G²LM|LIC
Gender, Growth, and Labour Markets in Low Income Countries Programme

WP

G²LM|LIC Working Paper No. 88 | December 2024

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

Kate Vyborny ® (World Bank)
Robert Garlick ® (Duke University and IZA)
Nivedhitha Subramanian ® (Bates College)
Erica Field (Duke University, BREAD, NBER and IZA)

IZA Institute of Labor Economics
Initiated by Deutsche Post Foundation

UKaid
from the British people

G²LM|LIC Working Paper No. 88 | December 2024

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

Kate Vyborny ® (World Bank)
Robert Garlick ® (Duke University and IZA)
Nivedhitha Subramanian ® (Bates College)
Erica Field (Duke University, BREAD, NBER and IZA)

G²LM|LIC
c/o IZA – Institute of Labor Economics
Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-510

Email: glm-contact@iza.org

I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

UKaid
from the British people

G²LM|LIC Working Paper No. 88 | December 2024

ABSTRACT

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

Understanding specific 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. Perhaps surprisingly, the marginal treatment-induced applications have approximately constant rather than decreasing returns. These results are consistent with a directed search model in which some jobseekers miss some high-return vacancies due to psychological costs of initiating applications. These findings show that small reductions in search costs can substantially improve search outcomes in environments with some relatively inactive jobseekers.

JEL Classification:
J20, J60, O10

Keywords:
behavioural economics, job search, search frictions, platform

Corresponding author:
Robert Garlick
Department of Economics
Duke University
213 Social Science Building
Box 90097
Durham NC 27708
USA
E-mail: robert.garlick@duke.edu

* We are grateful for valuable feedback from the editor and three anonymous reviewers; 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, Harmalah Khan, and Sahil Nisar 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 assistance for this research program. 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. Author order is randomized.

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 decisions (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, measured in terms of interview invitations, are approximately constant rather than decreasing. This raises 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. Thus, they have a wide range of education, from incomplete primary to graduate levels, and a wide range of baseline labor force attachment, from employed and searching to non-employed and non-searching. This sample breadth is unusual in experimental job search studies (Poverty Action Lab, 2022), partly because of the household listing and partly because using the platform requires only basic literacy, a simple phone, and almost no airtime, generating very few technological and pecuniary barriers to search.

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 the qualifications and preferences they report at sign-up. Users in the control group must call the platform to initiate applications, while users in the treatment group also receive a phone call inviting them to apply, so they do not need to initiate calls

[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.

to apply. The experimental design holds constant other parts of the search process: the phone call has negligible effects on pecuniary and time costs of applying, provides no direct encouragement or pressure to apply, and provides no extra information about vacancies. Hence, we interpret treatment as primarily reducing the psychological cost of initiation.

Our two key findings are that phone calls dramatically increase the job application rate, and that the average return to additional applications is roughly constant rather than decreasing. Treatment increases the share of jobseeker-vacancy matches receiving applications by seven-fold, from 0.2 to 1.5%.[2] Using treatment as an instrument for applying shows that marginal applications submitted due to treatment have a 5.9% probability of yielding interviews, which is neither substantively nor statistically different from the 6.3% probability for applications from the control group. 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. An additional experiment shows 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 complications around the exclusion and monotonicity conditions in our instrumental variables analysis.

The finding of roughly constant returns is surprising. We might expect jobseekers to prioritize applying to vacancies with the highest combination of expected interview probabilities and desirable attributes, and hence that extra applications would have decreasing returns. This behavior would be consistent with many models of 'directed' job search, reviewed by Wright et al. (2021). The constant returns finding by itself is consistent with canonical models of 'random' job search (e.g. Pissarides 2000) but we show later that our other results are inconsistent with random search.

To explain our two key findings, we propose a modified directed search model. As in many models, in each period jobseekers apply to those vacancies with expected return above the cost of applying. Our key assumption is that application costs vary – within jobseeker through time and/or between jobseekers – and can be large enough that some jobseekers submit no applications in periods in which they face high costs, even to high return vacancies. For example, a jobseeker may face a high psychological cost of initiating applications when they are stressed by illness, domestic responsibilities, or work. The phone call treatment reduces application costs, leading naturally to more applications. However, these 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 period without treatment – now apply to *some* vacancies. Because the second

[2] It is unsurprising that most matches do not generate applications. A match simply means the jobseeker qualified for the job and is interested in that occupation. In any search environment, jobseekers will apply to only a small subset of such jobs. The same pattern holds on some other platforms that economists have studied (Appendix A).

type of marginal applications can have higher returns than inframarginal applications, the average return to marginal applications – averaged across treated jobseekers facing high and low cost draws – can equal the return to inframarginal applications.

This framework shows how the common assumption of decreasing returns to additional search for *each individual jobseeker in each period* can lead to constant returns to additional search *averaged over jobseekers and periods*, provided some jobseekers are not actively searching in some periods. This model's predictions match both our two key findings and our secondary results about which jobseekers submit marginal applications and where they direct them.

Given the importance of application costs in the experiment and the model, we explore in detail what types of costs jobseekers face. We show that pecuniary and time costs of applying on our platform are low, and that additional experiments that directly reduce the pecuniary or time costs of applying have little effect on applications. This leaves psychological costs of initiating applications as the most likely category of cost addressed by the phone call treatment. Within this category, the existing literature suggests multiple candidates, including the cognitive 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), all of which can vary through time. Our findings and interpretation are consistent with research showing that eliminating the need to initiate decisions can raise financial and health investment, reviewed by DellaVigna (2009). Our key modeling assumption of heterogeneous psychological costs borrows from behavioral models that seek to explain low adoption of seemingly high-return investments (Carroll et al., 2009; Duflo et al., 2011).

We can test and reject several plausible alternative explanations. Perhaps most importantly, the constant returns finding by itself is consistent with prominent models of 'random' job search, the main alternative to directed search models. In these models, vacancies are homogeneous and jobseekers randomly choose where to apply (Pissarides, 2000). But random search models do not match other results we find: not only do we observe that jobseekers on our platform direct applications to vacancies with desirable attributes, but we also run an additional experiment to encourage random search that generates sharply *decreasing* returns to marginal applications.

We can also reject some specific behavioral explanations – reminders or explicit encouragement or pressure – because these 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.

We do not find evidence that this additional search has negative spillovers on other jobseekers. Individual jobseekers' interview probabilities are unaffected by competing against more treatment-induced applications. We treat 50% of jobseekers on the platform, increasing total search enough that large spillovers are possible. But our scope for detecting spillovers is limited by the fact that the majority of firms' job applications come from sources other than the platform.

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. 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 most platforms do not track data on job offers or employment.[3]

Our paper makes three contributions. First, by studying psychological costs of job search, we contribute to a nascent literature on behavioral job search, reviewed by Cooper & Kuhn (2020). Existing work shows patterns of job search that are consistent with present bias, motivated reasoning, and reference dependence (e.g. DellaVigna et al. 2017, 2022; Mueller & Spinnewijn 2022; Paserman 2008), but does not isolate the psychological costs of initiating job applications.[4]

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, there are few evaluations of policies designed to directly target behavioral factors, all of these focus on helping jobseekers make plans to search more, and none compares 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 different changes to the job search environment can produce substantially different impacts on search effort and different returns to search. Behavioral channels may be relevant for many other job search policies: 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.[5] However, research into these job search policies has not sought to pin down behavioral components.

Third, we provide a direct estimate of returns to additional search effort caused by reducing behavioral barriers. Returns to search effort, typically interpreted as job applications, are central

[3] Banfi et al. (2019), Belot et al. (2022b), Faberman & Kudlyak (2019), He et al. (2021), Marinescu (2017a), 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., 2023; 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).

[4] 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).

[5] 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. (2023), Belot et al. (2018, 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), Field et al. (2024), and Franklin (2017) study subsidies. Arni & Schiprowski (2019), Bolhaar et al. (2020), Lechner & Smith (2007), and Schiprowski (2020) study caseworkers.

to 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 variation 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 make this third contribution, we combine experimental variation in search costs with data on both individual applications and the outcomes of those applications. This is a very rare combination in the literature. Many papers study the effect on employment of search subsidies or requirements for 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 any 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.[6] The closest work to our own shows how additional policy-induced job applications affect unemployment duration (Arni & Schiprowski, 2019; Lichter & Schiprowski, 2021). While we do not observe administrative data on employment, we extend 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.[7]

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 and alternative interpretations in Section 4. Section 5 discusses spillover effects.

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.

[6]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 papers 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).

[7]Alfonso Naya et al. (2020), Behaghel et al. (2023), 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.

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). Gender is an important feature of Lahore's labor market (Gentile et al., 2023) but we do not focus on gender in this paper because all of our main experimental results hold for both women and men. Job search and matching platforms are a growing feature of Pakistan's labor market, particularly in major cities 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.[8] 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 (40% of the sample) followed by visiting establishments to ask about vacancies (23%) and applying formally (15%).[9] Only 4% had used a job search assistance program or online platform other than Job Talash. The average respondent had 7.9 years of work experience with an interdecile range of 0-16. Respondents' education levels also vary widely: 15% had no education, 15% had completed secondary school, and 25% had a university degree. 31% were female and the average age was 31, with an interdecile range of 20-45.

The study sample starts from a representative household listing, but only 5.3% of adults from this listing completed registration on Job Talash and became part of the study sample. In Table

[8] 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.

[9] The prevalence of network-based search matches patterns in other developing economies (Caria et al., 2024; Carranza & McKenzie, 2024). 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.

A.1, we compare our study sample to the population of Lahore, captured by both the official Labor Force Survey and our household listing. Our study sample is slightly younger, more male, more educated, less likely to be employed and much more likely to be searching for work. This selection means that our findings should not be extrapolated to the population of Lahore, but rather speak to a population who registers on a new job search platform. See Gentile et al. (2023) for additional discussion on selection from the household listing to registration on this platform.

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, location, 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.[10] 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.[11] 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.

[10] These summary statistics weight each vacancy by the number of jobseekers who match with the vacancy. We define a jobseeker $\times$ 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.

[11] 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).

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

	Control group mean & std. dev			Selection into applying in control group	Balance checks
	All	Never apply	Ever apply	Mean for ever apply − never apply [p-value]	Mean for treated − for control [p-value]
	(1)	(2)	(3)	(4)	(5)
Employed and searching	0.200	0.184	0.292	0.108	0.034
	(0.400)	(0.388)	(0.455)	[0.000]	[0.228]
Employed and not searching	0.141	0.148	0.097	-0.051	-0.028
	(0.348)	(0.355)	(0.296)	[0.000]	[0.256]
Searching and not employed	0.345	0.338	0.386	0.048	0.024
	(0.475)	(0.473)	(0.487)	[0.033]	[0.344]
Not searching and not employed	0.314	0.330	0.225	-0.105	-0.030
	(0.464)	(0.470)	(0.418)	[0.000]	[0.307]
Search method: network	0.397	0.379	0.506	0.127	0.032
	(0.489)	(0.485)	(0.500)	[0.000]	[0.476]
Search method: formal application	0.154	0.150	0.176	0.026	0.028
	(0.361)	(0.357)	(0.381)	[0.147]	[0.651]
Search method: asked at establishments	0.225	0.211	0.305	0.094	0.032
	(0.417)	(0.408)	(0.461)	[0.000]	[0.728]
Years of work experience	7.85	7.89	7.62	-0.27	-0.22
	(8.88)	(8.97)	(8.25)	[0.463]	[0.568]
Education: none	0.146	0.154	0.083	-0.071	-0.012
	(0.353)	(0.361)	(0.276)	[0.000]	[0.294]
Education: primary or some secondary	0.457	0.470	0.361	-0.109	-0.023
	(0.498)	(0.499)	(0.481)	[0.000]	[0.871]
Education: complete secondary	0.148	0.143	0.180	0.037	0.002
	(0.355)	(0.350)	(0.384)	[0.027]	[0.673]
Education: university degree	0.250	0.232	0.376	0.144	0.033
	(0.433)	(0.422)	(0.485)	[0.000]	[0.335]
CV: excellent score	0.093	0.092	0.098	0.006	0.084
	(0.291)	(0.289)	(0.298)	[0.812]	[0.868]
CV: good score	0.330	0.342	0.299	-0.043	0.032
	(0.471)	(0.475)	(0.459)	[0.281]	[0.970]
CV: average or lower score	0.576	0.566	0.603	0.037	-0.116
	(0.495)	(0.496)	(0.491)	[0.383]	[0.872]
Female	0.303	0.307	0.271	-0.036	0.022
	(0.460)	(0.461)	(0.445)	[0.063]	[0.329]
Age	30.7	31.0	28.7	-2.3	-0.5
	(9.7)	(9.8)	(9.1)	[0.000]	[0.307]
# matches sent by platform	113	107	154	47	-
	(121)	(120)	(119)	[0.000]	
# applications on platform	0.226	0.000	1.83	1.83	-
	(0.863)	(0.000)	(1.76)	[0.000]	
# interviews through platform	0.014	0.000	0.115	0.115	-
	(0.128)	(0.000)	(0.349)	[0.000]	

Notes: This table shows summary statistics for jobseekers' baseline characteristics and, in the last three rows, platform use characteristics. This table uses one observation per jobseeker. Column (1) shows the mean and standard deviation for the control group. Column (2) shows the mean and standard deviation for the control group sample of jobseekers who never applied to any job on the platform. Column (3) shows the mean and standard deviation for the control group sample of jobseekers who apply to at least one job on the platform. Column (4) shows the difference between the mean for the control group sample of jobseekers who apply to at least one job and those who never applied, 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 (5) 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 and include fixed effects for the strata within which treatment was randomized (see footnote 14). We leave column (5) 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.

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 the employer's preferences in only 18 and 13% of matches respectively.[12] 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

[12]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 $\approx$ 0.03 USD 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).

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.

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 might in principle 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, most matches do not generate applications: the average jobseeker submits only 0.23 applications and applies to 0.2% of matches they receive. The application count is unsurprisingly right-skewed: 74% of jobseekers submit zero applications and 5% submit more than 5 applications. Column 4 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 4.4. This application rate may strike some readers as surprisingly low. However, it is unsurprising that most matches do not generate applications. A match simply means the jobseeker qualified for the job and is interested in that occupation. In any search environment, jobseekers will apply to only a small subset of such jobs. The application rate is comparable to some other platforms in countries ranging from France to Mozambique. Furthermore, our sample deliberately includes people who were not actively searching at baseline and includes all registered platform users. In contrast, some studies of job search platforms restrict their samples to only "active" users, defined in various ways, which naturally generates much higher application rates. See Table A.3 for details on application rates on other search platforms.

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.[13]

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.3, 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 (Table 2, column 2, 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.3, panel A). This pattern also holds for the summary index of match value (Figure C.3, panel B).

