# Why Don't Jobseekers Search More? Barriers and Returns to Search on a Job Matching Platform*

Kate Vyborny[†] ⓡ Robert Garlick[‡] ⓡ Nivedhitha Subramanian[§] ⓡ Erica Field[¶]

November 11, 2023

**Abstract**

Understanding barriers to job search and returns to relaxing these barriers is important for economists and policymakers. An experiment that changes the default process for initiating job applications increases applications by 600% on a search platform in Pakistan. The marginal treatment-induced applications have approximately constant rather than decreasing returns, contrasting with intuitive job search models. These results are consistent with a model in which some jobseekers miss some high-return vacancies due to psychological costs of initiating applications. The finding of constant returns to marginal applications, combined with limited spillovers onto other jobseekers, raises the possibility of suboptimally low search effort.

*We are grateful for valuable feedback from seminar participants at Bowdoin, Duke, Maine, Oxford, and Virginia; from conference participants at AEA, AFE, CREB ADE, GLM-LIC, IZA Firm-Worker Matching, MIEDC, PacDev, SOLE, and WB Jobs and Development; and from Martin Abel, Michele Belot, Jenn Kades, Jeremy Magruder, Simon Quinn, Todd Sarver, Jeffrey Smith, Jessi Streib, Duncan Thomas, and Chris Woodruff. We thank Gustavo Acosta, Leonor Castro, Aiman Farrukh and Harmalah Khan for outstanding research assistance in preparing this draft. We thank the entire Job Talash team at CERP, in particular Maniha Aftab, Tehneiat Amjad Butt, Noor Anwar Ali, Rukhshan Arif Mian, Sarah Hussain, Hareem Fatima, Uzair Junaid, Lala Rukh Khan, Alieha Shahid, Mahin Tariq and Zoha Waqar for excellent research assistance and programmatic support over the course of this experiment. This study is made possible by the generous support of the American people through the United States Agency for International Development (USAID) through Development Innovation Ventures (DIV). We gratefully acknowledge funding for this project from J-PAL GEA. We are also thankful for funding support for the broader Job Talash research program from the Asian Development Bank, GLM-LIC, 3IE, IGC, IPA, J-PAL, PEDL, and the National Science Foundation (SES #1629317). This research received ethics approval from Duke University (#2019-0067). The experimental design, intervention, and primary outcomes are registered on the AEA's RCT registry at https://doi.org/10.1257/rct.5997-1.0. The authors' views do not necessarily represent the views of USAID or the US government.

[†]World Bank, kvyborny@worldbank.org

[‡]Duke University, robert.garlick@duke.edu

[§]Bates College, nsubrama@bates.edu

[¶]Duke University, BREAD, NBER, field.erica@duke.edu

# 1 Introduction

Job search is a central feature of labor markets, and search frictions can have important economic consequences. For instance, in macroeconomic models, frictional search can help to explain both employment levels and the productivity of firm-worker matches (Shimer, 2010). Microeconomic research has documented many specific job search frictions ranging from pecuniary search costs to incomplete information (e.g. Abebe et al. 2021a,b; Abel et al. 2019; Bandiera et al. 2021; Belot et al. 2018; Franklin 2017). Recent work has shown that behavioral factors such as present bias, reference-dependence, and motivated reasoning can also impact search effort (e.g. Cooper & Kuhn 2020; DellaVigna et al. 2022; Mueller & Spinnewijn 2022).

We study behavioral barriers to job search effort on a search and matching platform. The platform sends text messages about relevant new vacancies to jobseekers, who must call the platform to apply. Adding follow-up calls inviting jobseekers to immediately start applications, which reduces the initiative required to apply, substantially increases their propensity to apply. Moreover, returns to the additional applications are approximately constant rather than decreasing, raising the question of why jobseekers don't apply more in the absence of calls. To explain this, we propose a model with heterogeneous *psychological costs of initiating applications* that can be high enough to deter some applications to even high-return vacancies, resulting in suboptimally low search effort.

To generate experimental evidence on this search barrier, we work with a novel job search platform in Lahore, Pakistan.[1] Platform data allow us to observe all vacancy characteristics, job application decisions, application materials, and interview outcomes for roughly 1.1 million matches between vacancies and jobseekers. The 9,800 jobseekers are recruited from a city-wide representative household listing, and thus have a wide range of education, from incomplete primary school to graduate degrees, and a wide range of baseline labor force attachment, from employed and searching to non-employed and non-searching. Using the platform requires only basic literacy, a simple phone, and almost no airtime, generating very few technological and pecuniary barriers to search. This sample breadth is unusual in experimental job search studies (Poverty Action Lab, 2022).

Our main experimental treatment changes how jobseekers initiate applications on the platform, moving them from an active role to a passive role. Specifically, all users receive monthly text messages listing new vacancies that match their education, experience, and occupational preferences.[2] Users in the control group must call the platform to initiate applications. Users in the treatment group also receive a call inviting them to apply, so they do not need to initiate calls to apply. The experimental design holds constant many other parts of the search process: the phone call treatment has negligible effects on pecuniary and time costs of applying, provides no direct encouragement

[1] Job platforms have become a central feature of many labor markets. In Pakistan in 2021, Rozee, LinkedIn, and Bayt had respectively 9.5, 7.5, and 3 million users. Bayt had 39 million users in 2021 across the Middle East, North Africa, and South Asia. LinkedIn had $>10$ million users in 2022 in at least 8 developed and 10 developing countries.

[2] These matches are determined by information jobseekers report at sign-up, before treatment assignment.

or pressure to apply, and provides no extra information about the vacancies. Hence, we interpret the phone call treatment as a pure reduction in the psychological cost of initiating applications.

Our two key findings are that the phone call treatment dramatically increases the job application rate, and that the average return to the additional applications is roughly constant rather than decreasing. Treatment increases the share of applications to jobseeker-vacancy matches from 0.2 to 1.5%.[3] In terms of returns, using treatment as an instrument for applying shows that marginal treatment-induced applications have a 5.9% probability of yielding interviews. This is neither substantively nor statistically different from the 6.3% probability for applications from the control group, which are inframarginal to treatment by definition. This implies that returns to job search are roughly constant over this large increase in applications. The same pattern holds when we weight interviews by their desirability in terms of salary, hours, commuting, and non-salary benefits. We run an extra experiment to show that this finding is not explained by differences in the quality of jobseekers who submit marginal versus inframarginal applications. We also develop tests to show that the constant returns finding is robust to potential violations of the exclusion and monotonicity conditions in our instrumental variables analysis.

The finding of roughly constant returns clashes with an intuitive model in which jobseekers prioritize applying to vacancies with the highest combination of expected interview probabilities and desirable attributes, which implies that additional applications would have decreasing returns. Similar models are common in the 'directed' job search literature (Wright et al., 2021). Constant returns are consistent with canonical models of 'random' job search, where vacancies are homogeneous and jobseekers randomly choose where to apply (Pissarides, 2000). But random search models do not match other results we find: jobseekers on our platform do direct applications to vacancies with desirable attributes like higher salaries, and we run an additional experiment to encourage random search that generates sharply *decreasing* returns to marginal applications.

We present a simple model of job search to explain this pattern. Each jobseeker receives matches from the platform each month and applies to those with expected present return above the cost of applying. The model's key assumption is that the cost of applying includes a psychological cost of initiating applications that varies across jobseekers or through time, the same approach used in some behavioral investment models (Carroll et al., 2009; Duflo et al., 2011). Although we cannot pin down the exact source of this cost, the existing literature suggests multiple candidates, including the cost of paying attention to text messages and mentally processing their content (Gabaix, 2019), fear of applications being rejected (Köszegi et al., 2022), and present bias (Ericson & Laibson, 2019).

While we cannot firmly establish that this is the only possible explanation, we show that this model explains our findings better than plausible alternatives. First, the heterogeneous cost as-

[3] The control group's application rate is comparable to some other platforms that have been studied (Appendix A).

sumption explains a key pattern we see in the control group: some jobseekers in some months submit no applications, and miss even high-return matches. Second, the model explains why treatment increases job applications: switching communication with the platform from an active to a passive decision lowers the psychological cost of initiating applications, leading to more applications. This echoes research showing that eliminating the need to initiate decisions can increase financial and health investment (DellaVigna, 2009; DellaVigna & Malmendier, 2006; Madrian & Shea, 2001; Thaler & Benartzi, 2004). Third, the heterogeneous cost assumption explains how returns to inframarginal and treatment-induced marginal applications can be equal. Inframarginal applications come only from jobseekers facing low costs. Marginal applications come from two sources. Jobseekers facing already low costs apply to additional vacancies, which will have lower average returns than their inframarginal applications. And jobseekers facing high costs - who would not have applied to any vacancies in that month without treatment - now apply to some vacancies. The second type of marginal applications can have higher returns than inframarginal applications. Hence the return to marginal applications, averaged across treated jobseekers facing higher and lower costs, can equal the return to inframarginal applications.

We can also test and reject many alternative interpretations for our treatment effects. Perhaps most importantly, additional experiments show that directly reducing the pecuniary or time costs of applying has minimal effects on application behavior, validating our interpretation of treatment as reducing the psychological cost of initiating applications. Some specific behavioral explanations – encouragement, pressure, or reminders – are inconsistent with the platform design and results from additional experiments. Information- or belief-based explanations – e.g. more information about matches or higher perceived returns to applications – are also inconsistent with the platform design, results from additional experiments, and survey measures of beliefs.

Importantly, we do not find evidence that this additional search has negative spillovers on other jobseekers. We treat 50% of jobseekers on the platform, increasing total search enough that large spillovers are possible. Instead, we find that individual jobseekers' interview probabilities are unaffected by competing against more treatment-induced applications. We also find that jobseekers' treatment status does not affect their off-platform search behavior.

On the platform, search outcomes are measured using interviews and interview attributes. Interviews are an important search outcome because they are a necessary condition for job offers and impose non-trivial costs on both job applicants and firms. Hence their widespread use in some areas of labor economics such as audit studies. We are not powered to study employment effects at the scale of this experiment. Using interviews or even applications as final outcomes is relatively common in the literature studying search on platforms (e.g. Belot et al. 2018), either for power reasons or because many platforms do not track data on job offers or employment.[4]

[4] Banfi et al. (2019), Belot et al. (2022b), Faberman & Kudlyak (2019), He et al. (2021), Marinescu (2017a),

Our paper makes three contributions. First, by studying psychological job search costs, we contribute to a nascent literature on behavioral job search, reviewed by Cooper & Kuhn (2020). Existing work shows patterns of job search consistent with present bias, motivated reasoning, and reference dependence (DellaVigna & Paserman, 2005; DellaVigna et al., 2017, 2022; Mueller & Spinnewijn, 2022; Paserman, 2008), but does not isolate psychological costs of initiation.[5]

Second, our results have clear and novel policy implications for addressing behavioral barriers to search. Babcock et al. (2012) suggest multiple ways to harness behavioral economics to encourage and improve job search. However, evaluations of policies designed to directly target behavioral factors are rare, only evaluate interventions that help jobseekers to plan, and do not compare returns to marginal and inframarginal search (Abel et al., 2019; Caria et al., 2023; Sanders et al., 2019). We extend this work by running multiple field experiments to show how small, theory-informed policy changes to the job search environment can increase search without lowering returns, while other, only slightly different policy changes either do not increase search or produce decreasing returns. Many other job search policies might also work through behavioral channels: motivated reasoning might affect how jobseekers process and use new information, present bias and reference dependence might influence how jobseekers spend subsidies, and relationships between caseworkers and jobseekers might have behavioral components.[6] However, research into these forms of job search assistance has not sought to pin down behavioral components.

Third, we provide a direct estimate of returns to additional search effort harnessed by reducing behavioral barriers. Returns to search effort, typically interpreted as job applications, are a central feature of canonical job search models (Pissarides, 2000) and are important for evaluating policies such as search subsidies or search requirements for recipients of unemployment insurance. Direct estimates of returns to search are very rare, making it difficult to understand variation in the effects of search-related policies – e.g. is this due to different effects on search or returns to search? – or to design search promotion policies – e.g. how many applications should be subsidised?

To achieve this, we combine experimental variation in search costs with data on both individual applications and the outcomes of those applications, a very rare combination in the literature. For instance, many papers study the effect on employment of search subsidies or requirements for

and Marinescu & Wolthoff (2020) use applications as their final outcomes. Fewer papers use employment as a final outcome and these largely use administrative employment data in high-income countries (Behaghel et al., 2020; Belot et al., 2022a; Ben Dhia et al., 2022; Fernando et al., 2021; Marinescu & Skandalis, 2021). Online gig work platforms provide employment data, but for a very different type of work (e.g. Agrawal et al. 2015). A related set of papers study platform users using survey data but have limited data on platform use (e.g. Kelley et al. 2021; Wheeler et al. 2022).

[5] Related work studies links between job search and locus of control (Caliendo et al., 2015; McGee, 2015) and behavioral job search in labs (Brown et al., 2011; Falk et al., 2006a,b; Fu et al., 2019; McGee & McGee, 2016).

[6] Abebe et al. (2021a,b), Abel et al. (2020), Altmann et al. (2018, 2022), Bandiera et al. (2021), Bassi & Nansamba (2020), Beam (2016), Behaghel et al. (2020), Belot et al. (2018), Belot et al. (2022a), Boudreau et al. (2022), Carranza et al. (2021), Dammert et al. (2015), Kiss ⓡ al. (2023), Spinnewijn (2015), and Subramanian (2021) study information. Abebe et al. (2019, 2021a), Banerjee & Sequeira (2020), and Franklin (2017) study subsidies. Arni & Schiprowski (2019), Bolhaar et al. (2020), Lechner & Smith (2007), and Schiprowski (2020) study caseworkers.

receipt of government benefits, but do not observe actual search effort (reviewed by Card et al. 2010, 2018; Filges et al. 2015; Marinescu 2017b). A smaller, more recent literature studies the effect of search subsidies or requirements on online search effort, but without observing outcomes of search (Baker & Fradkin, 2017; Marinescu, 2017a; Marinescu & Skandalis, 2021). Other recent papers experimentally shift search strategies or search technologies, but do not isolate the role of search effort and mostly rely on low-frequency survey data that cannot link specific search actions to outcomes.[7] The closest work to our own shows directly how additional policy-induced job applications affect labor market outcomes, specifically unemployment duration (Arni & Schiprowski, 2019; Lichter & Schiprowski, 2021). While we do not observe administrative data on employment, we advance this work by using application-level data that allow us to describe how marginal and inframarginal search effort is directed, and to compare the outcomes of marginal and inframarginal applications. Our findings about how jobseekers direct applications to specific vacancies and miss applying to some high-return vacancies link to a growing literature on directed job search.[8]

Our findings of a positive and roughly constant return to search for individual jobseekers and a lack of spillovers on other jobseekers suggest the possibility of suboptimally low search effort. This is relevant to debates about possible spillovers or congestion effects from rising search effort, which has important implications for labor market policy.[9] Our results match those from studies using vacancy-level data from platforms to show that the numbers of interviews and offers do not respond to the number of applications (Fernando et al., 2021; Horton & Vasserman, 2021). One potential explanation for this pattern is that vacancy fill rates are $< 60\%$ on all three platforms, so more applications can raise the probability that some applicant meets the firm's reservation quality.

Finally, our paper showcases the methodological value of embedding multiple related experiments into a job search and matching platform to identify a specific barrier to search. This contributes to a growing literature using platforms as laboratories to study search and hiring behavior, although integrating findings from multiple experiments remains rare (Pallais & Sands, 2016).

In Section 2, we describe the context, sample, platform, and experimental design. In Section 3, we present the treatment effects on job applications and interviews and the implied effect of marginal job applications on interviews. We describe our preferred interpretation in Section 4 and show evidence against alternative interpretations in Section 5. Section 6 discusses spillover effects.

[7] See the preceding footnote for examples. In particular, our work differs from recent papers studying the effects of encouraging enrollment on job search platforms (e.g. Afridi et al. 2022; Chakravorty et al. 2023; Jones & Sen 2022). Joining a platform is a bundled experience that might shift factors ranging from wage expectations (Kelley et al., 2021) to information about specific vacancies (Wheeler et al., 2022). These have substantially different interpretations to our treatment, as does the effect of access to (faster) online job search (Bhuller et al., 2019; Chiplunkar & Goldberg, 2022; Gurtzgen et al., 2020; Hjort & Poulsen, 2019; Kuhn & Skuterud, 2004; Kuhn & Mansour, 2014).

[8] Alfonso Naya et al. (2020), Behaghel et al. (2020), Belot et al. (2018, 2022b), Kiss ® al. (2023), Gee (2019), He et al. (2021), and Marinescu & Wolthoff (2020) also study the role of information about vacancies in job applications.

[9] Blundell et al. (2004), Crepon et al. (2013), Doniger & Toohey (2022), Ferracci et al. (2014), Gautier et al. (2018), Johnston & Mas (2018), LaLive et al. (2022), and Lise et al. (2004) study search policy design with spillovers.

# 2 Economic Environment

## 2.1 Context

Our experiment takes place on Job Talash ("job search" in Urdu), a job search and matching platform in Lahore, created by our research partners at the Center for Economic Research in Pakistan. Lahore is a city of about 10 million with an adult labor force participation rate of 49% and employment rate of 47%, both substantially higher for men than women (Table A.1). Job search and matching platforms are a growing feature of Pakistan's labor market, particularly in major urban areas such as Lahore, as we describe in footnote 1.

## 2.2 Samples of Jobseekers and Firms

We recruited participants by conducting a household listing from a random sample of 356 enumeration areas across Lahore between October 2016 and September 2017. This provides a representative listing of 49,506 households and 182,585 adults. We invited each adult household member to sign up for the Job Talash platform and 46,571 expressed interest. The Job Talash call center called each of these people to collect information on their education, work experience, job search, and occupational preferences. The 9,838 people who completed sign-up comprise our main sample.

This sampling process is designed to include participants with different levels of education and labor market attachment, including those who are neither employed nor searching. This is relatively unusual in experimental work in labor economics.[10] This allows us to show that the search barrier we identify affects many different types of active and potential jobseekers. This breadth is also important because the distinction between non-employed searchers and non-searchers is loose and transient in many developing economies (Donovan et al., 2018).

Column 1 of Table 1 presents descriptive statistics for the control group in our study sample. At baseline, 20% of the sample were employed and searching through some channel other than Job Talash, 35% were searching but not employed, 14% were employed but not searching, and 31% were neither employed nor searching. Network search was the most common method, more than twice as common as applying directly and three times as common as visiting establishments to ask about vacancies.[11] Only 4% had used a job search assistance program or online platform other than Job Talash. The average respondent has 7.9 years of work experience with an interdecile range of 0-16. Respondents' education levels also vary widely: 15% have no education, 15% have

[10] Of the 29 experimental job search studies reviewed by Poverty Action Lab (2022), only 8 construct samples from household listings, while another 12 sample from unemployment registries and 4 from job search assistance services, whose participants are required or strongly encouraged to search.

[11] The prevalence of network-based search matches patterns in other developing economies (Government of Bangladesh, 2015; Government of South Africa, 2018; Government of Namibia, 2016). In Lahore's Labor Force Survey, direct applications are slightly more common than network-based search (Table A.1). But this survey does not measure on-the-job search, unlike our own survey.

completed secondary school, and 25% have a university degree. 31% are female and the average age is 30, with an interdecile range of 20-45. In Table A.1, we compare the study sample to the population of Lahore, captured by both the official Labor Force Survey and our household listing. Our sample is slightly younger, more male, more educated, less likely to be employed, and more likely to be searching.

We enrolled firms through a door-to-door listing in commercial areas of Lahore, described in more detail in Appendix A. We invited firms to list any current vacancies during enrollment and recontacted them several times each year to invite them to list more vacancies. For each vacancy, we collected the job title, occupation, salary, benefits, and hours. Vacancies cover a wide range of education and experience levels and occupations, including computer operator, makeup artist, salesperson, sweeper, security guard and HR manager. Column 1 of Table 2 shows that the average vacancy offers a monthly salary of 14,381 Pakistani Rupees (431 USD PPP) and is posted by a firm with 27 employees that hired 5.5 people in the last year.[12] At baseline, only 22% of firms had ever advertised a vacancy on a job search platform, while 67% had recruited through referrals, 35% from CVs dropped off by jobseekers, and 11% through newspapers or other traditional media.

### 2.3 Job Talash Platform

The Job Talash service is free to both jobseekers and firms. It requires only literacy and access to a phone with call and text message functionality. This allows broad access to the platform and easy scaling because 97% of urban households in Lahore's province have mobile phones (MICS, 2018).

After signing up, jobseekers are matched to each listed vacancy using a simple algorithm: the jobseeker must have at least the required years of education and experience, match any gender requirement, and have indicated interest in this occupation.[13] We refer to each jobseeker-vacancy pair, for which the respondent qualifies and has indicated interest in the occupation, as a *match*. We study 1,116,952 matches generated by the platform over four years. The average jobseeker received 113 matches (1.8 per month) with interdecile range 7-271.

Importantly, there is substantial heterogeneity in proxies for the quality of these jobseeker-vacancy matches. Column 1 of Table 2 shows summary statistics for match attributes in the control group. For example, the jobseeker has education and work experience that are an exact match for

[12]These summary statistics weight each vacancy by the number of jobseekers who match with the vacancy. We define a jobseeker × vacancy match in the next subsection. The mean salary offer is roughly 60% of the mean salary in the Labor Force Survey data for Lahore (Figure A.1) and roughly 60% of the mean salary for vacancies posted during the same period on Rozee, Pakistan's largest job search portal (Matsuda et al., 2019). However, this does not necessarily indicate negative selection into our sample of vacancies, as the Labor Force Survey data are not restricted to starting salaries and Rozee caters mainly to highly educated jobseekers.

[13]Of the vacancies listed on this platform, 20.2% are open only to women and 45.3% are open only to men. Explicitly gender-targeted job listings are common in Lahore's labor market and in other settings (Kuhn & Shen, 2013). We examine this in more detail in another paper (Gentile et al., 2023).

Table 1: Jobseeker Summary Statistics, Selection into Applications, and Balance Tests

| | (1) Mean \| T=0 (Std dev. \| T=0) | (2) Selection into application Mean \| T=0, A=1 − Mean \| T=0 [p-value] | (3) Balance checks Mean \| T=1 − Mean \| T=0 [p-value] |
|---|---|---|---|
| Employed and searching | 0.200 | 0.092 | 0.034 |
| | (0.400) | [0.000] | [0.228] |
| Employed and not searching | 0.141 | -0.044 | -0.028 |
| | (0.348) | [0.000] | [0.256] |
| Searching and not employed | 0.345 | 0.041 | 0.024 |
| | (0.475) | [0.033] | [0.344] |
| Not searching and not employed | 0.314 | -0.089 | -0.030 |
| | (0.464) | [0.000] | [0.307] |
| Search method: network | 0.397 | 0.109 | 0.032 |
| | (0.489) | [0.000] | [0.476] |
| Search method: formal application | 0.154 | 0.022 | 0.028 |
| | (0.361) | [0.147] | [0.651] |
| Search method: asked at establishments | 0.225 | 0.080 | 0.032 |
| | (0.417) | [0.000] | [0.728] |
| Years of work experience | 7.85 | -0.23 | -0.22 |
| | ( 8.88) | [0.463] | [0.568] |
| Education: none | 0.146 | -0.063 | -0.012 |
| | (0.353) | [0.000] | [0.294] |
| Education: primary or some secondary | 0.457 | -0.096 | -0.023 |
| | (0.498) | [0.000] | [0.871] |
| Education: complete secondary | 0.148 | 0.032 | 0.002 |
| | (0.355) | [0.027] | [0.673] |
| Education: university degree | 0.250 | 0.126 | 0.033 |
| | (0.433) | [0.000] | [0.335] |
| CV: excellent score | 0.093 | 0.005 | 0.084 |
| | (0.291) | [0.812] | [0.868] |
| CV: good score | 0.330 | -0.031 | 0.032 |
| | (0.471) | [0.281] | [0.970] |
| CV: average or lower score | 0.576 | 0.027 | -0.116 |
| | (0.495) | [0.383] | [0.872] |
| Female | 0.303 | -0.032 | 0.022 |
| | (0.460) | [0.063] | [0.329] |
| Age | 30.7 | -2.0 | -0.5 |
| | (9.7) | [0.000] | [0.307] |
| # matches sent by platform | 113 | 41 | - |
| | (121) | [0.000] | |
| # applications on platform | 0.226 | 1.599 | - |
| | (0.863) | [0.000] | |
| # interviews through platform | 0.014 | 0.101 | - |
| | (0.128) | [0.000] | |

Notes: This table shows summary statistics for jobseekers' baseline characteristics and, in the last three rows, platform use characteristics. Each unit of observation is a jobseeker × vacancy match, to align with the subsequent analysis in the paper. Column (1) shows the mean and standard deviation for the control group. Column (2) shows the difference between the mean for the control group sample of jobseekers who apply to at least one job and the mean of the full control group sample, along with the p-value for testing if this difference is zero. This shows how jobseekers who apply to jobs on the platform differ from jobseekers who do not apply to jobs on the platform. Column (3) provides balance tests by showing the difference between the mean for the treated sample and the mean for the control group sample, along with the p-value for testing if this difference is zero. This checks if the treated and control respondents have the same baseline characteristics on average. P-values are generated from regressions that use heteroskedasticity-robust standard errors clustered by jobseeker (the unit of treatment assignment) and include fixed effects for the strata within which treatment was randomized (see footnote 16). We leave column (3) blank for the final three rows because applications and interviews are post-treatment outcomes and the number of matches can be influenced by post-treatment actions, although we show in Section 3.1 that this influence is irrelevant for our main results.

the employer's preferences in only 18 and 13% of matches respectively.[14] Furthermore, 85% of jobseekers indicate interest in multiple occupations, with the median jobseeker interested in six occupations. These patterns show heterogeneity in how much firms might value jobseekers matched to their vacancies and how much jobseekers might value the vacancies to which they are matched. This creates the potential for heterogeneous returns to applications, which is important for interpreting our experimental results.

The platform collects new vacancy listings from firms every 1-2 months and sends jobseekers text message updates if they have matched to any vacancies in that month. See Figure A.2 for a sample text message. The text messages contain the job title, firm name, firm location, and salary of each match, along with the deadline to apply. Jobseekers only learn about vacancies to which they match, as the platform does not have a website that lists vacancies. Jobseekers on average receive a text every 2.8 months. Conditional on receiving any matches in that month, the average jobseeker receives 3.1 matches with interdecile range 1-6.

If a jobseeker wants to apply to any of these vacancies, the platform forwards her CV to the firm. (We describe the application process in Section 2.5.) The CVs are constructed by the platform by populating a template with respondent-specific demographics, education, and work experience, so there is no variation in CV design. The platform sends all applications to the firm in a packet at the application deadline, so application timing does not affect interview probability. If the firm wants to interview the jobseeker, they contact the jobseeker directly to arrange the interview. The Job Talash team surveys each firm a few weeks after the application deadline to ask which applicants they interviewed.

The platform design has two key advantages for our research, relative to other job search environments. First, we observe all information available to both sides of the market. We observe the same information about vacancies that jobseekers receive through the text messages, and the same information about jobseekers that firms receive through the CVs. We also gather a CV quality score from the hiring managers for a subset of jobs on the platform for the CVs of the 1,470 jobseekers matched to those jobs. Second, respondents see only the vacancies to which they match. This generates a well-defined jobseeker-vacancy unit of analysis that we use throughout the paper, and refer to as a *match*. This is not possible on platforms that allow unrestricted search, as every jobseeker can apply to any vacancy on the platform and the researcher may not observe which vacancies the jobseeker has seen. This makes it difficult to distinguish between vacancies a jobseeker sees but decides not to apply to and vacancies she has not seen at all.

[14] For each vacancy, the platform collects both the *required levels* and *preferred types* of education and experience. Jobseekers are only matched to vacancies if they have the required levels of experience and education, e.g. complete high school and five years of work experience. They can be matched even if they do not have the preferred types of education and experience, e.g., their work experience might be in a non-preferred field. We use the alignment between jobseekers' education and experience and vacancies' preferred types as a measure of match quality.

Table 2: Vacancy- and Match-level Summary Statistics and Selection into Applications

| | (1) Mean \| T=0 (Std dev. \| T=0) | (2) Selection into application Mean \| T=0, A=1 − Mean \| T=0 [p-value] |
|---|---|---|
| Salary | 14,381 | 6,576 |
| | (9,170) | [0.000] |
| Firm # employees | 26.6 | 61.7 |
| | (135) | [0.000] |
| Firm # vacancies in last year | 5.50 | 6.80 |
| | (12.2) | [0.000] |
| Exact education match \| vacancy requires high ed | 0.184 | -0.016 |
| | (0.387) | [0.542] |
| Exact experience match \| vacancy requires experience | 0.126 | 0.050 |
| | (0.331) | [0.016] |
| Gender preference aligned | 0.700 | -0.191 |
| | (0.458) | [0.000] |
| Short commute | 0.519 | 0.021 |
| | (0.500) | [0.329] |
| $V_{vm}$ index: proxies of value of vacancy to jobseeker | 0.016 | 0.226 |
| | (0.899) | [0.000] |
| Applied | 0.002 | 0.998 |
| | (0.045) | |
| Interviewed | 0.000 | 0.063 |
| | (0.011) | [0.000] |

Notes: This table shows summary statistics for vacancy- and match-level characteristics. Column (1) shows the mean and standard deviation for the control group sample. Column (2) shows the difference between the mean for the control group sample of matches that resulted in applications and the mean of the full control group sample of matches, along with the p-value for testing if this difference is zero. This shows how matches that lead to applications differ from other matches. P-values are generated from regressions that control for stratification block fixed effects and use heteroskedasticity-robust standard errors clustered by jobseeker. The p-value for 'Applied' in column (2) is omitted because the standard error is zero by definition for the mean application rate conditional on application. Salary is in Pakistani Rupees per month. 1 Rupee $\approx$ USD 0.03 in purchasing power parity terms during the study period. Exact education match is an indicator for an exact match between the employer's preferred field of educational specialization and the jobseeker's field. Exact experience match is an indicator for a match in which the jobseeker has experience in the same occupation as the vacancy. These two variables are only defined for vacancies that require respectively more than basic education and some experience. These two variables use employers' *preferred* education and experience, rather than the *required* education and experience used in the matching algorithm. The $V_{vm}$ index is an inverse covariance-weighted average of all the preceding rows, following Anderson (2008).

The set of matches jobseekers receive are based on information collected during platform sign-up. However, jobseekers can contact the platform to update their education, experience, or occupation preferences at any time, including after treatment occurs. They can also ask to pause or stop receiving matches. This can create a sample selection problem for the match-level dataset. But we show in Appendix B.4 that updates are rare, so there is little selection and correcting it does not affect our findings.

### 2.4 Platform Use

We highlight four important patterns of platform use, using the control group statistics in Tables 1 and 2. First, the application rate is low: the average jobseeker submits only 0.23 applications and applies to 0.2% of matches they receive. This is expected because our sample deliberately includes people who were not actively searching at baseline, on or off the platform. But it is comparable to rates on some other platforms in countries ranging from France to Mozambique (Table A.2). The application count is unsurprisingly right-skewed: 74% of jobseekers submit zero applications and 5% submit more than 5 applications. Column 2 of Table 1 shows that, within our sample, jobseekers who do and do not actively use the platform differ on baseline characteristics. We discuss what this implies for interpreting our experimental results in Section 4.5.

Second, the interview rate is low, but mainly because the application rate is low. The average jobseeker receives 0.014 interviews through the platform but each application has a 6.3% probability of generating an interview.[15]

Third, there is substantial variation in match value, and applications are directed to relatively high-value matches. For example, the standard deviation of monthly salary is roughly 9,200 Pakistani Rupees (275 USD PPP) and higher-salary vacancies get more applications (Table 2, column 2, row 1 and Figure C.2, panel A). At the match level, jobseekers are more likely to apply to vacancies where their work experience is a closer match (Table 2, column 2, row 5). Combining our available proxies for vacancy and match value in a single summary index shows that applications are substantially more likely for high-value matches (row 8). This confirms that jobseekers can and do apply to higher-value matches, rather than randomly picking where to apply from relatively homogeneous matches, as random search models assume.

Fourth, however, control group jobseekers miss applying to many high-value matches. For example, jobseekers apply to only 0.46% of the matches in the top quintile of their within-jobseeker salary distributions (Figure C.2, panel A). This pattern also holds for the summary index of match value (Figure C.2, panel B).

These patterns naturally motivate our research. On the one hand, the facts that job applications are rare, even to high-value matches, and that applications have reasonably high interview probabilities suggest that lowering application costs could lead to more applications and substantially more interviews. On the other hand, the facts that jobseekers seem to choose strategically where to apply and that pecuniary and time costs of applying are already very low suggest that additional applications could go to relatively low-value matches and yield few interviews. Our experiment is designed to adjudicate between these two possibilities, both by identifying returns to additional applications and by understanding which barriers deter additional applications in this setting.

