%	100%	0%	50%	100%
	(1)	(2)	(3)	(4)	(5)	(6)
Did not receive call	0	1	0	0	0	0
Did not receive text	4	4	3	4	3	6

Notes: This table shows parents' reasons for non-response on permission texts by information treatment (partial or full) and payment targeted to parents (0%, 50%, or 100% of total incentive).

B.3 Alternate specifications

Table B.5: All statistical tests, mapped to models of the household.

A. Unitary household

H1: Household completion is the same across θ or s^p.

Joint test across all 6 groups, which makes no assumptions on collective household model.

(1) $y_h = \beta_0 + \beta_1\{50, \text{Partial}\}_h + \beta_2\{50, \text{Partial}\}_h + \beta_3\{0, \text{Full}\}_h + \beta_4\{50, \text{Full}\}_h + \beta_5\{100, \text{Full}\}_h + \varepsilon_h$

$H_0 : \beta_1 = ... = \beta_5 = 0$ underpowered
H_a : violation of one equality

Classic test of income pooling under information asymmetry*.

(2) $y_h = \delta_0 + \delta_1\{100, \text{Partial}\}_h + \varepsilon_h$ if $\{0, \text{Partial}\}$ or $\{100, \text{Partial}\} = 1$

$H_0 : \delta_1 = 0$ powered
$H_a : \delta_1 \neq 0$

Test of information symmetry**

(3) $y_h = \eta_0 + \eta_1\{\text{Pooled}, \text{Full}\}_h + \varepsilon_h$ if $\{0, \text{Partial}\}$ or $\{s^p, \text{Full}\} = 1$

$H_0 : \eta_1 = 0$ marginal
$H_a : \eta_1 \neq 0$

B. Efficient collective household

H2: Under full information, parental consent does not vary with s^p.

(4) $y_p = \beta_0 + \beta_1\{50, \text{Partial}\}_h + \beta_2\{50, \text{Partial}\}_h + \beta_3\{0, \text{Full}\}_h + \beta_4\{50, \text{Full}\}_h + \beta_5\{100, \text{Full}\}_h + \varepsilon_p$

$H_0 : \beta_3 = \beta_4 = \beta_5$ underpowered
H_a : violation of one equality

H3: Under full information, the girl's decision to drop-out does not vary with s^p.

(5) $y_d = \beta_0 + \beta_1\{50, \text{Partial}\}_h + \beta_2\{50, \text{Partial}\}_h + \beta_3\{0, \text{Full}\}_h + \beta_4\{50, \text{Full}\}_h + \beta_5\{100, \text{Full}\}_h + \varepsilon_d$

$H_0 : \beta_3 = \beta_4 = \beta_5$ underpowered
H_a : violation of one equality

H4: Under full information, household completion does not vary with s^p.

(6) $y_h = \beta_0 + \beta_1\{50, \text{Partial}\}_h + \beta_2\{50, \text{Partial}\}_h + \beta_3\{0, \text{Full}\}_h + \beta_4\{50, \text{Full}\}_h + \beta_5\{100, \text{Full}\}_h + \varepsilon_h$

$H_0 : \beta_3 = \beta_4 = \beta_5$ underpowered
H_a : violation of one equality

Notes: *See Section 2.2, Remark 3 for discussion around the efficient collective household model and why we must evaluate the unitary model under information asymmetry. **See Section 2.2 Subsection 2.2.2 for discussion of why under the efficient collective household, we want to pool all treatment groups with full information.

Table B.6: Unitary household tests, robust to covariates

Dependent variable:	Completed (=1)			
Sample:	Full Sample		Parent Surveyed	
	(1)	(2)	(3)	(4)
50%, Partial (β_1)	0.099	0.10	0.13*	0.12
	[0.065]	[0.071]	[0.073]	[0.079]
100%, Partial (β_2)	0.15**	0.16**	0.18**	0.18**
	[0.066]	[0.067]	[0.072]	[0.073]
0%, Full (β_3)	0.084	0.10	0.11	0.12*
	[0.065]	[0.064]	[0.072]	[0.071]
50%, Full (β_4)	0.072	0.081	0.12	0.13
	[0.067]	[0.069]	[0.079]	[0.082]
100%, Full (β_5)	0.075	0.073	0.13*	0.11
	[0.062]	[0.067]	[0.071]	[0.077]
p-values: Appendix Table B.5				
(1) Joint test				
Boostrap	0.35	0.27	0.20	0.23
Fisher's exact	0.37	0.78	0.21	0.31
(2) Income pooling				
Boostrap	0.02	0.01	0.01	0.01
Fisher's exact	0.03	0.02	0.01	0.01
(3) Full information				
Boostrap	0.14	0.14	0.04	0.06
Fisher's exact	0.14	0.13	0.04	0.05
Strata FE		×		×
Baseline covariates		×		×
Control mean	0.16	0.16	0.16	0.16
Households (N)	489	489	416	416