These patterns naturally motivate our research. On the one hand, the facts that job applications

[13]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).

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.

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 *additionally* 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 5).[14] Treated jobseekers are called in a random order, starting as soon as the text messages are sent. 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 4.4 that our findings are robust to omitting rounds when they had access to more information.

[14] 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.

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 *psychological* 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 4.4 that we can rule out several other interpretations based on the platform design, additional experiments we run, and additional survey measures.

Both the treatment and control conditions on Job Talash have many similarities to other large job search platforms. On Job Talash and these platforms, users can choose to receive notifications about jobs that match their qualifications and preferences and can apply using platform-generated CV templates. On most other platforms, users submit applications online or using phone apps. These are different technologies to Job Talash's text-and-phone approach but they also allow scope for higher or lower psychological costs of initiating applications. For example, platforms can present information about matched jobs in ways that impose higher or lower attention costs. See Table A.2 for a more detailed comparison of application processes on different platforms.

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. These results imply roughly constant returns to the additional search effort induced by the treatment.

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. μ_b is a fixed effect for the stratification blocks within which

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).

treatment was randomized (see footnote 14). 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 are 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.[15]

The treatment effects on both applications and interview invitations are broad-based. Treatment substantially raises job application and interview rates 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.1). 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 Appendix B.2, 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.

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:

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

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

β 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 4, 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). As we discuss further below, this is not a consequence of low power.

Marginal and inframarginal applications also have equal returns measured in 'value-weighted'

[15] A few firms do not provide the list of jobseekers they interviewed. We assume no jobseekers matched to these vacancies are interviewed. Our main results are unchanged if we instead code these interview values as missing.

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.2).

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 creates scope for heterogeneous returns from applying to different types of matches. Furthermore, Table B.1 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 p.p. 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 returns may be substantially lower with very high search effort.[16]

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

[16] 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 only drop from 5.5 to 3.7%.

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.

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 areas of labor economics such as audit studies, highlighted in the review by Neumark (2018).

The disadvantage of platforms is that they do not generally provide data on employment outcomes, so evaluations relying on employment outcomes require off-platform data. To this end, we survey jobseekers about their employment and find a treatment effect on self-reported employment of 1 percentage point, with a standard error of 2 p.p (Table B.7, column 4). However, this estimate should be interpreted very cautiously for three reasons. First, the surveys take place on average 40 months after randomization, so they capture the effects of multiple years of ongoing exposure to treatment rather than immediate effects. Second, despite extensive tracking effort, the survey response rate is 47% and differs between treatment and control groups, which could produce sample selection bias. To address this, we randomize some features of the survey data collection, e.g., number of call attempts, and 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. Third, and perhaps most importantly, we are underpowered to study treatment effects on employment at the scale of this experiment. In particular, the 95% confidence interval on the estimated treatment effect covers -2.8 to 5.0 p.p. (Table B.7, column 4). Even with a 100% survey response rate, the minimum detectable effect size on employment would be 2.5 p.p. The phone call treatment increases the share of jobseekers receiving any interview invitations by 5.1 p.p., so an employment effect of 2.5 p.p. would be achieved in the possible but unlikely event that half of these jobseekers converted their additional interview into a job.[17]

These calculations suggest that the strength of light-touch treatments like this is the possibility of modestly raising employment rates on larger platforms at very low marginal costs. For example, Pakistan's Rozee has 9.5 million users, 1000 times the size of our platform. *If* a treatment like the one we study could raise the share of respondents getting interviews by the same 5.1 p.p. and *if* only 5% of these additional interviews converted into offers, that would lead to roughly 24,000 offers. As Kircher (2022) notes, many other studies of interventions on job search platforms either do not study employment effects at all (see examples in footnote 3) or use samples of hundreds of thousands of jobseekers to detect effects of 1 percentage point or smaller (e.g. Behaghel et al. 2023; Le Barbanchon et al. 2023).[18]

Our survey also asked jobseekers about their off-platform search behavior, after 40 months of treatment exposure. Treatment effects on any search, number of applications, and number of search methods used are negative and 5-15% of the control group means but with wide confidence intervals that include zero. So we again recommend great caution when interpreting the estimated effects (Tables B.7 and B.8).

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 appli-

[17] To derive this minimum detectable effect size, we assume: 80% test power, 5% test size, mean employment of 73% (equal to the control group's reported employment rate in the survey), and that covariates can absorb 10% of the outcome variation (roughly what we see in the interview invitation regressions). We preregistered employment and employment characteristics as trial outcomes because we did not know at the time (July 2020) how much COVID-19 would constrain platform operation, survey data collection, and hence power. We did not collect survey data on employment characteristics or proxies for match quality once COVID-19 made it apparent that we would not be powered to study treatment effects on these outcomes.

[18] The latter studies are based exclusively in high-income countries where data can be linked between government-run job search platforms and unemployment benefit registers. This is not currently feasible in any developing country, including the one we study.

cations. We then present several empirical results to support this framework, better understand the nature of application costs, and argue against alternative frameworks. We summarize the empirical analysis relatively briefly in this text and provide detailed explanations of the methods and results in Appendix C.

4.1 Conceptual Framework

Here we present a brief, intuitive discussion of our conceptual framework, with the formal model left to Appendix C.2. This paper's contribution is empirical rather than theoretical, so the framework is deliberately simple and stylized.[19] This framework shows how the common assumption of decreasing returns to marginal applications for *each individual jobseeker in each period* can lead to constant returns *averaged over jobseekers and periods*, provided some jobseekers are not actively searching in some periods.

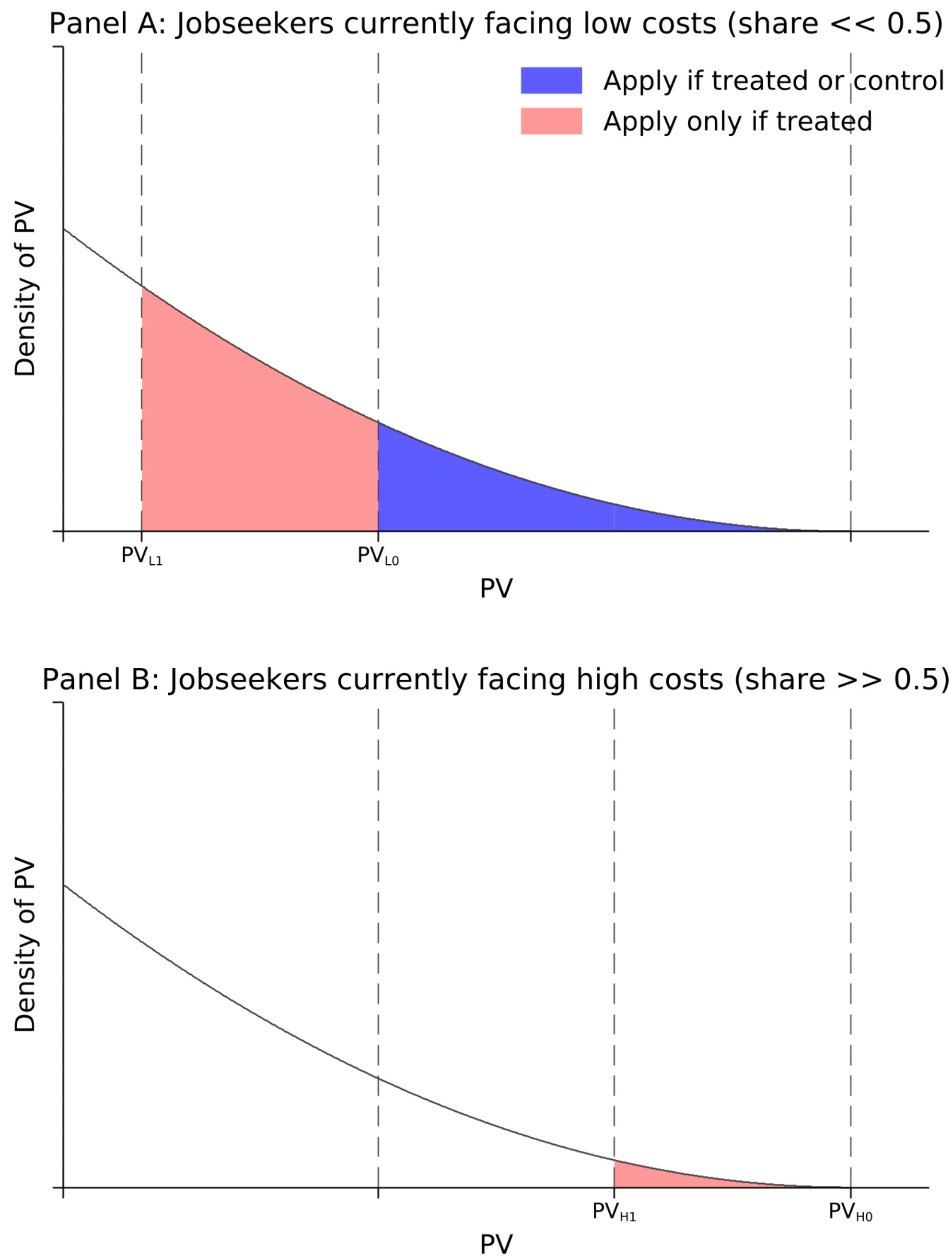
The platform sends each jobseeker a monthly batch of matches. We begin with a standard assumption (A1) that the jobseeker applies to all matches whose expected gross return, PV, exceeds the cost of applying. P is the probability of an interview conditional on applying. V is the gross value of getting an interview, which captures the expected present value of the flow of future utility from the interview, including the potential for a job offer.

Our key assumption (A2) is that the cost of applying varies within jobseeker through time and/or between jobseekers, and can be high enough that some untreated jobseekers choose not to apply in some matching rounds. We write this section in terms of costs that vary within jobseeker through time to simplify the writing but the full explanation of the framework in Appendix C.2 allows both types of cost variation. 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, there are two types of marginal applications induced by treatment. The first type of marginal applications comes 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 type of marginal applications comes 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.

[19] 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 miss some high-return matches due to frictions (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.2.

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.2 and explain that the framework does not require the simplifying assumption of only two cost types.[20]

This framework provides a clear economic interpretation of the LATE we estimate in Section 3.2: it is the average effect of applying on interview invitations, for applications sent due to a treatment-induced drop in the cost of applying. In this framework, marginal applications come from jobseekers who face higher costs of applying in the absence of treatment, relative to jobseekers submitting inframarginal applications. The constant returns finding shows that these higher costs are not associated with lower returns to applications.[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, either permanently or during some matching rounds. Second, *treatment and control group applications will go to vacancies with similar average values*, because treatment will induce applications to a mix of higher- and lower-return matches whose average value is similar to the control group.[22] Third, *treatment group applications will go to vacancies with more dispersed values*, as shown by the wider range of PV in the pink+blue region versus the blue-only region in Figure 1.

We show evidence consistent with all three predictions, summarized here with the detailed methods and results in Appendix C.4. We use two proxies for the value of each jobseeker $\times$ vacancy match: the salary, an admittedly narrow proxy but one that is easily interpretable and valued by all jobseekers; and the inverse covariance-weighted average of many positive attributes of a match (e.g. salary, benefits, commute distance) that is defined in Section 3.2. Figures C.2 and C.3 show evidence consistent with the first prediction: that many high-value matches receive few or no control group applications. For example, under half of all control group applications are sent to matches in the top quintile of each value proxy, and under 0.1% of matches in the top

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

[21] This echoes a finding in education research that costs of education and returns to education are weakly correlated over individuals in some applications (Card, 2001).

[22] Technically, this prediction holds in the special case of the framework where returns to marginal and inframarginal applications are equal, as we see in our data. When returns to marginal and inframarginal returns differ, then treatment and control applications may be sent to vacancies with different average values, as we explain in Appendix C.2.

quintile receive applications. Figures C.2 and C.3 also show evidence consistent with the second prediction: that treatment and control group applications will go to vacancies with similar average values. Specifically, they show that the shares of control group applications sent to each of quintiles 1, ..., 5 are equal to the shares of treatment group applications sent to each of these quintiles, with formal test results reported in the footnotes below the figures. As an additional test of the second prediction, we use the same proxies to calculate the mean values of matches that get applications and compare these between the treatment and control groups. Table C.5 shows that there are some differences between mean values of observed characteristics between treatment and control group applications but that these differences do not show consistently higher values in either group, consistent with the second prediction. For example, control group applications go to jobs that offer slightly higher salaries, slightly less flexible hours, similar values of the summary index V_{vm}, and similar latent interview probabilities (which we predict using a LASSO-based approach). Third, we use the same proxies to calculate the dispersion of values of matches receiving applications and compare these between the treatment and control group. Table C.6 shows that treatment raises the variance and lowers the 10^{th} percentiles for the value proxies, matching the model prediction of more dispersion in the treatment group.

4.3 Understanding Application Costs

The roughly constant returns to marginal applications shown in Section 3.2 and the patterns of matches that receive applications shown in Section 4.2 are both consistent with treatment reducing the cost of initiating applications. In this subsection, we report the results of several additional experiments designed to shed light on what types of application costs are likely to fall with treatment. We show more details on all methods and results in Appendix C.5. Each treatment in these experiments is assigned to a small share of the sample, and controlling for these assignments and their interactions has no impact on the estimated effects of the main phone call treatment (Table C.1).

First, *pecuniary costs* of applications are unlikely to explain our main results. Job applications on the platform are very cheap, as the control group can call the platform to apply for $< 1\%$ of the daily minimum wage, and mobile phone providers in Pakistan allow users to fund phone calls with short-term loans when they have no airtime credit. We also randomly select some control group jobseekers in one round to receive a text message reminder that they can ask the platform to call them back about a job posting, saving the cost of calling back entirely. This free callback reminder treatment has an effect one hundredth of the size of the effect of the main phone call treatment, suggesting a very small role for pecuniary costs of job search (Table C.7).

Similarly, *time costs* of applications are unlikely to explain our main results. We randomly offer some control group jobseekers in some rounds the option to text the platform and ask for a callback at a specific time of their choice. This eliminates time waiting on hold or connecting to a call center agent. The effect of this callback request treatment is one quarter the size of the main phone call and statistically significantly smaller. This suggests that while time costs matter, they matter much less than the overall effect of the phone call treatment (Table C.7).