[15] As a benchmark, Belot et al. (2018) find that 3.6% of job applications submitted on a Scottish platform generate interview invitations. Other studies of platform-based job search do not report this ratio. Studies of off-platform job search in developing economies find $> 10\%$ of applications generate interviews, although we might expect a higher ratio for more expensive off-platform search (Abebe et al., 2021a; Banerjee & Sequeira, 2020; Carranza et al., 2021).

### 2.5 Experimental Design and Interpretation

Our primary experiment varies a single element of communication with jobseekers in order to reduce the cost of applying for jobs on the platform: whether the platform initiates the application phone call or the jobseeker must do so. The platform sends text messages to all jobseekers, irrespective of treatment status, at the same time at the start of each monthly "matching round." The text messages list the job title, firm name, firm location, and salary of each match received by the jobseeker that month and tell jobseekers to call the Job Talash number by a stated deadline if they want to apply. The deadline is on average ten days after the text message, with some variation between matching rounds due to operational factors such as platform staff capacity. When a jobseeker calls the platform, they are offered a free call back on the same day to complete the application process. The financial cost of placing the call to initiate the application is a maximum of 5 Pakistani rupees (0.03 USD PPP, less than 1% of a day's earnings at minimum wage).

In the treatment condition, the call center *also* makes two attempts to phone each jobseeker and ask if they would like to initiate the application process. Roughly 50% of jobseekers are assigned to treatment for the duration of the experiment. Assignments are balanced on baseline jobseeker characteristics (Table 1, column 3).[16] Treated jobseekers are called in a random order, starting as soon as the text messages are sent and continuing until the deadline. Treatment is designed to minimize anticipation effects: treated jobseekers are told in initial matching rounds that they may not receive a phone call in every round, and should contact the call center if they wish to apply.

Importantly, the text message and phone call scripts contain identical information. The phone call scripts are also identical for the treatment and control groups. The only difference between the two is that the call center initiates the call for the treatment group. Call center agents are trained to not encourage or pressure jobseekers to apply at any moment during the call, and a supervisor audits the recording of at least one call per call center agent per matching round to ensure agents are following the script. Jobseekers can ask for more information about jobs on the calls but call center agents had access to no additional information in most matching rounds and we show in Section 5.1 that our findings are robust to omitting rounds when they had access to more information.

We interpret treatment as a reduction in the cost of applying for jobs on the platform. In principle, these costs might be monetary (of airtime to initiate a call), time (of waiting for their call to get answered), or psychological (e.g. cognitive costs of processing vacancy information or fear of rejection). However, the platform is already designed to minimize the monetary and time costs jobseekers incur to initiate applications, and we show in Section 4.3 that additional experiments further reducing monetary and time costs produce substantially smaller effects on applications. Hence the most plausible interpretation of the phone call treatment is a reduction in the *psycho-*

[16] Randomization took place within 82 strata based on the time that each geographic area completed household listing, platform sign-up, and the first round of matching.

*logical* cost of initiating an application. We develop this interpretation in more detail in Section 4.1, showing what this implies for treatment effects on applications and the returns to treatment-induced applications. We show in Section 5 that we can rule out several other interpretations based on the platform design, additional experiments we run, and additional survey measures.

# 3 Search Effort and Returns to Search

In this section we first show that the phone call treatment substantially increases the number of job applications and interviews. We then combine these results in a two-stage least squares framework to show that marginal applications submitted due to treatment yield interviews with roughly the same probability as inframarginal applications submitted without treatment, and yield interviews for vacancies of similar quality. This implies roughly constant returns to additional search effort.

## 3.1 Treatment Effects on Search Effort and Search Outcomes

We run all analysis at the level of the jobseeker $\times$ vacancy match. As described in Section 2, each jobseeker only learns about vacancies that match their occupational preferences, education, and work experience, so these matches provide a well-defined unit of observation. We first estimate:

$$Y_{jv} = T_j \cdot \Delta + \mu_b + \epsilon_{jv}, \tag{1}$$

$Y_{jv}$ is either an indicator for jobseeker $j$ applying to vacancy $v$ or an indicator for jobseeker $j$ being invited to an interview for vacancy $v$. $\mu_b$ is a fixed effect for the stratification blocks within which treatment was randomized (see footnote 16). We estimate heteroskedasticity-robust standard errors clustered by jobseeker, the unit of treatment assignment.

Treatment leads to a large increase in job applications. Treated respondents apply to 1.32 percentage points more matches with standard error 0.08 p.p. (Table 3, column 1). This effect is seven times larger than the control group's application rate of 0.18%. Treatment effects decline through time but remain positive for at least four years after jobseekers register for the platform. As a result, at the jobseeker level, treatment shifts the entire distribution of the number of applications to the right (Figure B.1). In particular, treatment increases the proportion of jobseekers who ever apply to a vacancy on the platform from 21 to 44%.

Treatment also increases the probability of getting an interview by 0.078 p.p. with a standard error of 0.009 p.p. (Table 3, column 2). This effect is nearly seven times larger than the control group's 0.012% share of jobseeker $\times$ vacancy matches that generate interviews. At the jobseeker level, treatment also shifts the entire distribution of the number of interview invitations to the right (Figure B.1). The interview data is collected from firms, not jobseekers, and firms are unaware of respondent-level treatment assignments. So using firm reports of interview invitations minimizes measurement error from experimenter demand effects.[17]

[17]A few firms do not provide the list of jobseekers they interviewed. We assume no jobseekers matched to these

Table 3: Treatment Effects on Job Search & Search Returns

| | (1) Apply | (2) Interview | (3) Int. $\times$ $V_{vm}$ | (4) Interview | (5) Int. $\times$ $V_{vm}$ |
|---|---|---|---|---|---|
| Phone call treatment | 0.01322 | 0.00078 | 0.00281 | | |
| | (0.00075) | (0.00009) | (0.00036) | | |
| Apply | | | | 0.05865 | 0.21283 |
| | | | | (0.00516) | (0.02151) |
| # matches | 1,116,952 | 1,116,952 | 1,116,952 | 1,116,952 | 1,116,952 |
| # jobseekers | 9831 | 9831 | 9831 | 9831 | 9831 |
| Mean outcome \| T = 0 | 0.00185 | 0.00012 | 0.00044 | 0.00012 | 0.00044 |
| Mean outcome \| T = 0, Apply = 1 | | | | 0.06290 | 0.23778 |
| p: IV effect = mean \| T = 0, Apply = 1 | | | | 0.647 | 0.501 |
| IV strength test: F-stat | | | | 312.8 | 312.8 |
| IV strength test: p-value | | | | 0.00000 | 0.00000 |

Notes: Column 1 shows the coefficient from regressing an indicator for job application on treatment assignment. Column 2 shows the coefficient from regressing an indicator for interview invitation on treatment assignment. Column 3 shows the coefficient from regressing an indicator for interview invitation weighted by a proxy index for the value of the vacancy to the jobseeker, $V_{vm}$, on treatment assignment. Column 4 shows the coefficient from regressing an indicator for interview invitation on job application, instrumented by treatment assignment. Column 5 shows the coefficient from regressing an indicator for interview invitation weighted by the proxy index $V_{vm}$ on job application, instrumented by treatment assignment. The proxy index $V_{vm}$ is an inverse covariance-weighted average (following Anderson 2008) constructed using vacancy-level characteristics log salary and indicators for offering any non-salary benefits, below-median working hours, and allowing flexible hours as well as indicators for the match-level characteristics of vacancy salary exceeding the jobseeker's expected salary, below-median commuting distance, the jobseeker's educational specialization exactly matching the vacancy's preference, and the jobseeker's work experience exactly matching the vacancy's preference. Anderson-style indices, by construction, have zero means and hence some negative values. But multiplying the interview invitation indicator by a negative value would not produce sensible results. Hence we recenter the index so it has strictly positive values. All regressions use one observation per jobseeker $\times$ vacancy match, include stratification block fixed effects, and use use heteroskedasticity-robust standard errors clustered by jobseeker, which are shown in parentheses. The p-value is for a test of equality between the IV treatment effect and the mean interview rate for control group applications. The first-stage F-statistic and p-value are for the test of weak identification from Kleibergen & Paap (2006).

The treatment effects on both applications and interview invitations are broad-based. Treatment substantially raises job application and interview rates for women and men, for people who were employed and not employed at baseline, searching and not searching at baseline, and with above- and below-median education and age (Table B.3). This suggests that the economic behavior driving the treatment effects, which we discuss in Section 4, occurs across many types of jobseekers.

The treatment effects on applications and interviews are robust to a range of checks we show in Appendices B.2 - B.4, including different ways of handling fixed effects, conditioning on baseline covariates, reweighting the data to give equal weight to each jobseeker rather than each jobseeker $\times$ vacancy match, and accounting for pauses in receiving matches that some jobseekers request.

vacancies are interviewed. Our key results are unchanged if we instead code these interview values as missing.

### 3.2 Returns to Inframarginal Search and Treatment-Induced Marginal Search

To evaluate the returns to search, we estimate the relationship between the treatment effects on applications and interviews using an instrumental variables approach. We estimate the system:

$$\text{Apply}_{jv} = T_j \cdot \alpha + \mu_b + \epsilon_{jv} \tag{2}$$

$$\text{Interview}_{jv} = \text{Apply}_{jv} \cdot \beta + \eta_b + \varepsilon_{jv} \tag{3}$$

$\beta$ recovers the local average effect of a treatment-induced application on the probability of an interview (LATE) under four conditions: treatment should be independent of all other factors influencing applications and interviews (independence), influence applications (strength), influence interviews only through applications (exclusion), and increase the probability of application for all respondents (monotonicity). The independence condition holds by random assignment and the preceding results show that the strength condition holds. We discuss potential complications with the monotonicity and exclusion conditions and how we address them at the end of this subsection.

Marginal applications submitted due to treatment have roughly the same return as inframarginal applications, measured in terms of interview invitations. The LATE estimate shows that the average treatment-induced application has a 5.9% probability of an interview invitation with standard error 0.5 (Table 3, column 3, row 2). This is very similar to the 6.3% mean interview probability for control group applications and we cannot reject equality of the probabilities (p = 0.647).

Marginal and inframarginal applications also have equal returns measured in 'value-weighted' interviews. This finding is important, as the return to an application, and the decision to apply, reflects both the probability of an interview $P$ and the value of an interview $V$. To show this, we construct a proxy index $V_{vm}$ for the value of each match a jobseeker receives: an inverse-covariance weighted average of positive attributes of the vacancy and match such as salary and commuting distance, defined in detail in the note below Table 3. We estimate the system (2)-(3), replacing the second stage outcome with an interaction between the interview invitation indicator and the proxy index. This gives us the local average treatment effect on $P \cdot V$. The returns to inframarginal and marginal search using this measure are again very similar: respectively 0.22 and 0.24, with p = 0.501 for the test of equality (Table 3, column 5). We repeat this value-weighting exercise using each individual proxy for interview value and fail to reject equality of marginal and inframarginal applications' value-weighted interview outcomes for all eleven proxies (Table B.1).

The finding of roughly constant returns on both interviews and value-weighted interviews is not a mechanical consequence of a matching algorithm or labor market that ensures homogeneous returns. Instead, as we explain in Section 2, most jobseekers are matched with vacancies from multiple occupations and with firms that prefer different types of work experience and education. This create scope for heterogeneous returns from applying to different types of matches. Furthermore, Table B.3 shows that the constant returns finding also holds for jobseekers with above-median education and who were employed at baseline. They match to a broader set of jobs, giving them more scope to direct applications widely, making the constant returns finding more surprising.

The finding of roughly constant returns is also not a consequence of low power. The return to marginal applications is precisely estimated, with a 95% confidence interval of 4.9 to 6.9 percentage points for interview invitations. Relative to the interview rate of 6.3% for inframarginal applications, we can reject decreases of more than 1.4 p.p. and increases of more than 0.6 p.p. Even the lower bound of the confidence interval implies a decrease of only 1.4/6.3 = 23% in the average interview probability over a 615% increase in the application rate, implying a slowly decreasing return to search effort. A similar pattern holds for the returns measured in value-weighted interviews. We do not, of course, claim that returns would be constant over all possible levels of search effort and acknowledge that they could be substantially lower at sufficiently high levels.[18]

Before proceeding, we briefly discuss an extensive battery of robustness checks on the constant returns finding, shown in detail in Appendices B.2 - B.4. First, we address the possibility that treatment increases applications from some jobseekers and decreases applications from others, which would violate the monotonicity condition used in our IV analysis. To do this, we derive a bound on the bias from violations of monotonicity in these data, following De Chaisemartin (2017). This implies that a bias-corrected LATE of applications on interviews is bounded between 4.5 and 5.9%. Second, we address the possibility that treatment affects both the quantity and quality of applications, which would complicate the exclusion restriction used in our IV analysis. All application content is sent by the Job Talash platform using template CVs. We show that treatment effects on measures of application quality that jobseekers can change by updating information used in their CV templates are close to zero. Third, we address the possibility that treatment affects which matches jobseekers receive, which would create a sample selection problem because we use each jobseeker $\times$ vacancy match as a unit of analysis. This can only occur if treatment causes jobseekers to update the information used to match them to vacancies: their occupational preferences, education, or experience. We show that treatment has little impact on updating this information and that our key results are unchanged when we use a sample consisting of the counterfactual set of matches that would have been generated in the absence of these updates. Fourth, we use a non-IV approach to compare the returns to marginal and inframarginal applications under different assumptions, which also generates similar estimates of returns. Finally, we show that our key findings are robust to different ways of handling fixed effects and conditioning on baseline covariates, including allowing interactions between treatment assignment and the fixed effects.

[18]As a very speculative back-of-the-envelope calculation, we can estimate a linear returns curve using the control group means and treatment effects for the application and interview rates. We can then use the estimated curve to extrapolate the marginal interview probability at even higher application rates. The estimated curve is relatively flat. For example, if the share of matches generating applications increased 25 fold, from 0.185% to 4.625%, then the linear extrapolation implies that the interview probability for the marginal application would drop from 5.5 to 3.7%.

We focus on interviews and value-weighted interviews as outcomes because these take advantage of the strengths of the platform we study. The platform gives us detailed data at the level of jobseeker $\times$ vacancy matches: all vacancy characteristics observed by the jobseeker, all jobseeker characteristics observed by the firm, application decisions, and interview invitations. These data allow us to precisely describe how search decisions are made and the consequences of those decisions up to the interview stage. Interviews are also a key search outcome because they are a necessary condition for job offers, impose non-trivial costs on both job applicants and firms, and provide learning opportunities for jobseekers - hence their widespread use as central outcomes in some areas of labor economics such as audit studies.

The disadvantage of platforms is that they do not generally provide employment outcomes, so evaluations relying on employment outcomes require off-platform data. Moreover, employment outcomes require even greater scale for statistical power. For these reasons, our study is not powered to study effects on employment at the scale of this experiment on this platform, like some other studies of platform-based job search such as the important work on search direction by Belot et al. (2018). The treatment effect on employment in a survey of jobseekers is 1 percentage point, with standard error 2 p.p (Table B.9).[19] Any employment effect is unlikely to be driven by treatment effects on off-platform search, which are negative but close to zero at the extensive and several intensive margins (Tables B.9 and B.10). While imprecise, these numbers suggest that treatments like this have the potential to generate substantial increases in the number of employed users for larger platforms. For example, Pakistan's Rozee has 9.5 million users, 1000 times the size of our platform.

## 4 Explaining Marginal Returns to Search

Our finding of roughly constant returns to job search raises a puzzle: why do jobseekers not apply to more jobs in the absence of treatment, especially given the seemingly low cost of applying on the platform? In this section, we develop a simple conceptual framework that can explain both the large treatment effect on applications and the roughly constant returns to treatment-induced applications. We show that this framework is also consistent with additional patterns in the treatment and control group data. Finally, we show that our results are unlikely to be explained by treatment effects on the pecuniary or time costs of applying, leaving treatment effects on the psychological costs of applying as the most plausible explanation. In Section 5, we show that several alternative

[19] The survey is conducted an average of 40 months after treatment. The survey response rate is 47% and differs between treatment and control groups, which might create a sample selection problem. To address this, we randomize some features of the survey data collection, e.g., number of call attempts. We use this to create instruments for a sample selection correction term, following DiNardo et al. (2021). We describe the selection correction method and how the randomized survey features influence response rates in detail in Appendix B.6. We preregistered employment as a trial outcome because we did not know at the time (July 2020) how much COVID-19 would constrain platform operation, data collection, and hence power.

frameworks are inconsistent with our results.

### 4.1 Conceptual Framework

Here we present a brief, intuitive discussion of our conceptual framework, with the formal model left to Appendix C.1. This paper's contribution is empirical rather than theoretical, so the framework is deliberately simple and stylized.[20]

The platform sends each jobseeker a monthly batch of matches. We begin with a standard assumption (A1) that the jobseeker applies to all matches with positive expected net return, i.e., matches whose expected gross return exceeds the cost of applying. The expected gross return is $P \cdot V$, where $P$ is the probability of an interview conditional on applying and $V$ is the gross value of getting an interview, which is a reduced-form expression for the expected present value of the flow of future utility from the interview.

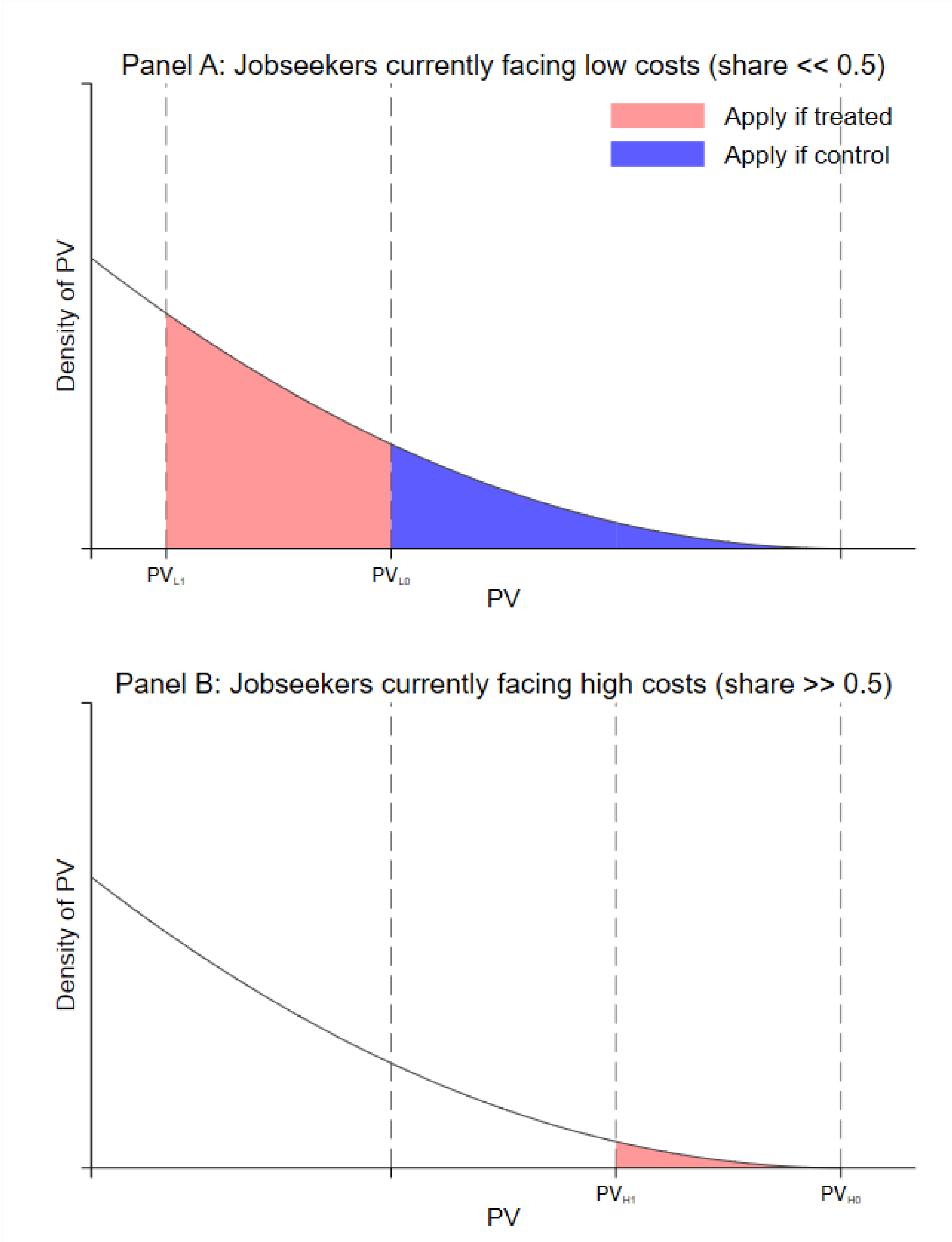
Our key assumption (A2) is that the cost of applying varies across jobseekers and/or through time, and can be high enough that some untreated jobseekers choose not to apply in some matching rounds. Figure 1 shows application behavior by untreated jobseekers under assumptions (A1) and (A2): jobseekers facing low costs in that month apply to matches with $PV$ above $PV_{L0}$ (the blue-shaded area in panel A), while jobseekers facing high costs apply to no matches (panel B).

In this framework, marginal applications induced by treatment can come from two sources. The first type of marginal applications come from jobseekers facing low costs at the time, who would apply to at least one match in that round even without treatment. Treatment lowers their cost of applying, so they apply to matches with $PV$ above $PV_{L1}$ (the pink-shaded area in panel A). These marginal applications have strictly lower returns than the inframarginal applications. The second group of marginal applications come from jobseekers facing high costs at the time, who would not apply to any matches in that round without treatment. Treatment lowers their cost of applying, so they apply to matches with $PV$ above $PV_{H1}$ (the pink-shaded area in panel B). These marginal applications will have higher returns than the inframarginal applications if the cost reduction due to treatment is small relative to the cost variation within the control group, i.e., $PV_{H1} > PV_{L0}$.

The treatment effect on applications and return to marginal applications are averages across these two types, weighted by their relative size. The large effect on applications relative to the control group mean suggests that many more jobseekers face high application costs at each time than low. The roughly equal returns to marginal and inframarginal applications can occur if the lower marginal return to applications from low-cost jobseekers (panel A) are offset by the potentially higher marginal return to applications from the more numerous high-cost jobseekers (panel B). We show this formally in Appendix C.1 and explain that the framework does not require the

[20]This framework has a similar spirit to recent models of 'partially directed search,' in which jobseekers want to apply to the highest-return matches but face costs of identifying which these are, leading them to miss some high-return matches (Lentz et al., 2022; Wu, 2021).

Figure 1: Application Decisions for Treated and Control Jobseekers Facing High versus Low Costs




Notes: This figure shows the application decisions for jobseekers facing low application costs at the time they receive matches (top panel) and jobseekers facing high application costs at the time they receive matches (bottom panel). The blue-shaded sections show the matches to which control group jobseekers apply. The pink-shaded sections show the additional matches to which treatment group jobseekers apply. For simplicity, we show only the right tail of the density of $PV$. We formally derive values for $PV_{H0}$, $PV_{H1}$, $PV_{L0}$, and $PV_{L1}$ in Appendix C.1.

simplifying assumption of only two cost types.[21]

### 4.2 Additional Tests of the Conceptual Framework

This framework delivers three additional predictions that we can test. First, *control group jobseekers will not apply to some high-value vacancies* because some of them face high application costs during some matching rounds. To test this, Figure 2 panel A shows the control group application rate by quintiles of the vacancy salary in blue. The application rate increases monotonically from the bottom to the top quintile, consistent with the idea that jobseekers value higher salaries. But under half of all control group applications are sent to top quintile matches, and under 0.1% of matches in the top quintile receive applications. This shows that control group jobseekers miss many high-value matches, consistent with the conceptual framework.

Second, the framework predicts that *treatment and control group applications will go to vacancies with similar average values*, as treatment will induce applications to a mix of higher- and lower-return matches whose average value is similar to the control group. To test this, Figure 2 panel A shows the control and treatment group application rates by quintiles of the vacancy salary in respectively blue and red. Treatment effects increase monotonically from the bottom to the top quintile. But the share of total applications sent to each quintile does not differ between treatment and control groups. To show this, we test whether the ratio of the treatment group application rate to the control group application rate is equal across all five quintiles and fail to reject the null hypothesis (p = $0.739$). Salary is not the only input into match value, but treatment and control group applications have similar values of other proxies for match value, including the index $V_{vm}$ we introduced in Section 3.2 that combines all observed proxies of vacancy and match value (Panel B).[22] Table C.1 reinforces this result using an extension of the complier or latent types analysis introduced in Section 4.5. This shows that matches receiving applications from treated and control jobseekers have similar mean values.

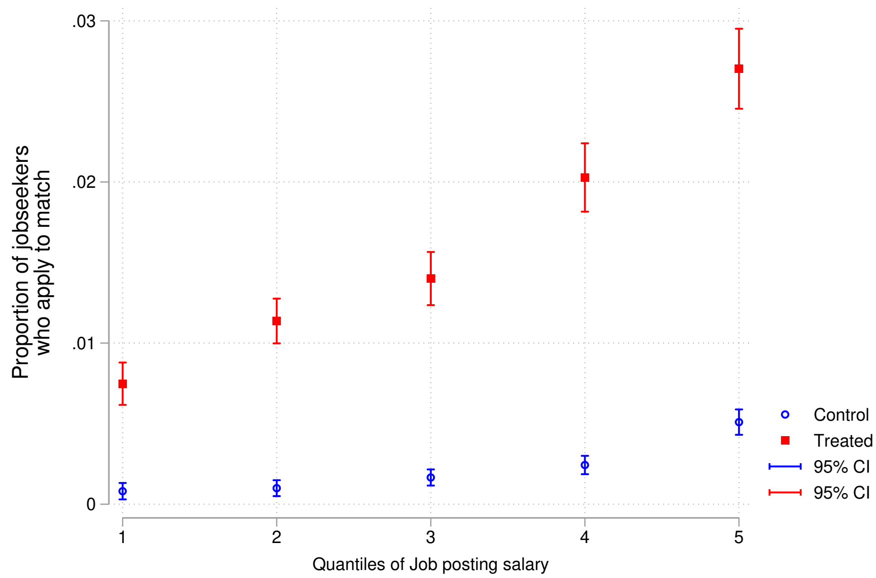
Third, in this framework, *treatment group applications will go to vacancies with more dispersed values*, as evident from the wider range of $PV$ in the pink versus the blue region in Figure 1. To test this, we estimate treatment effects on the variance and 10th percentile of log salary for matches that receive applications. Table C.2 shows that the salary variance is substantially higher in the treatment group and the 10th percentile is lower, consistent with treatment encouraging some applications to lower-value vacancies. The same pattern holds for the proxy index $V_{vm}$, although the treatment effects are not statistically significant for this proxy index. This shows that, consistent with the conceptual framework, marginal treatment-induced applications go to vacancies with the

[21]This framework allows the possibility of decreasing returns to marginal applications for treatments that decrease the application cost by more, leading to very large increases in application rates and $PV_{H1} < PV_{L0}$.

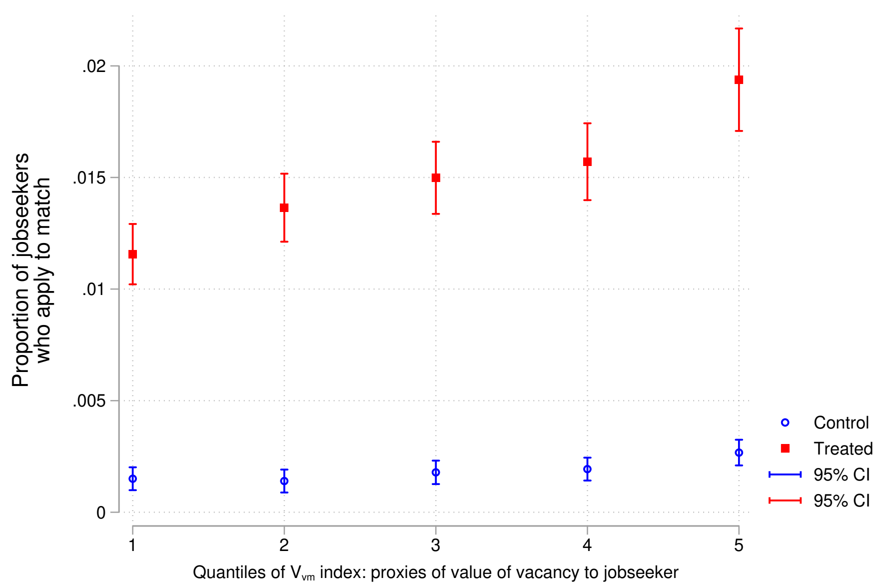
[22]Figures C.2 - C.5 show that the same patterns hold when we control for variation between jobseekers in the value of matches they receive and variation through time in the number of matches they receive.

Figure 2: Heterogeneous Treatment Effects by Value of Vacancy

Panel A: By salary




Panel B: By $V_{vm}$ index




Notes: This figure shows heterogeneous treatment effects of the phone call treatment on applications by quintiles of proxies for the value of the job posting to the jobseeker. Panel A uses the job posting salary as a value proxy and Panel B uses the $V_{vm}$ index described in Section 3.2 as a value proxy. The p-value for the equal ratios test is $0.739$ for Panel A and $0.911$ for Panel B. Results in both panels are conditional on stratification block fixed effects. Each observation is a jobseeker $\times$ vacancy match and the sample includes all matches. Solid vertical lines show 95% confidence intervals, constructed using heteroskedasticity-robust standard errors, clustering by jobseeker.

same average value as inframarginal applications but with more dispersed values.

The finding that marginal and inframarginal applications are directed to matches with equal average values rules out one possible explanation for the constant returns finding. If marginal applications were sent to vacancies with less desirable attributes than inframarginal applications, then they might face less competition and hence have higher interview probabilities. We instead find that marginal and inframarginal applications are sent to vacancies with equally desirable attributes and roughly equal interview probabilities, patterns inconsistent with the 'less competition' explanation and consistent with our preferred interpretation.

This framework provides a clear economic interpretation of the LATE we estimate in Section 3.2: it is the average effect of application on interview invitation, for applications sent due to a treatment-induced drop in the cost of applying. In this framework, marginal applications come from jobseekers who face average higher costs of applying in the absence of treatment, relative to jobseekers submitting inframarginal applications. The constant returns finding shows that returns to applications are similar for jobseekers facing relatively high versus low costs of applying at the time, in the absence of treatment. Section 4.5 showed that marginal and inframarginal applications come from observationally similar jobseekers. This may simply mean that the observed jobseeker characteristics are largely unrelated to the costs of and marginal returns to applying, a point we revisit in Section 4.4.

The roughly constant returns to marginal applications shown in Section 3.2 and the patterns of which matches receive applications shown in this section are all consistent with treatment helping jobseekers overcome costs of initiating applications. We next justify why we view psychological costs as a more plausible candidate than time or pecuniary costs.

### 4.3 Pecuniary and Time Costs of Job Applications

Pecuniary costs of applying are unlikely to explain our main results, based on the design of the platform and results from an additional experiment. Job applications on the platform are very inexpensive, even for jobseekers in the control group, who can call the platform and request an immediate free callback to minimize the cost of airtime. This call costs a maximum of 5 Pakistani rupees (USD 0.03, or less than 1% of a day's earnings at minimum wage). In addition, mobile phone providers in Pakistan offer small loan packages allowing customers to borrow 10-20 rupees of credit against a future top-up card, and the application period for each matching round stays open for at least a week, so a short-term zero balance is very unlikely to be a binding constraint.

We run an additional experiment to show that lowering the pecuniary cost of applying does not substantially increase the application rate. We randomly select some control group jobseekers to receive a text message reminder that they can ask the platform to call them back about a job posting, saving the cost of their calling the platform.[23] Column 1 of Table C.3 shows that this free callback reminder treatment has an effect one hundredth of the size of the effect of the main phone call treatment, and is statistically significantly different (p = 0.017).

Time costs of applying are also unlikely to explain our main results, based on the design of the platform and results from an additional experiment. The platform allows relatively quick applications even in the control group by allowing applications by phone and prescreening vacancies that match jobseekers' qualifications and interests. This saves time relative to many search methods such as traveling to submit applications in person or reviewing unscreened vacancies. Time costs of applying are slightly lower for the main treatment group because control group jobseekers wait for their call to be answered and for the call center operator to find their record in the system. But this time difference is very small: call records show that this takes approximately 4 minutes on average, compared to 10-24 minutes for completing the application process itself over the phone.