Notes: This table shows the effect of payment targeting and information symmetry on household completion in the full analysis sample ($N = 489$) and the parent survey sample ($N = 416$). Columns (1) and (2) report results from Regression 1 of Appendix Table B.5. Columns (2) and (4) include strata fixed effects and baseline covariates for whether the student is in high school, in the last year of school, household size, working status of the father and mother, high school graduation status of the father and mother, if there is a computer at home, if the student has their own cell phone, and if the student plans to work. The second panel reports p-values from a joint test of null effect across all treatment groups, the income pooling test, and the information symmetry test from panel A of Appendix Table B.5. Bootstrap results are generated from 10,000 draws. Fisher's exact p-values are generated from 10,000 re-randomized treatment groups with stratification.

Table B.7: Efficient collective household, robust to covariates

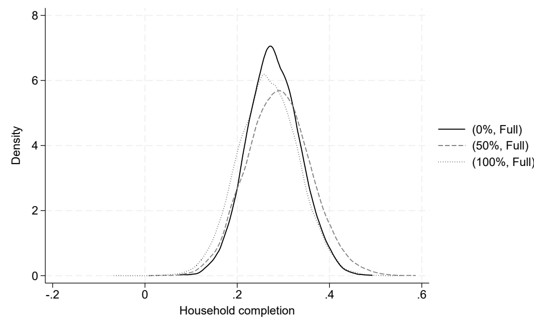
Dependent variable:	Completed (=1)		Parent consented (=1)		Girl's drop-out (=1)	
Sample:	Parent Surveyed				Parent Consented	
	(1)	(2)	(3)	(4)	(5)	(6)
50%, Partial (β_1)	0.13*	0.12	0.080	0.086	-0.12	-0.090
	[0.073]	[0.079]	[0.077]	[0.079]	[0.093]	[0.100]
100%, Partial (β_2)	0.18**	0.18**	0.072	0.086	-0.22**	-0.20**
	[0.072]	[0.073]	[0.076]	[0.078]	[0.092]	[0.094]
0%, Full (β_3)	0.11	0.12*	0.090	0.12	-0.11	-0.10
	[0.072]	[0.071]	[0.076]	[0.078]	[0.093]	[0.091]
50%, Full (β_4)	0.12	0.13	0.053	0.067	-0.15	-0.15
	[0.079]	[0.082]	[0.084]	[0.089]	[0.10]	[0.10]
100%, Full (β_5)	0.13*	0.11	0.083	0.086	-0.14	-0.12
	[0.071]	[0.077]	[0.077]	[0.078]	[0.092]	[0.098]
p-values: $\beta_3 = \beta_4 = \beta_5$						
Bootstrap	0.96	0.96	0.89	0.82	0.90	0.88
Fisher's exact	0.96	0.95	0.88	0.82	0.91	0.87
Strata FE		×		×		×
Baseline covariates		×		×		×
Control mean	0.16	0.16	0.69	0.69	0.75	0.75
Households (N)	416	416	416	416	313	313

Notes: This table shows the effect of payment targeting and information symmetry on household completion and parental consent in the parent survey sample ($N = 416$) as well as girls' drop-out in the parental consent sample ($N = 313$). Columns (1), (3), and (5) report results from Regressions 4, 5, and 6 of Appendix Table B.5, respectively. Columns (2), (4), and (6) include strata fixed effects and baseline covariates for whether the student is in high school, in the last year of school, household size, working status of the father and mother, high school graduation status of the father and mother, if there is a computer at home, if the student has their own cell phone, and if the student plans to work. Bootstrapped standard errors in parentheses. The second panel reports p-values from a joint test of equality in the full information group, derived from Panel B of Appendix Table B.5. Bootstrap results are generated from 10,000 draws. Fisher's exact p-values are generated from 10,000 re-randomized treatment groups with stratification.