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. As we discuss in the introduction, the existing literature suggests five types of psychological costs that might be reduced by the phone call treatment. First, control group jobseekers might ignore text messages due to *attention costs* (Gabaix, 2019). Second, control group jobseekers might not initiate applications due to *fear of rejection* (Köszegi et al., 2022). Third, *present bias* might lead control group jobseekers to repeatedly postpone applications until the deadline passes (Ericson & Laibson, 2019). Fourth, phone calls may function as *reminders* to apply. Fifth, phone calls may *encourage or pressure* jobseekers to apply. There is existing empirical evidence that some of these factors can influence job search decisions (e.g. DellaVigna & Paserman 2005; Zizzamia 2023), as well as a larger body of research reviewed by DellaVigna (2009) showing that eliminating or reducing the need to initiate decisions can raise financial and health investments. We show in Appendix C.2 how each of these factors can enter our model.

While we cannot pin down exactly how the phone call treatment reduces psychological costs of initiating applications, we can test and largely reject two of these five possible psychological explanations. First, treatment does not simply increase application rates by providing a *reminder effect* that offsets forgetfulness. We randomly send some control group jobseekers in some rounds a second text message listing their matched vacancies as a reminder. The effect of the reminder is one-fourteenth as large as the effect of the phone call (p-value of difference < 0.001, Table C.8). Furthermore, the phone call treatment has a smaller effect on the application rate when phone calls take place longer after text messages and when the window between text messages and application deadlines is shorter (Table C.9). This pattern is not consistent with an important role for reminder effects, as earlier calls and shorter application windows allow less time for forgetting and hence less scope for reminder effects.

Second, we find some evidence against an explanation that treatment increases applications because call center agents *encourage or pressure jobseekers to apply*. Agents are trained not to *explicitly* encourage or pressure jobseekers, and regular audits of call recordings verified that they followed their scripts. It remains possible that jobseekers feel *implicit* pressure to apply because they have been called or because they are interacting with a person. However, if treated respondents did feel pressure to apply when called, they could easily avoid this pressure by avoiding calls after the first few rounds of matching. Instead, we find that jobseekers who answer calls in the first few rounds of matching are actually more likely to answer calls in subsequent rounds, conditional on observed characteristics (Table C.10). Furthermore, treated and control jobseekers are equally likely to apply to the first job listed on their call or text message, showing that treated jobseekers do not simply apply to the first listed job to end a pressure-inducing call quickly (Figure C.7).

This collection of results implies that the phone call treatment produces more applications at roughly constant returns by reducing the psychological rather than the pecuniary or time costs of applying, and probably not by providing reminders, encouragement, or pressure. However, we acknowledge that we cannot pin down exactly which psychological cost(s) are reduced by the phone call treatment, so alternative explanations remain possible.

4.4 Evaluating Alternative Explanations

In this section we summarize evidence against five alternative explanations for our two key findings: that the phone call treatment substantially increases the job application rate and that the returns to these marginal applications are approximately constant, in terms of interview probabilities. We show more details on all methods and results in Appendix C.6. First, treatment-induced job applications are not sent to systematically *less competitive or desirable vacancies* than control group job applications. Instead, we have already shown in Section 4.2 that treatment and control group job applications are sent to equally desirable matches on average. This allows us to reject an alternative explanation in which marginal applications and inframarginal applications might have roughly equal returns if marginal applications are sent to vacancies that are both systematically worse matches for the jobseekers (leading to lower interview probabilities) and systematically less competitive (leading to higher interview probabilities).

Second, treatment-induced job applications do not come from *systematically better-qualified jobseekers* than control group applications. Instead, applications in the treatment and control groups come from jobseekers with roughly equal values of observed characteristics such as education, work experience, and CV quality scores (Table C.11). We also use a data-driven approach to estimate the latent probability that an application sent to each match by each matched jobseeker would yield an application, based on the jobseeker characteristics that both we and the firm observe. Matches that receive applications from the treatment and control groups have similar values of this latent probability, suggesting that treatment does not shift the selection of which types of jobseekers submit applications (Table C.11). Furthermore, our main findings hold when we control for time-invariant characteristics in two ways. We use an additional "crossover" experiment that randomly moves some control group jobseekers to the treatment group in some rounds, which allows us to estimate treatment effects conditional on jobseeker fixed effects (Table C.12).[23] And

[23] In particular, we cannot reject equality of the interview rates or quality-adjusted interview rates for inframarginal applications and marginal applications submitted due to the crossover treatment ($p > 0.480$). 16% of jobseekers have at least one match affected by this crossover treatment, allowing precise estimation of the treatment effect conditional on the fixed effects. But only 0.65% of matches are affected by this treatment, so it has almost no impact on our estimates of the overall treatment effect (Table C.13).

we repeat our main analysis controlling for observed baseline characteristics.[24] This allows us to reject an alternative explanation in which marginal and inframarginal applications might have roughly equal returns because 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.

Third, it is unlikely that phone calls provide *more information about specific jobs*. Call center agents are trained to read specific scripts with no additional information about jobs, general labor market conditions, or assessments of the jobseeker's prospects, and regular audits of call recordings verified that they followed their scripts. Additionally, in 80% of matching rounds, we gave call center agents no additional information about the jobs and our results are almost unchanged when we restrict analysis to these rounds (Table C.14). It is possible that jobseekers might be more likely to receive phone calls than text messages, perhaps if some text messages are blocked or go unread. However, when we survey jobseekers and ask 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 matches, with or without adjusting for survey non-response (Table C.15). This allows us to reject an alternative explanation in which the phone call treatment might provide information about specific jobs, leading to higher application rates and enabling jobseekers to target better-matched vacancies that have higher interview probabilities, keeping the returns to marginal applications as high as inframarginal applications.

Fourth, it is unlikely that the phone call treatment increases job application rates by *shifting jobseekers' beliefs about the value of applying*. It is possible that calls from a professional recruiting service might be taken as signals that platform firms are larger or wealthier and thus able to provide more benefits or opportunities for advancement (higher V), or that these are jobs to which the jobseeker is particularly well-matched and likely to get an interview (higher P). But this explanation does not match several patterns in the data. We directly test this by collecting data on jobseekers' beliefs about P and V and estimating treatment effects on these two belief measures.[25] Treatment effects on both these measures are close to zero, with or without adjusting for survey non-response (Table C.16). Furthermore, if phone calls influence job applications because a jobseeker views them as informative about the quality of a specific match, then phone calls should have larger effects on applications when the jobseeker views the phone call as unusual than when

[24] To show this, we repeat our analysis of the main experiment using a post-double selection LASSO to control for an extensive set of jobseeker baseline characteristics, following Belloni et al. (2014). Table B.3 shows that the point estimates and standard errors are almost identical.

[25] Translated from Urdu, these questions ask: "Suppose Job Talash sends you one hundred job ads in the next year. Based on your past experience with our job matching service, how many of these jobs do you think would be desirable for you?" and "Suppose you apply for all of these jobs that you think are desirable. How many do you think would make you an offer?" Our main treatment assignment is time-invariant, so these questions are asking jobseekers about jobs sent by the mode of communication used in their treatment group.

she views it as part of normal platform operations. But when control group jobseekers receive occasional phone calls as part of our within-jobseeker "crossover" experiment, their response is very similar to the main phone call treatment. See details on the experiment two paragraphs above and results in Table C.12. These patterns suggest that the phone call is unlikely to shift application decisions by signaling that these are unusually high-value matches. This allows us to reject an alternative explanation in which the phone call treatment might increase application rates by raising the perceived value of applying.

Fifth, the main experimental results do not arise because jobseekers search randomly, which might lead to constant returns to additional applications. Random job search may seem implausible but it is often assumed in canonical search models (e.g. Pissarides 2000) and may be plausible when jobseekers have very limited information about labor market conditions and match quality (Behaghel et al., 2023; Belot et al., 2018). However, random search is not consistent with the pattern we showed in Sections 2.4 and 4.2 that applications are directed toward vacancies with more desirable attributes. We also run an experiment to directly induce random job search and show that this produces different results to the phone call treatment. Specifically, in 20% of rounds we randomize the order in which vacancies are listed in both text messages and phone calls, which generates a random increase in the rate of applications to vacancies listed first. Unlike jobs to which individuals are encouraged to apply because of the phone call treatment, applications induced by jobs being listed first have decreasing rather than constant returns. The average interview probability for these marginal applications is 2.4%, substantially lower than the 6.3% for inframarginal applications (Table C.17). This allows us to reject an alternative explanation in which the phone call treatment generates marginal applications with constant returns because jobseekers are searching randomly, so marginal and inframarginal applications are sent to similar vacancies.

5 Spillover Effects

Increased search effort by some jobseekers may affect firms and other jobseekers. The sign of this effect is theoretically ambiguous. For firms, getting more applications can raise the probability of getting 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, competing against more applications can lead to crowd-out. But the magnitude of crowd-out may be small and offset if firms are more likely to hire when they get 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). This approach works well because the 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. On such platforms, it is difficult to define how much each user is competing with other users without a full model of the job search process.

We briefly summarize our methods and results here and provide many more details in Appendix D. Within each of the 1,340 vacancies, we estimate the share of jobseekers matched to that vacancy who are treated and the treatment effect on interview invitations. We create a vacancy-level dataset with these two statistics and regress the treatment effect on the treatment share, conditional on other vacancy-level characteristics. We find no evidence of a negative relationship: a vacancy exposed to the 75th percentile of the treatment rate rather than the 25th percentile would have a 0.018 percentage point higher treatment effect on interviews (standard error 0.011, $p = 0.096$). This shows that treatment effects on individual jobseekers' interview probabilities are not smaller when they face more treatment-induced competition, suggesting they do not face negative spillovers. Similarly, we find no evidence of negative spillovers when we allow the relationship to be nonlinear or allow spillovers to have different effects on treated and control group jobseekers.

These results suggest negligible between-jobseeker spillovers on interview invitations. However interpretation of these patterns is complex and requires some caveats. Spillovers might be negligible because firms report filling only 60% of the vacancies posted on the platform, so more applications might lead to more well-matched applicants and hence fill more vacancies, in line with findings by Fernando et al. (2021). But spillovers might also be negligible because firms report receiving 70% of applications from outside the platform so the majority of competition that jobseekers face is unaffected by treatment. Finally, spillovers may be very different at the interview stage versus the hiring stage, which we do not observe. See Appendix D for more discussion.

6 Conclusion

We show that job search effort can be dramatically increased by reducing the psychological cost of initiating job applications. Perhaps surprisingly, returns to the additional search effort, measured in terms of interview invitations, are constant rather than decreasing. 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 in that period without treatment. These findings show that small reductions in search costs can substantially improve search outcomes in environments with some relatively inactive jobseekers. This echoes findings that changing default options to avoid initiation costs can lead to economically significant increases in financial and health investments (DellaVigna, 2009). 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 depends on existing application volumes.

References

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.

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

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.

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 & T. Zuber (2023) "The Potential of Recommender Systems for 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*, forthcoming.

Caria, S., K. Orkin, A. Andrew, R. Garlick, R. Heath & N. Singh (2024) "Barriers to Search and Hiring in Urban Labour Markets," *VoxDevLit*, 10 (1).

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

Carranza, E. & D. McKenzie (2024) "Job Training and Job Search Assistance Policies in Developing Countries," *Journal of Economic Perspectives*, 38 (1), 221–244.

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, NBER.

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 Development*, 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. & 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.

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, IZA.

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.

Field, E., R. Garlick & K. Vyborny (2024) "Women's Mobility and Labor Supply: Experimental Evidence from Pakistan," Working paper.

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.

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.

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, NBER.

Hjort, J. & J. Poulsen (2019) "The Arrival of Fast Internet and Employment in Africa," *American Economic Review*, 109, 1032–1079.

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.

Kircher, P. (2022) "Job Search in the 21st Century," *Journal of the European Economic Association*, 20 (6), 2317–2352.

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.
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.
LaLive, R., C. Landais & J. Zweimuller (2022) "Market Externalities of Large Unemployment Insurance Extension Programs," *American Economic Review*, 105, 3564–3596.
Le Barbanchon, T., L. Hensvik & R. Rathelot (2023) "How can AI improve search and matching? Evidence from 59 million personalized job recommendations," Working Paper.
Lechner, M. & J. Smith (2007) "What is the value added by caseworkers?" *Labour Economics*, 14 (2), 135–151.
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.
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.
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.
Neumark, D. (2018) "Experimental Research on Labor Market Discrimination," *Journal of Economic Literature*, 56 (3), 799–866.
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.

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.

Zizzamia, R. (2023) "Ignorance is bliss? Rejection and discouragement in on-the-job search," Working Paper.

Appendices for Online Publication Only

Contents

A Additional Information about the Platform and Sample

Here we provide additional information and descriptive statistics about the sample and platform.

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 areas followed by a 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.

Platform information: Table A.2 compares the processes on Job Talash for registering, being notified about matched jobs, and applying for jobs to three other prominent job search platforms. This shows substantial overlap in the processes for learning about matched jobs and submitting applications to these jobs. The other three platforms also include search functions that allow users to learn about jobs outside the match notification system. Table A.3 compares the average monthly job application rate on this platform to other platforms studied by economists that report comparable statistics. This shows that the job application rate on Job Talash is comparable to some other job search platforms. All of the studies reporting the highest application rates consider only "active" platform users, rather than all users.

Figure A.2 shows a sample text message sent to jobseekers. Figure A.3 shows the exact communication process between the platform and jobseekers in the treatment and control groups, including the structure of the script for phone calls.