We run an additional experiment to show that lowering the time cost of applying has a modest effect on the application rate. We randomly offer some control group jobseekers the option to text the platform and ask for a callback at a specific time. This eliminates the differential wait time between the main treatment and control groups. Column 2 of Table C.3 shows that the effect of this callback request treatment is one quarter the size of the main phone call treatment, and is statistically significantly different (p = 0.002). This shows that time costs deter some applications in the control group but can explain only a small share of the effect of the main phone call treatment.

### 4.4 Psychological Costs of Job Applications

Given the limited role for pecuniary and time costs of applying, we view psychological costs of initiating applications as the most likely explanation for our main results. The existing literature suggests multiple types of psychological costs that might be reduced by the phone call treatment. It might reduce *attention costs* as treated jobseekers do not need to pay attention to text messages and deliberately set aside time to process their content and decide whether to apply (Gabaix, 2019). Control group jobseekers might not initiate applications due to *fear of rejection*, while the phone call allows treated jobseekers to apply 'in the moment' and spend less time anticipating rejection (Köszegi et al., 2022). *Present bias* might lead control group jobseekers to repeatedly postpone applications until the deadline passes, while the phone call gives treated jobseekers a reason to apply at that moment (Ericson & Laibson, 2019). We show in Appendix C.1 how each of these factors can enter our model. These explanations are consistent with research showing that eliminating or reducing the need to initiate decisions can raise financial and health investments (DellaVigna, 2009; DellaVigna & Malmendier, 2006; Madrian & Shea, 2001; Thaler & Benartzi, 2004).

We do not directly observe psychological costs of initiating applications in our data. Hence we

[23] Each treatment used to test a mechanism in Sections 4 and 5 is assigned to a small share of the sample. Controlling for these assignments and their interactions has no impact on the estimated effects of the main phone call treatment.

cannot test if the effects of the phone call treatment vary with psychological costs across jobseekers or through time. Observable characteristics that could be correlated with psychological costs, such as education, may also be correlated with returns to applications, so we do not view the heterogeneous treatment effects discussed in Section 3 as appropriate tests of the model.

While we cannot pin down exactly how the phone call treatment reduces psychological costs of initiating applications, we can test and reject two possible explanations. First, we show that treatment does not simply function as a reminder. Our main findings could be explained by a combination of procrastination and forgetfulness, as in Ericson (2017): some jobseekers postpone applications until near the deadline, forget to submit some applications, and hence miss some high-value matches. Phone calls might provide reminders that reduce the share of forgotten applications.

But the reminder interpretation is inconsistent with results from three mechanism tests. First, in a subset of rounds, we send a second text message as a reminder to a random subsample of control group jobseekers. Table C.4 shows that the effect of the reminder message is one-fourteenth as large as the effect of the phone call ($p < 0.001$). Second, we randomize the timing of the phone call within the application window. If the call functions as a reminder, treatment should have a larger effect for jobseekers called later, who have had more time to forget to apply. Instead, treatment effects are smaller when phone calls occur later (Table C.5, column 1). Third, treatment has a smaller rather than larger effect when there is a longer window between the initial text message and application deadline for non-experimental reasons such as staff capacity (column 2).

Notably, the fact that the treatment effect on applications declines with time to deadline is not only inconsistent with forgetting, but also consistent with time-varying psychological costs: if the window is longer, there is a greater chance that a jobseeker will face a low psychological cost of applying at some point during the window and hence apply. This evidence against a reminder interpretation does not rule out all roles for present bias or procrastination; these may be barriers to applications in ways that are not addressed by reminders.

We can also reject an explanation that treatment increases applications because call center agents encourage or pressure jobseekers to apply. First, platform staff are explicitly trained not to encourage or pressure jobseekers to apply, and regular audits of call recordings verified that they followed their scripts. Second, we find no evidence that treated jobseekers apply to the first job listed on the call at a higher rate than control jobseekers, which would be the lowest-cost way for jobseekers to respond to pressure to apply to something (Figure C.7). Similarly, if jobseekers applied more simply to avoid pressure from call center agents, we might expect marginal applications to go to poor matches and produce decreasing returns to search, which does not match our results. This evidence against encouragement effects does not rule out a role for fear of rejection, as the phone calls may reduce the cost of this fear by reducing time spent anticipating rejection.

We view psychological costs of initiating applications as a reasonable explanation for our main experimental findings given the model and additional tests in Sections 4.1 and 4.2, the evidence against pecuniary and time costs in the Section 4.3, the importance of psychological costs in prior research on related decision-making, and the evidence against alternative explanations we discuss in the next section. However, we acknowledge that we cannot pin exactly which psychological cost(s) are reduced by the phone call treatment, so alternative explanations remain possible.

### 4.5 Role of Jobseeker Selection

Random assignment means that there are no systematic differences between jobseekers in the treatment and control groups. But job applications are endogenous decisions, so there may be systematic differences between the treated and control jobseekers who submit applications. This would not bias any of the treatment effects that we estimate, but it would change the interpretation of the estimates. In particular, roughly constant returns might arise if each individual jobseeker experiences decreasing returns to additional search effort but treatment-induced applications come from jobseekers who are positively selected on education, experience, etc. Here we show that observed jobseeker characteristics do not explain the heterogeneity that we hypothesize in the model. This suggests that heterogeneity within jobseeker over time is a more likely explanation.

First, and perhaps most importantly, we reduce the scope for jobseeker selection by running a within-jobseeker version of our experiment. This uses a "crossover" design that randomly reassigns some control group jobseekers to the treatment group in some matching rounds. Only $0.65\%$ of matches are affected by this treatment, so it has minimal impact on the overall design. It allows us to replicate our main analysis with jobseeker fixed effects, using only within-jobseeker variation to identify treatment effects. Table B.2 shows that the results of this experiment are very similar to our main results. In particular, we cannot reject equality of the interview rates for inframarginal applications and marginal applications submitted due to the crossover treatment (p = 0.503).

Second, we control for jobseeker selection on observed characteristics. We repeat our analysis of the main experiment with controls using a post-double selection LASSO selecting from an extensive set of jobseeker baseline characteristics, following Belloni et al. (2014). Table B.5 shows that the point estimates and standard errors are almost identical.

Third, we show that marginal and inframarginal applications come from jobseekers with similar observed characteristics. This uses a complier or latent type analysis in a similar spirit to Abadie (2003) that we describe in Appendix C.2. This estimates the mean characteristics of jobseeker $\times$ vacancy matches that get applications only when the jobseeker is treated (marginal applications or 'complier matches') and matches that get applications whether or not the jobseeker is treated (inframarginal applications or 'always-taker' matches). Comparing these means shows if the average marginal application and average inframarginal application come from jobseekers with different observed characteristics. Table 4 shows that mean education and and CV quality scores (provided

Table 4: Comparing Observed Characteristics of Jobseekers Submitting Marginal and Inframarginal Applications

| | (1) Inframarginal applications | (2) Marginal applications | (3) Difference (p-value) |
|---|---|---|---|
| Years of education | 13.409 | 13.401 | -0.008 |
| | | | (0.989) |
| Years of work experience | 7.472 | 8.601 | 1.129 |
| | | | (0.102) |
| CV Score excellent | 0.297 | 0.295 | -0.002 |
| | | | (0.985) |
| CV Score good | 0.386 | 0.366 | -0.020 |
| | | | (0.826) |
| CV Score average or lower | 0.317 | 0.338 | 0.021 |
| | | | (0.793) |
| $\hat{P} \mid X_j$: Prob. interview \| jobseeker characteristics | 0.063 | 0.067 | 0.004 |
| | | | (0.179) |

Notes: Table shows the means of covariates for the inframarginal applications that are submitted without treatment (column 1) and marginal applications that are submitted due to treatment (column 2). Column 3 shows the difference between the covariate means for marginal and inframarginal applications with p-values in parentheses, estimated using heteroskedasticity-robust standard errors clustered by jobseeker. The unit of observation is the jobseeker $\times$ vacancy match. The predicted interview probabilities in the final row are estimated using a logit LASSO specification with the sample of applications from the control group jobseekers. The logit LASSO model is allowed to select from the following baseline jobseeker characteristics: completed CV, total # of occupational preferences selected, greater than median number of occupational preferences selected, age, education level indicators, years of work experience, currently studying, any work experience, female, female and married, female and has children, female and has a child age $< 5$, employed and searching, employed and not searching, searching and not employed, not employed and not searching, indicators for each reported job search method used, and expected salary less than 90th percentile of salaries the jobseeker is matched to on platform. The CV quality score variables are not included in the interview probability prediction because they are only observed for the 15% of jobseekers who are matched with vacancies for which the hiring managers shared their CV evaluations.

by firms, as discussed in Section 2.3) are almost identical for the marginal and inframarginal applications. Marginal applications come from jobseekers with slightly more work experience. But, as we note above, our main findings are unchanged when we control for experience.[24]

Fourth, we show that marginal and inframarginal applications come from jobseekers with similar latent interview probabilities. We do this by estimating latent interview probabilities using

[24]This complier analysis is conceptually different to a heterogeneous treatment effects approach, which compares the magnitude of treatment effects by values of observed characteristics. For interested readers, we also estimate heterogeneous treatment effects on applications by CV quality. The estimated effects are slightly but not statistically significantly larger for lower-quality CVs (Table B.4). This also suggests that marginal applications do not come from observably stronger jobseekers than inframarginal applications, and hence cannot explain the roughly equal returns to marginal and inframarginal applications.

a data-driven approach and then using the complier analysis approach described in the previous paragraph to compare latent probabilities between the inframarginal and marginal applications. Specifically, we first restrict the sample to the set of applications from control group jobseekers, i.e. jobseeker × vacancy matches with $T = 0$ and $Apply = 1$. We then regress *Interview* on a vector of jobseeker characteristics using a logit LASSO and predict $\hat{P}|X_j = \hat{Pr}(\textit{Interview} \mid \textit{Apply} = 1, X_j)$ for each jobseeker $j$. This is the probability the jobseeker will get an interview if she applies, given her observed characteristics.[25] The final row of Table 4 shows that the mean of this measure does not differ between marginal and inframarginal applications.

Taken together, these results show that the roughly constant return to treatment-induced job search is not explained by treatment changing patterns of jobseeker selection into applications. This suggests that heterogeneity within jobseeker over time is a more likely explanation.

# 5 Evaluating Alternative Explanations

## 5.1 Differential Access to Information

Our main findings could be driven by phone calls providing more information about specific jobs, both increasing application rates and enabling jobseekers to better target high-return vacancies. Although the text message and phone call scripts contain identical information about each vacancy, phone calls introduce the possibility that jobseekers request more information about the position from agents. However, in approximately 80% of matching rounds, no additional information beyond the content of the text message script was available to call center agents, and all our results hold when restricting the sample to these rounds (Table C.6).

Another possibility is that jobseekers are simply more likely to receive phone calls than text messages; for example, if text messages are sometimes blocked or simply go unread. To test this, we survey jobseekers to ask them if they remember receiving a recent job match from the platform by either phone call or text message. Treatment and control jobseekers are equally likely to report that they received a match, with or without the sample selection correction for survey non-response described in Section 3.2. See Table C.7 for results and measurement details. In addition, treatment effects do not differ between the 93% of respondents who indicated at registration that they are comfortable communicating with the platform by text message and the remaining 7%.

## 5.2 Changes in Perceived Returns To Search

The phone call treatment might increase jobseekers' perceived value of jobs on the platform if a call from a professional recruiting service signals that platform firms are larger or wealthier and

[25]This approach assumes that the relationship between interviews and observed characteristics does not differ for marginal and inframarginal applications, as we use the inframarginal applications for estimation and then predict out-of-sample to the marginal applications. This assumption is more reasonable in this application than many others because the platform observes and controls all information sent by the jobseeker to the firm.

thus able to provide more benefits or opportunities for advancement (higher $V$). Alternatively, it might signal to the jobseeker that the firm sees her as a good fit for the job (higher $P$).

We first note that higher perceived $P$ and $V$ should increase the application rate but are unlikely to generate constant returns to marginal search, particularly because we showed in Section 4.2 that treatment does not substantially change how jobseekers direct applications.[26] We also directly test this channel by collecting data on jobseekers' beliefs about $P$ and $V$ and estimating treatment effects on these two belief measures.[27] Table C.8 shows that both results are close to zero and not statistically significant. We view belief updating as a distinct explanation from fear of rejection, which refers to the utility cost of fearing that an application will be rejected, conditional on the belief that it will be rejected.

### 5.3 Random Search

If jobseekers apply to vacancies at random and the phone call treatment reduces the cost of applying, then treatment should increase the application rate and yield constant returns to marginal applications. Random job search may seem implausible. But it has been widely assumed in canonical search models, even if only as a simplifying benchmark (Pissarides, 2000). It may also be plausible given some empirical evidence that jobseekers have limited information about labor market conditions and match quality (Behaghel et al., 2020; Belot et al., 2018).

However, the random search framework does not match two additional results from our platform. First, we showed in Sections 2.4 and 4.2 that applications are directed toward vacancies with higher $V$, showing that applications are not random. Second, we run an additional experiment designed to induce random search in order to compare that to marginal search effort induced by the phone call treatment. Specifically, in 20% of rounds we randomize the order in which vacancies are listed on both text messages and phone calls, which encourages additional applications to the randomly-chosen vacancies that are listed first.[28]

Listing vacancies first produces more applications with decreasing, rather than constant, returns. Table C.9 shows that the probability of application is 0.5 p.p. higher for vacancies listed first

[26] Consider a simple static rule in which jobseekers apply to all vacancies with $\tilde{P} \cdot \tilde{V} > C$, for perceived interview probability $\tilde{P}$ and perceived interview value $\tilde{V}$. If the phone call treatment raises $\tilde{P}$ or $\tilde{V}$ for all vacancies, then jobseekers will apply to vacancies with lower $P \cdot V$, producing decreasing returns to additional search effort.

[27] We ask: "Suppose Job Talash sends you one hundred job ads over a year. Based on your past experience with our job matching service, how many of these ads do you think would be desirable for you?" and "Suppose you apply for all the jobs you think are desirable jobs. How many of those do you think would make you an offer?"

[28] Vacancies listed earlier might attract more applications because applying to them takes less time or because jobseekers interpret the ordering as a signal of job quality or attainability. Results from the analysis are similar if we use only the 20% of rounds with randomized order or use all rounds and control for firm fixed effects, as firm identifiers determined vacancy order in non-randomized rounds. Order of job listing is uncorrelated with job and jobseeker baseline characteristics conditional on these fixed effects. Results are similar when we compare only the first job to all subsequent jobs or include order indicators. We restrict the sample to jobseeker $\times$ round units in which the jobseeker matched with more than one vacancy, which is necessary for variation in vacancy order.

instead of second or later (column 1). Moreover, the average interview probability for marginal applications submitted because the vacancy was listed first is 2.4% (column 4), which is substantially lower than the average interview probabilities for both inframarginal applications and marginal applications submitted due to the phone call treatment. This contrast suggests that the main phone call treatment is not inducing random search, consistent with the fact that 69% of applications induced by the phone call are sent to vacancies listed second or later.

The result of this experiment emphasizes that the return to marginal search depends on which intervention causes the marginal search and how it is directed. The randomized order treatment causes marginal search that is roughly randomly directed and has sharply decreasing returns. The phone call initiation treatment causes marginal search that is directed in similar ways to inframarginal search and has roughly equal returns. This highlights that the constant returns finding is a consequence of the type of search induced by a particular intervention, not inherent to this labor market or these jobseekers.

# 6 Spillover Effects

Increased search effort by some jobseekers may affect firms and other jobseekers. For firms, the sign of this effect is theoretically ambiguous: receiving more applications can increase the probability of receiving an application from a well-matched applicant and hence making a hire, but it can also generate congestion costs if firms need to review many poorly-matched applications. For other jobseekers, spillover effects are unlikely to be positive: competing against more applications can lead to crowd-out. But the magnitude of crowd-out may be small and offset if firms increase total hiring when they receive more applications.

We can identify spillover effects using variation in the vacancy-level treatment rate: the share of users matched to each vacancy who are treated. This share is random because matches are determined by pre-treatment characteristics (education, work experience, and occupational preferences). Our approach is analogous to papers that study spillovers using variation in treatment intensity within geographic labor markets (e.g. Blundell et al. 2004; Gautier et al. 2018; LaLive et al. 2022). This approach works well because this platform's matching structure fully determines the set of platform users who can compete with each other for each vacancy. This approach is not feasible for jobseeker-facing experiments on most platforms, where users can search and apply for many different jobs. This makes it difficult to define how much each user is competing with other users without a full model of the job search process.

We first verify that the experiment generates enough variation across vacancies in the treatment rate to identify spillovers. The percentage of matches that are treated has interdecile range across vacancies of [0.38,0.55], interquartile range [0.43,0.52], and standard deviation 0.079 (shown in Figure D.1). Vacancies matched to fewer jobseekers mechanically have more dispersed treatment rates, due to small-sample variation. But even vacancies with above-median numbers of matched jobseekers have standard deviation 0.054 in their treatment rates.

We estimate spillover effects using two methods. Our first method tests whether jobseeker-level outcomes are sensitive to the fraction of competing jobseekers who are treated, closely following Crepon et al. (2013). We define $TR_{jv}$ as the fraction of jobseekers matched to vacancy $v$ who are treated, excluding jobseeker $j$. This measures the treatment rate for jobseekers potentially competing against $j$ at vacancy $v$. We use match-level data to regress interview invitations on jobseeker-level treatment status, the treatment rate defined above and their interaction:

$$\text{Interview}_{jv} = T_j \cdot \beta_1 + TR_{jv} \cdot \beta_2 + T_j \cdot TR_{jv} \cdot \beta_3 + \mathbf{X}_v \cdot \Lambda + \mu_b + \epsilon_{jv}, \tag{4}$$

where $\mathbf{X}_v$ contains the number of jobseekers matched to vacancy $v$ and vacancy-level factors that determine matches (e.g. occupation) and $\mu_b$ is a stratification block fixed effect. We cluster standard errors by both jobseeker and vacancy because treatment is assigned at the jobseeker level and most of the variation in $TR_{jv}$ is across vacancies. Finding $\beta_2 < 0$ would be evidence of negative spillover effects on control group jobseekers, as it would show lower interview probabilities when more competing jobseekers are treated. Finding $\beta_2 + \beta_3 < 0$ would be evidence of negative spillovers on treated jobseekers. This method has an intention-to-treat spirit, as it uses only information on treatment assignments and matches, not application decisions.

We do not find evidence of negative spillover effects using this first method. Estimates of $\beta_2$ and $\beta_3$ are both small and not statistically significant (Table 5, column 1). To interpret their magnitude, we consider the effect on a jobseeker's interview probability of moving from the 25th to 75th percentile of $TR_{jv}$, the treatment exposure rate. This effect is $0.006$ percentage points for a control group jobseeker (standard error $0.011$ p.p., p = $0.589$) and $-0.011$ p.p. for a treatment group jobseeker (standard error $0.017$ p.p., p = $0.511$). As a benchmark, the effect of a jobseeker's own treatment status on interview invitations is substantially larger: 0.078 p.p. (from Table 3).

Equation (4) imposes a linear relationship. But spillover effects might be nonlinear and only substantial at high treatment rates. To test this idea, we repeat the analysis replacing the vacancy-level treatment rate $TR_{jv}$ with indicators for the middle and top terciles of the treatment rate. These effects are again close to zero for control or treatment group jobseekers (column 2).

Our second method tests if vacancy-level treatment effects vary with vacancy-level treatment rates, closely following Ferracci et al. (2014). For each of the 1,340 vacancies, we estimate the treatment effect on interview invitations, $\Delta\text{Interview}_v$, and the treatment rate for matched jobseekers, $TR_v$. We use these vacancy-level data points to estimate

$$\Delta\text{Interview}_v = TR_v \cdot \alpha + \mathbf{X}_v \cdot \Lambda + \varepsilon_v, \tag{5}$$

conditional on the same vacancy-level covariates $\mathbf{X}_v$ as the previous analysis. Finding $\alpha < 0$

Table 5: Spillover Effects Between Jobseekers

| | Method 1: Match-level | | Method 2: Vacancy-level | |
|---|---|---|---|---|
| | Interview | | Interview effect | |
| | (1) | (2) | (3) | (4) |
| Treatment | 0.00196 | 0.00100 | | |
| | (0.00084) | (0.00021) | | |
| Treatment rate† | 0.00085 | | | |
| | (0.00158) | | | |
| Treatment X treatment rate† | -0.00248 | | | |
| | (0.00175) | | | |
| Treatment rate† : mid tercile | | 0.00019 | | |
| | | (0.00014) | | |
| Treatment rate† : top tercile | | 0.00019 | | |
| | | (0.00023) | | |
| Treatment X treatment rate† : mid tercile | | -0.00031 | | |
| | | (0.00026) | | |
| Treatment X treatment rate† : top tercile | | -0.00030 | | |
| | | (0.00026) | | |
| Treatment rate | | | 0.00196 | |
| | | | (0.00117) | |
| Treatment rate: middle tercile | | | | 0.00022 |
| | | | | (0.00021) |
| Treatment rate: top tercile | | | | 0.00050 |
| | | | | (0.00031) |
| Outcome mean | 0.0005 | 0.0005 | 0.0004 | 0.0004 |
| Exposure regressor mean | 0.4688 | | 0.4752 | |
| Exposure regressor SD | 0.0558 | | 0.0799 | |
| p: treated terciles equal | | 0.412 | | |
| p: control terciles equal | | 0.403 | | |
| p: terciles equal | | | | 0.245 |
| # observations | 1116446 | 1116446 | 1340 | 1340 |

Notes: This table shows the results of tests for spillovers between jobseekers on interview invitations. Column (1) shows results from regressing match-level interview invitations on own treatment status, the fraction of other jobseekers matched to the same vacancy who are treated, and their interaction. Column (2) shows results from a regression that replaces the fraction of other jobseekers who are matched to the same vacancy with terciles for the middle and top terciles of this fraction. The $p$-values below the regression output are for tests of no spillovers onto treated jobseekers ('p: treated terciles equal') and control jobseekers ('p: control terciles equal'). Column (3) shows results from regressing vacancy-level treatment effects on interview invitations on vacancy-level fractions of matches that are treated. Column (4) shows results from regressing vacancy-level treatment effects on interview invitations on the middle and top terciles of vacancy-level fractions of matches that are treated. The $p$-value below the regression output is for a test that the treatment effects do not vary with treatment rate ('p: terciles equal'). All regressions condition on firm size and sector and vacancy occupation, posted salary, education and experience requirements, and number of matched jobseekers. Columns (1) and (2) also condition on stratification block fixed effects. Heteroskedasticity-robust standard errors are shown in parentheses, clustered by jobseeker and vacancy in columns (1) and (2). Outcome and treatment rate means are for the full sample. Variables marked with † are leave-one-out averages that omit the jobseeker's own values.

would be evidence of negative spillover effects, as this would show a smaller treatment effect on each jobseeker's interview probability at vacancies receiving more treatment-induced applications.

We do not find evidence of negative spillover effects using this second method. Instead, we find a positive but small estimate of $\alpha$ (Table 5, column 3). To interpret the magnitude, we note that this coefficient implies that a vacancy exposed to the 75th percentile of the treatment rate $TR_v$ rather than the 25th percentile would have a $0.018$ percentage point higher treatment effect on interviews (standard error $0.011$ p.p., p = $0.096$). To test for a nonlinear relationship, we repeat this analysis replacing the vacancy-level treatment rate with indicators for the middle and top terciles of the treatment rate. These coefficients are again positive (column 4). Vacancies with top-tercile rather than bottom-tercile treatment rates have 0.05 p.p. higher treatment effects on interviews, although we cannot reject the null hypothesis that treatment effects are equal across all three terciles. A nonparametric regression of vacancy-level treatment effects on treatment rates also shows no evidence of negative spillover effects (Figure D.2).

The lack of negative spillovers is consistent with descriptive patterns in vacancy-level outcomes. If firms dislike congestion, then the relationship between application and interview numbers might be non-monotonic: a small increase in the number of applications might lead to more interview invitations and a large increase might lead to fewer interview invitations. At the extreme, a very high number of applications might lead firms to ignore all applications and make no interview invitations. Instead, vacancy-level regressions show that both the number of interviews and the probability of interviewing any jobseeker are monotonically increasing in the number of applications (Table D.1).

What might explain the negligible spillover effects we find? Our design cannot directly answer this question, but we suggest three possible explanations. First we note that spillover effects might be zero if firms hire more when they receive more applications above their reservation hiring quality. Carranza et al. (2021) and Fernando et al. (2021) show indirect evidence consistent with this mechanism. Second, firms in this context report filling only 60% of vacancies, so more offers need not mechanically lead to crowd-out. Third, application volumes on this platform are relatively low: the average vacancy receives only 0.8 applications from control group applicants and another 6 applications from treated applicants (with pooled interdecile range 0-18). Firms report in surveys that they get on average 30% of their total applications through the platform. Taking these factors together, it is possible that firms in this labor market receive too few suitable applications in the absence of treatment for crowd-out to be relevant, at least at the interview stage.

# 7 Conclusion

We show that job search effort can be dramatically increased by reducing the psychological cost of initiating job applications. Returns to the additional search effort are constant rather than decreasing, in contrast with many intuitive job search models. This pattern is consistent with a model in which marginal applications in any period are a mix of lower-return applications from jobseekers who would send some applications without treatment and higher-return applications from jobseekers who would not apply without treatment, at least at that period in time. This finding of constant returns, combined with limited spillovers on other jobseekers, suggests the possibility of suboptimally low search effort. This echoes findings that changing default options to avoid initiation costs can lead to economically significant increases in financial and health investments (DellaVigna, 2009; DellaVigna & Malmendier, 2006; Madrian & Shea, 2001; Thaler & Benartzi, 2004). Our findings are particularly striking because this is a platform designed to have minimal pecuniary, time, and technology barriers to use and hence to be broadly accessible to jobseekers in a low-resource setting. Yet psychological costs of initiating applications still present a significant barrier for jobseekers on the platform.

These findings link to a broader literature around the design of job search policy and platforms. The possibility that psychological costs lead to suboptimally low search effort has implications for policies such as using caseworkers to increase jobseekers' accountability and motivation, subsidising job search, requiring active search for unemployment insurance recipients, or automatically enrolling jobseekers in search assistance services (Card et al., 2010, 2018). Job search and matching platforms could also encourage search by simplifying the process of evaluating job listings or encouraging jobseekers to start applications, although the value of such design changes may be low on platforms with already-high application volumes.

## References


Abadie, A. (2003) "Semiparametric instrumental variable estimation of treatment response models," *Journal of Econometrics*, 113 (2), 231–263.

Abadie, A., S. Athey, G. Imbens & J. Wooldridge (2017) "When Should You Adjust Standard Errors for Clustering?" Working Paper 24003, National Bureau of Economic Research.

Abebe, G., S. Caria, M. Fafchamps, P. Falco, S. Franklin & S. Quinn (2021a) "Anonymity or Distance? Job Search and Labour Market Exclusion in a Growing African City," *Review of Economic Studies*, 88 (3).

Abebe, G., S. Caria, M. Fafchamps, P. Falco, S. Franklin, S. Quinn & F. Shilpi (2021b) "Matching Frictions and Distorted Beliefs: Evidence from a Job Fair Experiment," Working paper, University of Oxford.

Abebe, G., S. Caria & E. Ospina-Ortiz (2019) "The Selection of Talent: Experimental and Structural Evidence from Ethiopia," *American Economic Review*, 111 (6).

Abel, M., R. Burger, E. Carranza & P. Piraino (2019) "Bridging the Intention-Behavior Gap? The Effect of Plan-Making Prompts on Job Search and Employment," *American Economic Journal: Applied Economics*, 11 (2), 284–301.

Abel, M., R. Burger & P. Piraino (2020) "The Value of Reference Letters: Experimental Evidence from South Africa," *American Economic Journal: Applied Economics*, 12 (3), 40–71.

Afridi, A., Farzana ad Dhillon, S. Roy & N. Sangwan (2022) “Social Networks, Gender Norms and Women’s Labor Supply: Experimental Evidence using a Job Search Platform,” Working paper, Indian Statistical Institute.

Agrawal, A., J. Horton, N. Lacetera & E. Lyons (2015) “Digitization and the contract labor market: A research agenda,” in Goldfarb, A., S. Greenstein & C. Tucker eds. *Economic analysis of the digital economy*, 219–250: University of Chicago Press.

Alfonso Naya, V., G. Bied, P. Caillou, B. Crepon, C. Gaillac, E. Perennes & M. Sebag (2020) “Designing labor market recommender systems: the importance of job seeker preferences and competition,” Manuscript, LISN.

Altmann, S., A. Falk, S. Jäger & F. Zimmermann (2018) “Learning about Job Search: A Field Experiment with Job Seekers in Germany,” *Journal of Public Economics*, 164, 33–49.

Altmann, S., A. Glenny, R. Mahlstedt & A. Sebald (2022) “The Direct and Indirect Effects of Online Job Search Advice,” IZA Discussion Paper 15830.

Anderson, M. (2008) “Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects,” *Journal of the American Statistical Association*, 103 (484), 1481–1495.

Archibong, B., F. Annan, O. Okunogbe & I. Oliobi (2022) “Firm Culture: The Effects of Information Interventions on Gender Gaps in Online Labor Markets,” Working paper.

Arni, P. & A. Schiprowski (2019) “Job search requirements, effort provision and labor market outcomes,” *Journal of Public Economics*, 169, 65–88.

Attanasio, O., A. Kugler & C. Meghir (2011) “Subsidizing Vocational Training for Disadvantaged Youth in Colombia: Evidence from a Randomized Trial,” *American Economic Journal: Applied Economics*, 3 (3), 188–220.

Babcock, L., W. J. Congdon, L. F. Katz & S. Mullainathan (2012) “Notes on behavioral economics and labor market policy,” *IZA Journal of Labor Policy*, 1, 2.

Baker, S. R. & A. Fradkin (2017) “The Impact of Unemployment Insurance on Job Search: Evidence from Google Search Data,” *The Review of Economics and Statistics*, 99, 756–768.

Bandiera, O., V. Bassi, R. Burgess, I. Rasul, M. Sulaiman & A. Vitali (2021) “The Search for Good Jobs: Evidence from a Six-year Field Experiment in Uganda,” Working paper.

Banerjee, A. & S. Sequeira (2020) “Spatial Mismatches and Imperfect Information in the Job Search,” Discussion Paper 14414, Centre for Economic Policy Research.

Banfi, S., S. Choi & B. Villena-Roldan (2019) “Deconstructing Job Search Behavior,” Working Paper 3323545, Social Science Research Network.

Banfi, S., S. Choi & B. Villena-Roldán (2022) “Sorting on-line and on-time,” *European Economic Review*, 146.

Bassi, V. & A. Nansamba (2020) “Screening and Signaling Non-Cognitive Skills: Experimental Evidence from Uganda,” Manuscript, University of Southern California.

Beam, E. (2016) “Do Job Fairs Matter? Experimental Evidence from the Philippines,” *Journal of Development Economics*, 120, 32–40.

Behaghel, L., S. Dromundo, M. Gurgand & Y. Hazard (2020) “Directing job search: a large scale experiment,” Manuscript, Paris School of Economics.

Belloni, A., V. Chernozhukov & C. Hansen (2014) “Inference on treatment effects after selection among high-dimensional controls,” *The Review of Economic Studies*, 81 (2), 608–650.

Belot, M., P. Kircher & P. Muller (2018) “Providing Advice to Jobseekers at Low Cost: An Experimental Study on Online Advice,” *Review of Economic Studies*, 86 (4), 1411–1447.

——— (2022a) "Do the Long-term Unemployed Benefit from Automated Occupational Advice during Online Job Search?", Working paper, Cornell.

——— (2022b) "How Wage Announcements Affect Job Search—A Field Experiment," *American Economic Journal: Macroeconomics*, 14 (4), 1–67.

Ben Dhia, A., B. Crepon, E. Mbih, L. Paul-Delvaux, B. Picard & V. Pons (2022) "Can a Website Bring Unemployment Down? Experimental Evidence from France," Working paper.

Bhuller, M., A. Kostøl & T. Vigtel (2019) "How Broadband Internet Affects Labour Market Matching," Working paper.

Blundell, R., M. C. Dias, C. Meghir & J. Van Reenen (2004) "Evaluating the Employment Impact of a Mandatory Job Search Program," *Journal of the European Economic Association*, 2 (4), 569–606.

Bolhaar, J., N. Ketel & B. van der Klaauw (2020) "Caseworker's discretion and the effectiveness of welfare-to-work programs," *Journal of Public Economics*, 183, 104080.

Boudreau, L., R. Heath & T. McCormick (2022) "Migrants, Experience, and Working Conditions in Bangladeshi Garment Factories," Working paper.