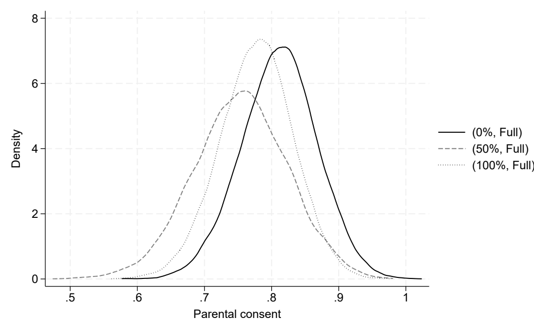
B.4 Null effect under full information

Figure B.5: Boostrapped distribution of group means under full information

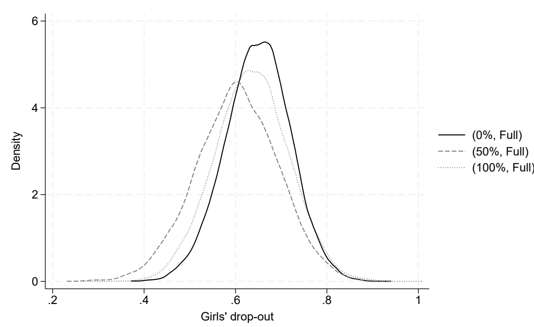
Panel A. Household completion *(N=416)*



Panel B. Parental consent *(N=416)*



Panel C. Girls' drop-out *(N=313)*



Notes: We generate distributions of the probability of household completion, parental consent, and girls' drop-out using 10,000 bootstrapped draws. The regression specification is Regression (1) from Appendix Table B.5 and includes stratum fixed effects and baseline covariates for whether the student is in high school, in the last year of school, household size, working status of the father and mother, high school graduation status of the father and mother, if there is a computer at home, if the student has their own cell phone, and if the student plans to work.

Table B.8: Effect sizes of inefficiencies under the collective household

Paper	Evaluation	Inefficiency size
Udry (1996)	Husband/Wife	6%
Choukhmane et al. (2023)	Husband/Wife	24%
Lowe and McKelway (2024)	Husband/Wife	50%
Schaner (2015)	Husband/Wife	52%
Bursztyn and Coffman (2012)	Parent/Child	6%
Bergman (2021)	Parent/Child	25%
Baird et al. (2011)	Parent/Child	43%
de Walque and Valente (2023)	Parent/Child	54-75%

Notes: This table reports quantified effect sizes of household inefficiency in relevant economics papers. This list is by no means exhaustive.

B.5 Alternative Models

B.5.1 Collective Household Model with Parental Nudging

Following the timeline of household bargaining in Section 2.2, 2.2.2, 2.2.2, after parents give consent, they can nudge daughters to complete the program so that the household receives the 3,000 PKR. We assume that it is cost-less for parents to nudge daughters and they exert more influence on daughtersâĂŹ actions when the public return $(\theta(s-s^p)+s^p)$ is higher. By completing the program, daughters get indirect utility v^d from complying to parents and v^d is scaled by a function of the public return, denoted by $\tau(\theta(s-s^p)+s^p)$. Under full information, the total indirect utility $\tau(s)v^d$ does not vary by s^p. Under partial information, the indirect utility $\tau(\theta(s-s^p)+s^p)v^d$ is increasing in s^p.

We rewrite Equation 2.2 of the girl's decision function for program completion to include the indirect utility term.

$$I^d_{\text{completes}}\{u^d+\tau(\theta(s-s^p)+s^p)v^d+(1-\theta)(s-s^p)+(1-\lambda)(\theta(s-s^p)+s^p)>c^d\} \tag{B.1}$$

We generate a new upper bound on payment sharing by isolating λ from Equation B.1.

$$\lambda < \frac{u^d+\tau(\theta(s-s^p)+s^p)v^d+s-c^d}{\theta(s-s^p)+s^p} = \lambda'_{ub} \tag{B.2}$$

We combine λ_{lb} of Equation 2.3 from the parent's problem and λ'_{ub} of Equation B.2 from the daughter's problem to generate the new support of λ that will incentivize the household to complete the program.