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

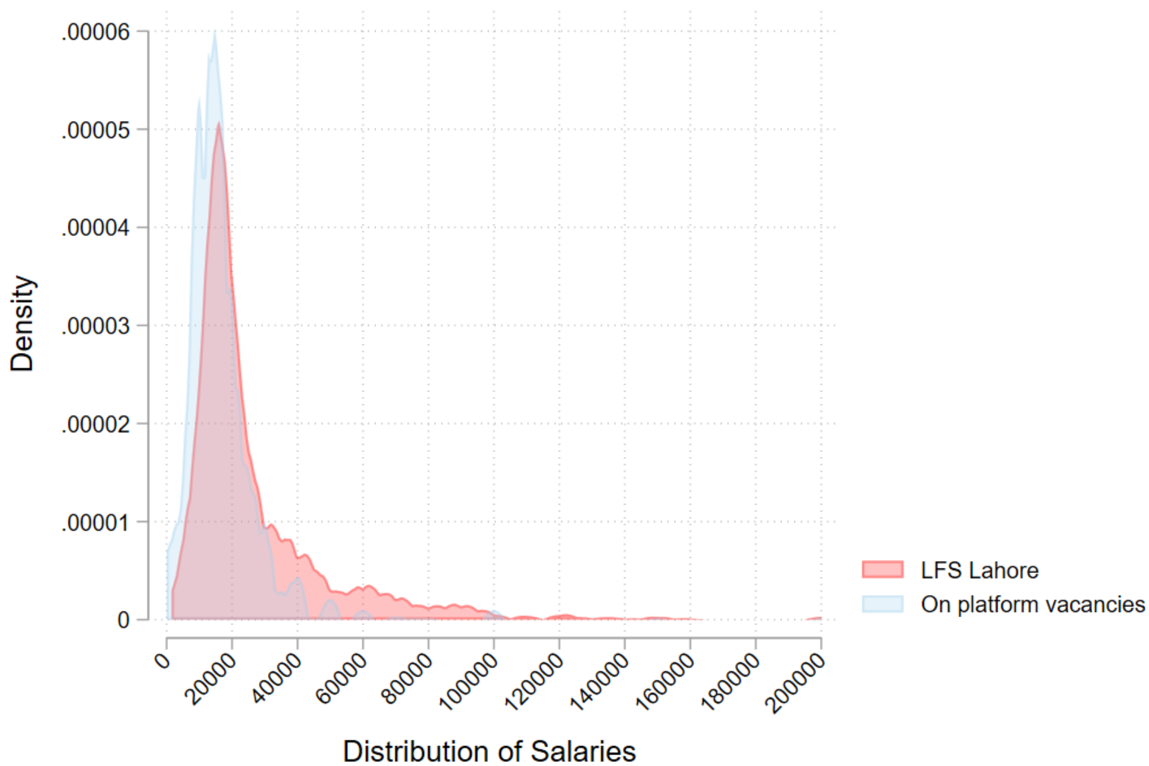
	LFS Pakistan	LFS Lahore	LFS Lahore Reweighted	HH listing Sample	Experimental Sample
	(1)	(2)	(3)	(4)	(5)
Age	33.9	33.8	32.7	32.6	30.7
	(11.6)	(11.6)	(11.0)	(11.1)	(9.5)
Highest Education Level					
Less than Intermediate/High School	0.853	0.679	0.700	0.706	0.491
Completed Intermediate/High School	0.073	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.

Table A.2: Registration and Job Application Processes on Job Talash and Other Job Search Platforms

Platform	Job Talash (control group)	LinkedIn	Rozee	Indeed
Registration process	Complete phone call with the platform that asks about demographics, education, work experience, and occupational preferences. No fee for registration.	Complete registration on the website that requires contact information, location, education, and occupation preferences, with the option of adding more information later. No fee for registration. Can upload CV.	Complete registration on the website that requires contact information, gender, education, work experience, and occupation preferences. No fee for registration. Can upload CV.	Complete registration on the website that requires contact information, location, and gender. Can also provide information on education, work experience, and skills or upload a CV. No fee for registration.
Notification process	Notified about jobs that match education, experience, occupational preferences. Sent by text message.	Notified about jobs that match preferred job title and location. Sent by email or in the app.	Notified about jobs that match preferred experience, salary, location and optional keywords. Sent by email.	Notified about jobs that match preferred job title, salary, location and work schedule. Sent by email or in the app.
Job application process	Phone platform and ask them to send your template CV to the jobs you're interested in. No fee to apply.	If the job allows applications via LinkedIn: submit contact information, upload CV, and for some jobs answer additional job-specific questions. Otherwise redirected to the company website. No fee to apply.	If the job allows applications via Rozee: confirm contact information is correct, upload CV or submit platform-generated CV, and for some jobs answer additional job-specific questions. Otherwise redirected to the company website. No fee to apply.	Confirm contact information is correct, upload CV or submit platform-generated CV, and for some jobs answer additional job-specific questions. No fee to apply.
Other platform notes		Largest online professional networking site in the world by number of users	Largest online job search platform in Pakistan by number of users	Largest employment website in the world by number of visitors.

Notes: This table compares the processes on Job Talash for registering, being notified about matched jobs, and applying for jobs to three other prominent job search platforms. This shows substantial overlap in the processes for learning about matched jobs and submitting applications to these jobs. The other three platforms also include search functions that allow users to learn about jobs outside the match notification system.

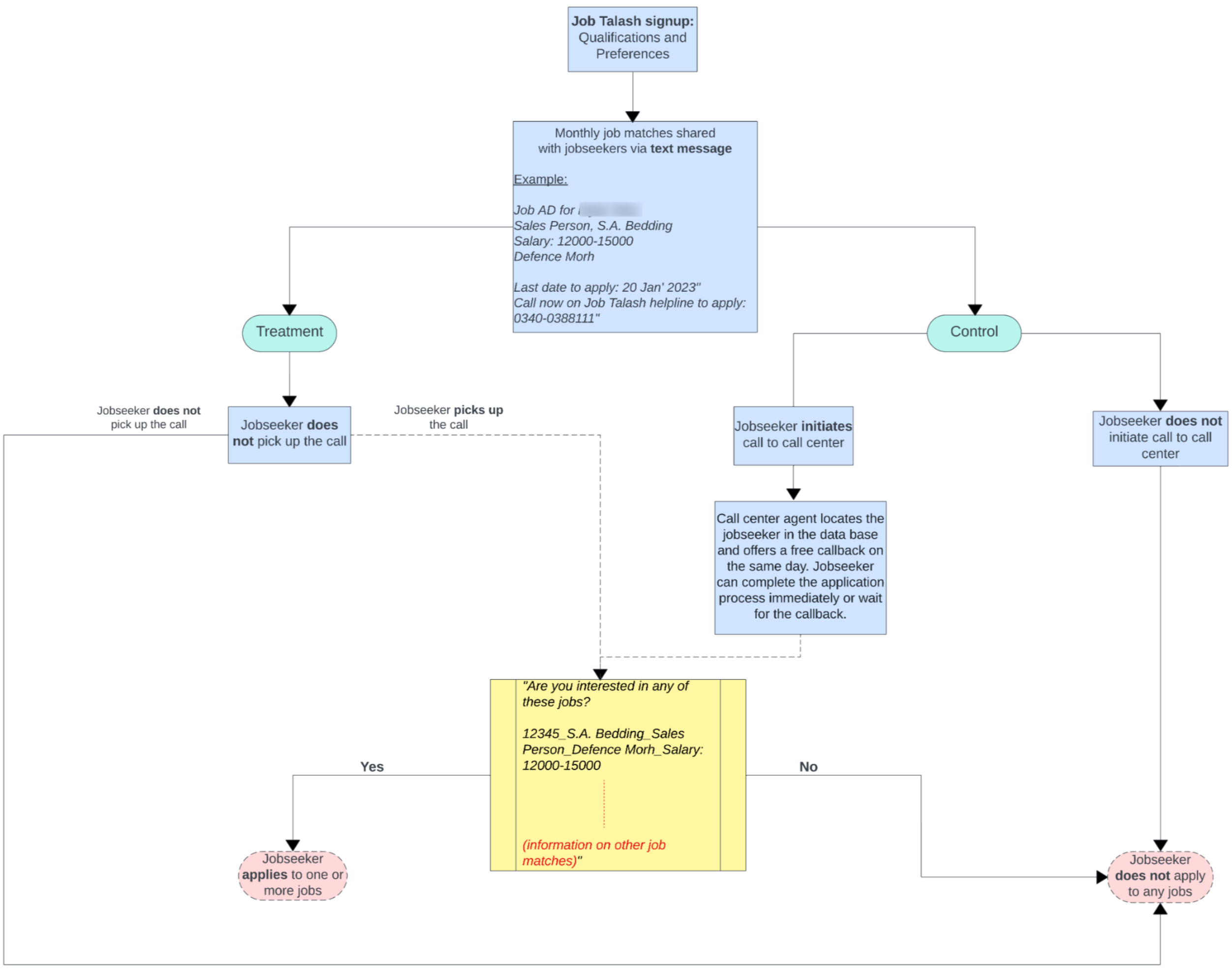
Table A.3: Job Application Rates on Search and Matching Platforms

Study	Country	Platform	Mean apps per user per month	Notes
Behaghel et al. (2023)	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.
He et al. (2023)	China	Not stated	28.8	Sample excludes ≈ 99% of users because they are "inactive."

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

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 Effort and Returns to Search

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

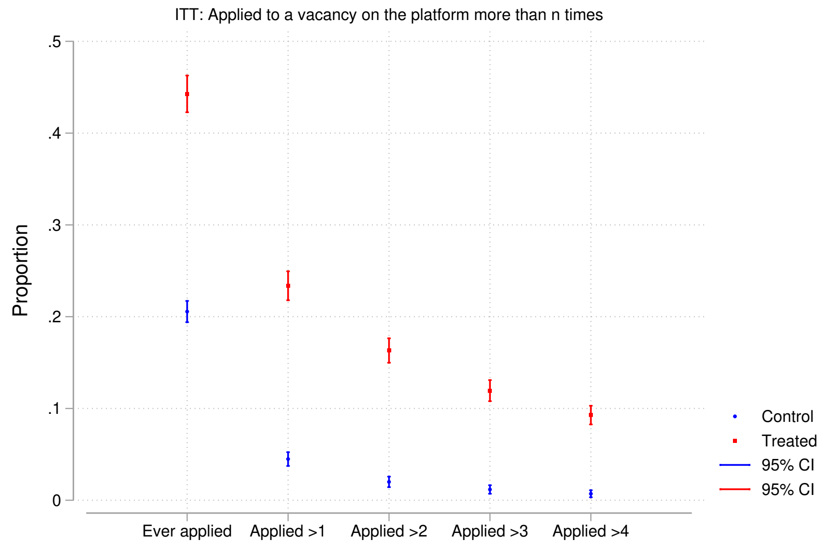
Figure B.1 shows treatment effects on the number of times each jobseeker applies to and is invited to an interview for a job. This figure shows that treatment raises the probability of submitting K applications and getting L interviews for all K and for L ≤ 4.

Table B.1 shows heterogeneous treatment effects by baseline employment, search status, education, and age. Treatment substantially increases both applications and interviews for all eight subgroups (panels A and B). The marginal return to additional applications ranges from 4.9 to 7.3% across the subgroups (panel C). We fail to reject constant returns to marginal search for any of the subgroups after adjusting for multiple hypothesis testing (sharpened $q > 0.30$).

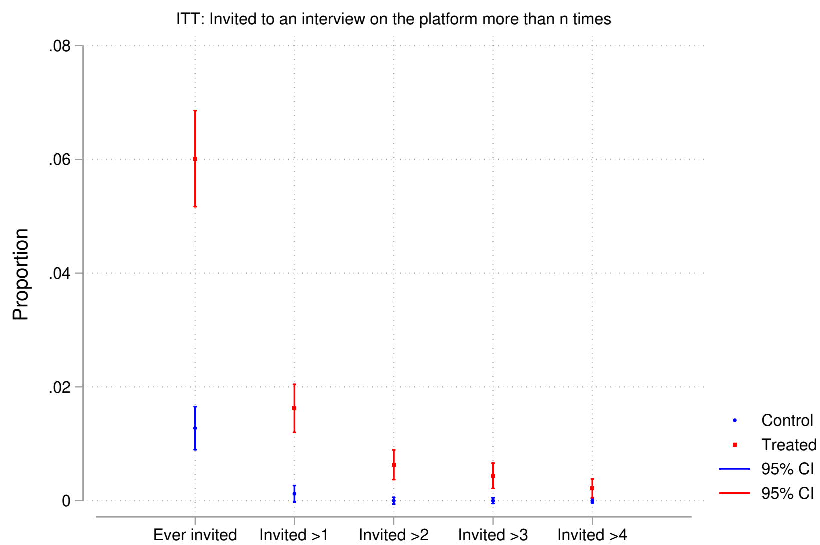
Table B.2 shows treatment effects on interview probabilities weighted by different proxies for interview value, such as salary. This includes all components of the proxy index for interview value discussed in Section 3.2 and some combinations of proxies, e.g., commute-cost-adjusted salary in column 4 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.

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: Heterogeneous Treatment Effects

Group 1 vs. 0	**Employed vs. Unemployed**	**Searching vs. Not Searching**	**Less vs. More than High School**	**Less vs. More than 30 Years Old**
Panel A: Applications				
	(1)	(2)	(3)	(4)
Effect on Group = 1	0.01276	0.01566	0.01469	0.01425
	(0.00097)	(0.00109)	(0.00103)	(0.00093)
Effect on Group = 0	0.01356	0.01174	0.01161	0.01165
	(0.00091)	(0.00117)	(0.00084)	(0.00091)
p: (Effect on Group = 1) = (Effect on Group = 0)	0.48063	0.00367	0.00778	0.01800
Mean Outcome | T = 0, Group = 1	0.00161	0.00252	0.00265	0.00206
Mean Outcome | T = 0, Group = 0	0.00202	0.00121	0.00120	0.00157
Panel B: Interview Invitations				
Effect on Group = 1	0.00069	0.00090	0.00071	0.00089
	(0.00011)	(0.00013)	(0.00011)	(0.00011)
Effect on Group = 0	0.00084	0.00085	0.00085	0.00060
	(0.00011)	(0.00014)	(0.00011)	(0.00010)
p: (Effect on Group = 1) = (Effect on Group = 0)	0.27097	0.74611	0.30252	0.02092
Mean Outcome | T = 0, Group = 1	0.00010	0.00018	0.00010	0.00011
Mean Outcome | T = 0, Group = 0	0.00013	0.00006	0.00013	0.00013
Panel C: IV				
Effect on Group = 1	0.05453	0.05783	0.04871	0.06227
	(0.00665)	(0.00630)	(0.00603)	(0.00596)
Effect on Group = 0	0.06172	0.07266	0.07330	0.05206
	(0.00625)	(0.00876)	(0.00715)	(0.00677)
p: (Effect on Group = 1) = (Effect on Group = 0)	0.35689	0.10855	0.00253	0.17554
Mean Outcome | T = 0, Apply = 1, Group = 1	0.05941	0.07143	0.03683	0.05293
Mean Outcome | T = 0, Apply = 1, Group = 0	0.06494	0.05085	0.10997	0.08040
p: Effect = Mean Outcome | Group = 1	0.78064	0.14681	0.02912	0.07441
p: Effect = Mean Outcome | Group = 0	0.73344	0.27590	0.23010	0.38474
# matches	1,116,160	921,011	1,116,952	1,116,952
Proportion in Group = 1	0.41427	0.58115	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, high school or higher education at baseline in column 3 (the sample median level of education), and age under 30 years old at baseline in column 4. 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.2: 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.