Brown, M., C. J. Flinn & A. Schotter (2011) "Real-Time Search in the Laboratory and the Market," *American Economic Review*, 101 (2), 948–74.

Caliendo, M., D. Cobb-Clark & A. Uhlendorff (2015) "Locus of Control and Job Search Strategies," *The Reivew of Economics and Statistics*, 97 (1), 88–103.

Card, D. (2001) "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems," *Econometrica*, 69 (5), 1127–1160.

Card, D., J. Kluve & A. Weber (2010) "Active Labour Market Policy Evaluations: A Meta-Analysis," *The Economic Journal*, 120, F452–F477.

——— (2018) "What Works? A Meta-Analysis of Recent Active Labor Market Program Evaluations," *Journal of the European Economic Association*, 16 (3), 894–931.

Caria, S., G. Gordon, M. Kasy, S. Quinn, S. Shami & A. Teytelboym (2023) "An Adaptive Targeted Field Experiment: Job Search Assistance for Refugees in Jordan," *Journal of the European Economic Association*.

Carranza, E., R. Garlick, K. Orkin & N. Rankin (2021) "Job Search and Hiring with Limited Information about Workseekers' Skills," Manuscript, Duke University.

Carroll, G., J. CHoi, D. Laibson, B. Madrian & A. Metrick (2009) "Optimal Defaults and Active Decisions," *Quarterly Journal of Economics*, 124 (4), 1639–1674.

Chakravorty, B., A. Y. Bhatiya, C. Imbert, M. Lohnert, P. Panda & R. Rathelot (2023) "Impact of the COVID-19 Crisis on India's Rural Youth: Evidence from a Panel Survey and an Experiment," *World Development*, 168.

Chiplunkar, G. & P. K. Goldberg (2022) "The Employment Effects of Mobile Internet in Developing Countries," Working Paper 30741, National Bureau of Economic Research.

Cooper, M. & P. Kuhn (2020) "Behavioral Job Search," in Zimmermann, K. F. ed. *Handbook of Labor, Human Resources and Population Economics*, 1–22, Cham: Springer International Publishing.

Crepon, B., E. Duflo, M. Gurgand, R. Rathelot & P. Zamora (2013) "Do Labor Market Policies Have Displacement Effects? Evidence from a Clustered Randomized Experiment," *Quarterly Journal of Economics*, 128 (2), 531–580.

Dammert, A., J. Galdo & V. Galdo (2015) "Integrating Mobile Phone Technologies into Labor-market Intermediation: a Multi-treatment Experimental Design," *IZA Journal of Labor and De-*

*velopment*, 4.
De Chaisemartin, C. (2017) "Tolerating defiance? Local average treatment effects without monotonicity," *Quantitative Economics*, 8, 367–396.
DellaVigna, S. (2009) "Psychology and Economics: Evidence from the Field," *Journal of Economic Literature*, 47 (2), 315–72.
DellaVigna, S., J. Heining, J. Schmieder & S. Trenkle (2022) "Evidence on Job Search Models from a Survey of Unemployed Workers in Germany," *Quarterly Journal of Economics*, 137, 1181–1232.
DellaVigna, S., A. Lindner, B. Reizer & J. F. Schmieder (2017) "Reference-Dependent Job Search: Evidence from Hungary," *The Quarterly Journal of Economics*, 132, 1969–2018.
DellaVigna, S. & U. Malmendier (2006) "Paying Not to Go to the Gym," *American Economic Review*, 96 (3), 694–719.
DellaVigna, S. & M. D. Paserman (2005) "Job Search and Impatience," *Journal of Labor Economics*, 23, 527–588.
DiNardo, J., J. Matsudaira, J. McCrary & L. Sanbonmatsu (2021) "A Practical Proactive Proposal for Dealing with Attrition: Alternative Approaches and an Empirical Example," *Journal of Labor Economics*, 39, S507–S541.
Doniger, C. & D. Toohey (2022) "These Caps Spilleth Over: Equilibrium Effects of Unemployment Insurance," FEDS Working Paper 2022-74.
Donovan, K., J. Lu, T. Schoellman et al. (2018) "Labor Market Flows and Development," in *2018 Meeting Papers*, 976, Society for Economic Dynamics.
Duflo, E., M. Kremer & J. Robinson (2011) "Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya," *The American Economic Review*, 101, 2350–2390.
Ericson, K. (2017) "On the Interaction of Memory and Procrastination: Implications for Reminders, Deadlines, and Empirical Estimation," *Journal of the European Economic Association*, 15 (3), 692–719.
Ericson, K. M. & D. Laibson (2019) "Intertemporal Choice," in Bernheim, B. D., S. DellaVigna & D. Laibson eds. *Handbook of Behavioral Economics - Foundations and Applications 2*, 2 of Handbook of Behavioral Economics: Applications and Foundations 1, 1–67: North-Holland.
Faberman, J. & M. Kudlyak (2019) "The Intensity of Job Search and Search Duration," *American Economic Journal: Macroeconomics*, 3 (11), 327–357.
——— (2006b) "Self-Confidence and Search," IZA Discussion Paper 2525.
Falk, A., D. Huffman & U. Sunde (2006a) "Do I Have What It Takes? Equilibrium Search with Type Uncertainty and Non-Participation," Discussion paper 2531, Institute for the Study of Labor.
Fernando, N., N. Singh & G. Tourek (2021) "Hiring Frictions in Urban Labor Markets: Experimental Evidence from India," Working paper, Notre Dame.
Ferracci, M., G. Jolivet & G. J. van den Berg (2014) "Evidence of Treatment Spillovers Within Markets," *Review of Economics and Statistics*, 96, 812–823.
Filges, T., G. Smedslund, A. D. Knudsen & A. K. Jørgensen (2015) "Active Labour Market Programme Participation for Unemployment Insurance Recipients: A Systematic Review," *Campbell Systematic Reviews*, 11, 1–342.
Franklin, S. (2017) "Location, Search Costs and Youth Unemployment: Experimental Evidence from Transport Subsidies," *The Economic Journal*, 128 (614), 2353–2379.
Fu, J., M. Sefton & R. Upward (2019) "Social Comparisons in Job Search," *Journal of Economic*

*Behavior & Organization*, 168, 338–361.
Gabaix, X. (2019) "Behavioral Inattention," in Bernheim, D., S. DellaVigna & D. Laibson eds. *Handbook of Behavioral Economics*, 261–343: Elsevier.
Garlick, R. & J. Hyman (2022) "Quasi-Experimental Evaluation of Alternative Sample Selection Corrections," *Journal of Business and Economic Statistics*, 40 (3), 101–125.
Gautier, P., P. Muller, B. van der Klaauw, M. Rosholm & M. Svarer (2018) "Estimating Equilibrium Effects of Job Search Assistance," *Journal of Labor Economics*, 36 (4), 1073–1125.
Gee, L. (2019) "The More You Know: Information Effects on Job Application Rates in a Large Field Experiment," *Management Science*, 65 (5), 2077–2094.
Gentile, E., N. Kohli, N. Subramanian, Z. Tirmazee & K. Vyborny (2023) "A Leaky Pipeline? Decomposing the gender gap in job search in urban Pakistan," Working paper.
Government of Bangladesh (2015) "Bangladesh Labor Force Survey 2015."
Government of Namibia (2016) "Namibia Labor Force Survey 2016."
Government of South Africa (2018) "South Africa Quarterly Labour Force Survey 2013-2018."
Gurtzgen, N., L. Benjamin, L. Pohlan & G. van den Berg (2020) "Does Online Search Improve the Match Quality of New Hires?", Working paper.
He, H., D. Neumark & Q. Weng (2021) ""I Still Haven't Found What I'm Looking For": Evidence of Directed Search from a Field Experiment," Working Paper 28660, National Bureau of Economic Research.
Heckman, J. (1974) "Shadow Prices, Market Wages, and Labor Supply," *Econometrica*, 42 (4), 679–694.
Hjort, J. & J. Poulsen (2019) "The Arrival of Fast Internet and Employment in Africa," *American Economic Review*, 109, 1032–1079.
Horton, J. & S. Vasserman (2021) "Job-Seekers Send Too Many Applications: Experimental Evidence and a Partial Solution," Manuscript, MIT.
Imbens, G. & D. Rubin (2015) *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*: Cambridge University Press.
Johnston, A. C. & A. Mas (2018) "Potential Unemployment Insurance Duration and Labor Supply: The Individual and Market-Level Response to a Benefit Cut," *Journal of Political Economy*, 126, 2480–2522.
Jones, S. & K. Sen (2022) "Labour Market Effects of Digital Matching Platforms: Experimental Evidence from Sub-Saharan Africa," 07.
Kelley, E. M., C. Ksoll & J. Magruder (2021) "How do Online Job Portals Affect Employment and Search Outcomes? Evidence from India," Working paper.
Kiss, A. ⓡ R. Garlick ⓡ K. Orkin ⓡ L. Hensel (2023) "Jobseekers' Beliefs about Comparative Advantage and (Mis)Directed Search," Working paper, Duke University.
Kleibergen, F. & R. Paap (2006) "Generalized Reduced Rank Tests Using the Singular Value Decomposition," *Journal of Econometrics*, 133 (1), 97–126.
Köszegi, B., G. Loewenstein & T. Murooka (2022) "Fragile Self-Esteem," *Review of Economic Studies*, 89, 2026–2060.
Kremer, M., G. Rao & F. Schilbach (2019) "Behavioral Development Economics," in Bernheim, B. D., S. DellaVigna & D. Laibson eds. *Handbook of Behavioral Economics*, 2, 345–458.
Kudlyak, M., D. Lkhagvasuren & R. Sysuyev (2013) "Systematic job search: New evidence from individual job application data," Working paper.
Kuhn, P. & H. Mansour (2014) "Is Internet Job Search Still Ineffective?" *The Economic Journal*,

124 (581), 1213–1233.
Kuhn, P. & K. Shen (2013) "Gender Discrimination in Job Ads: Evidence from China," *The Quarterly Journal of Economics*, 128 (1), 287–336.
Kuhn, P. & M. Skuterud (2004) "Internet Job Search and Unemployment Durations," *American Economic Review*, 94 (1), 218–232.
Laibson, D. (1997) "Golden Eggs and Hyperbolic Discounting," *The Quarterly Journal of Economics*, 112 (2), 443–478.
LaLive, R., C. Landais & J. Zweimuller (2022) "Market Externalities of Large Unemployment Insurance Extension Programs," *American Economic Review*, 105, 3564–3596.
Lechner, M. & J. Smith (2007) "What is the value added by caseworkers?" *Labour Economics*, 14 (2), 135–151.
Lee, D. (2009) "Trimming, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects," *Review of Economic Studies*, 76, 1071–1102.
Lentz, R., J. Maibom & E. Moen (2022) "Competitive or Random Search?", Working Paper.
Lichter, A. & A. Schiprowski (2021) "Benefit duration, job search behavior and re-employment," *Journal of Public Economics*, 193, 104–326.
Lise, J., S. Seitz & J. Smith (2004) "Equilibrium Policy Experiments and the Evaluation of Social Programs," Working paper 10283, National Bureau of Economic Research.
Madrian, B. C. & D. F. Shea (2001) "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior," *The Quarterly Journal of Economics*, 116, 1149–1187.
Marbach, M. & D. Hangartner (2020) "Profiling Compliers and Noncompliers for Instrumental-Variable Analysis," *Political Analysis*, 28 (3), 435–444.
Marinescu, I. (2017a) "The general equilibrium impacts of unemployment insurance: Evidence from a large online job board," *Journal of Public Economics*, 150, 14–29.
——— (2017b) "Job Search Monitoring and Assistance for the Unemployed," *IZA World of Labor*, 380.
Marinescu, I. & D. Skandalis (2021) "Unemployment Insurance and Job Search Behavior," *The Quarterly Journal of Economics*, 136, 887–931.
Marinescu, I. & R. Wolthoff (2020) "Opening the Black Box of the Matching Function: The Power of Words," *Journal of Labor Economics*, 38 (2), 535–568.
Martins, P. S. (2017) "Clicking towards Mozambique's New Jobs: A research note," Working Papers 85, Queen Mary, University of London, School of Business and Management, Centre for Globalisation Research.
Matsuda, N., T. Ahmed & S. Nomura (2019) "Labor Market Analysis Using Big Data: The Case of a Pakistani Online Job Portal," Policy Research Working Paper 9063, World Bank.
McGee, A. (2015) "How the Perception of Control Influences Unemployed Job Search," *Industrial and Labor Relations Review*, 68 (1), 184–211.
McGee, A. & P. McGee (2016) "Search, effort, and locus of control," *Journal of Economic Behavior and Organization*, 126, 89–101.
MICS (2018) "Multiple Indicator Cluster Survey Punjab," Bureau of Statistics Punjab, Planning & Development Board, Government of the Punjab.
Mueller, A. I. & J. Spinnewijn (2022) "Expectations Data, Labor Market and Job Search," in Bachmann, R., G. Topa & W. van der Klaauw eds. *Handbook of Economic Expectations*: ScienceDirect.
Pakistan Bureau of Statistics (2018-2019) "Quarterly Labour Force Survey."

Pallais, A. & E. G. Sands (2016) "Why the referential Treatment? Evidence from field experiments on referrals," *Journal of Political Economy*, 124 (6), 1793–1828.

Paserman, M. D. (2008) "Job Search and Hyperbolic Discounting: Structural Estimation and Policy Evaluation," *The Economic Journal*, 118, 1418–1452.

Pissarides, C. A. (2000) *Equilibrium Unemployment Theory*: MIT Press, 2nd edition.

Poverty Action Lab (2022) "Reducing Search Barriers for Job Seekers," 01.

Sanders, M., G. Briscese, R. Gallagher, A. Gyani, S. Hanes, E. Kirkman & O. Service (2019) "Behavioural insight and the labour market: evidence from a pilot study and a large stepped-wedge controlled trial," *Journal of Public Policy*, 41, 42–65.

Schiprowski, A. (2020) "The Role of Caseworkers in Unemployment Insurance: Evidence from Unplanned Absences," *Journal of Labor Economics*, 38 (4), 1189–1225.

Shimer, R. (2010) *Labor Markets and Business Cycles*: Princeton University Press.

Spinnewijn, J. (2015) "Unemployed but Optimistic: Optimal Insurance Design with Biased Beliefs," *Journal of the European Economic Association*, 13 (1), 130–167.

Subramanian, N. (2021) "Workplace Attributes and Women's Labor Supply Decisions: Evidence from a Randomized Experiment," Working paper.

Thaler, R. H. & S. Benartzi (2004) "Save More Tomorrow™: Using Behavioral Economics to Increase Employee Saving," *Journal of Political Economy*, 112, S164–S187.

Wheeler, L., R. Garlick, E. Johnson, P. Shaw & M. Gargano (2022) "LinkedIn(to) Job Opportunities: Experimental Evidence from Job Readiness Training," *American Economic Journal: Applied Economics*, 14, 101–125.

Wright, R., P. Kircher, B. Julien & V. Guerrieri (2021) "Directed Search and Competitive Search Equilibrium: A Guided Tour," *Journal of Economic Litterature*, 59 (1), 90–148.

Wu, L. (2021) "Partially Directed Search in the Labor Market," Working Paper.

Zhu, Y. (2021) "Phase transition of the monotonicity assumption in learning local average treatment effects," arXiv:2103.13369 [econ, math, stat].

# Appendices for Online Publication Only

## Contents

# A Additional Information about the Platform and Sample

This appendix provides additional descriptive statistics about the platform and the sample.

**Firm sample**: We listed a representative sample of 10,000 firms across the metropolitan area, using a similar approach as described in Section 2.3 for individual respondents, i.e. a cluster-randomized selection of Enumeration Blocks followed by listing of all firms in each selected block. A team of enumerators presents the Job Talash service to firms, offering them the opportunity to enroll to list vacancies immediately or later. We also promote the service publicly and include firms who self-select to sign up. Approximately 1,200 firms have signed up across these two samples. The majority of firms responding across both channels have never advertised jobs on any public platform, and usually recruit through networks. These firms are recontacted several times a year to invite them to post additional vacancies on the platform. Any firm can also call Job Talash to post a job at any time. Approximately 20 firms post jobs with the service per month, with approximately half posting at least one job over the course of the experiment.

**Jobseeker sample**: We use secondary data to compare our experimental samples of jobseekers and job ads to representative samples. Table A.1 compares our experimental sample of jobseekers (column 5) and all respondents in our household listing exercise (column 4) to data from Pakistan's Labor Force Survey for the entire country (column 1), the city of Lahore (column 2), and the city of Lahore reweighted to match the distribution of age, gender, and education as the experimental sample (column 3). Figure A.1 compares the distribution of salaries for vacancies posted on the platform to the distribution of salaries for the Lahore subsample of Pakistan's Labor Force Survey (Pakistan Bureau of Statistics, 2018-2019). These distributions should be compared with caution, as the former covers vacancies and the latter covers filled jobs, including jobs where incumbent workers have substantial experience with that firm.

Table A.2 compares the average monthly job application rate on this platform to other platforms studied by economists that report comparable statistics.

Figure A.2 shows a sample text message sent to jobseekers.

## Table A.1: Summary Statistics for Experimental and External Comparison Samples

**Panel A - Full Sample**

| | LFS Pakistan | LFS Lahore | LFS Lahore Reweighted | HH Listing Sample | Experimental Sample |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Female | 0.511 | 0.493 | 0.315 | 0.496 | 0.315 |
| Age | 34.0 | 34.0 | 30.3 | 33.2 | 30.5 |
| | (11.8) | (11.7) | (9.5) | (11.5) | (9.8) |
| Highest education level | | | | | |
| Less than Intermediate/High School | 0.825 | 0.692 | 0.592 | 0.708 | 0.593 |
| Completed Intermediate/High School | 0.088 | 0.141 | 0.146 | 0.121 | 0.146 |
| More than Intermediate/High School | 0.087 | 0.167 | 0.263 | 0.154 | 0.262 |
| Employed | 0.547 | 0.471 | 0.593 | 0.397 | 0.335 |
| Not employed and available for work | 0.030 | 0.022 | 0.036 | N/A | 0.319 |
| Searching | N/A | N/A | N/A | N/A | 0.569 |
| Searching and not employed | 0.015 | 0.017 | 0.031 | N/A | 0.319 |
| Applied to prospective employer | 0.007 | 0.009 | 0.018 | N/A | 0.123 |
| Checked at work sites, factories, markets, etc. | 0.005 | 0.006 | 0.011 | N/A | 0.088 |
| Sought assistance from friends, relatives, others | 0.006 | 0.008 | 0.016 | N/A | 0.237 |
| Placed or answered advertisements | 0.003 | 0.003 | 0.007 | N/A | 0.075 |
| Registered with an employment agency | 0.001 | 0.001 | 0.003 | N/A | 0.030 |
| Took other steps | 0.003 | 0.002 | 0.005 | N/A | 0.005 |

**Panel B - Female Sample**

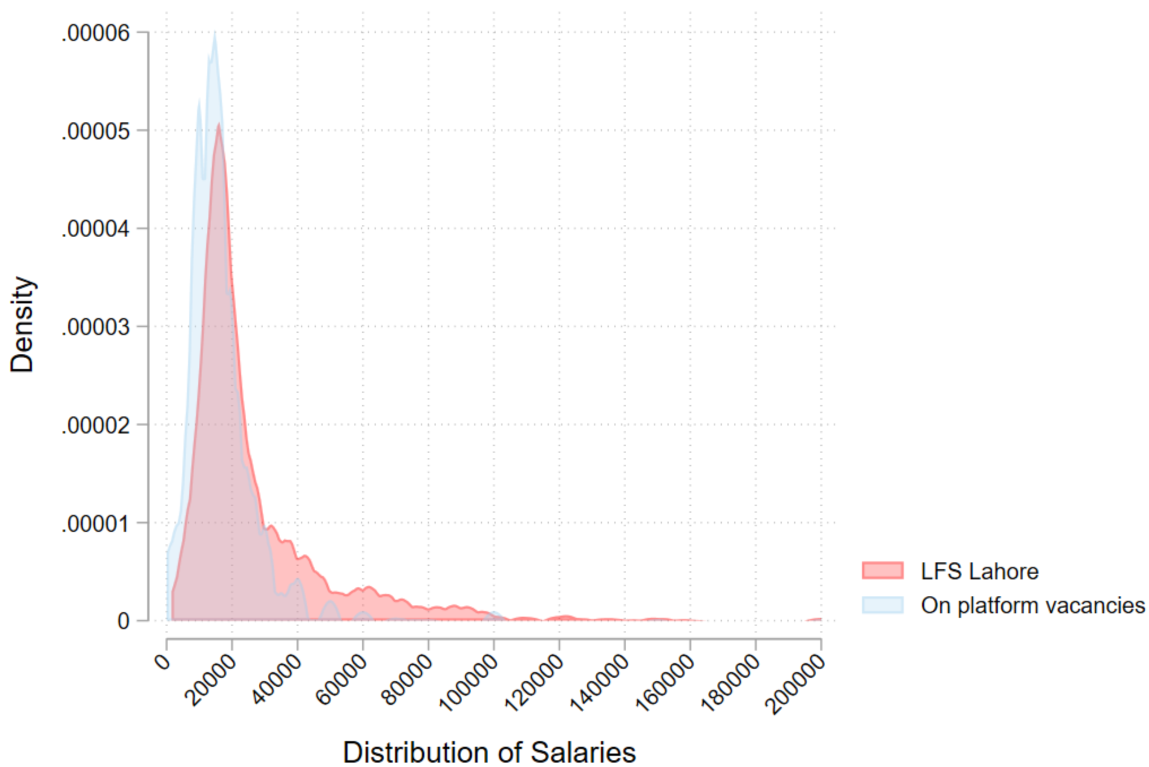
| | LFS Pakistan | LFS Lahore | LFS Lahore Reweighted | HH listing Sample | Experimental Sample |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Age | 33.9 | 33.8 | 32.7 | 32.6 | 30.7 |
| | (11.6) | (11.6) | (11.0) | (11.1) | (9.5) |
| Highest Education Level | | | | | |
| Less than Intermediate/High School | 0.853 | 0.679 | 0.700 | 0.706 | 0.491 |
| Completed Intermediate/High School | 0.073 | 0.148 | 0.130 | 0.127 | 0.144 |
| More than Intermediate/High School | 0.074 | 0.173 | 0.170 | 0.159 | 0.365 |
| Employed | 0.242 | 0.098 | 0.100 | 0.081 | 0.178 |
| Not employed and available for work | 0.034 | 0.014 | 0.015 | N/A | 0.322 |
| Searching | N/A | N/A | N/A | N/A | 0.446 |
| Searching and not employed | 0.011 | 0.009 | 0.009 | N/A | 0.322 |
| Applied to prospective employer | 0.004 | 0.004 | 0.005 | N/A | 0.101 |
| Checked at work sites, factories, markets, etc. | 0.001 | 0.002 | 0.002 | N/A | 0.057 |
| Sought assistance from friends, relatives, others | 0.004 | 0.003 | 0.003 | N/A | 0.240 |
| Placed or answered advertisements | 0.002 | 0.000 | 0.000 | N/A | 0.066 |
| Registered with an employment agency | 0.001 | 0.001 | 0.001 | N/A | 0.026 |
| Took other steps | 0.004 | 0.000 | 0.000 | N/A | 0.004 |

**Panel C - Male Sample**

| | LFS Pakistan | LFS Lahore | LFS Lahore Reweighted | HH Listing | Experimental Sample |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Age | 34.4 | 34.4 | 33.0 | 33.3 | 30.4 |
| | (12.2) | (11.9) | (11.3) | (11.4) | (9.9) |
| Highest education level | | | | | |
| Less than Intermediate/High School | 0.797 | 0.705 | 0.730 | 0.720 | 0.640 |
| Completed Intermediate/High School | 0.103 | 0.134 | 0.117 | 0.118 | 0.146 |
| More than Intermediate/High School | 0.100 | 0.160 | 0.153 | 0.152 | 0.214 |
| Employed | 0.865 | 0.832 | 0.834 | 0.713 | 0.408 |
| Not employed and available for work | 0.026 | 0.031 | 0.032 | N/A | 0.317 |
| Searching | N/A | N/A | N/A | N/A | 0.625 |
| Searching and not employed | 0.020 | 0.025 | 0.026 | N/A | 0.317 |
| Applied to prospective employer | 0.009 | 0.013 | 0.014 | N/A | 0.131 |
| Checked at work sites, factories, markets, etc. | 0.008 | 0.010 | 0.010 | N/A | 0.101 |
| Sought assistance from friends, relatives, others | 0.008 | 0.014 | 0.015 | N/A | 0.236 |
| Placed or answered advertisements | 0.004 | 0.005 | 0.005 | N/A | 0.078 |
| Registered with an employment agency | 0.002 | 0.001 | 0.002 | N/A | 0.032 |
| Took other steps | 0.003 | 0.005 | 0.004 | N/A | 0.005 |

Notes: Table compares the sample of jobseekers in this study (column 5) to several external benchmarks: the country (column 1), Lahore district, where the study takes place (column 2), and people in Lahore in the eligible age range for the study, reweighted with propensity scores to approximate the experimental sample on age, education, and sex (column 3). The table also compares the jobseekers in this study (column 5) to an internal benchmark, the Lahore representative household listing from which the experimental sample was recruited (column 4). The external benchmarks are calculated from the Labour Force Survey (LFS) 2018 using post-stratification weights provided by Pakistan Bureau of Statistics. Standard deviations are shown in parentheses for all continuous variables. Cells with 'N/A' mean that measure was not collected for that sample. The LFS only asked non-employed respondents about search.

Figure A.1: Salary Distribution for Experimental and External Comparison Sample




Notes: Figure shows the distribution of monthly salaries reported in the Labor Force Survey for Lahore in 2018 (red distribution, slightly to the right) and the distribution of salaries for vacancies posted on the platform (blue distribution, slightly to the left). Salary values greater than 200,000 have been top-coded at 200,000. Salaries are reported in Pakistani Rupees per month. 1 Rupee ≈ USD 0.03 in purchasing power parity terms during the study period.

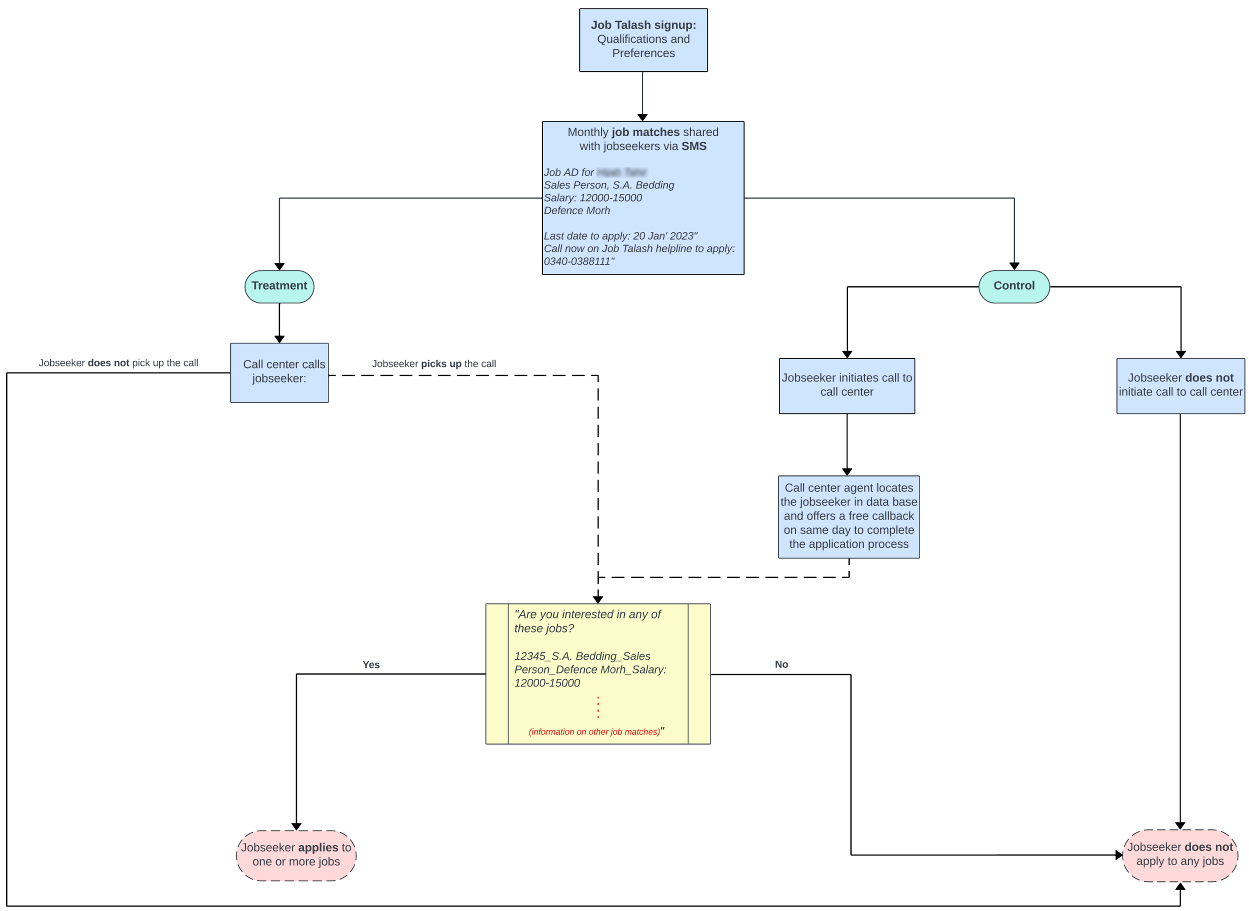
Figure A.2: Sample Text Message in English (Actual Messages are Sent in Urdu)

Table A.2: Job Application Rates on Search and Matching Platforms

| Study | Country | Platform | Mean apps per user per month | Notes |
|---|---|---|---|---|
| Behaghel et al. (2020) | France | La Bonne Bôite | 0.02 | |
| Martins (2017) | Mozambique | emprego.co.mz | 0.03 | |
| Wheeler et al. (2022) | South Africa | LinkedIn | 0.03 | |
| This paper | Pakistan | Job Talash | 0.03 | 74% of users do not submit any applications. |
| Archibong et al. (2022) | Nigeria | Not stated | 0.12 | |
| Ben Dhia et al. (2022) | France | Bob Emploi | 0.16 | |
| Banfi et al. (2022) | United States | careerbuilder.com | 0.18 | |
| Gee (2019) | Multiple countries | LinkedIn | 0.19 | |
| Marinescu & Skandalis (2021) | France | Not stated | 0.30 | 69% of users do not submit any applications. |
| Banfi et al. (2019) | Chile | trabajando.com | 1.22 | |
| Kelley et al. (2021) | India | Shikari | 1.25 | |
| Matsuda et al. (2019) | Pakistan | Rozee | 3.33 | Sample excludes users who submitted 0 applications. |
| Kudlyak et al. (2013) | United States | SnagAJob | 3.60 | Sample excludes users who submitted 0 applications. |
| Belot et al. (2018) | Scotland | Not stated | 4.40 | People in the sample were paid to use the platform. |
| Faberman & Kudlyak (2019) | United States | SnagAJob | 7.64 | Sample excludes users who submitted 0 applications. |

This table shows job application rates for users of job search and matching platforms in published and working papers. All application rates are converted into monthly, although different papers use periods ranging from 4 weeks to multiple years. The final column notes that some papers exclude users who submit zero applications during the period of study from their sample. Some other papers restrict their sample to 'active' or 'recently active' users but do not define what this means.

Figure A.3: Information Structure for Phone Call Treatment and Control Jobseekers




Notes: This flowchart shows the structure of how information flows for the phone call treatment and control jobseekers. The only difference between the two is that the former receives a phone call from the platform, whereas the latter initiates the call to the platform. The content in the text message and the phone call scripts are identical for both groups.

# B Additional Analysis on Search Effects and Returns to Search

## B.1 Average & Heterogeneous Effects on Interview- and Application-Related Outcomes

Figure B.1 shows treatment effects on the number of times each jobseeker applies to and is invited to an interview for a job. This figure shows that treatment raises the probability of submitting K applications and getting L interviews for all K and for $L \leq 4$.

Table B.1 shows treatment effects on interview probabilities weighted by different proxies for interview value, such as salary. This includes all components of the proxy index for interview value discussed in Section 3.2 and some combinations of proxies, e.g., commute-cost-adjusted salary in column 4 combines information from salary in column 1 and commute time in column 3. We show both intention-to-treat and two-stage least squares estimates but the latter are economically easier to interpret. We fail to reject equality of marginal and inframarginal returns for all eleven proxies. This supports the argument that returns to marginal treatment-induced search are roughly constant, by examining multiple possible measures of the value of search outcomes.