$$\lambda'_{ub} - \lambda_{lb} = \frac{u^d + u^p + \tau(\theta(s - s^p) + s^p)v^d + s - (c^d + c^p)}{\theta(s - s^p) + s^p} \tag{B.3}$$

Under partial information ($\theta \in [0, 1]$), the effect of incentive allocation s^p on completion is ambiguous because both the numerator and the denominator in Equation B.3 are increasing in s^p. When a greater amount of payment is assigned to parents, they exert a stronger influence on daughtersâĂŹ actions to ensure the household can attain the reward. Consequently, daughters gain indirect utility by completing the program but lose own monetary incentive to finish it. Our experimental results imply that the positive effect of parental nudging dominates the negative effect of daughters losing incentive for completion.

Under full information ($\theta = 1$), this model with parental nudging generates the same predictions as the original model.

This adjusted model fits our experimental results on household completion very well, however, we are unable to directly test hypotheses on girls' drop-out decisions under partial information in our data because we only observe their decisions conditional on parental consent. See Remark 1 of Section 2.2.

B.5.2 Transfer Frictions

We consider an alternative model with frictions on the ability to split the stipend according to the sharing rule. Suppose that $\alpha^d, \alpha^p \in [0,1]$ represent daughters' and parents' beliefs about the other agent's credibility of complying to the sharing rule ex-post. Recall that in the friction-less environment, $s^p + \theta(s - s^p)$ is the public return that daughters and parents negotiate on. Given α^d, α^p, the available amount for negotiation from parents' perspective is $s^p + \theta\alpha^d(s - s^p)$ and that for daughters is $\alpha^p s^p + \theta(s - s^p)$. Hence, the new parents' problem is as follows.

$$I^p_{\text{consent}}\{u^p + \lambda(\theta\alpha^d(s - s^p) + s^p) > c^p\}$$

Rearranging inside the indicator function:

$$\lambda > \frac{c^p - u^p}{\theta\alpha^d(s - s^p) + s^p} = \lambda_{lb}$$

Holding θ constant, the lower-bound is decreasing in s^p. Hence, this model predicts that in the full information group, parental consent should be increasing in s^p. Since girls' drop-out decisions and completion are conditional on parental consent which varies with s^p, we cannot derive direct test for these two outcomes using our data.

Our results find no difference in parental consent in the full information groups and hence we reject the model with parental beliefs on transfer frictions.

REFERENCES

Abebe, Girum, A 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," *The Review of Economic Studies*, 2021, *88* (3), 1279–1310.

_ , _ , _ , _ , _ , _ , and Forhad Shilpi, "Matching frictions and distorted beliefs: Evidence from a job fair experiment," *The Economic Journal*, 2025, *135* (671), 2089–2121.

Agrawal, Ajay, Nicola Lacetera, and Elizabeth Lyons, "Does Standardized Information in Online Markets Disproportionately Benefit Job Applicants from Less Developed Countries?," *Journal of Labor Economics*, 2016, *34* (S1), S225–S268.

Alfonsi, Livia, Mary Namubiru, and Sara Spaziani, "Meet Your Future: Experimental Evidence on the Labor Market Effects of Mentors," Technical Report Working Paper No. 87, GÂšLM|LIC, IZA, Bonn 2024.

_ , Oriana Bandiera, Vittorio Bassi, Robin Burgess, Imran Rasul, Munshi Sulaiman, and Anna Vitali, "Tackling youth unemployment: Evidence from a labor market experiment in Uganda," *Econometrica*, 2020, *88* (6), 2369–2414.

Ashraf, Nava, "Spousal Control and Intra-household Decision Making: An Experimental Study in the Philippines," *American Economic Review*, September 2009, *99* (4), 1245–1277.

— **, B. Kelsey Jack, and Emir Kamenica**, "Information and Subsidies: Complements or Substitutes?," *Journal of Economic Behavior & Organization*, 2013, *88*, 133–139.

— **, Erica Field, and Jean Lee**, "Household Bargaining and Excess Fertility: An Experimental Study in Zambia," *American Economic Review*, July 2014, *104* (7), 2210–2237.

Baird, Sarah, Craig McIntosh, and Berk Özler, "Cash or Condition? Evidence from a Cash Transfer Experiment," *The Quarterly Journal of Economics*, November 2011, *126* (4), 1709–1753.