B.2 Robustness Checks

Table B.3 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 Section 4.4 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.3: 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:

$$\beta_{2SLS} = \frac{\mathbf{P_C} \cdot \mathbf{\Delta I_C} - \mathbf{P_D} \cdot \mathbf{\Delta I_D}}{\mathbf{P_C} - \mathbf{P_D}}. \tag{4}$$

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 (4) as

$$\mathbf{\Delta I_C} = \frac{\beta_{2SLS} \cdot (\mathbf{P_C} - \mathbf{P_D}) + \mathbf{P_D} \cdot \mathbf{\Delta I_D}}{\mathbf{P_C}} = \frac{\beta_{2SLS} \cdot \beta_{S1} - \mathbf{P_D} \cdot \mathbf{\bar{I}_D}}{\beta_{S1} + \mathbf{P_D}}, \tag{5}$$

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 (5) 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

$$0 \leq \mathbf{P_D} \leq \bar{A}_0. \tag{6}$$

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

$$0 \leq \bar{\mathbf{I}}_{\mathbf{A}} \leq 1 \tag{7}$$

$$0 \leq \bar{\mathbf{I}}_{\mathbf{D}} \leq 1. \tag{8}$$

Evaluating equation (5) 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: $(\bar{\mathbf{I}}_{\mathbf{A}} \cdot \mathbf{P_A} + \bar{\mathbf{I}}_{\mathbf{D}} \cdot \mathbf{P_D})/(\mathbf{P_A} + \mathbf{P_D})$. Recalling that $\mathbf{P_D} + \mathbf{P_A} = \bar{A}_0$ and rearranging terms gives

$$\mathbf{P_D} \cdot \left(\bar{\mathbf{I}}_{\mathbf{D}} - \bar{\mathbf{I}}_{\mathbf{A}}\right) = \bar{A}_0 \cdot \left(\bar{I}_0 - \bar{\mathbf{I}}_{\mathbf{A}}\right). \tag{9}$$

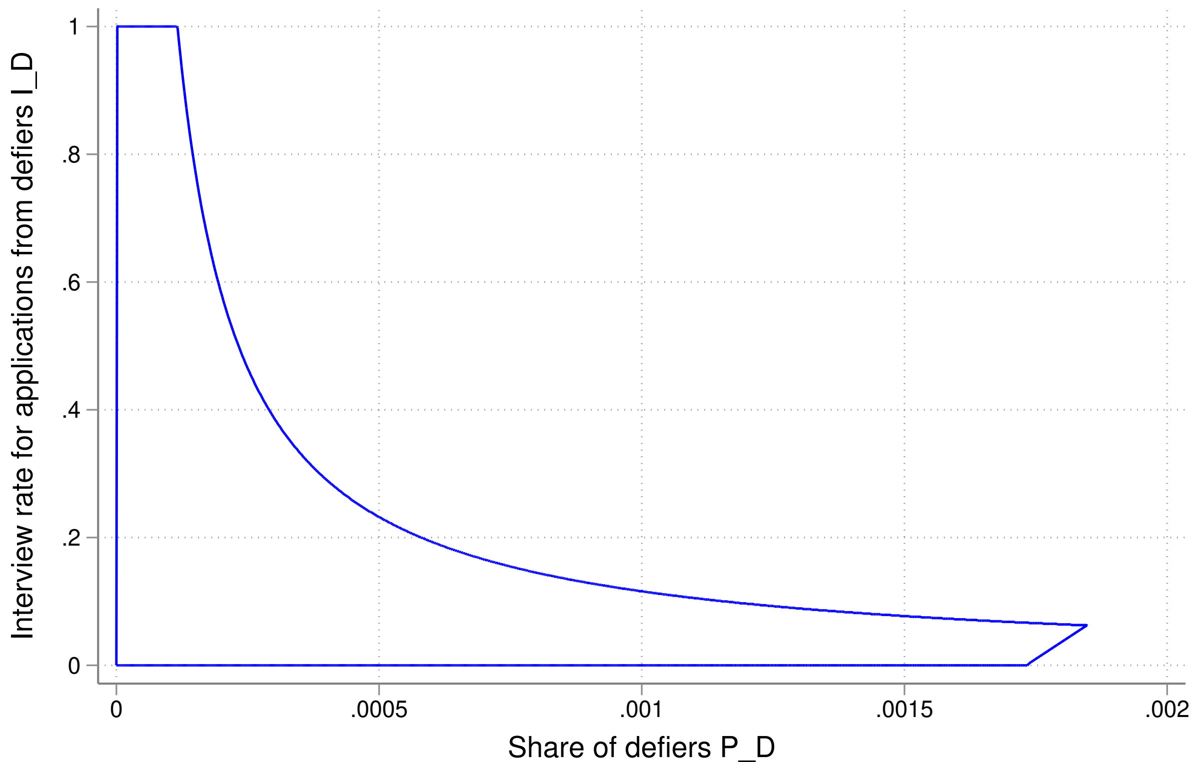
Combining (5), (6), (7), (8), and (9) gives a system of two equality restrictions and three inequality restrictions in which $\mathbf{\Delta I_C}$ depends on three unknown quantities: $\bar{\mathbf{I}}_{\mathbf{D}}$, $\bar{\mathbf{I}}_{\mathbf{A}}$, and $\mathbf{P_D}$. This does not allow us to point identify $\mathbf{\Delta I_C}$ but allows us to obtain a lower bound.

Fourth, we consider each value P_D satisfying (6), solve for the set of values of $\bar{\mathbf{I}}_{\mathbf{D}}$ and $\bar{\mathbf{I}}_{\mathbf{A}}$ 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 $\bar{\mathbf{I}}_{\mathbf{D}}$ in solid blue. When the share of defiers is small, only condition (7) binds on $\bar{\mathbf{I}}_{\mathbf{D}}$. As the share of defiers increases, the maximum feasible value of $\bar{\mathbf{I}}_{\mathbf{D}}$ shrinks to stop the left-hand side of equation (9) from becoming so large that it can only be satisfied by a value of $\bar{\mathbf{I}}_{\mathbf{A}}$ that violates condition (8). As the share of defiers approaches $\bar{A}_0$ and hence the share of always-takers approaches zero, $\bar{\mathbf{I}}_{\mathbf{D}}$ 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 β_{2SLS}. This matches the intuitive interpretation of equation (4): 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.4, 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.4, 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.5, 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.4: Treatment Effects on Non-Application Measures of Platform Use

	(1) # pref. updates	(2) Any pref. update	(3) Completed CV	(4) Added educ.	(5) 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.5: 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.4, column 3). But replicating our main analysis for respondents who completed their entire CV at registration replicates our main findings (Table B.5, 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 for these matches as zeros. 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.5, 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.6, 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.6, column 2, row 3).

Table B.6: Alternative Test for Constant Returns to Search

	(1) Interview	(2) Interview $\times$ V_{vm} index
Treatment effect	0.00078	0.00281
	(0.00009)	(0.00036)
Constant-returns implied effect	0.00083	0.00314
	(0.00005)	(0.00018)
Difference	-0.00006	-0.00033
	(0.00007)	(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.7 and B.8 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.7: 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.8: 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 documents, 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.9).

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.10. 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).[26]

[26] 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.9: 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.7), specific off-platform search activities (Table B.8), receipt of calls/text messages (Table C.15), and beliefs about jobs on the platform (Table C.16). 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.10: 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 off-platform search (column 2), a binary question for any employment (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 Explaining Marginal Returns to Search

C.1 Overview

This appendix provides detailed methods and results to support the argument in Section 4 of the paper – "Explaining Marginal Returns to Search." For each of the additional experiments described in this appendix, only a very small share of the matches in the sample are treated. For our analysis of the main phone call treatment, we do not take these treatment assignments into account. Table C.1 shows that including indicators for each different treatment in our main regressions has almost no impact on the estimated effects of the main phone call treatment.

To help the reader navigate this appendix, we provide three tables on the next three pages that summarize all analyses in this appendix. Table C.2 summarizes the additional tests of our conceptual framework. These tests are covered in Appendix C.4. They are based on data patterns in the control group or comparisons of the control and treatment groups from our main experiment. Table C.3 summarizes all additional experiments and additional analyses we use to understand the nature of application costs (e.g. pecuniary, time, psychological). This content is covered in detail in Appendix C.5. Table C.4 summarizes the evidence we use to evaluate alternative explanations. This evidence is covered in detail in Appendix C.6. Each table lists the research question, method, finding, corresponding appendix section/table/figure containing more details, numbers of jobseekers and matches involved, and - for the additional experiments only – the unit of randomization (jobseeker, jobseeker $\times$ round, or jobseeker $\times$ vacancy) and share of units randomized.

Table C.1: Treatment Effects on Job Search & Search Returns, Controlling for All Other Treatments

	(1) Apply	(2) Interview	(3) Int. $\times V_{vm}$	(4) Interview	(5) Int. $\times V_{vm}$
Phone call treatment	0.01338	0.00078	0.00283		
	(0.00075)	(0.00009)	(0.00036)		
Apply				0.05842	0.21179
				(0.00511)	(0.02125)
# 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				0.627	0.482
IV strength test: F-stat				315.6	315.6
IV strength test: p-value				0.00000	0.00000

Notes: This table shows that the estimates in Table 3 are unchanged when we include in each regression indicator variables for the random subset of control group matches that were assigned to any other treatment in a subset of rounds. All other sample definitions, regression specifications, and inference methods are identical to Table 3.

Table C.2: Additional Tests of the Conceptual Framework – Covered in Detail in Appendix C.4

Question	Method				Result & Interpretation	Appendix
	Type	Description	# jobseekers	# matches		
All analysis in this table is based on two proxies for match value for all jobseeker $\times$ vacancy matches: salary, and the index of all positive attributes defined in Section 3.2.						
Do control group jobseekers miss applying to some high-value matches?	Additional analysis	Divide all matches into quintiles in multiple different ways: using the distribution of value proxies over all matches, using the within-jobseeker distribution of value proxies over all matches, and using the distribution of value proxies at the jobseeker $\times$ matching round level rather than match level. Calculate the share of matches in each quintile that receive applications from control and treatment group jobseekers.	9,831	1,116,952	Higher-value matches receive more applications but many high-value matches receive no applications: $<$0.1% of top-quintile matches receive applications from the control group and $<$50% of all control group applications are sent to top-quintile matches. Consistent with the conceptual framework's prediction that control group jobseekers miss applying to many high-value matches.	Appendix C.4, Figures C.2-C.6
Are treatment-induced job applications sent to matches with systematically similar attributes to matches that the control group apply to?	Additional analysis		9,831	1,116,952	The share of all treatment group applications that go to the matches in the Qth quintile of value is equal to the share of control group applications that go to the matches in the Qth quintile of value, averaged across all quintiles 1, ..., 5. Shows that treatment induces more job applications but that these are not sent to systematically higher- or lower-value matches than control group applications. Consistent with the conceptual framework's prediction that treatment-induced applications should go to matches with similar mean values to those in the control group.	Appendix C.4, Figures C.2-C.6
Are treatment-induced job applications sent to matches with systematically similar attributes to matches that the control group apply to?	Additional analysis	Calculate the means of the value proxies for matches that get applications from the treatment group. Compare these to the means for matches that get applications from the control group.	2,567	9,088	Marginal and inframarginal job applications are sent to jobs with similar means for the value proxies. Consistent with the conceptual framework's prediction that treatment-induced applications should go to matches with similar mean values to those in the control group.	Appendix C.4, Table C.5
Are treatment-induced job applications sent to matches with wider variance of attributes than matches that the control group apply to?	Additional analysis	Calculate the distributions of the value proxies for matches that get applications from the treatment group. Compare these to the distributions for matches that get applications from the control group.	2,567	9,088	The variances of the value proxies are higher in the treatment than the control group. The 10th percentiles of the value proxies are lower in the treatment than the control group. Consistent with the conceptual framework's prediction that treatment-induced applications should go to matches with more dispersed values than in the control group.	Appendix C.4, Table C.6

Table C.3: Understanding Application Costs – Covered in Detail in Appendix C.5

Question	Method						Result & Interpretation	Appendix
	Type	Description	# jobseekers	# matches	unit randomized	% units treated		
Does lowering the pecuniary cost of applying increase the application rate by as much as the main treatment?	Additional experiment	For some control group jobseekers in one round, include additional information with their usual text message list of new matches: a reminder that they can apply by sending the platform a text message to request a free callback. This avoids the pecuniary cost of initiating a phone call.	3,081	9,421	jobseeker × round	50%	This treatment has an effect on the job application rate that is one hundredth as large as the effect of the main phone call treatment. The two effects are statistically significantly different ($p = 0.017$). Implies that pecuniary costs of applying are unlikely to explain our main results.	Appendix C.5.1, Table C.7
Does lowering the time cost of applying increase the application rate by as much as the main treatment?	Additional experiment	For some control group jobseekers in some rounds, include additional information with their usual text message list of new matches: a notification that they can apply by sending the platform a text message to request a free callback at a specific time of their choice. This avoids the need to wait for their call to be answered when they call the platform and hence saves time.	3,634	29,985	jobseeker × round	33%	This treatment has an effect on the job application rate that is one quarter as large as the effect of the main phone call treatment. The two effects are statistically significantly different ($p = 0.002$). Implies that time costs of applying are unlikely to explain our main results.	Appendix C.5.1, Table C.7
Does the phone call treatment increase application rates by functioning as a reminder?	Additional experiment	For some control group jobseekers in some rounds, send a second text message reminding them about their new matches that month. This provides a reminder.	3,634	29,985	jobseeker × round	33%	This treatment has an effect on the job application rate that is one-fourteenth as large as the effect of the phone call treatment. Implies that reminder effects are unlikely to explain our main results.	Appendix C.5.2, Table C.8
	Additional experiment	Estimate how the phone call treatment effect varies with the randomized length of time between the initial text message and phone call.	6,430	530,734	jobseeker × round	NA	The phone call treatment has a smaller effect when there is more time between the text message and phone call. If phone calls worked as reminders, then their effects should be larger when this time gap is larger, because jobseekers have more time to forget to apply. Implies that the ”phone calls as reminders” interpretation is relatively unlikely.	Appendix C.5.2, Table C.9
	Additional analysis	Estimate how the phone call treatment effect varies with the length of time between the initial text message and the job application deadline.	9,011	1,005,463	NA	NA	The phone call treatment has a smaller effect when there is more time between the text message and application deadline. If phone calls worked as reminders, then their effects should be larger when this time gap is larger, because jobseekers have more time to forget to apply. Implies that the ”phone calls as reminders” interpretation is relatively unlikely.	Appendix C.5.2, Table C.9
Do phone calls pressure jobseekers to apply for jobs?	Additional analysis	Test if treatment group jobseekers are more likely to apply for the first job listed on the call and text message, which would be a way to satisfy pressure to apply as quickly as possible.	1,897	6,673	NA	NA	Treatment and control group jobseekers submit equal shares of applications to the first listed job (31%). Hence it is unlikely that treated jobseekers are applying more to simply avoid pressure to apply.	Appendix C.5.3, Figure C.7
	Additional analysis	Test if treatment group jobseekers who answer calls in early rounds of the experiment are less likely to answer calls in later rounds, which might happen if they want to avoid pressure during calls to apply for jobs.	3,534	13,840	NA	NA	Jobseekers who answer calls in early rounds are more likely to answer calls in later rounds, conditional on observed characteristics and using multiple different definitions of ”answered” and ”early”. Hence it is unlikely that jobseekers actively avoid calls to avoid pressure to apply. Hence it is unlikely that treatment increases applications through pressure to apply.	Appendix C.5.3, Table C.10