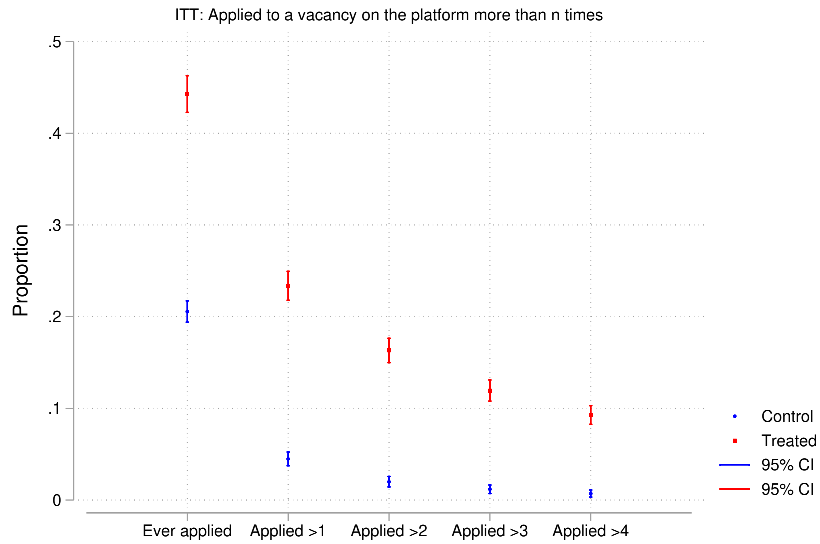
Table B.2 shows treatment effects on applications and interviews using the within-jobseeker through-time randomization. This supports the argument in Section 4.5 that constant returns to search are not explained by different types of jobseekers applying for jobs in the treatment and control groups.

Table B.3 shows heterogeneous treatment effects by baseline employment, search status, gender, education, and age. This shows that we cannot reject constant returns to search for a broad range of different types of jobseekers, although some of the subgroup estimates are imprecise enough that we do not recommend confident interpretation.

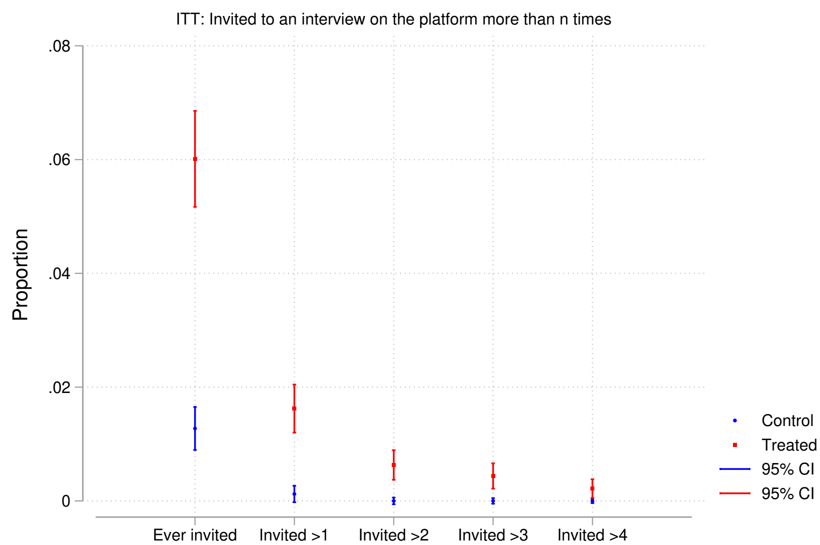
Table B.4 shows that treatment effects on applications do not differ by by quality scores that an employer assigned to CVs of jobseekers who applied to vacancies posted by that employer. This supports the argument in Section 4.5 that constant returns to search are not explained by different types of jobseekers applying for jobs in the treatment and control groups.

Figure B.1: Treatment Effects on Jobseeker-level Numbers of Applications and Interviews

Panel A: Number of Applications




Panel B: Number of Interviews




Notes: This figure shows treatment effects on the number of job applications submitted and number of interview invitations received. All estimates are from regressions of the number of applications or interview invitations on treatment assignment and stratification block fixed effects, using jobseeker-level data and the sample of all jobseekers. Solid vertical lines show 95% confidence intervals, constructed using heteroskedasticity-robust standard errors.

## Table B.1: Treatment Effects on Attributes of Marginal Interviews

| | Interview × | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | ln(Salary) | High salary | ln(Salary net commute cost) | Short commute | ln(Hourly salary) | Short hours | Flexible hours | Any benefits | Exact Match Ed. | Exact Match Exp. | Gender pref. aligned |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
| **Panel A - Treatment effects on interviews** | | | | | | | | | | | |
| Phone call treatment | 0.00781 | 0.00013 | 0.00672 | 0.00043 | 0.00354 | 0.00053 | 0.00066 | 0.00063 | 0.00007 | 0.00011 | 0.00049 |
| | (0.00091) | (0.00005) | (0.00078) | (0.00007) | (0.00042) | (0.00008) | (0.00008) | (0.00009) | (0.00003) | (0.00003) | (0.00006) |
| **Panel B - Treatment effects on interviews, instrumented by treatment** | | | | | | | | | | | |
| Apply | 0.60144 | 0.00909 | 0.55242 | 0.03548 | 0.32023 | 0.04705 | 0.05427 | 0.06646 | 0.00531 | 0.00737 | 0.03688 |
| | (0.05347) | (0.00343) | (0.04881) | (0.00474) | (0.02810) | (0.00552) | (0.00500) | (0.00703) | (0.00226) | (0.00175) | (0.00399) |
| # matches | 1,035,492 | 916,456 | 1,025,683 | 1,071,306 | 973,646 | 1,057,231 | 1,065,870 | 964,515 | 1,116,952 | 1,050,857 | 1,116,952 |
| # jobseekers | 9830 | 7194 | 9731 | 9813 | 9827 | 9828 | 9831 | 8999 | 9831 | 9831 | 9831 |
| Mean outcome \| T = 0 | 0.00120 | 0.00001 | 0.00107 | 0.00008 | 0.00054 | 0.00008 | 0.00010 | 0.00011 | 0.00001 | 0.00003 | 0.00008 |
| Mean outcome \| T = 0, Apply = 1 | 0.64568 | 0.00319 | 0.58095 | 0.04449 | 0.30800 | 0.04632 | 0.05392 | 0.08783 | 0.00365 | 0.01367 | 0.04376 |
| p: IV effect = mean \| T = 0, Apply = 1 | 0.645 | 0.130 | 0.749 | 0.283 | 0.799 | 0.935 | 0.969 | 0.116 | 0.600 | 0.109 | 0.359 |
| IV strength test: F-stat | 302.6 | 242.3 | 264.4 | 269.1 | 234.4 | 241.6 | 272.9 | 172.6 | 312.8 | 331.1 | 312.8 |
| IV strength test: p-value | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 |

Notes: Each column in panel A shows the coefficient from regressing an indicator for interview invitation weighted by a proxy of job quality on treatment assignment. Each column in panel B shows the coefficient from regressing an indicator for interview invitation weighted by a proxy of job quality on an indicator for application, instrumented by treatment assignment. All regressions include stratification block fixed effects. The unit of observation is the jobseeker × vacancy match. The sample is all matches. Heteroskedasticity-robust standard errors are shown in parentheses, clustering by jobseeker. Mean outcomes are for the control group. The proxies for job quality used in columns (1) to (11) are ln(posted salary), a binary variable indicating the expected salary being less than 90th percentile of salaries the jobseeker is matched to on the platform, ln(posted salary net of commute cost), a binary variable indicating a short commute (less than median distance), ln(hourly posted salary), a binary variable indicating less than median working hours, a binary variable indicating whether the firm ever allows employees in this position to work flexible hours, a binary variable indicating any benefits offered by the vacancy, a binary variable indicating whether the jobseeker has an exact match of educational specialization for the job advert, a binary variable indicating whether the jobseeker has an exact match of work experience for the job, and a binary variable indicating whether the job advert states preferring candidates from the jobseeker's gender.

Table B.2: Treatment Effects on Job Search & Search Returns Using Jobseeker Fixed Effects

| | (1) Apply | (2) Interview | (3) Int. $\times$ $V_{vm}$ | (4) Interview | (5) Int. $\times$ $V_{vm}$ |
|---|---|---|---|---|---|
| Randomly assigned to treatment in round t | 0.00764 | 0.00064 | 0.00251 | | |
| | (0.00066) | (0.00028) | (0.00116) | | |
| Apply | | | | 0.08356 | 0.32831 |
| | | | | (0.03421) | (0.14188) |
| # matches | 1,116,735 | 1,116,735 | 1,116,735 | 1,116,735 | 1,116,735 |
| # jobseekers | 9614 | 9614 | 9614 | 9614 | 9614 |
| Mean outcome \| T = 0 | 0.00185 | 0.00011 | 0.00042 | 0.00011 | 0.00042 |
| Mean outcome \| T = 0, Apply = 1 | | | | 0.06007 | 0.22598 |
| p: IV effect = mean \| T = 0, Apply = 1 | | | | 0.503 | 0.480 |
| IV strength test: F-stat | | | | 133.1 | 133.1 |
| IV strength test: p-value | | | | 0.00000 | 0.00000 |
| JS FE | Yes | Yes | Yes | Yes | Yes |
| Round FE | Yes | Yes | Yes | Yes | Yes |

Notes: All specifications are identical to those in Table 3 except that the treatment indicator varies both through time and between jobseekers. Column 1 shows the coefficient from regressing an indicator for job application on treatment assignment. Column 2 shows the coefficient from regressing an indicator for interview invitation on treatment assignment. Column 3 shows the coefficient from regressing an indicator for interview invitation weighted by a proxy index for the value of the vacancy to the jobseeker, $V_{vm}$, on treatment assignment. Column 4 shows the coefficient from regressing an indicator for interview invitation on job application, instrumented by treatment assignment. Column 5 shows the coefficient from regressing an indicator for interview invitation weighted by the proxy index $V_{vm}$ on job application, instrumented by treatment assignment. The proxy index is defined in the note to Table 3. All regressions use one observation per jobseeker $\times$ vacancy match, include jobseeker and round fixed effects, and use heteroskedasticity-robust standard errors clustered by jobseeker, which are shown in parentheses. The p-value is for a test of equality between the IV treatment effect and the mean interview rate for control group applications. The first-stage F-statistic and p-value are for the test of weak identification from Kleibergen & Paap (2006).

Table B.3: Heterogeneous Treatment Effects

| **Group 1 vs. 0** | **Employed vs. Unemployed** | **Searching vs. Not Searching** | **Female vs. Male** | **Less vs. More than High School** | **Less vs. More than 30 Years Old** |
|---|---|---|---|---|---|
| **Panel A: Applications** | | | | | |
| | (1) | (2) | (3) | (4) | (5) |
| Effect on Group = 1 | 0.01276 | 0.01566 | 0.01044 | 0.01469 | 0.01425 |
| | (0.00097) | (0.00109) | (0.00099) | (0.00103) | (0.00093) |
| Effect on Group = 0 | 0.01356 | 0.01174 | 0.01421 | 0.01161 | 0.01165 |
| | (0.00091) | (0.00117) | (0.00088) | (0.00084) | (0.00091) |
| p: (Effect on Group = 1) = (Effect on Group = 0) | 0.48063 | 0.00367 | 0.00148 | 0.00778 | 0.01800 |
| Mean Outcome \| T = 0, Group = 1 | 0.00161 | 0.00252 | 0.00263 | 0.00265 | 0.00206 |
| Mean Outcome \| T = 0, Group = 0 | 0.00202 | 0.00121 | 0.00165 | 0.00120 | 0.00157 |
| **Panel B: Interview Invitations** | | | | | |
| Effect on Group = 1 | 0.00069 | 0.00090 | 0.00097 | 0.00071 | 0.00089 |
| | (0.00011) | (0.00013) | (0.00016) | (0.00011) | (0.00011) |
| Effect on Group = 0 | 0.00084 | 0.00085 | 0.00070 | 0.00085 | 0.00060 |
| | (0.00011) | (0.00014) | (0.00009) | (0.00011) | (0.00010) |
| p: (Effect on Group = 1) = (Effect on Group = 0) | 0.27097 | 0.74611 | 0.11351 | 0.30252 | 0.02092 |
| Mean Outcome \| T = 0, Group = 1 | 0.00010 | 0.00018 | 0.00011 | 0.00010 | 0.00011 |
| Mean Outcome \| T = 0, Group = 0 | 0.00013 | 0.00006 | 0.00012 | 0.00013 | 0.00013 |
| **Panel C: IV** | | | | | |
| Effect on Group = 1 | 0.05453 | 0.05783 | 0.09253 | 0.04871 | 0.06227 |
| | (0.00665) | (0.00630) | (0.01149) | (0.00603) | (0.00596) |
| Effect on Group = 0 | 0.06172 | 0.07266 | 0.05033 | 0.07330 | 0.05206 |
| | (0.00625) | (0.00876) | (0.00505) | (0.00715) | (0.00677) |
| p: (Effect on Group = 1) = (Effect on Group = 0) | 0.35689 | 0.10855 | 0.00028 | 0.00253 | 0.17554 |
| Mean Outcome \| T = 0, Apply = 1, Group = 1 | 0.05941 | 0.07143 | 0.04194 | 0.03683 | 0.05293 |
| Mean Outcome \| T = 0, Apply = 1, Group = 0 | 0.06494 | 0.05085 | 0.07116 | 0.10997 | 0.08040 |
| p: Effect = Mean Outcome \| Group = 1 | 0.78064 | 0.14681 | 0.05283 | 0.02912 | 0.07441 |
| p: Effect = Mean Outcome \| Group = 0 | 0.73344 | 0.27590 | 0.00301 | 0.23010 | 0.38474 |
| # matches | 1,116,160 | 921,011 | 1,116,952 | 1,116,952 | 1,116,952 |
| Proportion in Group = 1 | 0.41427 | 0.58115 | 0.22850 | 0.46970 | 0.58101 |

Notes: Panel A shows the coefficients from regressing an indicator for job application on treatment assignment, stratification block fixed effects, an indicator for a group that varies between columns, and the interaction between the treatment assignment and the group indicator. Panel B shows the coefficient from regressing an indicator for interview invitation on the same right-hand side variables. The relevant group is: employed at baseline in column 1, searching at baseline in column 2, female in column 3, high school or higher education at baseline in column 4, and age under 30 years old at baseline in column 5. The unit of observation is the jobseeker × vacancy match. The sample in each of the columns varies due to item non-response in the baseline survey. Heteroskedasticity-robust standard errors are shown in parentheses, clustering by jobseeker.

Table B.4: Heterogeneous Treatment Effects by Employer-Scored CV Quality

| | Apply | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Phone call treatment | 0.02223 | 0.02212 | 0.01839 | 0.01949 |
| | (0.00302) | (0.00299) | (0.00671) | (0.00691) |
| CV: excellent score | -0.00128 | -0.00215 | -0.00306 | -0.01154 |
| | (0.00233) | (0.00244) | (0.00426) | (0.00642) |
| CV: good score | 0.00085 | 0.00042 | 0.00521 | 0.00245 |
| | (0.00108) | (0.00111) | (0.00549) | (0.00557) |
| CV: excellent score × Phone call treatment | -0.00755 | -0.00737 | 0.00675 | 0.00309 |
| | (0.00613) | (0.00603) | (0.00937) | (0.00948) |
| CV: good score × Phone call treatment | -0.00671 | -0.00710 | -0.00629 | -0.00717 |
| | (0.00373) | (0.00376) | (0.01017) | (0.01022) |
| # matches | 122946 | 122946 | 1982 | 1980 |
| # vacancies | 334 | 334 | 6 | 6 |
| # jobseekers | 1477 | 1477 | 1021 | 1021 |
| Mean outcome \| T = 0 | 0.00342 | 0.00342 | 0.00226 | 0.00227 |
| P-value \| $\beta_4$ + $\beta_5$ = 0 | 0.18583 | 0.16046 | 0.51627 | 0.65816 |
| Grader FE | No | Yes | No | Yes |
| Sample of vacancies | Selected and Similar Occupations | Selected and Similar Occupations | Selected Vacancies Only | Selected Vacancies Only |

Notes: The table shows the heterogeneous treatment effects on applications by CV quality with and without Grader fixed effects. Unit of observation: jobseeker × vacancy match. Specification in all columns consist of regressing an indicator for job application on treatment assignment, dummies for CV quality excellent and good, and interaction of treatment assignment with CV quality excellent and good. Omitted category: "average" or lower score. 759 out of 1477 jobseekers' CVs were scored by graders for both of the selected vacancies. In these cases, we use the mean of the two scores for Columns (1) and (2); and the grade corresponding to the selected vacancy in columns (3) and (4). "Selected" jobs include the six enumerator/call center jobs for which the recruiting managers were grading the CVs. "Similar occupations" consist of the following codes: Receptionist/Front Desk Officer/Telephone Operator, Sales/Marketing Officer, Computer Operator, Customer Service Officer/Enumerator, Telemarketing Officer/Call Centre Agent and Data Entry Operator. All specifications include stratification block fixed effects. Grader fixed effects only included for specifications in columns (2) and (4). Heteroskedasticity-robust standard errors, clustered by jobseeker, reported in parentheses.

### B.2 Robustness Checks

Table B.5 shows that our main findings from Table 3 are robust to alternative sets of conditioning variables, weighting, and clustering. Column 1 shows results from our preferred specification; column 2 includes interactions between treatment and the fixed effects, following the recommendation by Imbens & Rubin (2015); column 3 drops stratification block fixed effects. Results are similar across the three specifications: the effect on applications ranges from 1.28 to 1.34 percentage points and the marginal applications have a mean interview probability between 5 and 5.9%. We also show results conditioning on jobseeker-level covariates in column 4, vacancy- and match-level covariates in column 5, and all three sets of covariates in column 6. All covariates are selected using a post-double selection LASSO, following Belloni et al. (2014). The effect on applications ranges from 1.33 to 1.34 percentage points and the marginal applications have a mean interview probability between 5.9 and 6.8%. The findings in columns 4, 5, and 6 reinforce our argument in Sections 4.5 and 4.2 that the main findings are not driven by treatment effects on which jobseekers use the platform or where they direct applications.

Our main analysis uses one observation per match. This gives higher weight to jobseekers who get more matches, due to their occupational preferences, educational qualifications, or work experience. We repeat our main analysis weighting the data by the inverse number of matches received by each jobseeker, which assigns equal weight to each jobseeker and makes it easier to compare results to jobseeker-level analysis using survey data. Column (7) shows that the weighted treatment effect on applications is slightly higher (1.83 percentage points), which means that jobseekers who receive fewer matches are more responsive to treatment. The weighted treatment effect on interviews increases by a slightly smaller margin, leading to a 4.6% probability of converting marginal applications into interviews. This is slightly lower than the unweighted result but is not statistically significantly different to the unweighted result or the interview probability for control group applications, with or without weights.

Our main findings are also robust to two alternative ways of estimating the standard errors: clustering by enumeration areas used for household listing (column 8) and clustering by both jobseeker and vacancy (column 9). The former approach follows a recommendation from Abadie et al. (2017) and is appropriate for conducting inference about all enumeration areas around Lahore, not only the enumeration areas we randomly chose for our sample. The latter approach is arguably conservative, because treatment is randomized within vacancy, but it allows for the fact that applications are correlated with vacancies across jobseekers.

Table B.5: Robustness of Main Results to Alternative Controls, Weighting, and Clustering

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|
| **Panel A - Treatment effects on applications** | | | | | | | | | |
| Phone call treatment | 0.01322 | 0.01275 | 0.01342 | 0.01335 | 0.01331 | 0.01342 | 0.01835 | 0.01323 | 0.01322 |
| | ( 0.00075) | ( 0.00000) | ( 0.00056) | ( 0.00076) | ( 0.00079) | ( 0.00081) | ( 0.00121) | ( 0.00075) | ( 0.00100) |
| **Panel B - Treatment effects on interviews** | | | | | | | | | |
| Phone call treatment | 0.00078 | 0.00070 | 0.00075 | 0.00078 | 0.00091 | 0.00092 | 0.00085 | 0.00078 | 0.00078 |
| | ( 0.00009) | ( 0.00000) | ( 0.00006) | ( 0.00009) | ( 0.00010) | ( 0.00011) | ( 0.00011) | ( 0.00009) | ( 0.00013) |
| **Panel C - Application effects on interviews, instrumented by treatment** | | | | | | | | | |
| Apply | 0.05865 | 0.05033 | 0.05569 | 0.05873 | 0.06804 | 0.06846 | 0.04657 | 0.05866 | 0.05865 |
| | ( 0.00516) | ( 0.00735) | ( 0.00381) | ( 0.00519) | ( 0.00586) | ( 0.00590) | ( 0.00579) | ( 0.00521) | ( 0.00895) |
| # matches | 1116959 | 1116959 | 1116959 | 1100035 | 968936 | 955107 | 1116959 | 1116115 | 1116959 |
| # jobseekers | 9838 | 9838 | 9838 | 9630 | 9836 | 9628 | 9838 | 9825 | 9838 |
| # vacancies | 1343 | 1343 | 1343 | 1343 | 1217 | 1217 | 1343 | 1343 | 1343 |
| Fixed effects | Y | N | N | Y | Y | Y | Y | Y | Y |
| Fixed effects interactions | N | Y | N | N | N | N | N | N | N |
| Jobseeker-level controls | N | N | N | Y | N | Y | N | N | N |
| Vacancy-level & match-level controls | N | N | N | N | Y | Y | N | N | N |
| Weights | N | N | N | N | N | N | Y | N | N |
| Clustering | JS | JS | JS | JS | JS | JS | JS | EA | JS & V |

Notes: This table shows treatment effects on key outcomes using different regression specifications. Column 1 shows results for the default sample and regression specification, which includes stratification block fixed effects and either treatment assignment (Panels A-B) or application instrumented by treatment assignment (Panel C). Column 2 includes interactions between treatment and the fixed effects (and instrument in panel C) and estimates the treatment effect as the average of the treatment * fixed effect interactions weighted by the relative sizes of the stratification blocks (following Imbens & Rubin 2015). Column 3 excludes stratification block fixed effects. Column 4, 5 and 6 include respectively, jobseeker-level controls; advert- and match-level controls; and jobseeker-, advert-, and match-level controls. The controls are selected using a post-double-selection LASSO (following Belloni et al. 2014). The LASSO model is allowed to select from the following characteristics: at the jobseeker level, age of the jobseeker, gender of the jobseeker, whether the jobseeker is married at baseline, whether the jobseeker is married and has kids at baseline, whether the jobseeker has above-median education, whether the jobseeker has any work experience at baseline, jobseeker's years of work experience, and whether the jobseeker selects many occupational categories at baseline; at the match and vacancy level, high salary relative to respondent's matches, high salary relative to all matches, high number of years of experience required relative to all matches, and jobseeker has an exact match of work experience for the job. Column 7 weights observations by the jobseeker-level inverse number of matches so each jobseeker receives the same weight. Column 8 uses the same specification used in Column 1. Heteroskedasticity-robust standard errors shown in parentheses. Column 1 - 7 include standard errors clustered by jobseeker. Column 8 includes standard errors clustered by the enumeration area of the jobseeker. Column 9 includes standard errors two-way clustered by jobseeker and vacancy. Sample sizes vary slightly across columns due to non-response affecting covariates. All units of observation are at the jobseeker $\times$ vacancy match.

### B.3 Addressing Possible Violations of the IV Monotonicity Assumption

Researchers using instrumental variables to study treatment effects commonly make a monotonicity assumption. In our context, this monotonicity assumption is that the phone call treatment weakly increases the probability of application in all matches. Under this assumption all matches are either compliers, which lead to applications if and only if they are treated; always-takers, which lead to applications irrespective of treatment status; or never-takers, which do not lead to applications irrespective of treatment status. Under this assumption no matches are defiers, matches that lead to applications if and only if they are *not* treated. Note that these types are defined at the match level: the same jobseeker may be a complier in some matches, always-taker in some matches, and a never-taker in other matches.

This monotonicity assumption allows us to interpret our two-stage least squares estimate as the average treatment effect of applications on interview invitations for compliers, typically called the local average treatment effect (LATE).

If there are some defiers, two-stage least squares does not recover a well-defined treatment effect. The coefficient in a two-stage least squares regression with one binary instrument and one binary endogenous variable recovers the difference between the treatment effect on compliers and the treatment effect on defiers, weighted by their shares in the population. Define $\mathbf{P_j}$ as the population share of type $j$ and $\mathbf{\Delta I_j}$ as the treatment effect on interviews for type $j$. We use bold text to show that these quantities are unknown and follow this convention throughout the argument. Using this notation:

$$\beta_{2SLS} = \frac{\mathbf{P_C} \cdot \mathbf{\Delta I_C} - \mathbf{P_D} \cdot \mathbf{\Delta I_D}}{\mathbf{P_C} - \mathbf{P_D}}. \tag{6}$$

If the share of defiers is zero, as assumed in most empirical papers, then $\beta_{2SLS} = \mathbf{\Delta I_C}$.

If the share of defiers is not zero, we can bound the treatment effect on compliers $\mathbf{\Delta I_C}$ using a six-step argument that we adapt from De Chaisemartin (2017) and Zhu (2021). First, we note that the treatment effect on interviews for defiers, $\mathbf{\Delta I_D}$, is defined as $\mathbb{E}\left[I|T=1, \text{Defier}\right] - \mathbb{E}\left[I|T=0, \text{Defier}\right]$. The first term is zero because treated defiers, by definition, do not send applications and hence cannot get interviews. The second term is the mean interview rate for applications from untreated defiers, which we denote by $\bar{I}_D$. Hence we can rewrite equation (6) as

$$\mathbf{\Delta I_C} = \frac{\beta_{2SLS} \cdot (\mathbf{P_C} - \mathbf{P_D}) + \mathbf{P_D} \cdot \mathbf{\Delta I_D}}{\mathbf{P_C}} = \frac{\beta_{2SLS} \cdot \beta_{S1} - \mathbf{P_D} \cdot \mathbf{\bar{I}_D}}{\beta_{S1} + \mathbf{P_D}}, \tag{7}$$

where $\beta_{S1} = \mathbf{P_C} - \mathbf{P_D}$ is the coefficient from a first stage regression of application on treatment.

Second, we note that all unknown quantities in equation (7) can be bounded. Control group matches yield applications if and only if those matches are defiers or always-takers. Hence the mean application rate in the control group, which we denote by $\bar{A}_0$, equals $\mathbf{P_D} + \mathbf{P_A}$. This yields the inequality restriction

$$0 \leq \mathbf{P_D} \leq \bar{A}_0. \tag{8}$$

$\mathbf{\bar{I}_D}$ is the mean value of a binary variable. The same is true of $\mathbf{\bar{I}_A}$, the mean interview rate for applications from always-takers. Hence

$$0 \leq \mathbf{\bar{I}_A} \leq 1 \tag{9}$$

$$0 \leq \mathbf{\bar{I}_D} \leq 1. \tag{10}$$

Evaluating equation (7) in light of these three inequalities show that $\mathbf{\Delta I_C} \leq \beta_{2SLS}$, with equality when $\mathbf{P_D} = 0$, i.e. two-stage least squares recovers LATE when there are no defiers. This gives us an upper bound on $\mathbf{\Delta I_C}$. To derive the lower bound, we proceed to the next steps.

Third, we note again that any application in the control group must come from an always-taker or a defier. Hence the mean interview rate for applications submitted from control group matches, which we denote by $\bar{I}_0$, is the average of rates for always-takers and defiers weighted by their relative population shares: $(\mathbf{\bar{I}_A} \cdot \mathbf{P_A} + \mathbf{\bar{I}_D} \cdot \mathbf{P_D})/(\mathbf{P_A} + \mathbf{P_D})$. Recalling that $\mathbf{P_D} + \mathbf{P_A} = \bar{A}_0$ and rearranging terms gives

$$\mathbf{P_D} \cdot \left(\mathbf{\bar{I}_D} - \mathbf{\bar{I}_A}\right) = \bar{A}_0 \cdot \left(\bar{I}_0 - \mathbf{\bar{I}_A}\right). \tag{11}$$

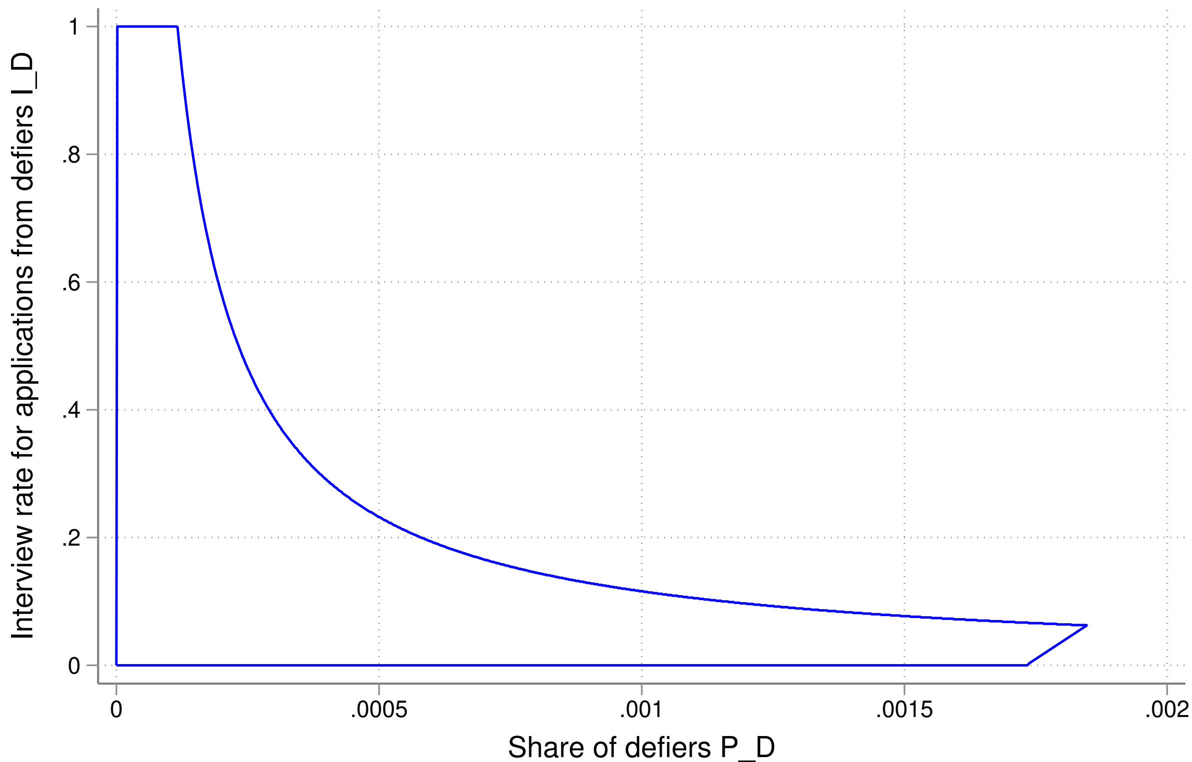
Combining (7), (8), (9), (10), and (11) gives a system of two equality restrictions and three inequality restrictions in which $\mathbf{\Delta I_C}$ depends on three unknown quantities: $\mathbf{\bar{I}_D}$, $\mathbf{\bar{I}_A}$, and $\mathbf{P_D}$. This does not allow us to point identify $\mathbf{\Delta I_C}$ but allows us to obtain a lower bound.

Fourth, we consider each value $P_D$ satisfying (8), solve for the set of values of $\mathbf{\bar{I}_D}$ and $\mathbf{\bar{I}_A}$ consistent with all the restrictions, and then solve for the set of values of $\mathbf{\Delta I_C}$ consistent with all the restrictions. Let $\{\mathbf{\Delta I_C}\}_{\mathbf{P_D}}$ denote this set of feasible values.

Figure B.2 shows, for each possible value of the share of defiers $\mathbf{P_D}$, the set of feasible values of $\mathbf{\bar{I}_D}$ in solid blue. When the share of defiers is small, only condition (9) binds on $\mathbf{\bar{I}_D}$. As the share of defiers increases, the maximum feasible value of $\mathbf{\bar{I}_D}$ shrinks to stop the left-hand side of equation (11) from becoming so large that it can only be satisfied by a value of $\mathbf{\bar{I}_A}$ that violates condition (10). As the share of defiers approaches $\bar{A}_0$ and hence the share of always-takers approaches zero, $\mathbf{\bar{I}_D}$ must approach $\bar{I}_0$ and the feasible set approaches a point.