Bandiera, Oriana, Vittorio Bassi, Robin Burgess, Imran Rasul, Munshi Sulaiman, and Anna Vitali, "The search for good jobs: evidence from a six-year field experiment in Uganda," *Journal of labor economics*, 2025, *43* (3), 885–935.

Banerjee, Abhijit and Sandra Sequeira, "Learning by searching: Spatial mismatches and imperfect information in Southern labor markets," *Journal of Development Economics*, 2023, *164*, 103111.

Baptista, Dulce, Richard Freund, and Rafael Novella, "Entrepreneurial skills training for online freelancing: Experimental evidence from Haiti," *Economics Letters*, 2023, *232.*

Barach, Moshe and John J. Horton, "How Do Employers Use Compensation History? Evidence from a Field Experiment," *Journal of Labor Economics*, 2019, *37* (1), S25–S46.

Bassi, Vittorio and Aisha Nansamba, "Screening and signalling non-cognitive skills: experimental evidence from Uganda," *The Economic Journal*, 2022, *132* (642), 471–511.

Becker, Gary S., "A Theory of Social Interactions," *Journal of Political Economy*, November 1974, *82* (6), 1063–1093. Publisher: The University of Chicago Press.

Becker, Gary S, *A treatise on the family: Enlarged edition*, Harvard university press, 1991.

Bergman, Peter, "Parent-Child Information Frictions and Human Capital Investment: Evidence from a Field Experiment," *Journal of Political Economy*, January 2021, *129* (1), 286–322.

Berry, J., "Child Control in Education Decisions: An Evaluation of Targeted Incentives to Learn in India," *Journal of Human Resources*, October 2015, *50* (4), 1051–1080.

Bertrand, Marianne, Bruno CrÃľpon, Stefano Caria, and Natalie Valent, "Improving Job SeekersâĂŹ Employment and Earnings Through Credible Skills Signals," Policy Insight, The Abdul Latif Jameel Poverty Action Lab (J-PAL) 2024. Last updated November 2024.

Breza, Emily and Supreet Kaur, "Labor Markets in Developing Countries," *Annual Review of Economics*, 2025, *17.*

_ , _ , and Nandita Krishnaswamy, "Social norms as a determinant of aggregate labor supply," Technical Report, National Bureau of Economic Research 2019.

_ , _ , and Yogita Shamdasani, "The morale effects of pay inequality," *The Quarterly Journal of Economics*, 2018, *133* (2), 611–663.

Browning, Martin and Pierre-André Chiappori, "Efficient intra-household allocations: A general characterization and empirical tests," *Econometrica*, 1998, pp. 1241–1278.

_ , Pierre-Andre Chiappori, and Valerie Lechene, "Collective and unitary models: A clarification," *Review of Economics of the Household*, 2006, *4*, 5–14.

Bursztyn, Leonardo, Alessandra L. González, and David Yanagizawa-Drott, "Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia," *American Economic Review*, October 2020, *110* (10), 2997–3029.

_ and Lucas C. Coffman, "The Schooling Decision: Family Preferences, Intergenerational Conflict, and Moral Hazard in the Brazilian Favelas," *Journal of Political Economy*, June 2012, *120* (3), 359–397.

Caria, Stefano, Kate Orkin, Alison Andrew, Robert Garlick, Rachel Heath, and Niharika Singh, "Barriers to search and hiring in urban labour markets," *VoxDevLit*, 2024, *10* (1), 1–46.

Carranza, Eliana, Robert Garlick, Kate Orkin, and Neil Rankin, "Job Search and Hiring with Limited Information about Workseekers' Skills," *American Economic Review*, 2022, *112* (11), 3547–3583.

Castilla, Carolina and Thomas Walker, "Is ignorance bliss? The effect of asymmetric information between spouses on intra-household allocations," *American Economic Review*, 2013, *103* (3), 263–268.

Cefala, Luisa, Pedro Naso, Michel Ndayikeza, and Nicholas Swanson, "Under-training by Employers in Informal Labor Markets: Evidence from Burundi," Technical Report, Working paper 2024.

Chakravorty, Bhaskar, Wiji Arulampalam, Apurav Yash Bhatiya, Clément Imbert, and Roland Rathelot, "Can information about jobs improve the effectiveness of vocational training? Experimental evidence from India," *Journal of Development Economics*, 2024, *169*, 103273.