Table C.4: Evaluating Alternative Mechanisms – Covered in Detail in Appendix C.6

Question	Method						Result & Interpretation	Appendix
	Type	Description	# jobseekers	# matches	unit randomized	% units treated		
Are treatment-induced job applications sent to matches with similar attributes to matches to which the control group applies?		See Table C.2 row 1 for methods used in this analysis.					Treatment and control applications are sent to jobs with similar value proxies. Hence it is unlikely that the types of jobs to which applications are sent affects the relative returns to marginal versus inframarginal applications.	Appendix C.6.1
Do treatment-induced job applications come from jobseekers with similar attributes to job applications from the control group?	Additional analysis	For each match that receives an application, calculate the characteristics of the applicant (e.g. education, work experience, CV quality, predicted interview probability) and compare these between the treatment and control groups.	2,567	9,088	NA	NA	Applications in the treatment and control groups come from observationally similar jobseekers. Hence it is unlikely that jobseeker selection into applications affects the relative returns to marginal versus inframarginal applications.	Appendix C.6.2, Table C.11
	Additional analysis	Randomly select some control group jobseekers in some matching rounds to get phone calls so we can estimate treatment effects conditional on jobseeker fixed effects. Test if the effects of this temporary treatment are equal to the permanent treatment.	3,113	14,366	jobseeker × round	50%	The temporary and permanent treatment have very similar effects on job application rates and very similar returns to marginal job applications. Hence it is unlikely that time-invariant characteristics are different for jobseekers submitting applications in the treatment and control groups.	Appendix C.6.2, Table C.12
Do phone calls provide more information about specific jobs than text messages?	Additional analysis	Restrict to the 80% of matching rounds where call centre agents had no additional information about specific jobs, so it was impossible for them to provide specific information.	9,603	801,922	NA	NA	Treatment effects are very similar in the full sample and the 80% of matching rounds where call center agents could provide no additional information. Also note that call center agents are trained to use specific scripts that give no additional information to jobseekers, and audits of calls confirm that they followed scripts. Hence it is unlikely that the phone call treatment provides additional information about specific jobs.	Appendix C.6.3, Table C.14
Are jobseekers more likely to receive phone calls than text messages?	Survey	Phone jobseekers (separately from the matching phone calls) and ask if they received a match within a specific period. Test if the probability of reporting that they received a match differs between the treatment and control groups.	1,955	NA	NA	NA	Jobseekers in the treatment and control groups are equally likely to recall receiving matches from the platform, with or without adjusting for survey non-response. Implies that the treatment effect on applications is not caused by jobseekers being more likely to receive phone calls that text messages. But survey non-response means this result is uncertain.	Appendix C.6.3, Table C.15
Does the phone call treatment change jobseekers' beliefs about the value of jobs or the probability of job offers?	Survey	Phone jobseekers (separately from the matching phone calls) and ask their beliefs about jobs on the platform: the share of jobs that are desireable for them and the share of desireable jobs that would give them a job offer if they applied. Test if their beliefs differ between the treatment and control groups.	2,003	NA	NA	NA	Treatment has near-zero impact on jobseekers' beliefs about the quality of jobs on the platform or probability of getting an offer, with or without adjusting for survey non-response. Hence it is unlikely that treatment increases job applications by changing the perceived returns to applying. But survey non-response means this result is uncertain.	Appendix C.6.4, Table C.16
	Additional experiment	Randomly select some control group jobseekers in some matching rounds to get phone calls. Test if the effects of this temporary treatment are different to the permanent treatment. The temporary treatment seems more unusual to jobseekers, so it might have larger effects on their beliefs and hence application decisions.	3,113	14,366	jobseeker × round	50%	The temporary and permanent treatment have very similar effects on job application rates and very similar returns to marginal job applications. Hence it is unlikely that phone calls shift application decisions by signaling that these are unusually high-value matches.	Appendix C.6.4, Table C.12 (also discussed in Appendix C.6.2)
Do jobseekers choose randomly where to apply, so the expected outcome of each application is equal and returns to marginal and inframarginal applications are equal?	Additional analysis	Test if matches with more desirable attributes (e.g., salary, benefits, commute distance) are more likely to get applications.	2,567	9,088	NA	NA	Jobseekers are more likely to apply to matches with more desirable attributes, in both the treatment and control groups.	Appendix C.6.5, Table C.5 (also discussed in Appendix C.4)
	Additional experiment	Randomize the order that matches are listed in text messages and phone calls. Estimate the treatment effects of being listed first.	9,255	938,284	match	100%	Jobs listed first get substantially more applications. But the marginal return to these applications is $<1/2$ as large as the marginal return to applications induced by the main phone call treatment. Shows that marginal applications to random jobs have decreasing returns, which is inconsistent with random search.	Appendix C.6.5 , Table C.17

C.2 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 $PV > \frac{C}{\delta\beta}$.

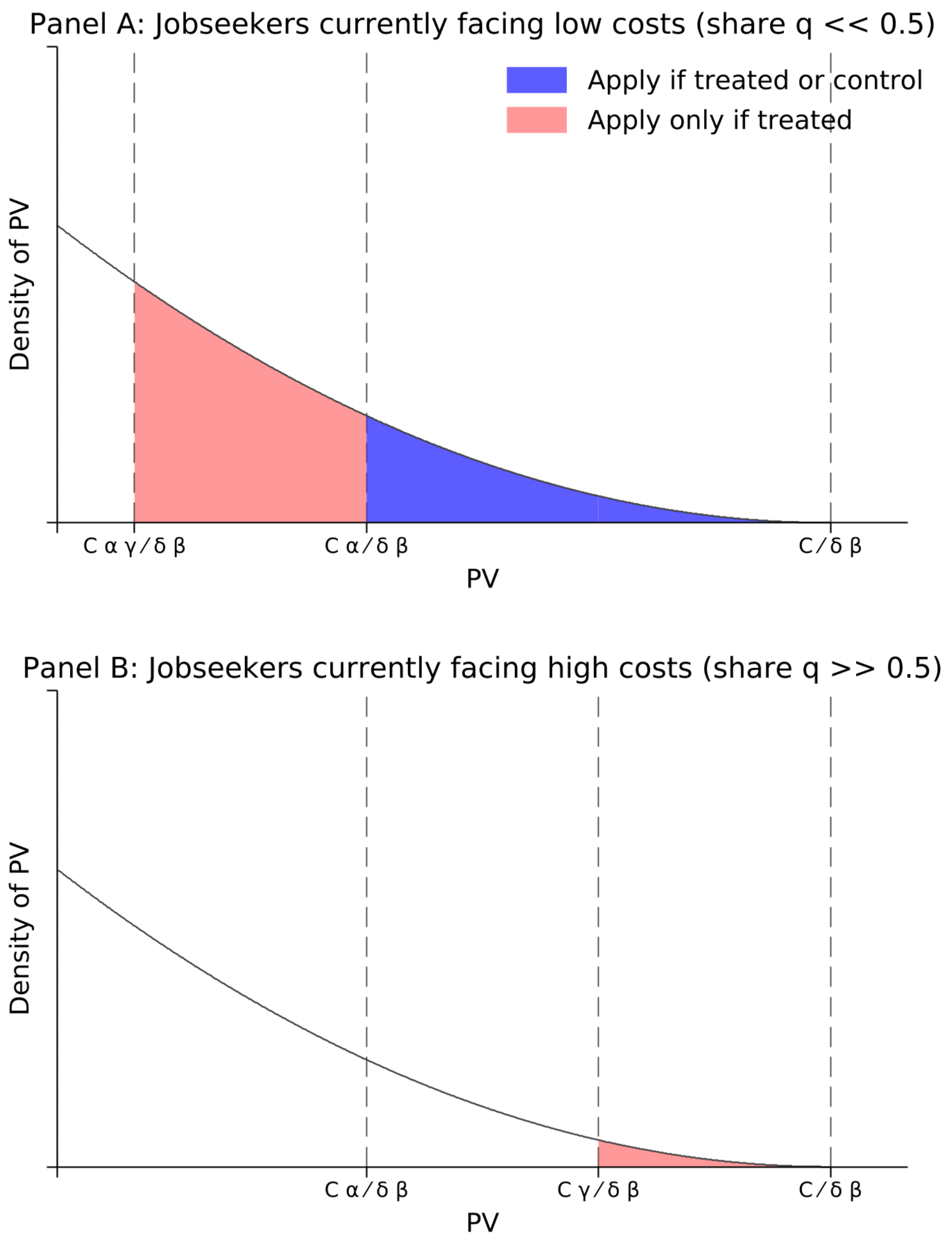
We can introduce heterogeneous application costs into this framework in multiple ways. We begin by assuming that in each month share q of jobseekers face the low application cost $C\alpha < C$, while the remaining share $1 - q$ of jobseekers face the high application cost C. This mimics the dynamic investment 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 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

Figure C.1: Application Decisions for Treated & 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. 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.

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.[27]

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.[28]

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 α 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).

[27] 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.

[28] 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 γ 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.3 Methods for Complier / Latent Type Analysis

Several analyses we report in subsequent appendices are based on comparisons of the characteristics of marginal job applications submitted due to treatment and inframarginal job applications that would be submitted in the absence of treatment. In this appendix we describe the method used to identify the characteristics of marginal and inframarginal job applications, which is adapted from Marbach & Hangartner (2020).[29]

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. *Note that type is defined at the level of the jobseeker $\times$ vacancy match.* Hence a jobseeker can submit a marginal application to one match – they would apply to that job if and only if treated – and submit an inframarginal application to another match – they would apply if assigned to treatment or control. And a vacancy can receive a marginal application from one jobseeker and an inframarginal application from another jobseeker.

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 μ_X^{AT} and μ_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.

[29] This method is a special case of the κ-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.

C.4 Additional Tests of the Conceptual Framework

The conceptual framework delivers three additional predictions that we can test. In this appendix, we describe each prediction, explain how we test it, and show that the test results are consistent with the predictions of the conceptual framework. This is the detailed version of the argument summarized in Section 4.2 of the paper.

Prediction 1: Control group jobseekers will not apply to some high-value matches. This prediction follows from the framework's assumption that some jobseekers face high enough application costs during some matching rounds that they will apply to no vacancies in these rounds. To test this, we consider two proxies for the value of a match: the salary, an admittedly narrow proxy but one that is easily interpretable and valued by all jobseekers; and the inverse covariance-weighted average of many positive attributes of a match (e.g. salary, benefits, commute distance) that is defined in Section 3.2. For each of the two proxies, we divide matches into quintiles and calculate the share of matches in each quintile that receive applications from control group jobseekers.

We find robust evidence that control group jobseekers miss applying to many high-value matches. Figure C.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. Panel B shows the same pattern for quintiles based on the index of match value. Figure C.3 shows the same pattern using the within-jobseeker between-vacancy distributions of salary and the index, which accounts for the fact that different jobseekers may receive matches with systematically different values.

We provide a further test that takes advantage of the way the platform often sends multiple matches to a jobseeker simultaneously. 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. Figure C.4 shows that 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 groups, 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.

Prediction 2: Treatment and control group applications will go to matches with similar average values. This prediction follows from the framework's prediction that when returns to marginal and inframarginal applications are roughly equal, then treatment will induce applications to a mix of higher- and lower-return matches whose average value is similar to the control group.[30]

We find robust evidence that treatment and control group applications go to matches with similar average values. To test this, Figure C.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). We see a similar pattern over many different ways of defining the value of matches, in each case reporting the results of the equal ratio tests in the footnotes below each figure. Figure C.2 Panel B shows the same pattern for quintiles based on the index of match value. Figure C.3 shows the same pattern using the within-jobseeker between-vacancy distributions of salary and the index. Figure C.5 shows a similar pattern when we collapse the data to the batch level (jobseeker $\times$ matching round), rather than the match level (jobseeker $\times$ vacancy). Finally, Figure C.6 shows a similar pattern when we measure the value of a batch based on the highest-value match in the batch, rather than the average values of the matches in the batch.

We provide a further test of this prediction using the complier / latent type method introduced in Appendix C.3. This allows us to estimate the mean characteristics of matches receiving marginal applications (i.e. applications submitted if and only if the jobseeker is treated) and of matches receiving inframarginal applications (i.e. applications submitted even if the jobseeker is not treated). Table C.5 shows that there are some differences between mean values of observed characteristics between 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 find no difference between the mean values of the summary index V_{vm} between marginal and inframarginal applications.

We extend the test in the preceding paragraph to show that marginal and inframarginal applications also go to matches with similar probabilities of yielding interviews. To do so, we estimate latent interview probabilities using a data-driven approach and then use the latent type method

[30]Technically, this prediction holds in the special case of the framework where returns to marginal and inframarginal applications are equal, as we see in our data. When returns to marginal and inframarginal returns differ, then treatment and control applications may be sent to vacancies with different average values, as we explain in Appendix C.2.

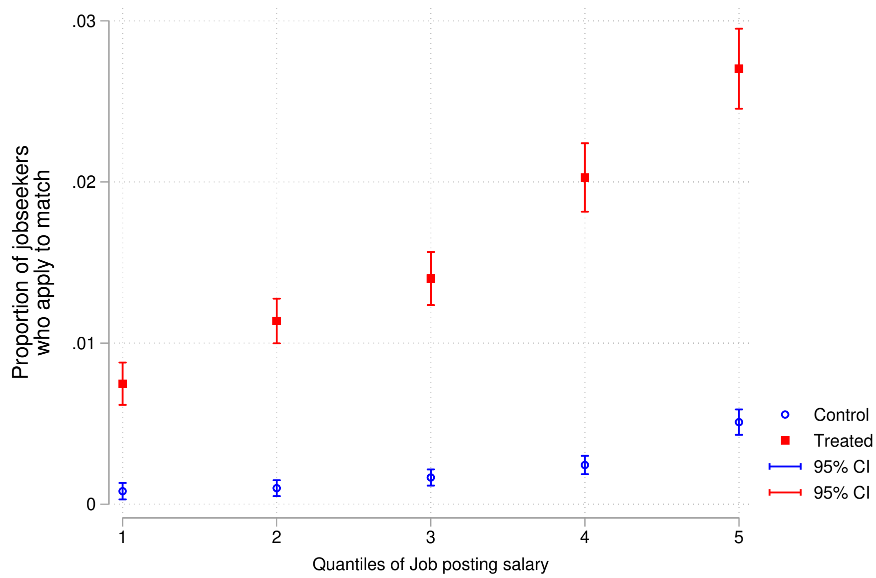
described in the previous paragraph to compare latent probabilities between the inframarginal and marginal applications. Specifically, we first restrict the sample to the set of applications from control group jobseekers, i.e. jobseeker $\times$ vacancy matches with *T* = 0 and *Apply* = 1. We then regress *Interview* on a vector of jobseeker, vacancy, and match characteristics using a logit LASSO and predict $\hat{P}|X_{jvm} = \hat{Pr}(\textit{Interview} \mid \textit{Apply} = 1, X_{jvm})$ for each match. This is the probability that an application to that match will produce an interview, given the observed characteristics of the jobseeker, vacancy, and match.[31] Table C.5's bottom panel shows that the mean probability is similar between marginal and inframarginal applications when estimated using only 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 final row of Table C.5 shows that the means of this omnibus proxy for marginal and inframarginal applications do not differ.