Fifth, we construct the union of feasible sets $\{\mathbf{\Delta I_C}\}_{\mathbf{P_D}}$ over all values of $\mathbf{P_D}$, which we define as $\{\mathbf{\Delta I_C}\}$. The maximum value of $\mathbf{\Delta I_C}$ in this set occurs when $\mathbf{P_D} = \mathbf{0}$ and is simply $\beta_{2SLS}$. This matches the intuitive interpretation of equation (6): if there are no defiers, then the monotonicity assumption automatically holds, and hence two-stage least squares recovers the treatment effect on interviews for defiers. The minimum value of $\mathbf{\Delta I_C}$ occurs as $\mathbf{P_D}$ approaches its maximum value of $\bar{A}_0$, i.e. when there are no always-takers and all control group applications come from defiers,

Figure B.2: Bounding the Local Average Treatment Effect Without Monotonicity




Notes: The blue solid line covers the values of the share of defiers $P_D$ and the interview rate for applications sent by defiers $I_D$ that are feasible, given the data-based restrictions derived in this section.

and hence $\mathbf{\Delta I_D}$ approaches $\bar{I}_0$. Note that $\mathbf{\Delta I_C}$ is undefined at $\mathbf{P_D} = \bar{A}_0$ because there are no compliers at that point. So the lower bound is defined by the limit as $\mathbf{P_D}$ approaches $\bar{A}_0$.

Using the estimated values of $\bar{A}_0 = 0.00185$, $\beta_{S1} = 0.01322$, $\bar{I}_0 = 0.06290$, and $\beta_{2SLS} = 0.5865$ from Table 3 yields a lower-bound estimate of 0.045461 for the average treatment effect on compliers. The bounded set for $\mathbf{\Delta I_C}$ thus equals [0.0455,0.0587], with a width of only 1.32 percentage points.

### B.4 Addressing Possible Complications around the IV Exclusion Assumption

In our application, the exclusion assumption is that treatment assignment affects interview invitations only through job applications. This is mechanically true, in the sense that interviews are only possible through job applications. Here we address three possibilities that might complicate interpretation of this assumption, without necessarily violating it. Our findings are robust to accounting for each of the three possibilities.

**Treatment effects on matches received**: Participants receive matches based on their education, work experience, and occupational preferences. Roughly 11% of control group respondents change job preferences after sign-up and treatment decreases this by 1.8 percentage points (Table B.6, column 2). Treatment has small effects that are not statistically significant on the probabilities of adding educational qualifications or work experience to the CV (Table B.6, columns 4-5).

These changes might in principle lead to treatment effects on the set of matches received by participants, leading to treatment-control differences in the samples used for analysis. We test whether our results are sensitive to this concern by constructing the set of matches that each respondent would have obtained if they had retained their original job preferences; we code applications and interviews as zeros for the counterfactual subset of these matches respondents did not actually receive, and estimate treatment effects in this sample. We do the same exercise with the original education and work experience information. The treatment effects on both applications and interviews are mechanically lower in these hypothetical samples. The returns to marginal and inframarginal applications range from 6.5 to 6.6% across all of these counterfactual samples, again showing roughly constant returns to marginal search effort (Table B.7, Panel C, columns 2-4).

**Treatment effects on application content**: Treatment might shift the content of job applications as well as the quantity of job applications. This is a standard concern with research designs based on instruments that shift quantities. For example, instruments that shift the cost of education

Table B.6: Treatment Effects on Non-Application Measures of Platform Use

| | (1)<br># pref. updates | (2)<br>Any pref. update | (3)<br>Completed CV | (4)<br>Added educ. | (5)<br>Added work exp. |
|---|---|---|---|---|---|
| Phone call treatment | -0.07087 | -0.01919 | 0.02494 | 0.02058 | -0.00510 |
| | (0.04183) | (0.00663) | (0.00896) | (0.00496) | (0.00370) |
| # jobseekers | 9823 | 9823 | 9823 | 9823 | 9823 |
| Mean outcome \| T = 0 | 0.56337 | 0.10633 | 0.15343 | 0.03558 | 0.02911 |

Notes: This table shows treatment effects on measures of platform use other than job applications: number of updated occupation preferences (column 1), an indicator for updating any occupation preference (column 2), completing their on-platform CV (column 3), adding more education information to their CV (column 4), or adding more work experience to their CV (column 5). Each column shows the coefficient from regressing the relevant outcome on treatment assignment, stratification block fixed effects, and fixed effects for the timing of the jobseeker follow-up surveys used to collect CV-related information. The unit of observation is the jobseeker. The sample is all jobseekers. Heteroskedasticity-robust standard errors are shown in parentheses.

Table B.7: Sensitivity of Treatment Effects to Accounting for Changes in Jobseeker Profile and Preferences on Platform

**Panel A - Treatment effects on applications**

| | Apply | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Phone call treatment | 0.01324 | 0.01078 | 0.01026 | 0.01077 | 0.01524 | 0.01578 |
| | (0.00075) | (0.00067) | (0.00065) | (0.00067) | (0.00111) | (0.00085) |
| # matches | 1,112,181 | 1,194,533 | 1,176,749 | 1,190,180 | 696,951 | 1,000,180 |
| # jobseekers | 9025 | 8925 | 8995 | 8927 | 5743 | 9646 |
| Mean outcome \| T = 0 | 0.00185 | 0.00154 | 0.00154 | 0.00155 | 0.00210 | 0.00199 |
| Sample | Full sample | Hypothetical matches w/initial preferences | Hypothetical matches w/initial edu & exp | Hypothetical matches w/initial preferences & edu & exp | Completed CV at baseline | Excluding matches during stops |

**Panel B - Treatment effects on interviews**

| | Interview | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Phone call treatment | 0.00078 | 0.00071 | 0.00066 | 0.00070 | 0.00113 | 0.00093 |
| | (0.00009) | (0.00008) | (0.00008) | (0.00008) | (0.00014) | (0.00010) |
| # matches | 1,112,188 | 1,194,533 | 1,176,749 | 1,190,180 | 696,951 | 1,000,180 |
| # jobseekers | 9025 | 8925 | 8995 | 8927 | 5743 | 9646 |
| Mean outcome \| T = 0 | 0.00012 | 0.00010 | 0.00010 | 0.00010 | 0.00016 | 0.00013 |
| Sample | Full sample | Hypothetical matches w/initial preferences | Hypothetical matches w/initial edu & exp | Hypothetical matches w/initial preferences & edu & exp | Completed CV at baseline | Excluding matches during stops |

**Panel C - Application effects on interviews, instrumented by treatment**

| | Interview | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Apply | 0.05902 | 0.06559 | 0.06451 | 0.06545 | 0.07405 | 0.05899 |
| | (0.00519) | (0.00579) | (0.00596) | (0.00580) | (0.00688) | (0.00501) |
| # matches | 1,112,181 | 1,194,533 | 1,176,749 | 1,190,180 | 696,951 | 1,000,180 |
| # jobseekers | 9025 | 8925 | 8995 | 8927 | 5743 | 9646 |
| Mean outcome \| T = 0 | 0.00012 | 0.00010 | 0.00010 | 0.00010 | 0.00016 | 0.00013 |
| Mean outcome \| T = 0, Apply = 1 | 0.06296 | 0.06566 | 0.06542 | 0.06465 | 0.07713 | 0.06290 |
| p: IV effect = mean \| T = 0, Apply = 1 | 0.67138 | 0.88933 | 0.80689 | 0.87800 | 0.28300 | 0.67046 |
| IV strength test: F-stat | 308.5 | 258.6 | 246.2 | 261.1 | 187.0 | 342.6 |
| IV strength test: p-value | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 |
| Sample | Full sample | Hypothetical matches w/initial preferences | Hypothetical matches w/initial edu & exp | Hypothetical matches w/initial preferences & edu & exp | Completed CV at baseline | Excluding matches during stops |

Notes: This table shows how treatment effects change (a) when we repeat our main analyses holding fixed jobseekers' initial occupational preferences, education, and experience so jobseekers' updates to these measures cannot influence the matches they receive, and (b) when dropping matches during periods in which the jobseeker requested a stop. Column 1 uses the sample of actual matches jobseekers receive, replicating the results in Table 3. Column 2 uses the sample of matches that jobseekers would have received if they did not update their occupational preferences. Column 3 uses the sample of matches that jobseekers would have received if they did not update their education or work experience. Column 4 uses the sample of matches that jobseekers would have received if they did not update occupational preferences, education, or experience. For all matches in columns 2, 3, and 4 that jobseekers did not actually receive, both application and interview are coded as zeros. Column 5 uses the sample of matches of jobseekers who completed their CVs at baseline. Column 6 uses the sample of matches during periods in which the jobseeker did not request to pause/stop getting matches.

Panels A and B and show the coefficients from regressing respectively invitations an indicator for job application and an indicator for interview invitation on treatment assignment. Panel C shows the coefficient from regressing an indicator for interview invitation on job application, instrumented by treatment assignment. The sample size for columns 1-4 in this table is slightly smaller than in the main treatment effects table due to some missing values for preference, education or experience data. All regressions include stratification block fixed effects. The unit of observation is the jobseeker $\times$ vacancy. Heteroskedasticity-robust standard errors are shown in parentheses, clustered by jobseeker.

may shift both the quantity and quality of education attained, complicating interpretation of any 'return to education' estimated in these designs (Card, 2001).

However, as discussed in Section 2, our platform allows us to observe everything that the firm observes about the jobseeker and that the jobseeker observes about the firm prior to the interview invitation. Jobseekers do not receive contact information for firms before firms reach out to invite them to an interview, so it is unlikely that jobseekers could send additional information to firms.

Thus we can test directly for quality effects. The most obvious proxy for quality is CV completion, as firms are less likely to view CVs with missing fields positively. Treated candidates are 2.5 percentage points more likely than control candidates to complete missing fields on their on-platform CV after registering, mainly due to adding educational information rather than adding work experience (Table B.6, column 3). But replicating our main analysis for respondents who completed their entire CV at registration replicates our main findings (Table B.7, column 5). Treatment effects on both applications and interviews and the return to education are all slightly higher in this sample. But the returns to marginal and inframarginal applications remain very similar to each other, respectively 7.4 and 7.7%.

**Treatment effects on platform engagement**: Respondents can ask to stop being sent matches temporarily or permanently. Treatment increases the probability of requesting a pause or stop by roughly 12 percentage points. This is partly because treatment shifts people from passive disengagement (ignoring text messages) to active disengagement (asking to stop calls and text messages). Our main analysis retains matches from jobseekers who request stops and codes applications and interviews as zeros for these matches. As a sensitivity check, we can instead drop observations from jobseekers during periods when they have requested stops. This mechanically slightly increases treatment effects on applications and interviews (Table B.7, column 6). But the returns to marginal and inframarginal applications are respectively 5.9% and 6.3% in this sample, almost identical to the full sample.

**Alternative Approach to Testing Constant Returns to Search**: We show evidence consistent with constant returns to search using an alternative method that makes slightly different assumptions to the instrumental variables method in the main paper. This method is adapted from Attanasio et al. (2011) and Carranza et al. (2021). We first estimate the treatment effect on the application probability multiplied by the control group's mean interview:application ratio, which we call the *CR-implied effect*. This quantity captures the increase in job interviews that would occur if treatment shifted interviews only by shifting the quantity of job applications, but had no effect on the probability of converting job applications into interviews. Under constant returns, the CR-implied effect should equal the average effect of treatment on the interview probability, a hypothesis we can test directly.

The CR-implied effect and average effect of treatment on interviews are very similar. Multiplying the 1.322 percentage point effect on application probability and the 0.0629 ratio of interviews to applications in the control group yields a CR-implied effect of 0.083 p.p., with standard error 0.05 p.p. (Table B.8, column 1, row 2). This is almost identical to the treatment effect on interviews of 0.078 p.p (column 1, row 1). The 0.006 p.p. difference between them is both small and not significantly different to zero, with standard error 0.007 p.p. (column 1, row 3). The CR-implied effect and average effect of treatment on 'value-weighted' interviews are also similar. Recall that our main measure of value-weighted interviews from Section 3 is the interview indicator multiplied by an inverse covariance-weighted average of the eight proxies for the value of the interview. For this measure, the CR-implied effect and average effect differ by only 0.0003 with standard error 0.0003, roughly 10% of the average effect (Table B.8, column 2, row 3).

Table B.8: Alternative Test for Constant Returns to Search

| | (1)<br>Interview | (2)<br>Interview $\times$ $V_{vm}$ index |
|---|---|---|
| Treatment effect | 0.00078<br>(0.00009) | 0.00281<br>(0.00036) |
| Constant-returns implied effect | 0.00083<br>(0.00005) | 0.00314<br>(0.00018) |
| Difference | -0.00006<br>(0.00007) | -0.00033<br>(0.00028) |
| # matches | 1,116,952 | 1,116,952 |
| # jobseekers | 9831 | 9831 |
| Mean outcome $\mid$ T = 0, Apply = 1 | 0.06290 | 0.23778 |

This table compares treatment effects on interviews (row 1) to the treatment effects on applications multiplied by the mean interview:application ratio in the control group (row 2). Under constant returns, these two quantities will be identical. Hence we name the effect in row 2 the 'CR-implied effect.' Each column shows results for a different outcome: interviews in column 1 and interviews multiplied by an inverse covariance-weighted average of eleven proxies for the value of an interview in column 2. The proxies are defined in the note to Table 3. The unit of observation is the jobseeker $\times$ vacancy match. The sample is all matches. All regressions include stratification block fixed effects. Heteroskedasticity-robust standard errors are shown in parentheses, clustering by jobseeker.

### B.5 Treatment Effects on Employment and Off-Platform Search

Tables B.9 and B.10 show treatment effects on employment and off-platform search, reported in a survey of jobseekers. They show that the effect on employment is positive and effects on search measures are negative, but all of these are small and not statistically significant.

Appendix B.6 explains how we construct the selection correction terms shown in even-numbered columns of these tables.

Table B.9: Treatments Effects on Off-Platform Search and Work

| | Any Off- Platform Search | | Any Work | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Phone call treatment | -0.00780 | -0.01078 | 0.00179 | 0.01081 |
| | (0.01630) | (0.02072) | (0.01587) | (0.02002) |
| # jobseekers | 4327 | 9823 | 4643 | 9823 |
| # jobseekers answered \| T = 0 | 2445 | 2445 | 2587 | 2587 |
| # jobseekers answered \| T = 1 | 1882 | 1882 | 2056 | 2056 |
| Mean outcome \| T = 0 | 0.26667 | 0.26667 | 0.73328 | 0.73328 |
| Adjusted for non-response | No | Yes | No | Yes |
| IV strength test: F-stat | | 170.381 | | 132.783 |
| IV strength test: p-value | | 0.000 | | 0.000 |

Notes: This table shows treatment effects on off-platform search and work. The outcome in columns (1) and (2) is an indicator for whether the jobseeker reported searching for work in the last 14 or 30 days, excluding job applications through the Job Talash platform. The outcome in columns (3) and (4) is an indicator for whether the jobseeker reported working in the last 14 or 30 days. Recall periods of 14 or 30 days are randomly assigned. Each outcome is regressed on an indicator for treatment assignment and stratification block fixed effects. Columns (2) and (4) include selection adjustment terms for survey non-response described in Appendix B.6 and using the method proposed by DiNardo et al. (2021). They use as instruments random assignment to receiving two additional call attempts, a heads-up text message before the call, a monetary incentive for answering the call and finishing the survey, and early call attempts. The unit of observation is the jobseeker. The IV strength tests are for joint tests that all the instruments have zero coefficients in the first stage. All specifications include stratification block fixed effects. Standard errors shown in parentheses. For columns without non-response adjustments, these are heteroskedasticity-robust. For columns with non-response adjustments, these are estimated using 500 iterations of a nonparametric bootstrap.

Table B.10: Treatment Effects on Off-Platform Search (Intensive Margin)

| | Off- Platform Applications | | % Search Methods Used | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Phone call treatment | -0.18882 | -0.21265 | -0.01300 | -0.01031 |
| | (0.14812) | (0.18591) | (0.01082) | (0.01390) |
| # jobseekers | 2715 | 9823 | 1646 | 9644 |
| # jobseekers responded \| T = 0 | 1565 | 1565 | 951 | 951 |
| # jobseekers responded \| T = 1 | 1150 | 1150 | 695 | 695 |
| Mean outcome \| T = 0 | 1.24281 | 1.24281 | 0.09976 | 0.09976 |
| Adjusted for non-response | No | Yes | No | Yes |
| IV strength test: F-stat | | 146.121 | | 65.303 |
| IV strength test: p-value | | 0.000 | | 0.000 |

Notes: This table shows treatment effects on specific off-platform search behaviors. The outcome in columns (1) and (2) is the number of applications submitted off-platform in the last 30 days and in columns (3) and (4) is the share of the following 7 search methods the respondent reported using: searching for clients, preparing CV or other related document, seeking assistance from friends or relatives, visiting employers, searching in newspaper/-magazine/social media, contacting some organization, and other steps. Each outcome is regressed on an indicator for treatment assignment and stratification block fixed effects. Odd-numbered columns include selection adjustment terms for survey non-response as described in Section B.6, following DiNardo et al. (2021). The unit of observation is the jobseeker. The first-stage F-statistics jointly test the strength of the four excluded instruments. Standard errors shown in parentheses. For columns without non-response adjustments, these are heteroskedasticity-robust. For columns with non-response adjustments, these are estimated using 500 iterations of a nonparametric bootstrap. Sample size varies across columns because we randomly rotated which questions about intensive margin search were included in each survey.

## B.6 Adjusting for Selection into Survey Response

We survey jobseekers about their off-platform search, employment, and beliefs about the platform and use this in parts of our analysis. The survey response rates are $53.3$ and $42.7\%$ for jobseekers in respectively the phone call control and treatment groups. This means that the treated and control group *survey respondents* might be systematically different, even though randomization ensures no systematic differences between the treated and control group *jobseekers*. However, reassuringly, survey responders and non-responders have almost identical job application rates (Table B.11).

In the presence of survey non-response, average treatment effects on outcomes are not identified without further assumptions. We use a selection adjustment method proposed by DiNardo et al. (2021) that permits identification under weaker assumptions than most other methods. To implement this method, we deliberately randomize features of the survey data collection: the order in which respondents are called, the number of call attempts made to each respondent, whether respondents get text message alerts before phone calls, and whether respondents are offered financial incentives. This allows us to use a selection correction in the spirit of Heckman (1974): we regress off-platform search or employment on treatment and a selection correction term, estimated from a first stage regression of survey response on treatment and the randomised survey features.

DiNardo et al. (2021) show that this approach recovers the population average treatment effect under four assumptions: the survey features are randomized, the survey features do not directly influence outcomes, the survey features influence the probability of response, and the error distribution for the outcome and selection models are jointly normally distributed. The first assumption holds by design. The second assumption is only violated if people are more likely to misreport under some survey features than others, which we view as unlikely but is not testable. The third assumption is testable and holds, as we show below. The fourth assumption is, like all distributional assumptions, arbitrary. But if it fails, this approach still recovers an average treatment effect for the subset of respondents who switch their survey response status in response to variation in the instruments (analogous to compliers in a LATE analysis).

We show the first-stage relationship between the randomized survey features and the response rate for each type of survey question in Table B.12. There are four types of survey questions: any off-platform work, any off-platform search, the proportion of specific search activities done, and beliefs about jobs on the platform. The instruments have a strong impact on the probability of response for all four types of survey questions, shown in the columns. Extra call attempts are the most important instrument, raising the probability of response by 6-10 percentage points with standard errors below 1 p.p. for each four question types. We can strongly reject the null hypothesis that their coefficients are jointly zero ($p < 0.001$ and $F \in [79, 152]$ across the four models).[29]

[29]The common rules-of-thumb for instrument strength, e.g. $F > 10$, are not directly applicable here. They are designed for two-stage least squares estimation rather than the control function estimation we use. Nonetheless,

Table B.11: Comparing Platform Use for Survey Respondents and Non-Respondents

| | (1) Ever applied | (2) Ever invited | (3) # applications | (4) # interviews |
|---|---|---|---|---|
| Ever answered survey | 0.00116 | 0.00990 | 0.02509 | 0.01539 |
| | (0.00977) | (0.00409) | (0.06718) | (0.00754) |
| # jobseekers | 9824 | 9824 | 9824 | 9824 |
| Mean outcome \| Never answered survey | 0.32093 | 0.03351 | 0.91574 | 0.04737 |
| Prop. ever answered survey | 0.36818 | 0.36818 | 0.36818 | 0.36818 |

Notes: This table tests whether survey response is related to different dimensions of platform use as measured by administrative data. Ever answered survey is defined as a dummy equal to 1 if a jobseeker was ever successfully reached for a 20% regular or bonus call, and reached the first module of questions. The unit of observation is the jobseeker. Heteroskedasticity-robust standard errors in parentheses.

We report treatment effects both with and without adjustments for survey responses for all analyses based on survey responses: any off-platform search or employment (Table B.9), specific off-platform search activities (Table B.10), receipt of calls/text messages (Table C.7), and beliefs about jobs on the platform (Table C.8). Adjusting for selection generally makes little difference.

Many researchers instead focus on bounding a different parameter: the average treatment effect in the subpopulation that responds irrespective of treatment status, following Lee (2009). This approach does not require instruments but the bounds are too wide in our setting to be informative.

We can implement a nonparametric version of the DiNardo et al. (2021) method that has a similar spirit to Lee bounds. In this implementation, we split jobseekers into cells based on the combination of randomized survey features they are assigned (e.g. extra call attempts, early call, no survey incentive, no text message in advance). We then select 'response-balanced cells': cells where response rates are balanced between treatment and control groups. Using only the response-balanced cells allows unbiased estimation of average treatment effects for the subpopulation of jobseekers who respond to the survey when they are assigned these specific combinations of survey responses. Intuitively, this approach uses the instruments to identify subpopulations where response rates are balanced between treatment and control groups, collapsing the Lee-style bounds to a single point. This has a similar approach to Lee's suggestion to use covariates to tighten bounds, with the added advantage that we use randomized instruments rather than non-random covariates. Using this approach yields similar point estimates to the main parametric analysis. But using only the response-balanced cells leads to larger standard errors, so we do not emphasize these results.

the statistically strong relationship between response and the instruments is reassuring. As an alternative metric for evaluating instrument strength, following Garlick & Hyman (2022), we note that the instruments jointly shift the probability of responding by at least 9 percentage points in each of the four models. For example, a jobseeker is 12.8 percentage points more likely to complete the beliefs module if she gets extra call attempts, no pre-call text message alerts, and no financial incentive than if she gets a pre-call text message alert, a financial incentive, and no extra call attempts.

Table B.12: Effect of Randomized Survey Features on Probability of Answering Survey Modules

| | Respondent answered survey module on: | | | |
|---|---|---|---|---|
| | Beliefs | Search | Work | Intensive-Margin Search |
| | (1) | (2) | (3) | (4) |
| Many call attempts | 0.09597 | 0.10968 | 0.10369 | 0.06479 |
| | (0.00805) | (0.00969) | (0.00977) | (0.00747) |
| Text message before call | 0.00918 | 0.01204 | 0.01894 | 0.00288 |
| | (0.01342) | (0.01640) | (0.01650) | (0.01237) |
| Incentive | -0.00179 | -0.02066 | -0.02746 | -0.00672 |
| | (0.01339) | (0.01628) | (0.01636) | (0.01229) |
| Text message before call × Incentive | -0.03933 | -0.04929 | -0.03915 | -0.01974 |
| | (0.02246) | (0.02723) | (0.02734) | (0.02068) |
| Assigned early call | | -0.00926 | -0.01824 | |
| | | (0.02051) | (0.02063) | |
| # jobseekers | 9824 | 9824 | 9824 | 9824 |
| Mean outcome | 0.21241 | 0.44089 | 0.47262 | 0.16791 |
| IV strength test: F-stat | 149.907 | 145.690 | 129.027 | 79.075 |
| IV strength test: p-value | 0.000 | 0.000 | 0.000 | 0.000 |

Notes: This table shows the effect of randomized survey features on the probability that jobseekers answer each survey module. We use these estimates to construct selection correction terms for all analyses using survey data, following DiNardo et al. (2021). The outcomes are indicators for ever answering: the survey module about beliefs (column 1), a binary question for any employment (column 2), a binary question for any off-platform search (column 3), and the survey module about intensive-margin off-platform search (column 4). We ask the two binary questions on every call attempt. For a subset of calls, we randomly select one of the beliefs module or the intensive-margin off-platform search module to ask. The randomized features are extra survey call attempts (row 1), a text message telling the respondent that they will be called soon (row 2), an incentive payment of 100 Pakistani Rupees for answering the call (row 3), and assignment to be called early in the survey operation (row 5). We include the interaction between the text message and survey incentive (row 4) because these are directly cross-randomized in the same set of call attempts. The early call attempts were only randomized for a subset of calls that did not include the belief or intensive-margin search questions, so we omit this feature from the regression models shown in columns (1) and (4). All regressions include a treatment indicator and stratification block fixed effects. Heteroskedasticity-robust standard errors shown shown in parentheses. The bottom two rows of the table report results for testing if the coefficients on the randomized survey features are jointly equal to zero.

# C Additional Analysis on Mechanisms

## C.1 Conceptual Framework Appendix

This appendix provides a more formal version of the conceptual framework from Section 4.1.

The platform sends each jobseeker a monthly batch of vacancies that match their education, experience, and occupational preferences. We define $P_{jv}$ as the probability that jobseeker $j$ gets an interview for vacancy $v$ conditional on applying to the vacancy and $V_{jv}$ as the value of an interview. $V_{jv}$ is a reduced-form measure of the present risk-adjusted value of the flow of future utility from the interview. We define $C_{jv}$ as the cost to jobseeker $j$ of applying to vacancy $v$. We omit the $jv$ subscript in the remainder of this section for simplicity. The expected gross return to applying is $PV\delta\beta$, where the quasi-hyperbolic discounting term $\delta\beta$ with $\delta, \beta, \leq 1$ (following Laibson 1997) reflects the fact that interviews occur after applications and allows for the possibility that jobseekers are present biased. We make the natural assumption that jobseekers apply to all jobs where the expected net present value of applying is positive. This is assumption (A1) in the main paper text and can be written as $PV\delta\beta - C > 0$ or $PV > \frac{C}{\delta\beta}$.

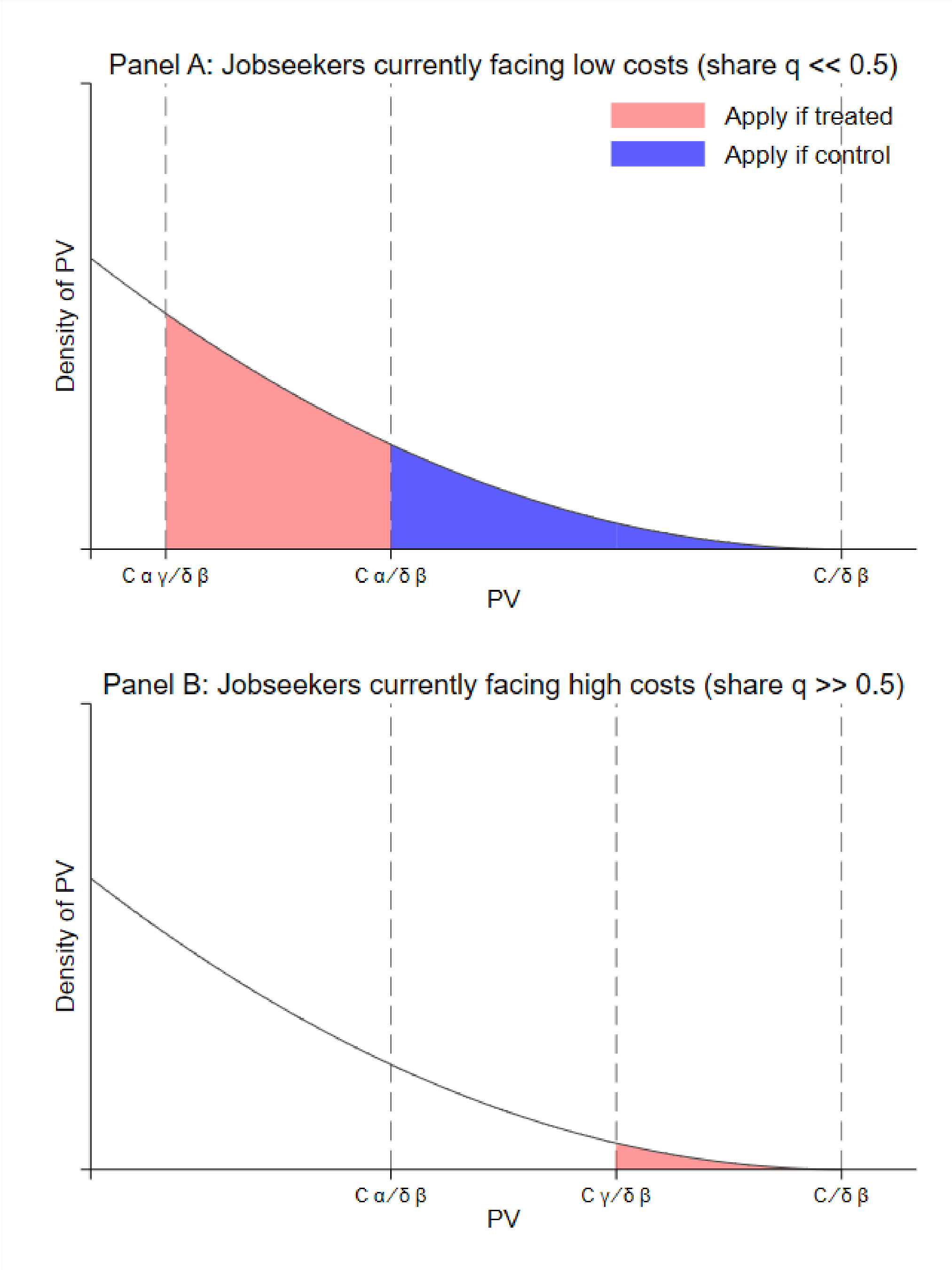
We can introduce heterogeneous application costs into this framework in multiple ways. We begin by assuming that in each month share $q$ of jobseekers face the low application cost $C\alpha < C$, while the remaining share $1 - q$ of jobseekers face the high application cost $C$. This mimics the dynamic investment model of Carroll et al. (2009), used to study retirement contributions. From the model's perspective, it does not matter if this cost is time-varying, with jobseekers moving between low- and high-cost status each month, or time-invariant, with some jobseekers facing permanently high costs and others facing permanently low costs.

We assume that the costs and returns are such that low-cost jobseekers apply to at least one match and high-cost jobseekers apply to no matches. This is assumption (A2) in the main paper text. Formally, this means that $PV > \frac{C\alpha}{\delta\beta}$ for some matches, so low-cost jobseekers apply to some jobs, and $PV < \frac{C}{\delta\beta}$ for all matches, so high-cost jobseekers apply to no jobs. This assumption matches the empirical patterns in the control group: some jobseekers submit applications but many jobseekers never apply or apply in only some periods. We assume costs are either high or low but all predictions of the framework hold with continuously distributed heterogeneity, provided this leads some jobseekers to apply for no vacancies in some periods.

Figure C.1 shows application behavior under these assumptions. In the top panel, low-cost jobseekers apply to the blue-shaded section of the density of $PV$. In the bottom panel, high-cost jobseekers apply to none of their matches. The figure shows identical densities of $PV$ for the two types of jobseekers but the framework's qualitative predictions hold with different densities.

Treatment lowers the application cost, reducing $C$ by a factor $\gamma \in (0, 1)$. Treated low-cost jobseekers apply if $PV > \frac{C\alpha\gamma}{\delta\beta}$. Because $\gamma < 1$, these applications must have lower expected

Figure C.1: Application Decisions for Treated and Control Facing High Versus Low Costs




Notes: This figure shows the application decisions for jobseekers facing low application costs at the time they receive matches (top panel) and jobseekers facing high application costs at the time they receive matches (bottom panel). The blue-shaded sections show the matches to which control group jobseekers apply. The pink-shaded sections show the additional matches to which treatment group jobseekers apply. For simplicity, we show only the right tail of the density of $PV$. This figure is identical to Figure 1 in the main paper, except that the horizontal axis labels show the values for the decision cutoff rules implied by the model.

returns than applications submitted by untreated low-cost jobseekers. These applications go to matches in the pink-shaded section in the top panel of Figure 1. Treated high-cost jobseekers apply if $PV > \frac{C\gamma}{\delta\beta}$, shown in the pink-shaded section in the bottom panel. If $\gamma > \alpha$, i.e., if the treatment-induced drop in application costs is small relative to the natural variation in costs, then these treated high-cost jobseekers' bar for applying is higher than $\frac{C\alpha}{\delta\beta}$, the untreated low-cost jobseekers' bar for applying. This shows the core intuition of the model: marginal applications induced by treatment come from a mix of low-cost jobseekers, whose applications have returns lower than the inframarginal applications, and high-cost jobseekers, whose applications have returns higher than the inframarginal applications if $\gamma > \alpha$. Averaged over these two types of jobseekers, marginal applications can have equal returns to inframarginal applications.[30]

This framework can also explain the large treatment effect on the application rate. The control group's low application rate suggests that the share $q$ of low-cost jobseekers in each month is $<< 0.5$. When $q$ is low, most treatment-induced marginal applications come from high-cost jobseekers, so the treatment effect on the application rate will exceed the control group application rate.[31]

This setup matches some of the potential psychological costs of initiating applications that we discuss in Section 4.1. For example, jobseekers facing the low application cost $C\alpha$ might have lower costs of paying attention to text messages and evaluating the matches, perhaps because they have fewer competing demands for their attention that month.