Choukhmane, Taha, Lucas Goodman, and Cormac O'Dea, "Efficiency in Household Decision Making: Evidence from the Retirement Savings of US Couples," Technical Report, National Bureau of Economic Research 2023.

Cui, Jingyi, "Bidding for Reputation," November 2025.

Das, Narayan, Md.R̃aied Arman, Anindita Bhattacharjee, Mohima Gomes, and Sakib Mahmood, "Tackling Women's Unemployment: Evidence from an Online Freelancing Experiment in Bangladesh," Working Paper SSRN 4843957, SSRN Electronic Journal 2024. Posted 27 May 2024.

Datta, Namita, Chen Rong, Sunamika Singh, Clara Stinshoff, Nadina Iacob, Natnael Simachew Nigatu, Mpumelelo Nxumalo, and Luka Kli-

maviciute, "Working Without Borders: The Promise and Peril of Online Gig Work," Report, World Bank, Washington, DC 2023.

de Walque, Damien and Christine Valente, "Incentivizing school attendance in the presence of parent-child information frictions," *American Economic Journal: Economic Policy*, 2023, *15* (3), 256–285.

Dean, Joshua and Seema Jayachandran, "Changing Family Attitudes to Promote Female Employment," *AEA Papers and Proceedings*, May 2019, *109*, 138–142.

Donni, Olivier and José Alberto Molina, "Household collective models: Three decades of theoretical contributions and empirical evidence," 2018.

Donovan, Kevin, Will Jianyu Lu, and Todd Schoellman, "Labor market dynamics and development," *The Quarterly Journal of Economics*, 2023, *138* (4), 2287–2325.

Duflo, Esther, "Grandmothers and granddaughters: old-age pensions and intrahousehold allocation in South Africa," *The World Bank Economic Review*, 2003, *17* (1), 1–25.

_ **and Christopher Udry**, "Intrahousehold Resource Allocation in Cote d'Ivoire: Social Norms, Separate Accounts and Consumption Choices," Technical Report, National Bureau of Economic Research May 2004. Working Paper No.10498.

Erdem, Tülin, Michael P. Keane, and Baohong Sun, "A Dynamic Model of Brand Choice When Price and Advertising Signal Product Quality," *Marketing Science*, 2008, *27* (6), 1111–1125.

Fazio, Maria Victoria, Richard Freund, and Rafael Novella, "Do Entrepreneurial Skills Unlock Opportunities for Online Freelancing? Experimental Evidence from El Salvador," *Journal of Development Economics*, 2025, *172*, 103363.

Hannane, Jonas, "Advertising in online labor markets: A signal of freelancer quality?," 2024.

Heath, Rachel and Xu Tan, "Intrahousehold Bargaining, Female Autonomy, and Labor Supply: Theory and Evidence from India," *Journal of the European Economic Association*, 2020, *18* (4), 1928–1968.

Hilbert, Martin R. and Kangbo Lu, "The Online Job Market Trace in Latin America and the Caribbean," Project Document LC/TS.2020/83, United Nations Economic Commission for Latin America and the Caribbean (ECLAC), Santiago 2020.

Huang, Yangguang, Chenyang Li, and Si Zuo, "Price signaling and reputation building: evidence from a service platform," *HKUST Business School Research Paper*, 2025, (2022-062).

International Labour Organization, "World Employment and Social Outlook 2021: The Role of Digital Labour Platforms in Transforming the World of Work," Flagship Report, International Labour Organization, Geneva 2021.

Jayachandran, Seema, "Social Norms as a Barrier to WomenâĂŹs Employment in Developing Countries," *IMF Economic Review*, September 2021, *69* (3), 576–595.

Judd, Kenneth L and Michael H Riordan, "Price and quality in a new product monopoly," *The Review of Economic Studies*, 1994, *61* (4), 773–789.

K, Varun, "Negotiations and Undercutting in Low Wage Labor Markets," Technical Report, Working paper 2025.

Kala, Namrata and Elizabeth Lyons, "The Effects of Digital Surveillance and Managerial Clarity on Performance," Technical Report, National Bureau of Economic Research 2025.

Kässi, Otto and Vili Lehdonvirta, "Do Microcredentials Help New Workers Enter the Market? Evidence from an Online Labor Platform," *Journal of Human Resources*, 2022.