Prediction 3: Treatment group applications will go to matches with more dispersed values. This prediction follows from the framework's prediction that marginal applications that are submitted by jobseekers who are already submitting inframarginal applications in that period will go to matches with lower values, as evidenced by the wider range of PV in the in the pink+blue region than the pink-only region of Figure 1. To test this, we estimate treatment effects on the variance and 10th 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.6 shows treatment raises the variance and lowers the 10^{th} percentile for both log salary and the proxy index V_{vm} that combines multiple proxies for match and vacancy value, although treatment effects for the index are not statistically significant. 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.

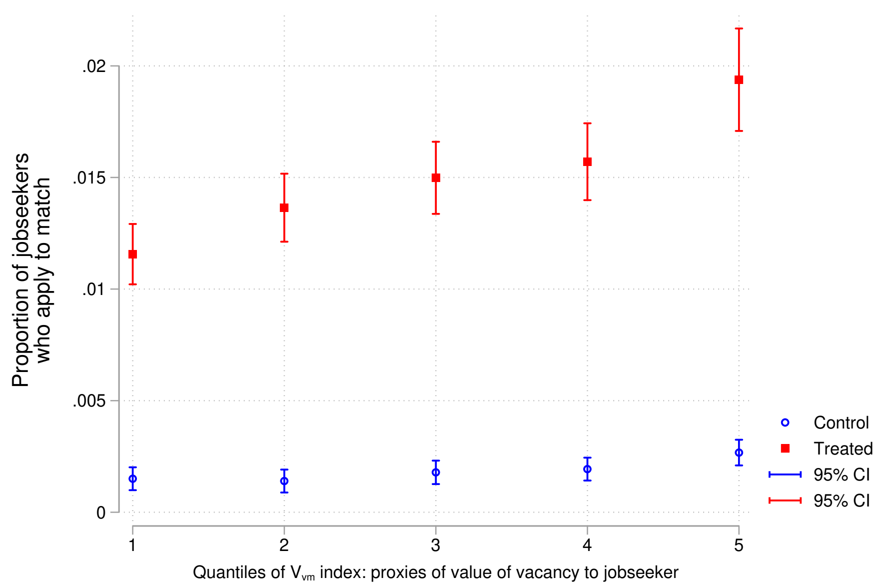
[31]This approach assumes that the relationship between interviews and observed characteristics does not differ for marginal and inframarginal applications, as we use the inframarginal applications for estimation and then predict out-of-sample to the marginal applications. This assumption is more reasonable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

Figure C.2: Heterogeneous Treatment Effects on Applications by Quintiles of Match Values

Panel A: Using Salary as a Proxy for Value



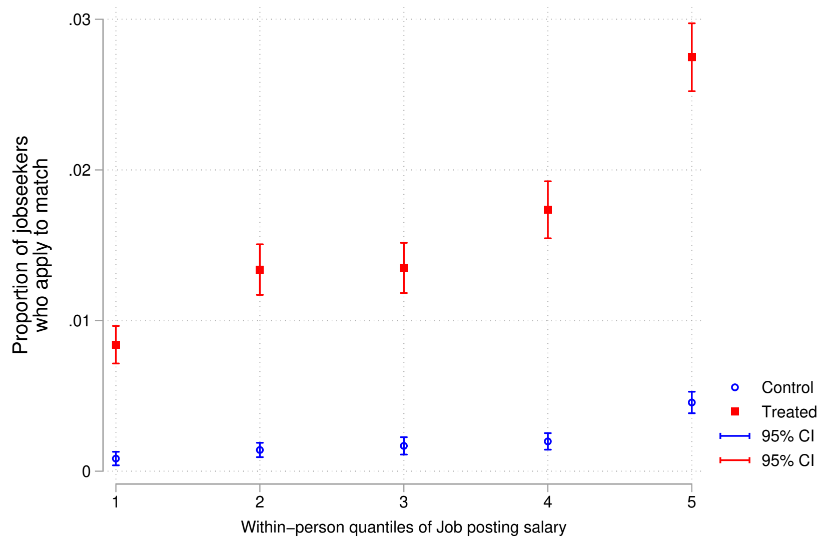
Panel B: Using the V_{vm} index as a Proxy for Value



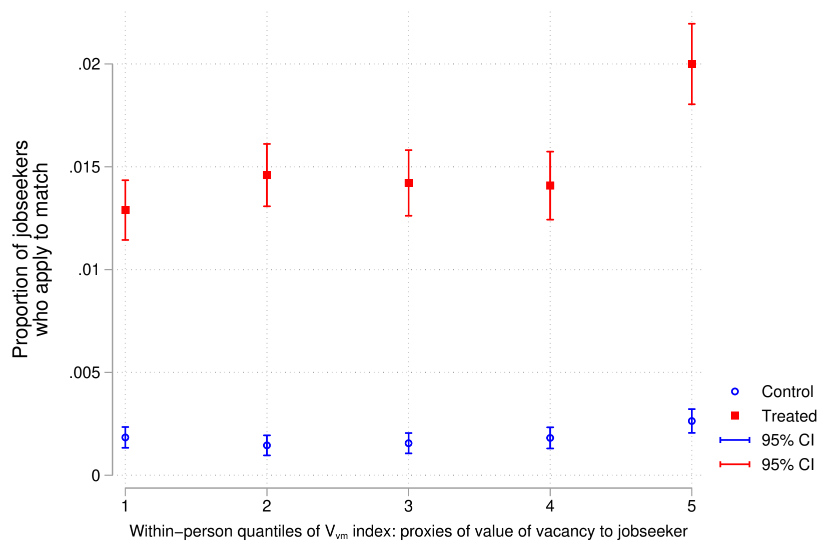
Notes: This figure shows heterogeneous treatment effects of the phone call treatment on applications by quintiles of proxies for the value of the jobseeker $\times$ vacancy match to the jobseeker. Panel A uses the job posting salary as a value proxy and Panel B uses the V_{vm} index described in Section 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.

Figure C.3: Heterogeneous Treatment Effects on Applications by Quintiles of the Within-Jobseeker Distribution of Match Values

Panel A: Using Salary as a Proxy for Value



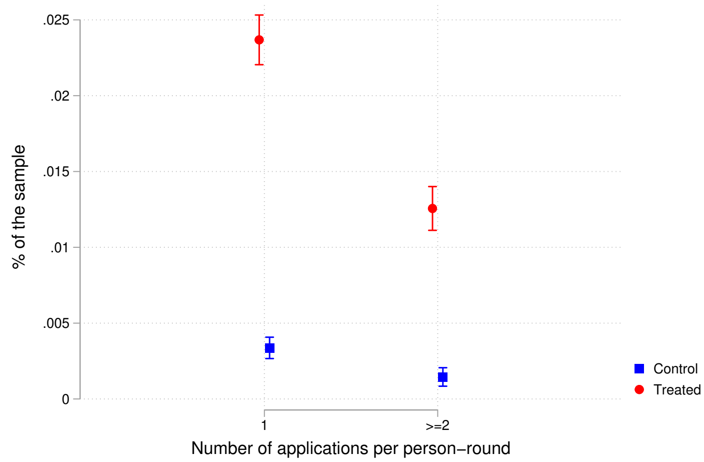
Panel B: Using the V_{vm} index as a Proxy for Value



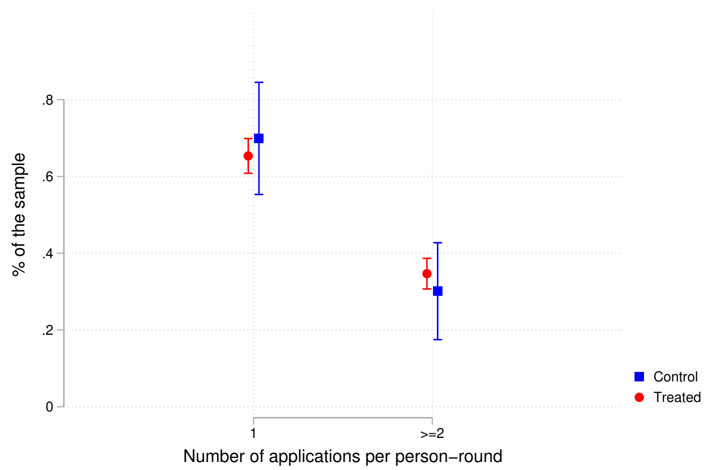
Notes: This figure shows heterogeneous treatment effects of the phone call treatment on applications by quintiles of proxies for the value of the jobseeker $\times$ vacancy match to the jobseeker. Panel A shows heterogeneity by job posting salary, defining quintiles based on the distribution of salary 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.4: Treatment Effects on the Number of Applications per Jobseeker $\times$ Matching Round

Panel A: Treatment Effects on Each Positive Number of Applications



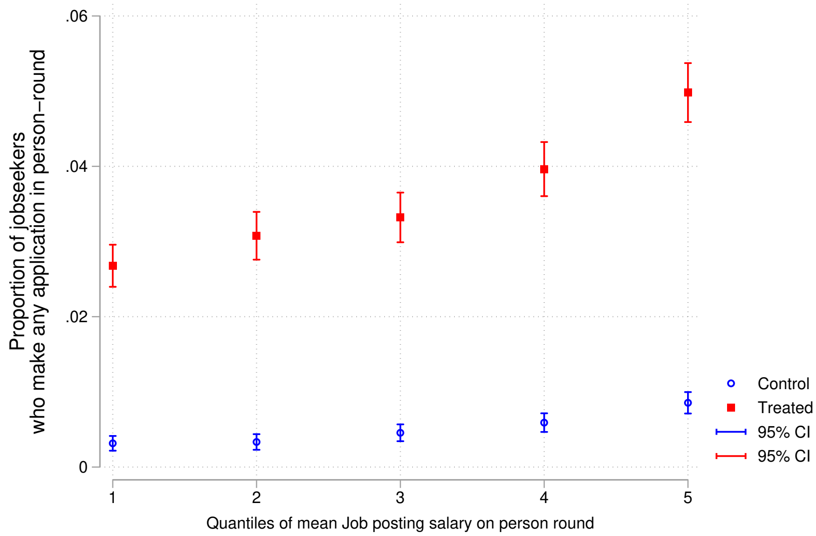
Panel B: Treatment Effects on Each Positive Number of Applications, Scaled by Pr(>0 Applications)



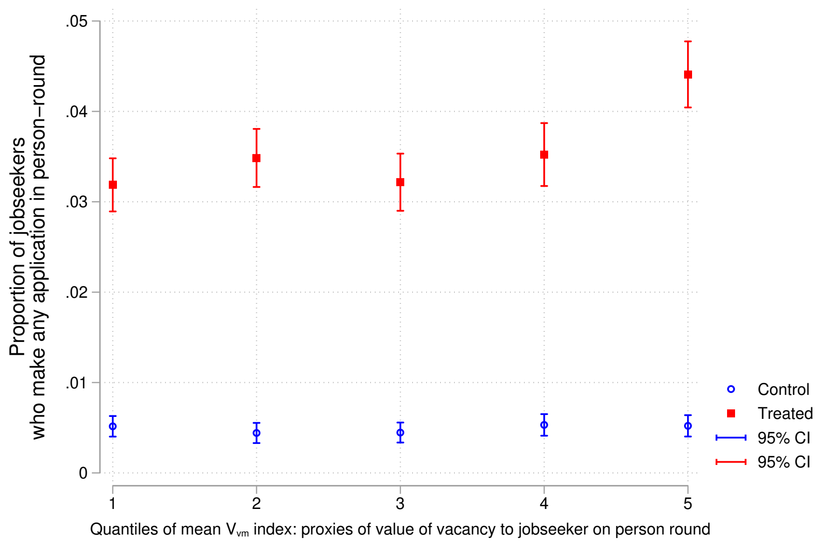
Notes: This figure shows treatment effects on the number of applications submitted per jobseeker $\times$ round. Estimation uses one observation per person-round, restricts the sample to jobseeker-rounds with at least two matches (65% of the data), conditions on stratification block fixed effects, and 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 Quintiles of Mean Batch Values

Panel A: Using Salary as a Proxy for Value



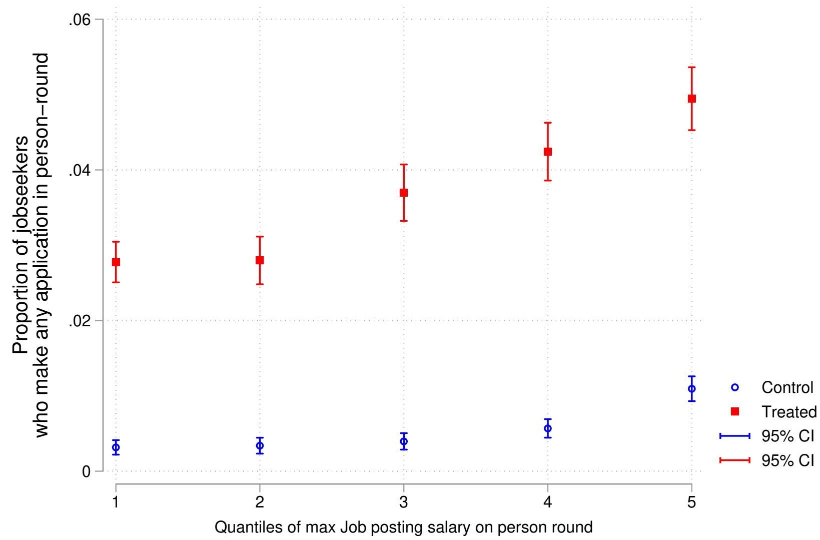
Panel B: Using the V_{vm} index as a Proxy for Value