We can adapt the model slightly to better align with other potential psychological costs of initiating applications. For example, we can adapt the model to align with varying present bias by assuming all jobseekers face application cost $C$ but that in each month share $q$ of jobseekers are time-consistent and have $\beta = 1$, while the remaining share $1 - q$ of jobseekers are present-biased and have $\beta < 1$. This delivers identical decision rules to those derived above with $\alpha$ replaced by $1/\beta$. This approach mimics the dynamic investment model that Duflo et al. (2011) use to study farmers' fertilizer investment. This version of the model matches the data if the share of present-biased jobseekers is high in each period, which is consistent with multiple studies finding relatively high rates of present bias, reviewed by Kremer et al. (2019).

[30] Formally, the mean average return in the control group is $\mathbb{E}\left[PV|PV > \frac{C\alpha}{\delta\beta}\right]$, while the average return in the treatment group is a weighted average of $\mathbb{E}\left[PV|PV > \frac{C\alpha\gamma}{\delta\beta}\right]$ for low-cost jobseekers and $\mathbb{E}\left[PV|PV > \frac{C\gamma}{\delta\beta}\right]$ for high-cost jobseekers. Under our assumption that $\gamma \in (\alpha, 1)$, the second and third expectations are respectively lower and higher than the mean return for control group jobseekers. The second and third expectations have weights $q \cdot Pr\left(PV > \frac{C\alpha\gamma}{\delta\beta}\right)$ and $(1-q) \cdot Pr\left(PV > \frac{C\gamma}{\delta\beta}\right)$ respectively. If the density of $PV$ is strictly continuous, there exists a share $q$ of low-cost jobseekers that equalizes the average returns to control and treated applications.

[31] Formally, the control group application rate is $q \cdot Pr\left(PV > \frac{C\alpha}{\delta\beta}\right)$. The treatment group application rate is $q \cdot Pr\left(PV > \frac{C\alpha\gamma}{\delta\beta}\right) + (1-q) \cdot Pr\left(PV > \frac{C\gamma}{\delta\beta}\right)$. The first term in the treatment group application rate is already larger than the control group application rate because $\gamma$ is defined to be $< 1$. Figure C.1 shows this. The probability in the second term in the treatment group application rate is lower than the probability in the control group application rate under our assumption that $\gamma > \alpha$. But the second term can still be substantially higher than the control group application for low values of $q$.

### C.2 How are Marginal and Inframarginal Applications Directed?

This appendix provides additional results about how marginal and inframarginal applications are directed and explains the complier or latent type method used in sections 4.5 and 4.

**Heterogeneous Treatment Effects by Vacancy Value:** We show several figures confirming the pattern documented in Section 4 that marginal and inframarginal applications are sent to vacancies with similar values to jobseekers. Figure 2 showed heterogeneous treatment effects by two proxies for the value of a vacancy – salary and an index of multiple value proxies – to show that the share of applications sent to high-value vacancies does not differ between treatment and control groups. Figure C.2 replicates this using the within-jobseeker between-vacancy distributions of salary and the index, showing loosely the same pattern. Figure C.3 replicates Figure 2 using the value of interviews rather than vacancies, showing the same pattern. Recall that we test if the share of applications sent to high-value vacancies differs between treatment and control groups by testing if the ratio of the treatment group to control group application rate is equal over the five quintiles shown in each figure. Test results are reported in the note below each figure.

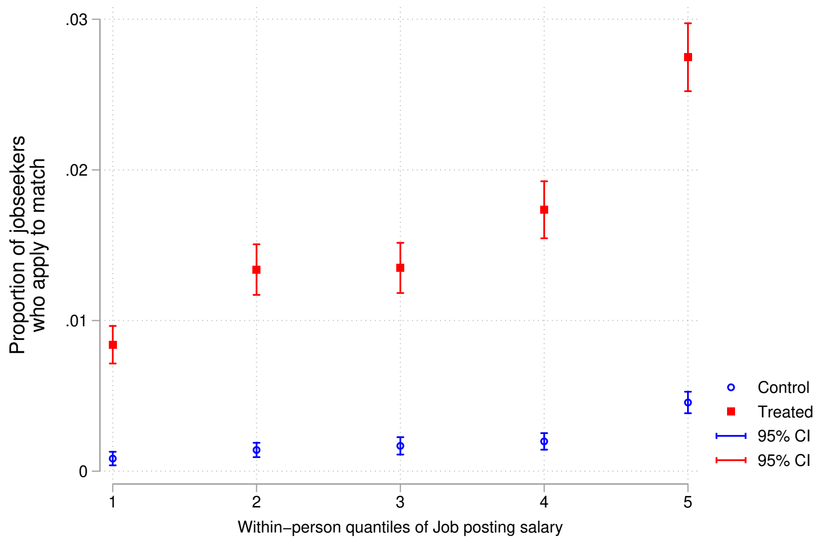
Matches are sent to jobseekers roughly every month, as part of a matching round. Any jobseeker who has received multiple matches in that round receives a *batch* of multiple matches. Roughly two thirds of matches are sent in batches and one third are sent individually. We use this structure to show in the next three figures that treatment and control group jobseekers apply to observationally similar batches of vacancies as well as to similar vacancies. This is consistent with the conceptual framework.

Figure C.4 shows how the phone call treatment shifts the number of applications that respondents make in each of these rounds. Panel A shows the full distribution, while Panel B shows the distribution conditional on a positive number of applications. The conditional distributions are similar between treatment and control group, with confidence intervals fully overlapping. This shows that the entire treatment effect on applications comes from the shift from applying to zero vacancies in a given round to a positive number of applications. This pattern is consistent with the conceptual framework: some jobseekers miss applying to some batches of matches due to temporarily high present bias or psychological application costs. If, instead, treatment shifted some jobseekers from making one to making two or more applications within a batch of matches, this would not be explained by a reduction in the psychological cost of initiating applications.

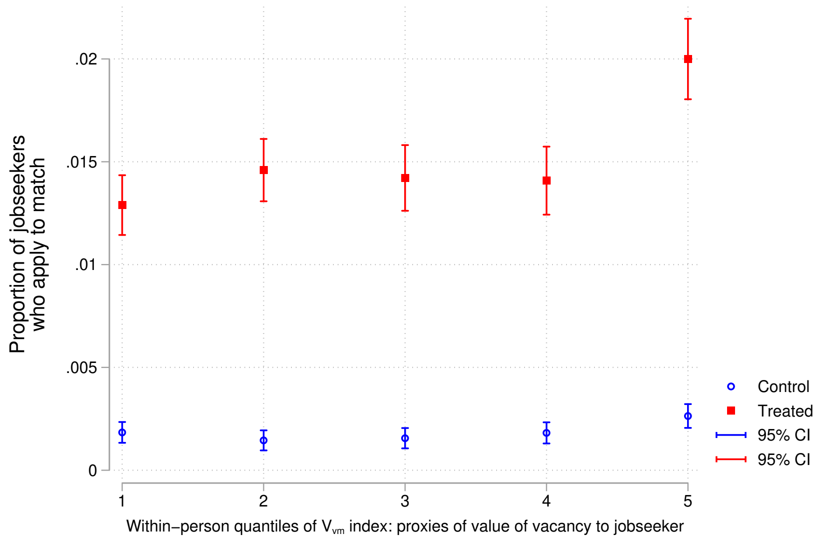
Figure C.5 shows heterogeneous treatment effects collapsing the data to the level of the matching round, replicating the results in Figure 2. Finally, Figure C.6 repeats this analysis measuring the value of a round based on the highest-value vacancy rather than average over the vacancies in the round. Results are similar across all approaches, showing that treatment and control group jobseekers apply to observationally similar batches of matches.

Figure C.2: Heterogeneous Treatment Effects on Applications by Vacancy Value Using Within-Jobseeker Variation in Value

Panel A: Match level: Quintiles of salary within-jobseeker



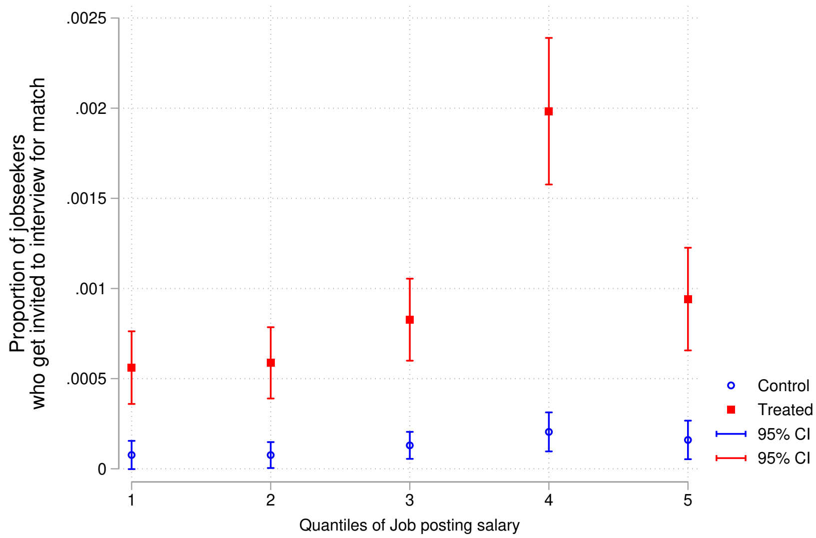

Panel B: Match level: Quintiles of $V_{vm}$ index within-jobseeker



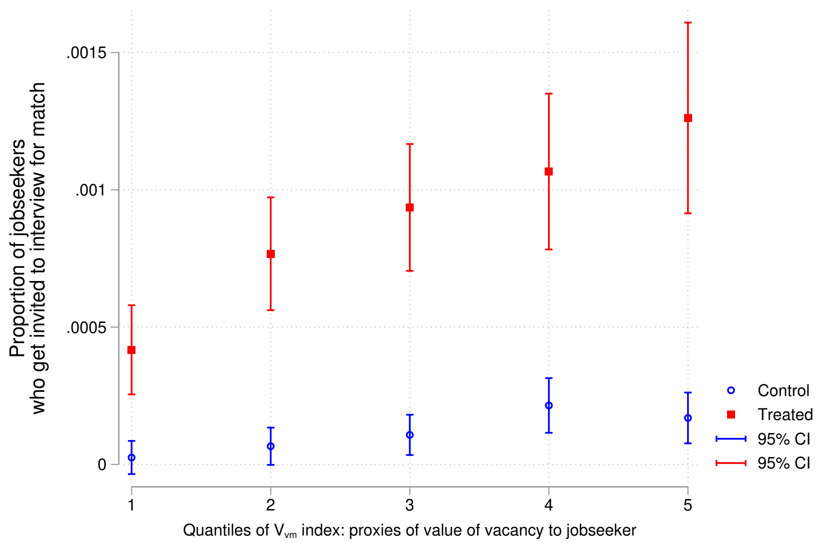

Notes: This figure shows heterogeneous treatment effects of the phone call treatment on applications by proxies for the value of the job posting. Panel A shows heterogeneity by job posting salary, defining quintiles based on the distribution of salary within-jobseeker. Panel B shows heterogeneity by the $V_{vm}$ index described in Section 3.2, again defining quintiles based on the distribution of salary within-jobseeker. The p-values for testing that the share of applications submitted to each quintile is equal between treatment groups is $0.652$ in Panel A and $0.444$ in Panel B. The unit of observation is the jobseeker $\times$ vacancy match. Results in both panels are conditional on stratification block fixed effects. Solid vertical lines show 95% confidence intervals, constructed using heteroskedasticity-robust standard errors, clustering by jobseeker.

Figure C.3: Heterogeneous Treatment Effects on Interviews by Vacancy Value

Panel A: Match level, by posted salary



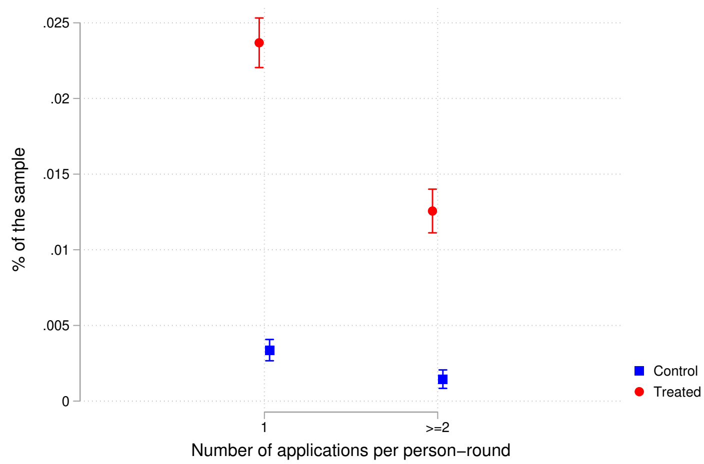

Panel B: Match level, by $V_{vm}$ index



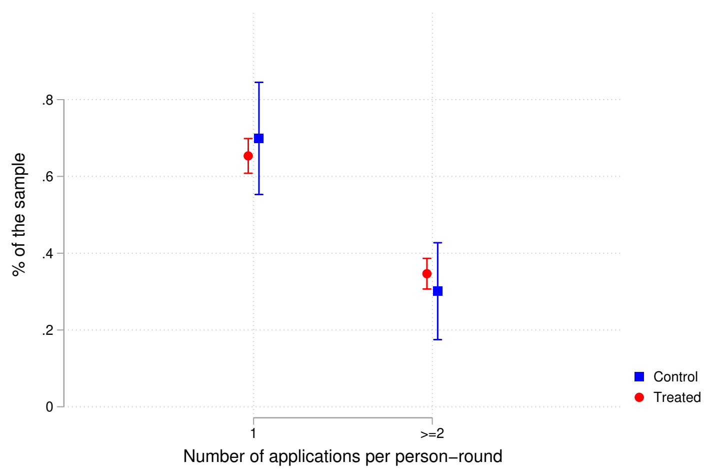

Notes: This figure shows heterogeneous treatment effects of the phone call treatment on interviews by proxies for the value of the job posting. Panels A and B show heterogeneity by job posting salary and $V_{vm}$ index described in Section 3.2 using the within-jobseekers between-vacancy distribution. The p-values for testing that the share of applications submitted to each quintile is equal between treatment groups is $0.984$ in Panel A and $0.950$ in Panel B. The unit of observation is the jobseeker $\times$ vacancy match. Results in both panels are conditional on stratification block fixed effects. Solid vertical lines show 95% confidence intervals, constructed using heteroskedasticity-robust standard errors, clustered by jobseeker.

Figure C.4: Treatment Effects on the Number of Applications per Jobseeker × Matching Round

Panel A: Treatment Effects on Each Positive Number of Applications



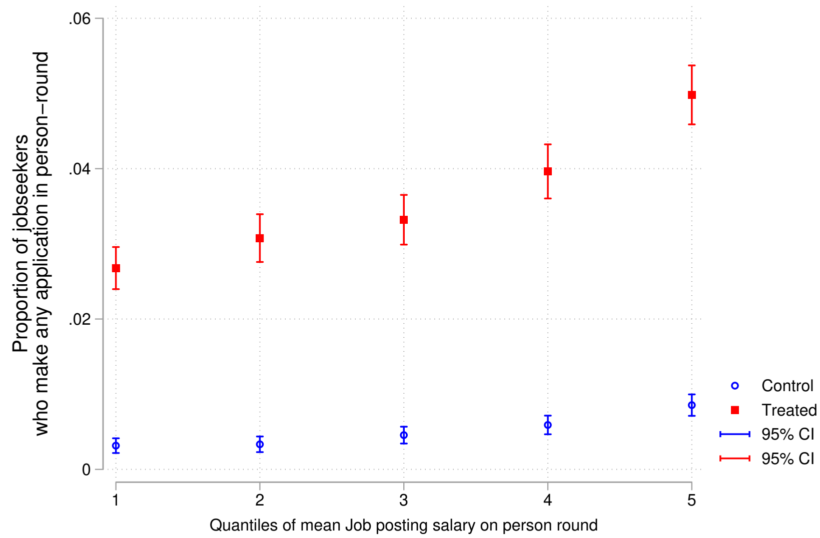

Panel B: Treatment Effects on Each Positive Number of Applications, Scaled by Pr($>0$ Applications)



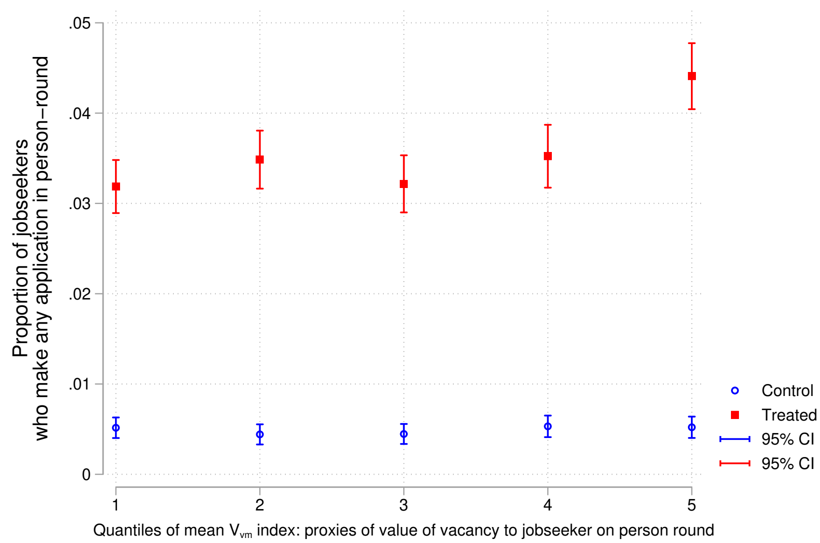

Notes: This figure shows treatment effects on the number of applications submitted per jobseeker × round. Estimation uses one observation per person-round, restricts the sample to jobseeker-rounds with at least two matches (65% of the data), conditions on stratification block fixed effects, and uses standard errors clustered by jobseeker. In Panel B, each estimate is multiplied by the probability of submitting $>0$ applications so that the estimated effects for 1 and $>2$ applications sum to one within each of the treatment and control groups. This allows us to show that treatment increases the number of job applications purely by increasing the number of rounds to which applications are submitted, rather than shifting the number of applications submitted within rounds to which jobseekers apply anyway.

Figure C.5: Heterogeneous Treatment Effects on Applications by Vacancy Value Using Jobseeker $\times$ Matching Round Level Data

Panel A: Jobseeker $\times$ round level, by salary



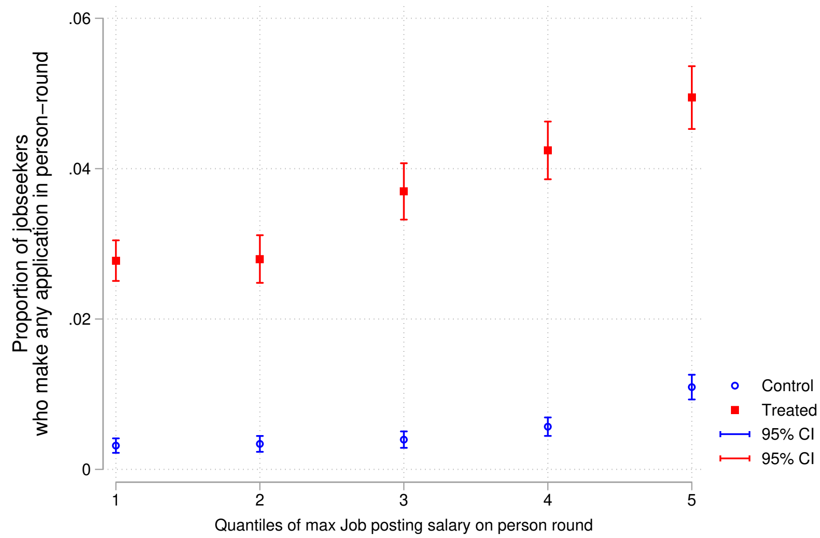

Panel B: Jobseeker $\times$ round level, by $V_{vm}$ index



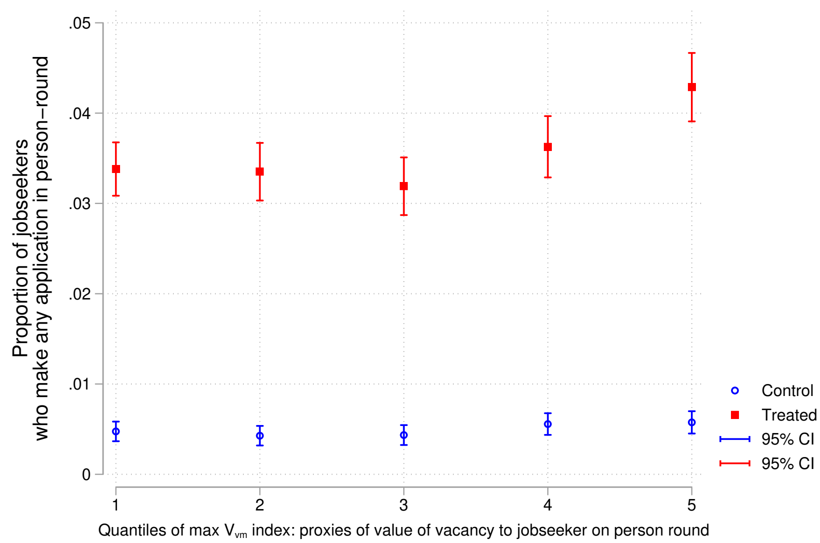

Notes: This figure shows heterogeneous treatment effects of the phone call treatment on applications by proxies for the value of the job posting. All analysis in this figure uses one jobseeker $\times$ round as a unit of observation, averaging over the values of the vacancies in that unit. Panels A and B show heterogeneity by job posting salary and $V_{vm}$ index described in Section 3.2 using the within-jobseekers between-vacancy distribution. The p-values for testing that the share of applications submitted to each quintile is equal between treatment groups is 0.302 in Panel A and 0.226 in Panel B. Results in both panels are conditional on stratification block fixed effects. Solid vertical lines show 95% confidence intervals, constructed using heteroskedasticity-robust standard errors, clustered by jobseeker.

Figure C.6: Heterogeneous Treatment Effects on Applications by Vacancy Value Using Jobseeker $\times$ Matching Round Data

Panel A: Person-round level: maximum salary instead of mean




Person-round level: maximum $V_{vm}$ index instead of mean




Notes: This figure shows heterogeneous treatment effects of the phone call treatment on applications by proxies for the value of the job posting. All analysis in this figure uses one jobseeker $\times$ round as a unit of observation, based on the maximum value of the vacancies within that unit. Panels A and B show heterogeneity by job posting salary and $V_{vm}$ index described in Section 3.2 using the within-jobseekers between-vacancy distribution. The p-values for testing that the share of applications submitted to each quintile is equal between treatment groups is $0.062$ in Panel A and $0.961$ in Panel B. Results in both panels are conditional on stratification block fixed effects. Solid vertical lines show 95% confidence intervals, constructed using heteroskedasticity-robust standard errors, clustering by jobseeker.

**Method for Complier / Latent Type Analysis:** We argue that the marginal job applications submitted due to treatment and inframarginal job applications submitted in the absence of treatment have similar characteristics, in terms of jobseeker characteristics (Section 4.5, Table 4) and vacancy/match characteristics (Section 5). In this appendix we describe the method used to support this argument, which is adapted from Marbach & Hangartner (2020).[32]

In the standard language of instrumental variable analysis, inframarginal applications are submitted by 'always-taker' types and marginal applications are submitted by 'complier' types. 'Never-taker' types do not submit applications by definition and there are no 'defier' types under the standard monotonicity assumption. We cannot observe the latent type of each individual match. But all applications submitted to untreated matches are by definition inframarginal. Hence the population share of inframarginal applications is $\mu^{AT} = \mathbb{E}[\text{Apply} \mid \text{Treat} = 0]$ and the mean value of each covariate $X$ for inframarginal applications is $\mu_X^{AT} = \mathbb{E}[X|\text{Apply} = 1, \text{Treat} = 0]$.

All applications submitted to treated matches are by definition either marginal or inframarginal. The treatment group's mean application rate is $\mathbb{E}[\text{Apply} \mid \text{Treat} = 1]$, so the population share of marginal applications is $\mu^C = \mathbb{E}[\text{Apply} \mid \text{Treat} = 1] - \mu^{AT}$. The mean value for each covariate $X$ in the treatment group is the average of the mean values for compliers and always-takers, weighted by their relative frequency: $\mathbb{E}[X|\text{Apply} = 1, \text{Treat} = 1] = \frac{\mu^{AT} \cdot \mu_X^{AT} + \mu_X^C \cdot \mu^C}{\mu^{AT} + \mu^C}$. Hence the mean value of each covariate $X$ for inframarginal applications is $\mu_X^C = \frac{(\mu^{AT} + \mu^C) \cdot \mathbb{E}[X|\text{Apply}=1,\text{Treat}=1] - \mu^{AT} \cdot \mu_X^{AT}}{\mu^C}$.

We can estimate $\mu_X^{AT}$ and $\mu_X^C$ for each covariate $X$ using combinations of sample averages and estimate the standard errors using the Delta method. We include stratification block fixed effects in all estimation and cluster standard errors by jobseeker.

**Results from Complier / Latent Type Analysis:** Table C.1 shows that there are some differences between mean values of observed characteristics marginal and inframarginal applications but these differences do not show consistently higher values for marginal or for inframarginal applications. For example, marginal applications are directed to jobs that offer slightly lower salaries, but are more likely to offer flexible hours. We summarize these measures by constructing an inverse covariance-weighted average of the value proxies, $V_{vm}$, and find no difference between the mean values of this index for marginal and inframarginal applications.

Latent interview probabilities are another proxy for the value of each application; these also do not differ on average between marginal and inframarginal applications. To show this, we estimate the latent probability that each match would yield an interview if an application were submitted using the same method introduced in Section 4.5, but now incorporating vacancy and match level characteristics into the prediction model. The mean probability is similar between marginal and in-

[32] This method is a special case of the $\kappa$-weighting method proposed by Abadie (2003). We do not need to use Abadie's more general method because this special case works for the problem we study – covariate means for compliers with a binary treatment and binary instrument.

framarginal applications when estimated using only these vacancy- and match-level characteristics or also including jobseeker characteristics. Finally, we interact each latent interview probability measure with the value index to create an omnibus proxy for $PV$. The means for marginal and inframarginal applications do not differ.

These patterns show that marginal and inframarginal applications are sent to vacancies with similar values to jobseekers, consistent with the conceptual framework. As a final check, we replicate our main analysis conditional on vacancy- and match-level characteristics and confirm that the estimated treatment effect on applications and return to marginal applications are unchanged (Table B.5, columns 1, 5, and 6).

Table C.1: Comparing Observed Characteristics of Inframarginal and Marginal Job Applications

| | (1) Inframarginal applications | (2) Marginal applications | (3) Difference (p-value) |
|---|---|---|---|
| **Firm characteristics** | | | |
| Leave one out ratio of firm interviews to applications (on platform) | 0.061 | 0.058 | -0.003 |
| | | | (0.583) |
| Firm baseline ratio of interviews to applications (off-platform) | 0.705 | 0.738 | 0.033 |
| | | | (0.053) |
| Firm # employees | 88.347 | 43.720 | -44.627 |
| | | | (0.001) |
| Firm # vacancies in last year | 12.301 | 9.096 | -3.205 |
| | | | (0.002) |
| **Vacancy characteristics** | | | |
| Ln(posted salary) | 9.848 | 9.704 | -0.143 |
| | | | (0.000) |
| $<$ median working hours | 0.600 | 0.582 | -0.018 |
| | | | (0.489) |
| Allows employees to work flexible hours | 0.697 | 0.803 | 0.106 |
| | | | (0.000) |
| Offers any benefits | 0.767 | 0.758 | -0.010 |
| | | | (0.624) |
| **Match characteristics** | | | |
| Exact education match \| vacancy requires high ed | 0.168 | 0.265 | 0.097 |
| | | | (0.008) |
| Exact experience match \| vacancy requires experience | 0.176 | 0.166 | -0.011 |
| | | | (0.684) |
| Short commute | 0.540 | 0.456 | -0.083 |
| | | | (0.002) |
| Gender preference aligned | 0.509 | 0.570 | 0.062 |
| | | | (0.012) |
| **Predicted interview probabilities and value of vacancy** | | | |
| $\hat{P} \mid X_{vm}$: Prob. interview \| vacancy and match characteristics | 0.063 | 0.063 | 0.001 |
| | | | (0.874) |
| $\hat{P} \mid X_{jvm}$: Prob. interview \| jobseeker, vacancy and match characteristics | 0.063 | 0.065 | 0.002 |
| | | | (0.575) |
| $V_{vm}$ index: proxies of value of vacancy to jobseeker | 0.242 | 0.253 | 0.011 |
| | | | (0.853) |
| $\hat{P} \mid X_{jvm} \times$ ln(posted salary) | 0.632 | 0.656 | 0.024 |
| | | | (0.552) |
| $\hat{P} \mid X_{jvm} \times V_{vm}$ index | 0.231 | 0.234 | 0.004 |
| | | | (0.810) |

Notes: Table shows the means of covariates for the inframarginal applications that are submitted irrespective of treatment status (column 1) and marginal applications that are submitted only if treated (column 2). Column 3 shows the difference between the covariate means for marginal and inframarginal applications. p-values reported in parentheses in column 3 are estimated using heteroskedasticity-robust standard errors clustered by jobseeker. The unit of observation is the jobseeker × vacancy match. Exact education match is an indicator for an exact match between the employer's preferred field of educational specialization and the jobseeker's field; this variable is conditional on vacancies requiring high education. Exact experience match is an indicator for a match in which the jobseeker has experience in the same occupation as the vacancy; this variable is conditional on vacancies requiring experience.

$\hat{P}$: All predicted interview probabilities have been estimated using logit LASSO specification, using applications from control group jobseekers. The logit LASSO model is allowed to select from the following characteristics. At the match level, high salary relative to respondent's matches; high salary relative to all matches; short commute (below median distance); jobseeker is overqualified relative to firm's minimum and preferential experience or educational requirements; jobseeker has an exact match of educational specialization for the job advert; jobseeker has an exact match of work experience for the job; and the job advert states preferring candidates from the jobseeker's gender. At the vacancy and firm level: industry classifications; vacancy occupation codes; work days for the vacancy; number of employees; total # of vacancies opened by the firm in the last year reported at baseline; minimum and maximum salary offered for the vacancy; ln(salary net of commute cost); ln(hourly salary); commute cost; vacancy offers a written employment contract; vacancy offers a permanent employment contract; total # of benefits offered by the vacancy; any benefits offered by vacancy; less than median working hours; whether the firm allows its employees to work flexible hours multiple times a week, once a week, multiple times a month, once a month, once after every few months or not at all; whether the firm is open to hiring women for the vacancy, number of positions to be filled; minimum years of experience and education required; any education required; any experience required; preferred years of experience; preferred years of experience in the same sector; firm provides pick and drop transport services to all, some or no employees; firm is located in a commercial, industrial or residential area; firm used web platform to advertise a vacancy at baseline; firm used third party outsourcing to advertise a vacancy at baseline; firm used newspaper to advertise a vacancy at baseline; whether CV drop-off was allowed at the firm's location at baseline; whether the firm reached out to its contacts to advertise a vacancy at baseline; whether the firm ever used newspaper to advertise a vacancy on platform or off platform at baseline; whether the firm ever used web platforms to advertise a vacancy on platform or off platform at baseline; whether the firm ever used third party outsourcing to advertise a vacancy on platform or off platform at baseline; years of education required for a vacancy posted by firm at baseline; an indicator for whether the firm either has no female employees and has no interest in hiring them, has no female employees but is open to hiring them, or has some female employees; total # of vacancies listed by the firm on platform; and firm baseline ratio of interviews to applications.

$V_{vm}$ index: is an inverse covariance-weighted average constructed using vacancy and match level characteristics, defined in the note to Table 3.

### C.3 Additional Mechanisms Results

We explain in Section 4.2 that the conceptual framework predicts that the value of matches receiving applications should have a wider range in the treatment group. To test this, we estimate treatment effects on the variance, 10th percentile, and 25th percentile of log salary for matches that receive applications, using a nonparametric bootstrap clustered by jobseeker to obtain standard errors on these treatment effects. Table C.2 shows treatment raises the variance and lowers the 10th and 25th percentiles for both log salary and the proxy index $V_{vm}$ that combines multiple proxies for match and vacancy value. This is consistent with the framework's prediction that marginal treatment-induced applications should go to vacancies with the same average value as inframarginal applications but more dispersed values.