Kelley, Erin M, Christopher Ksoll, and Jeremy Magruder, "How do digital platforms affect employment and job search? Evidence from India," *Journal of Development Economics*, 2024, *166*, 103176.

Kiss, Andrea, Robert Garlick, Kate Orkin, and Lukas Hensel, "Jobseekers' Beliefs about Comparative Advantage and (Mis)Directed Search," Working Paper CSAE-WPS-2023-11, Centre for the Study of African Economies (CSAE), University of Oxford 2023.

Levitt, Steven, John List, and Sally Sadoff, "The Effect of Performance-Based Incentives on Educational Achievement: Evidence from a Randomized Experiment," Technical Report, NBER Working Paper 22107 2016.

Lowe, Matt and Madeline McKelway, "Coupling Labor Supply Decisions: An Experiment in India," Technical Report, Working Paper 2024.

McKelway, Madeline, "WomenâĂŹs Employment in India: Intra-Household and Intra-Personal Constraints," Technical Report, Working Paper 2021.

Mel, Suresh De, David McKenzie, and Christopher Woodruff, "Are Women More Credit Constrained? Experimental Evidence on Gender and Microenterprise Returns," *American Economic Journal: Applied Economics*, 2009, *1* (3), 1–32.

Milgrom, Paul and John Roberts, "Price and advertising signals of product quality," *Journal of political economy*, 1986, *94* (4), 796–821.

Munro, Alistair, "Intra-Household Experiments: A Survey," *Journal of Economic Surveys*, 2018, *32* (1), 134–175. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/joes.12196.

Ndayikeza, Michel Armel, "Underemployment of College Graduates: Is Doing Anything Better Than Doing Nothing?," *Journal of Development Economics*, 2025, *174*, 103459.

Online Labour Observatory, "Online Labour Observatory," http://onlinelabourobservatory.org/ 2025. Accessed: 2025-09-18.

Pakistan Today, "Senate passes resolution for equipping women, girls with digital skills," 2023. https://www.pakistantoday.com.pk/2023/04/03/senate-passes-resolution-for-equipping-women-girls-with-digital-skills/, retrieved June 2023.

Pallais, Amanda, "Inefficient Hiring in Entry-Level Labor Markets," *American Economic Review*, 2014, *104* (11), 3565–3599.

Rao, Akshay R. and Kent B. Monroe, "The Effect of Price, Brand Name, and Store Name on BuyersâĂŹ Perceptions of Product Quality: An Integrative Review," *Journal of Marketing Research*, 1989, *26* (3), 351–358.

Roussille, Nina, "The Role of the Ask Gap in Gender Pay Inequality," *Quarterly Journal of Economics*, 2024, *139* (3), 1557–1610.

Samuelson, Paul A., "Social Indifference Curves," *The Quarterly Journal of Economics*, February 1956, *70* (1), 1–22.

Schaner, Simone, "Do opposites detract? Intrahousehold preference heterogeneity and inefficient strategic savings," *American Economic Journal: Applied Economics*, 2015, *7* (2), 135–174.

Stiglitz, Joseph E., "Imperfect Information in the Product Market," in Richard Schmalensee and Robert D. Willig, eds., *Handbook of Industrial Organization*, Vol. 1, Elsevier, 1989, chapter 13, pp. 769–847.

Tirole, Jean, *The Theory of Industrial Organization*, Cambridge, MA: MIT Press, 1988.

Udry, Christopher, "Gender, agricultural production, and the theory of the household," *Journal of political Economy*, 1996, *104* (5), 1010–1046.

UNICEF, "Bridging the Gender Digital Divide: Challenges and an Urgent Call for Action for Equitable Digital Skills Development," 2023. https://data.unicef.org/resources/ictgenderdivide/, retrieved June 2023.

Wathieu, Luc and Marco Bertini, "Price as a Stimulus to Think: The Case for Willful Overpricing," *Marketing Science*, 2007, *26* (1), 118–129.

Wolinsky, Asher, "Prices as signals of product quality," *The review of economic studies*, 1983, *50* (4), 647–658.

World Bank, "Inflation, consumer prices (annual %) - Pakistan," 2022. https://data.worldbank.org/indicator/FP.CPI.TOTL.ZG?locations=PK, retrieved June 2023.

World Skills Clock, "World Skills Clock Pakistan," 2023. https://skillsclock.io/, retrieved June 2023.