Table C.2: Treatment Effects on Dispersion of Value of Matches Receiving Applications

| | Ln(Salary) | | $V_{vm}$ index | |
|---|---|---|---|---|
| | Variance | 10th pctile | Variance | 10th pctile |
| | (1) | (2) | (3) | (4) |
| Control | 3.13 | 9.2 | 0.926 | 2.37 |
| | (0.468) | (0.018) | (0.060) | (0.036) |
| Treatment | 5.18 | 8.99 | 0.964 | 2.34 |
| | (0.223) | (0.000) | (0.032) | (0.014) |
| Treatment effect | 2.06 | -0.223 | 0.038 | -0.025 |
| | (0.527) | (0.018) | (0.067) | (0.038) |

Notes: This table shows how treatment changes the dispersion of the value of vacancies that receive applications, testing the model prediction that treatment should raise this dispersion. The table columns show dispersion statistics – variance and 10th percentile – of two proxies for vacancy value – log monthly salary and the index $V_{vm}$ of vacancy- and match-level proxies for vacancy value defined in the note to Table 3. The table rows show the levels of these dispersion statistics for the treatment and control groups and the treatment effect. Standard errors are estimated using 1000 iterations of a nonparametric bootstrap, clustering by jobseeker.

**Pecuniary and time costs:** Here we show results for the mechanism experiments described in Section 4.3. Column 1 of Table C.3 compares the effects of our main phone call initiation treatment to the effects of a randomized text message reminder that the jobseeker can ask the platform to call them back about a job posting. The free callback reminder treatment has an effect one hundredth of the size of the effect of the main phone call treatment, and the two effects are statistically significantly different (p = .017).

Column 2 of Table C.3 compares the effects of our main phone call initiation treatment to the effects of randomly offering some control group jobseekers the option to text the platform and ask for a callback at a specific time. This eliminates the differential wait time between the main treatment and control groups. This callback request treatment has an effect one quarter of the size of the effect of the main phone call treatment, and the two effects are statistically significantly different (p = .002). Each column uses only the set of jobseeker $\times$ vacancy matches from rounds in which the relevant feature was randomized.

Table C.3: Mechanism Experiment: Treatment Effects on Applications of Reductions in Pecuniary and Time Costs

| | Apply | |
|---|---|---|
| | (1) | (2) |
| Phone call treatment$_j$ | 0.00342 | 0.00226 |
| | (0.00145) | (0.00047) |
| Free callback salience treatment$_{jt}$ | 0.00003 | |
| | (0.00012) | |
| Callback request treatment$_{jt}$ | | 0.00059 |
| | | (0.00029) |
| # matches | 13126 | 54135 |
| # jobseekers | 4423 | 7004 |
| Mean outcome \| T = 0 | 0.00000 | 0.00030 |
| P-value for equality of treatments | 0.01742 | 0.00235 |
| Round FE | Yes | Yes |

Notes: Column 1 sample includes matches from jobseekers in the standard phone call treatment arm, jobseekers randomized into a free callback reminder, and the control group (mutually exclusive), from one round during which the mechanism experiment was active. Column 2 sample includes matches in the standard phone call treatment arm, a callback request treatment randomized at the person-round level, and the control group (mutually exclusive), from three rounds in which the experiment was active. The unit of observation is the jobseeker $\times$ vacancy match. Results are conditional on stratification block and round fixed effects. Heteroskedasticity-robust standard errors, clustered by jobseeker, are shown in parentheses.

**Reminder effects:** Here we show results for the mechanism experiments and non-experimental analysis relating to reminder effects discussed in Section 4.4. Table C.4 shows the effect of a reminder text message sent to a random subsample of control group jobseekers at the same time that the treatment group jobseekers receive calls. If reminder effects explain our results, this should have a similar effect to that of the phone call treatment. The effect of the reminder message is one-fourteenth as large as the effect of the phone call treatment in the same matching rounds and statistically significantly smaller ($p < 0.001$).

Table C.4: Mechanism Experiment: Treatment Effects on Applications of Reminder Text Messages

| | (1)<br>Apply |
|---|---|
| Phone call treatment | 0.00224 |
| | (0.00046) |
| Reminder text message treatment | 0.00016 |
| | (0.00015) |
| # matches | 54152 |
| # jobseekers | 7013 |
| Mean outcome \| T = 0 | 0.00010 |
| P-value for equality of treament | 0.00003 |

Notes: Table shows coefficients from regressing an indicator for job application on phone call treatment and eligibility for the reminder text message treatment. Sample includes matches in the standard phone call treatment arm, a reminder text message treatment which was randomized at the person-round level, and the control group (mutually exclusive), from three matching rounds during which the mechanism experiment was active. The phone call control group jobseekers eligible for the "crossover" treatment are coded as treated for the phone call treatment. The unit of observation is the jobseeker $\times$ vacancy match. The regression includes stratification block fixed effects. Heteroskedasticity-robust standard errors are shown in parentheses, clustered by jobseeker.

Table C.5, Column 1, shows the effect of the timing of the phone call made to the treatment group. We randomize the order in which we call jobseekers for the phone call treatment, within the application window between the text message job alert and the deadline for job applications. If reminder effects explain our results, we expect that the treatment should have a stronger effect for jobseekers called later within this window, as they will have had more time to forget to apply. Instead, we find that the later the phone call made to the jobseeker, the smaller the treatment effect on applications. This suggests that reminder effects do not explain our results.

Table C.5, Column 2, tests for heterogeneous treatment effects by the duration between the job alert text message and the application deadline. This duration is not randomly assigned, but varies due to logistical factors such as the number of call center agents on staff at the time of the matching round. We interact the duration of this window with treatment, controlling for quarter fixed effects to address variation over time in these logistical factors. Table C.5, Column 2 shows the results. If reminder effects explained our results, we would expect treatment to have a larger effect when there is a longer application window, as jobseekers will have had more time to forget to apply. The results show that treatment has a smaller effect when the window is longer, again suggesting reminder effects do not explain our results.

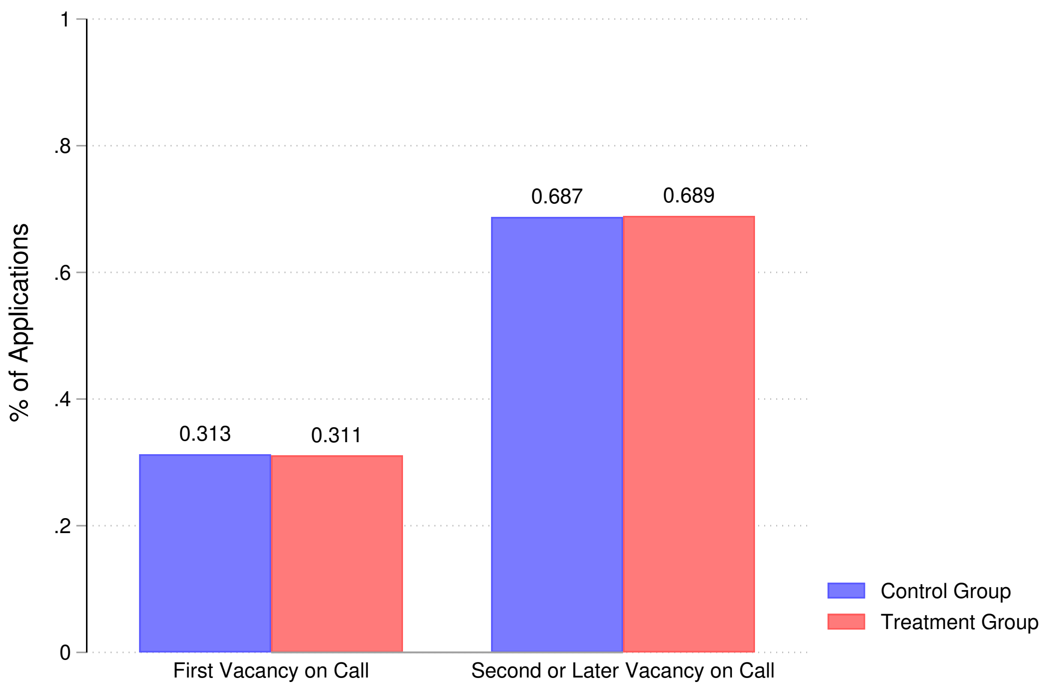
Table C.5: Mechanism Analysis: Treatment Effects on Applications by Timing of Phone Call and Length of Application Window

| | Apply | |
|---|---|---|
| | (1) | (2) |
| Phone call treatment | 0.01379 | 0.01616 |
| | (0.00090) | (0.00100) |
| Phone call treatment × Days between job alert and first call assigned to jobseeker | -0.00018 | |
| | (0.00010) | |
| Days between job alert and deadline | | 0.00005 |
| | | (0.00002) |
| Phone call treatment × Days between job alert and deadline | | -0.00072 |
| | | (0.00004) |
| # matches | 1116952 | 1005463 |
| # jobseekers | 9831 | 9011 |
| Mean outcome \| T = 0 | 0.00185 | 0.00135 |
| Round FE | Yes | No |
| Quarter FE | No | Yes |

Notes: Column (1) shows coefficients from regressing an indicator for job application on phone call treatment and its interaction with days between job alert and first call assigned to the jobseeker. This variable is coded as zero for jobseekers in the control group. Column (2) shows coefficients from regressing an indicator for job application on phone call treatment, days between job alert and deadline, and the interaction of phone call treatment and days between job alert and deadline. The sample size varies as the records of deadlines were lost from some early matching rounds. All regressions include stratification block fixed effects. The unit of observation is the jobseeker × vacancy. Heteroskedasticity-robust standard errors are shown in parentheses, clustered by the jobseeker.

**Pressure to apply:** Here we show results for the mechanism analysis relating to pressure and encouragement described in Section 4.4. Figure C.7) shows the proportion of applications that are directed to the first vacancy listed on the phone call versus later vacancies listed. If pressure were responsible for the main treatment effects on applications, we would expect to see treatment group jobseekers applying to the first vacancy listed on the call at a higher rate than control group jobseekers. Instead, 31% of applications go to the first vacancy listed on the call in both the treatment and control groups. To help contextualize this result, we note that 22% of all vacancies are listed first on the call. So jobseekers are disproportionately likely to apply to first-listed vacancies, but this pattern does not differ between treated and control jobseekers.

Figure C.7: Proportion of Applications by Order in which Vacancies are Listed




Notes: This figure shows the proportion of applications that jobseekers make to the first vacancy mentioned on the call versus vacancies mentioned second or later on the call. Sample consists of all applications (jobseeker $\times$ vacancy matches in which *Apply* = 1) in person-rounds in which the jobseeker receives at least two matches.

**Differential information receipt:** Here we show evidence that our main findings are unlikely to be explained by treatment group jobseekers receiving more information about matches than control group jobseekers.

Section 5.1 asked if treatment group jobseekers might acquire more information about vacancies by asking call center agents for more information during calls. To test this, we estimate treatment effects using only matching rounds where the call center agents had access to no additional information about the vacancies. This covers 80% of matching rounds containing 72% of all matches. Table C.6 shows that our key findings still hold in this test: treatment increases the application rate by more than 600% and we cannot reject equal returns for marginal and inframarginal applications.

Section 5.1 also asked if the phone call treatment might increase application rates because jobseekers were more likely to receive phone calls than text messages. To test this, we survey respondents and ask if they have received matches by phone call and/or text message in the previous 14 or 30 days (recall period randomized). Table C.7 shows treatment effects on respondents' reports of receiving information about job matches. Column (1) shows that treated jobseekers are not more likely to report receiving a match. Column (2) shows that the same pattern holds when we use the selection correction method described in Section B.6.

Most jobseekers were not actually sent matches in the last 14 or 30 days, because they match to an average of less than one vacancy per matching round. This explains why the control group mean for reporting receiving matches is only 39%. To account for this pattern, we estimate treatment effects on reporting receiving matches controlling for actually being sent a match in at least one of the last two matching rounds. Columns (3) and (4) show that these treatment effects remain close to zero. They also show that jobseekers sent matches are 26 percentage points more likely to report receiving matches, a reassuring check on the quality of the survey data.[33]

[33] We do not expect 100% of jobseekers sent matches in the last two matching rounds to report receiving them, for two reasons. First, the recall periods cover 14 or 30 days before the survey, while the two matching rounds cover roughly 60 days on average. Second, some measurement error in recall is natural, either because jobseekers forget they received matches (and hence underreport receiving matches) or because they forget the exact date they received them (which might lead to overreporting or underreporting receiving matches).

Table C.6: Treatment Effects on Job Search & Search Returns Excluding Matching Rounds when Call Center Agents had More Information About Vacancies

| | (1) Apply | (2) Interview | (3) Int. $\times$ V | (4) Interview | (5) Int. $\times$ V |
|---|---|---|---|---|---|
| Phone call treatment | 0.01739 | 0.00083 | 0.00303 | | |
| | (0.00097) | (0.00009) | (0.00038) | | |
| Apply | | | | 0.04767 | 0.17435 |
| | | | | (0.00433) | (0.01764) |
| # matches | 801922 | 801922 | 801922 | 801922 | 801922 |
| # jobseekers | 9603 | 9603 | 9603 | 9603 | 9603 |
| Mean outcome \| T = 0 | 0.00208 | 0.00008 | 0.00030 | 0.00008 | 0.00030 |
| Mean outcome \| T = 0, Apply = 1 | | | | 0.03708 | 0.14428 |
| p: IV effect = mean \| T = 0, Apply = 1 | | | | 0.18247 | 0.35332 |
| IV strength test: F-stat | | | | 320.1 | 320.1 |
| IV strength test: p-value | | | | 0.00000 | 0.00000 |

Notes: This table repeats the analysis reported in Table 3 excluding the 20% of matching rounds when the call center agents had more information available about each vacancy and could provide that information to jobseekers. The results show that returns to marginal applications are still roughly constant when jobseekers cannot use the phone calls to get more information about the vacancies.

Column 1 shows the coefficient from regressing an indicator for job application on treatment assignment. Column 2 shows the coefficient from regressing an indicator for interview invitation on treatment assignment. Column 3 shows the coefficient from regressing an indicator for interview invitation weighted by a proxy index for the value of the vacancy to the jobseeker, $V_{vm}$, on treatment assignment. Column 4 shows the coefficient from regressing an indicator for interview invitation on job application, instrumented by treatment assignment. Column 5 shows the coefficient from regressing an indicator for interview invitation weighted by the proxy index $V_{vm}$ on job application, instrumented by treatment assignment. The proxy index $V_{vm}$ is an inverse covariance-weighted average (following Anderson 2008) constructed using vacancy-level characteristics log salary and indicators for offering any non-salary benefits, below-median working hours, and allowing flexible hours as well as indicators for the match-level characteristics of vacancy salary exceeding the jobseeker's expected salary, below-median commuting distance, the jobseeker's educational specialization exactly matching the vacancy's preference, and the jobseeker's work experience exactly matching the vacancy's preference. Anderson-style indices, by construction, have zero means and hence some negative values. But multiplying the interview invitation indicator by a negative value would not produce sensible results. Hence we recenter the index so it has strictly positive values.

All regressions use one observation per jobseeker $\times$ vacancy match, include stratification block fixed effects, and use use heteroskedasticity-robust standard errors clustered by jobseeker, which are shown in parentheses. The p-value is for a test of equality between the IV treatment effect and the mean interview rate for control group applications. The first-stage F-statistic and p-value are for the test of weak identification from Kleibergen & Paap (2006).

Table C.7: Mechanism Analysis: Treatment Effects on Recalling Receiving Matches

| | Respondent reported receiving matches | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Phone call treatment | -0.01533 | 0.00111 | -0.00691 | 0.00373 |
| | (0.02357) | (0.03499) | (0.02277) | (0.03302) |
| Platform sent match in last 2 rounds | | | 0.25577 | 0.2597 |
| | | | (0.02085) | (0.02175) |
| # jobseekers | 2177 | 14069 | 2177 | 14069 |
| # responses $\mid$ T = 0 | 978 | 978 | 978 | 978 |
| # responses $\mid$ T = 1 | 1199 | 1199 | 1199 | 1199 |
| Mean outcome $\mid$ T = 0 | 0.38753 | 0.38753 | 0.38753 | 0.38753 |
| IV strength test: F-stat | | 57.845 | | 70.838 |
| IV strength test: p-value | | 0.000 | | 0.000 |
| Adjusted for non-response | No | Yes | No | Yes |

Notes: This table shows treatment effects on the probability that respondents report receiving matches from the platform from either a phone call or a text message. The recall period is randomized to 14 or 30 days. All jobseekers who responded to the survey were asked these questions, even if the platform did send them a recent match. Each outcome is regressed on an indicator for treatment assignment, an indicator for a 30-day recall period, and stratification block fixed effects. Even-numbered columns include selection adjustment terms for survey non-response described in Section B.6, following DiNardo et al. (2021). The first-stage F-statistics jointly test the strength of the four excluded instruments. The regressions in columns (3) and (4) control for an indicator equal to one if the platform sent a match to the jobseeker in the last 2 rounds, which cover roughly 2 months. Standard errors shown in parentheses. For columns without non-response adjustments, these are heteroskedasticity-robust and clustered by jobseeker. For columns with non-response adjustments, these are estimated using 500 iterations of a nonparametric bootstrap, clustered by jobseeker. The unit of observation is a survey response, as some jobseekers were surveyed twice, which explains why the sample sizes in columns (2) and (4) are larger than the number of jobseekers in the study. Only 0.6% of jobseekers complete two surveys.

**Beliefs about returns to search on the platform:** Section 5.2 introduced the possibility that treatment shifts application rates by changing jobseekers' beliefs about returns to search on the platform. We directly test if the phone call treatment shifts beliefs about $P$ and $V$ by surveying jobseekers. We ask: "Suppose Job Talash sends you one hundred job ads over a year. Based on your past experience with our job matching service, how many of these ads do you think would be desirable for you?" and "Suppose you apply for all the jobs you think are desirable jobs. How many of those do you think would make you an offer?"[34] We use a jobseeker-level version of equation (1) to estimate treatment effects on these two belief measures. Table C.8 shows that treatment does not shift jobseekers' answers to either of these questions. Jobseekers in the control group on average think that they will receive an offer from 43% of jobs they are interested in; the phone call treatment increases this by 1 percentage point (standard error 1.8. p.p.). Jobseekers in the control group on average think that 32% of the vacancies on the platform would be desirable for them; the phone call treatment decreases this by 0.5 p.p. (standard error 1.6 p.p.). The even-numbered columns show that results are similar when we adjust for survey non-response using the same method introduced described in Appendix B.6. The survey data indicate that treatment does not increase respondents' perceptions of the average values of $V$ or $P$ on the platform, and hence cannot explain the large treatment effects on applications.

Table C.8: Mechanism Analysis: Beliefs About Potential Returns to Search on Job Talash Platform

| | % desirable jobs respondent believes would make an offer (P) | | % of jobs respondent believes desirable (V) | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Phone call treatment | -0.01082 | -0.02583 | -0.00662 | 0.00164 |
| | (0.01775) | (0.02089) | (0.01593) | (0.01861) |
| # jobseekers | 2003 | 9483 | 2081 | 9483 |
| # jobseekers answered \| T = 0 | 1191 | 1191 | 1238 | 1238 |
| # jobseekers answered \| T = 1 | 812 | 812 | 843 | 843 |
| Mean outcome \| T = 0 | 0.42681 | 0.42681 | 0.31339 | 0.31339 |
| Adjusted for non-response | No | Yes | No | Yes |
| IV strength test: F-stat | | 145.679 | | 140.017 |
| IV strength test: p-value | | 0.000 | | 0.000 |

Notes: This table shows treatment effects on beliefs collected as part of jobseeker followup surveys. Each outcome is regressed on an indicator for treatment assignment and stratification block fixed effects. Columns (2) and (4) include selection adjustment terms for survey non-response as described in Section B.6, following DiNardo et al. (2021). The unit of observation is the jobseeker. The first-stage F-statistics jointly test the strength of the four excluded instruments. Standard errors shown in parentheses. For columns without non-response adjustments, these are heteroskedasticity-robust. For columns with non-response adjustments, these are estimated using 500 iterations of a nonparametric bootstrap.

[34] We measure beliefs about offer probabilities for jobs to which the respondent would consider applying, because shifting beliefs about $P$ for jobs the jobseeker would not consider should not influence their application decisions.

**Returns to random search:** Table C.9 shows the treatment effects of listing a vacancy first during the application phone call and hence encouraging applications to randomly chosen vacancies, discussed in Section 5.3. Listing the vacancy first substantially increases the application rate (column 1) but produces decreasing returns to search (columns 4 and 5) that are not statistically significantly different to zero.

Table C.9: Treatment Effects of Lowering Cost of Applying to Randomly Chosen Vacancies

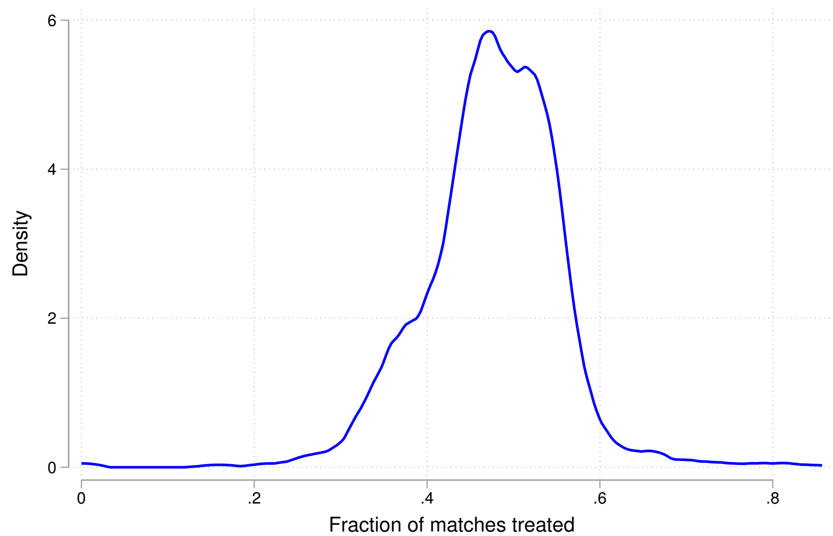
| | (1) Apply | (2) Interview | (3) Int. $\times$ $V_{vm}$ | (4) Interview | (5) Int. $\times$ $V_{vm}$ |
|---|---|---|---|---|---|
| Vacancy listed first in batch on phone call | 0.00440 | 0.00011 | 0.00042 | | |
| | (0.00065) | (0.00009) | (0.00033) | | |
| Apply | | | | 0.02437 | 0.09491 |
| | | | | (0.02052) | (0.07590) |
| # matches | 938,284 | 938,284 | 938,284 | 938,284 | 938,284 |
| # jobseekers | 9255 | 9255 | 9255 | 9255 | 9255 |
| # vacancies | 1317 | 1317 | 1317 | 1317 | 1317 |
| Mean outcome \| T = 0 | 0.00627 | 0.00039 | 0.00143 | 0.00039 | 0.00143 |
| Mean outcome \| T = 0, Apply = 1 | | | | 0.06287 | 0.22851 |
| p: IV effect = mean \| T = 0, Apply = 1 | | | | 0.07675 | 0.10859 |
| IV strength test: F-stat | | | | 45.17 | 45.17 |
| IV strength test: p-value | | | | 0.00000 | 0.00000 |

Notes: This table shows the effect of varying the relative marginal cost of applying to an individual vacancy within a round, by changing the order in which vacancies are listed on the application phone call. Column 1 shows the coefficient from regressing an indicator for job application on an indicator equal to 1 for a vacancy that is listed first in the call to the jobseeker during the round and 0 otherwise. Column 2 shows the coefficient from regressing an indicator for interview invitation on an indicator for vacancy listed first in the call. Column 3 shows the coefficient from regressing an indicator for interview invitation weighted by a proxy index for the value of the vacancy to the jobseeker, $V_{vm}$, on an indicator for vacancy listed first in the call. Column 4 shows the coefficient from regressing an indicator for interview invitation on job application, instrumented by vacancy listed first on the call. Column 5 shows the coefficient from regressing an indicator for interview invitation weighted by the proxy index for $V_{vm}$ on job application, and instrumented by vacancy listed first on the call. See the note below Table 3 for a definition of $V_{vm}$. The p-value is for a test of equality between the IV treatment effect and the mean interview rate for control group applications. The first-stage F-statistic and p-value are for the test of weak identification from Kleibergen & Paap (2006). All columns: The sample is restricted to jobseeker- rounds with $\geq 2$ matches, which includes 84% of all matches in the full sample. For the first part of the study, vacancy order was not fully randomized and varied by the first digit of the firm ID and subsequently. For the remainder of the study, vacancy order was randomized within the sets of high- and low-priority matches for the jobseeker based on relevant experience. As a result, all these regressions control for the first digit of firm ID and its interaction with the time period when job order was/was not randomized. The unit of observation is the jobseeker × vacancy match. Heteroskedasticity-robust standard errors are shown in parentheses, with two-way clustering by the jobseeker and vacancy. Mean outcomes are for the control group, i.e. vacancies listed second or later on the telephone call. The proportion of applications submitted to the first vacancy is 0.31.

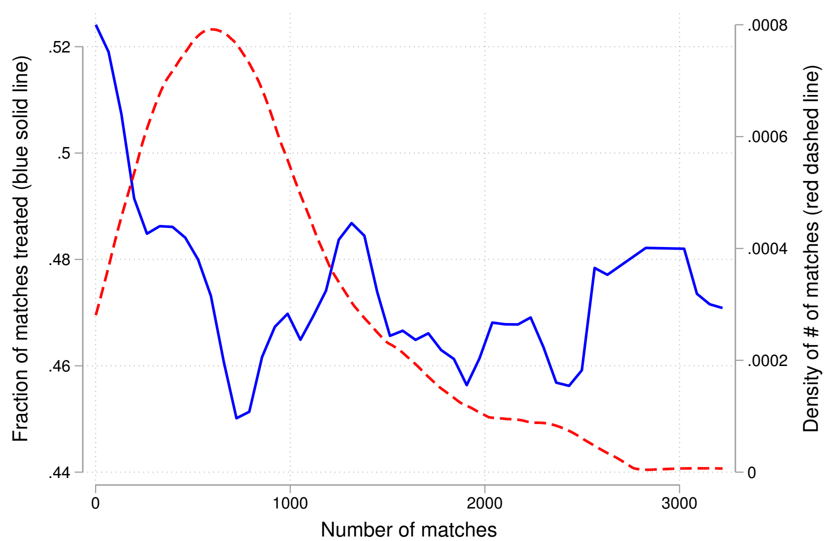
# D Additional Analysis on Spillovers

Figure D.1: Variation in Treatment Rate Between Vacancies

Panel A: Density of Vacancy-Level Treatment Rate

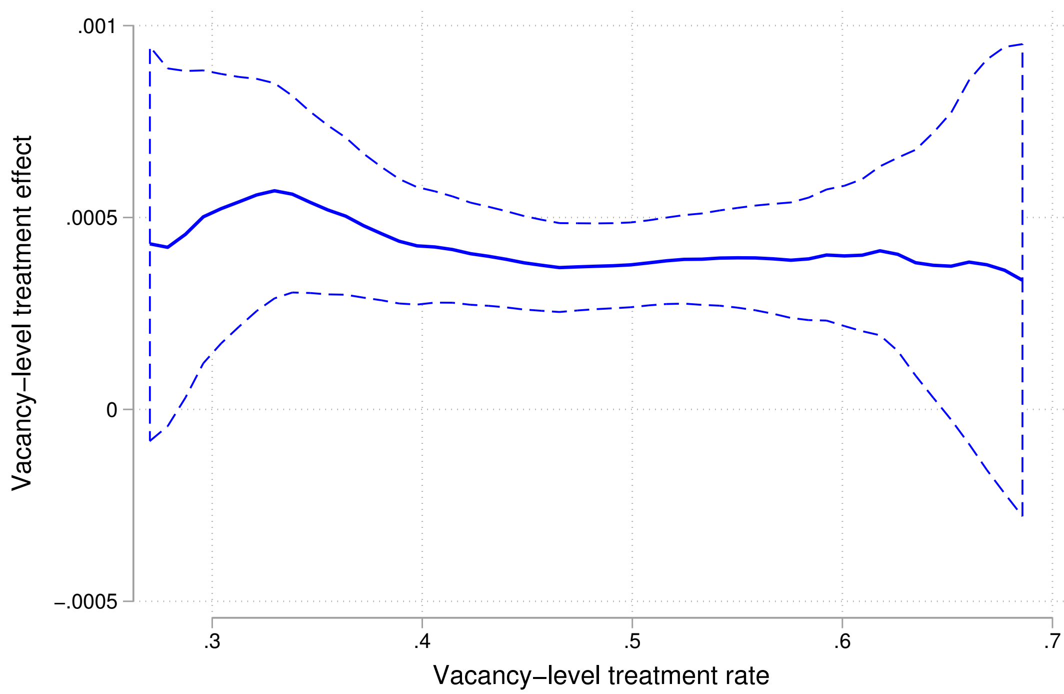



Panel B: Mean of Vacancy-Level Treatment Rate by Number of Matches




Notes: This figure shows the variation between vacancies in the fraction of matched jobseekers who are treated. This variation is used to identify the spillovers analysis in Section 6. Panel A shows the density of treatment rates at the vacancy level. Panel B shows the results from a local linear regression of vacancy-level treatment rate against the number of jobseekers matched to each vacancy (solid blue line). This panel demonstrates that the vacancy-level treatment rate is not systematically related to vacancy size. It also shows the density of vacancy size (dashed red line) to illustrate the available variation.

Figure D.2: Relationship between Vacancy-Level Treatment Effects on Interviews and Treatment Rates




Notes: Figure shows the relationship between vacancy-level treatment effects on interviews and treatment rates, as a test for spillover effects on interview invitations. The figure is constructed by estimating the treatment effect on interview invitations separately for each of the 1,340 vacancies, estimating the share of jobseekers matched to each vacancy who are treated, and then regressing the former quantity on the latter using local linear regression. The dashed lines show 95% confidence intervals. The relatively flat slope of this regression is evidence against spillover effects: it shows that jobseekers' treatment effects on interviews do not depend on the share of other jobseekers matched to the vacancy who are treated, even though a higher treatment rate leads to more applications.

Table D.1: Descriptive Analysis of Application-Interview Relationship at the Vacancy Level

| | # applications | # interviews | | | Any interview | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| # matches | 0.01254 | -0.00000 | 0.00001 | -0.00012 | -0.00003 | -0.00003 |
| | (0.00285) | (0.00013) | (0.00038) | (0.00014) | (0.00003) | (0.00003) |
| Treatment rate | 14.38843 | | | | | |
| | (6.94709) | | | | | |
| # applications | | 0.01336 | 0.01215 | | 0.00102 | |
| | | (0.00429) | (0.02937) | | (0.00115) | |
| # applications: mid tercile | | | | 0.28726 | | 0.09396 |
| | | | | (0.05329) | | (0.02173) |
| # applications: top tercile | | | | 0.73644 | | 0.06900 |
| | | | | (0.14354) | | (0.02610) |
| Outcome mean | 6.77629 | 0.38852 | 0.38852 | 0.38852 | 0.12528 | 0.12528 |
| IV strength test: F-stat | | | 4.290 | | | |
| IV strength test: p-value | | | 0.039 | | | |
| p: terciles equal | | | | 0.000 | | 0.000 |
| # vacancies | 1340 | 1340 | 1340 | 1340 | 1340 | 1340 |

Notes: This table shows the relationship between the number of applications and interviews at the vacancy level, to contextualize the spillovers analysis in Section 6. Column (1) shows that vacancies get more applications if they are matched to more jobseekers and if more of these jobseekers are treated. Column (2) shows that vacancies that get more applications issue more interview invitations. Column (3) shows that the positive relationship between applications and interviews persists when we instrument the number of applications with the fraction of matched jobseekers who are treated, although the instrument is relatively weak and the second stage estimate is imprecise. Column (4) replicates column (2) but replaces the number of interviews with indicators for the middle and top terciles of the number of applications. Columns (5) and (6) replicate columns (2) and (4) but replace the number of interviews with an indicator for conducting any interviews as an outcome. Columns (2) and (4) - (6) provide non-experimental evidence against congestion effects: when the number of applications gets very high, firms do not issue fewer interview invitations or decline to interview any applicants. All regressions condition on firm size and sector and on vacancy occupation, salary, education and experience requirements, and number of matched jobseekers. The unit of observation is the vacancy. Heteroskedasticity-robust standard errors shown in parentheses.