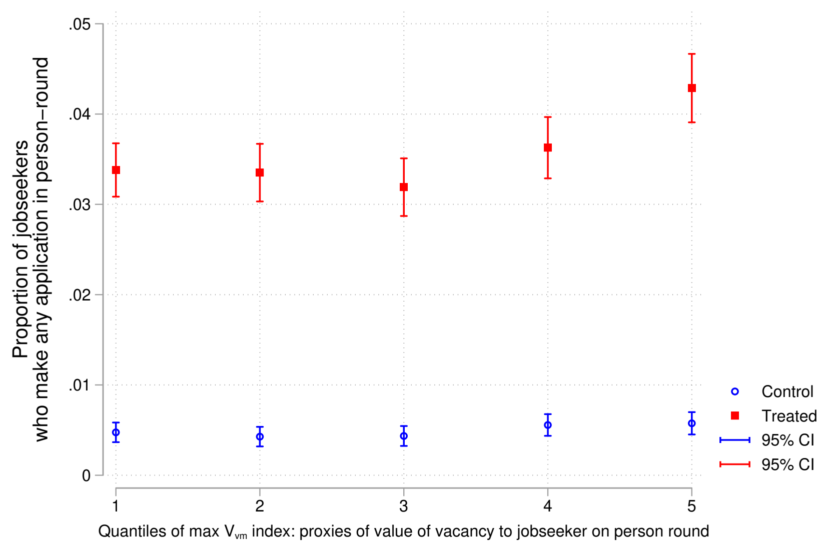
Notes: This figure shows heterogeneous treatment effects of the phone call treatment on applications by quintiles of proxies for the value of the jobseeker $\times$ vacancy match to the jobseeker. All analysis in this figure uses one batch (i.e. one jobseeker $\times$ matching round) as a unit of observation, *averaging over the values of the matches* in that batch. Panels A and B show heterogeneity by job posting salary and V_{vm} index described in Section 3.2 using the within-jobseeker 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 Quintiles of Maximum Batch Values

Panel A: Using Salary as a Proxy for Value



Panel B: Using the V_{vm} index as a Proxy for Value



Notes: This figure shows heterogeneous treatment effects of the phone call treatment on applications by quintiles of proxies for the value of the jobseeker $\times$ vacancy match to the jobseeker. All analysis in this figure uses one batch (i.e. one jobseeker $\times$ matching round) as a unit of observation, *taking the maximum of match values* in that batch. 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.

Table C.5: 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 $\mid$ vacancy requires high ed	0.168	0.265	0.097
			(0.008)
Exact experience match $\mid$ 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 $\mid$ vacancy and match characteristics	0.063	0.063	0.001
			(0.874)
$\hat{P} \mid X_{jvm}$: Prob. interview $\mid$ 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.

Table C.6: 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.

C.5 Understanding Application Costs

This appendix provides detailed evidence about the types of application costs that are reduced by the main phone call treatment, expanding on the summary in Section 4.3 of the paper.

C.5.1 Pecuniary and Time Costs

Here we show results for the mechanism experiments described in Section 4.3 that reduce pecuniary and time costs of job applications on the platform. Column 1 of Table C.7 compares the effects of our main phone call 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. This reminds the jobseeker that they can apply at near-zero pecuniary cost. 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, suggesting a negligible role for pecuniary costs.

Column 2 of Table C.7 compares the effects of our main phone call 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, suggesting a limited role for time costs. Each column uses only the set of jobseeker $\times$ vacancy matches from rounds in which the relevant feature was randomized.

Table C.7: 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.

C.5.2 Reminder Effects

Here we show results for the additional experiment and non-experimental analysis relating to reminder effects discussed in Section 4.3 of the paper. In principle, the phone call treatment's positive effect on job applications be explained by a combination of procrastination and forgetfulness, as in Ericson (2017): some jobseekers may 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 this reminder interpretation is inconsistent with the results from three mechanism tests. First, in three matching rounds, we send a second text message as a reminder to a random subsample of control group jobseekers, at the same time that the treatment group jobseekers get called. Table C.8 shows the effect of this reminder text 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.8: Treatment Effects on Applications of Reminder Text Messages

	(1) 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.

Second, we randomize the timing of the phone call within the period between sending the job alert text message and the application deadline. 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, Table C.9 column 1 shows that treatment effects are smaller when phone calls occur later.

Third, we use non-experimental variation in the length of time between the job alert text message and the application deadline. This length of time 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.9 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 more time to forget to apply. Instead, treatment has a smaller effect when the window is longer.

Table C.9: 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 is smaller in column (2) because the records of deadlines were not retained 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.

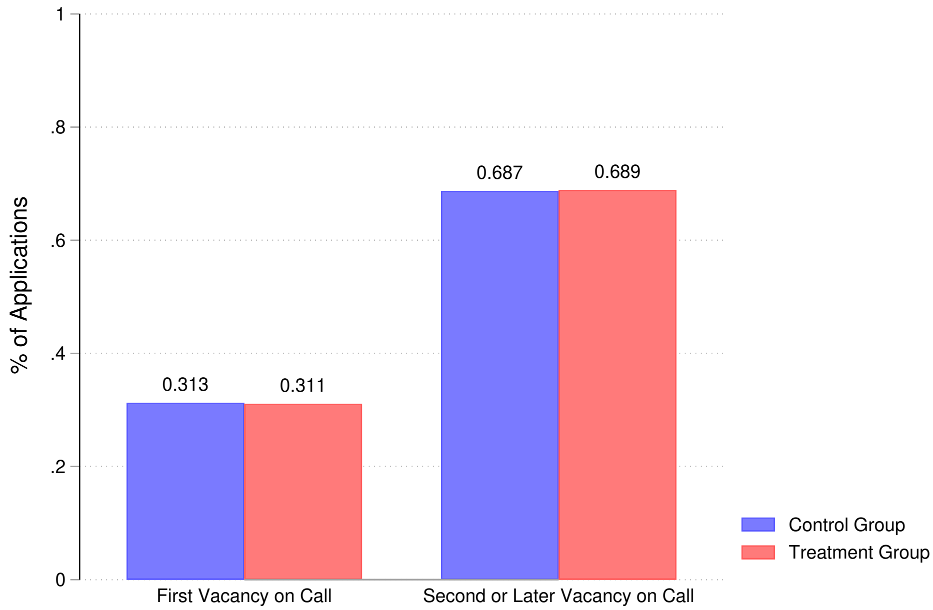
C.5.3 Encouragement and Pressure to Apply

Here we show results for the mechanism analysis relating to pressure and encouragement described in Section 4.3. In principle, the phone call treatment's positive effect on job applications might be explained by jobseekers experiencing encouragement or pressure to apply from the phone calls.

But this explanation is inconsistent with three patterns in our data. First, call center agents are trained not to encourage or pressure jobseekers to apply, and regular audits of call recordings verified that they followed their scripts. This shows that jobseekers do not experience *explicit* encouragement or pressure. It remains possible that jobseekers feel *implicit* pressure because they have been called or because they are interacting with a person. But, second, if jobseekers feel implicit pressure, they do not respond in the lowest-cost way, which would be to apply for the first job listed on the call to end the call as quickly as possible. Instead, both treated and control group jobseekers send 31% of their job applications to vacancies listed first on their phone call or text message (Figure C.7). 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.

Third, to further evaluate the potential role of implicit pressure, we study the within-jobseeker time series of responses to calls. If treated respondents did experience pressure to apply when called, they could easily avoid this pressure by not answering calls. In this scenario, jobseekers who answer calls in early rounds when they are learning about what it's like to interact with the platform might be less likely to answer calls in subsequent rounds. To test for this potential negative relationship, we restrict the sample to the first 6 months of the experiment and regress an indicator for 'answered the call in round T' on an indicator for 'answered the call in round $T-1$,' controlling for observed jobseeker characteristics. Table C.10 column 1 shows jobseekers who picked up the previous call are 20 p.p. *more* likely to pick up the current call, rather than less likely. The relationship is very similar when we restrict to an even shorter period, 3 months (in column 2). Jobseekers could alternatively avoid pressure by answering phone calls and then saying they are unavailable to talk and immediately ending the call. But repeating this analysis using indicators for 'answered the call and was willing to talk' shows similarly positive relationships in the first 3 and 6 months of the experiment (column 3 & 4).

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 × vacancy matches in which *Apply* = 1) in person-rounds in which the jobseeker receives at least two matches.

Table C.10: Response Persistence Between Periods

	Pickup Call		Available	
	(1)	(2)	(3)	(4)
Pickup call$_{t-1}$	0.20433	0.22598		
	(0.00946)	(0.03032)		
Available$_{t-1}$			0.18895	0.19330
			(0.00945)	(0.02776)
# matches	13840	1474	13840	1474
# jobseekers	3534	922	3534	922
# jobseekers with outcome variation	2158	170	1773	181
Mean outcome | Pickup$_{t-1}$ = 0	0.35023	0.44396		
Mean outcome | Available$_{t-1}$ = 0			0.20996	0.32934
No of Months	First 6 months	First 3 months	First 6 months	First 3 months

Notes: Columns (1) and (2) show the relationship between call pickup in the previous and current matching rounds. "Pickup call" = 1 if anyone answered the phone call. Columns (3) and (4) show the relationship between availability in the previous and current matching rounds. "Available" = 1 if the target jobseeker was reached and was available to listen to their job matches. The unit of observation is jobseeker × round. The sample in all columns is restricted to jobseekers assigned to the phone call treatment group. Additionally, columns (1) and (3) focus on the first six months of the experiment, while columns (2) and (4) focus on the first three months. All specifications control for randomization block fixed effects and the following baseline jobseeker characteristics: age, gender, married, child, above-median education, any work experience, years of work experience, and number of occupations selected. Heteroskedasticity-robust standard errors, clustered by jobseeker, are reported in parentheses.

C.6 Evaluating Alternative Explanations

This appendix provides detailed evidence about the five alternative explanations for the effects main phone call treatment that we evaluate and largely reject. This expands on the summary in Section 4.4 of the paper.

C.6.1 What Types of Vacancies Receive Marginal & Inframarginal Applications?

The returns to marginal applications submitted due to the phone call treatment depend in part on the characteristics of vacancies and matches to which these are submitted. In principle, it might be possible that marginal applications and inframarginal applications might have roughly equal returns if marginal applications are sent to vacancies that are both systematically worse matches for the jobseekers (leading to lower interview probabilities) and systematically less competitive (leading to higher interview probabilities).

However, we show in Appendix C.4 that marginal and inframarginal applications are sent to vacancies and matches with very similar observed characteristics, including predicted probabilities of yielding interviews. This suggests this explanation is unlikely.

C.6.2 Which Jobseekers Submit Marginal & Inframarginal Applications?

The returns to marginal applications submitted due to the phone call treatment depend in part on the characteristics of jobseekers who submit each type of applications. In principle, it might be possible that marginal and inframarginal applications have roughly equal returns because 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.

However, in this appendix we show that jobseekers who submit marginal and inframarginal applications have similar time-invariant characteristics using four methods. First, marginal and inframarginal applications come from jobseekers with similar observed characteristics. To show this, we use the same latent type analysis introduced in Appendix C.3. Table C.11 shows that mean education and CV quality scores (provided by firms, as discussed in Section 2.3) are almost identical for the jobseekers submitting marginal and inframarginal applications. Marginal applications come from jobseekers with slightly more work experience. But, as we show below, our main findings are unchanged when we control for experience.

Second, marginal and inframarginal applications come from jobseekers with similar latent interview probabilities. To show this, we estimate latent interview probabilities using a data-driven approach and then 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 $\times$ 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. The final row of Table C.11 shows that the mean of this measure does not differ between marginal and inframarginal applications.

Third, our main findings hold when we include jobseeker fixed effects. To show this, we run a "crossover" experiment that randomly reassigns some control group jobseekers to the treatment group in some matching rounds. It allows us to replicate our main analysis with jobseeker fixed effects, using only within-jobseeker variation through time to identify treatment effects. Table C.12 shows that the results of this experiment are similar to our main results. In particular, we cannot reject equality of the interview rates or quality-adjusted interview rates for inframarginal applications and marginal applications submitted due to the crossover treatment ($p > 0.480$). 16% of jobseekers have at least one match affected, allowing precise estimation of the crossover treatment effect conditional on the fixed effects. But only 0.65% of matches are affected by this treatment, so it has almost no impact on our estimates of the overall treatment effect (Table C.13).

Fourth, controlling for time-invariant observed jobseeker characteristics leaves our main results unchanged, including the constant returns finding. To show this, we repeat our analysis of the main experiment using a double selection LASSO to control for an extensive set of jobseeker baseline characteristics, following Belloni et al. (2014). The point estimates and standard errors are almost identical (Table B.3, columns 1, 4 – 6).

These four results show that the roughly constant return to treatment-induced job search is not explained by treatment changing patterns of jobseeker selection into applications.

Table C.11: 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.

Table C.12: Treatment Effects on Job Search & Search Returns Using Jobseeker Fixed Effects

	(1) Apply	(2) Interview	(3) Int. × V_{vm}	(4) Interview	(5) Int. × 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: This table shows the results from replicating the analysis in Table 3 but including jobseeker and round fixed effects and replacing the time-invariant treatment indicator with an indicator variable for jobseeker × vacancy matches assigned to the "crossover" phone call treatment, which reassigned some control group jobseekers in some rounds to the phone call treatment group. This indicator is used as a regressor in columns (1) – (3) and as an instrument for applications in columns (4) – (5). Other than the jobseeker fixed effects and use of the crossover treatment indicator, all other sample definitions, regression specifications, and inference methods are identical to Table 3. Findings are qualitatively similar to those from analyzing the time-invariant phone call treatment.

Table C.13: Treatment Effects on Job Search & Search Returns, Controlling for Crossover Matches

	(1) Apply	(2) Interview	(3) Int. × V_{vm}	(4) Interview	(5) Int. × V_{vm}
Phone call treatment	0.01333	0.00078	0.00285		
	(0.00075)	(0.00009)	(0.00036)		
Apply				0.05886	0.21380
				(0.00512)	(0.02133)
# 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				0.662	0.517
IV strength test: F-stat				314.7	314.7
IV strength test: p-value				0.00000	0.00000

Notes: This table shows that the estimates in Table 3 are unchanged when we include in each regression an indicator variable for the random subset of control group matches that are are reassigned to the treatment group in a subset of rounds. All other sample definitions, regression specifications, and inference methods are identical to Table 3.

C.6.3 Does Treatment Provide More Information About Matches?

In principle, the phone call treatment might provide information about specific jobs, leading to higher application rates and enabling jobseekers to target better-matched vacancies.

However, in this appendix we show four pieces of evidence that are not consistent with this mechanism. First, call center agents are trained to read precise scripts that contain the same information as in the text messages. Second, agents do not provide additional information about general labor market conditions or assessments of the individual jobseeker's prospects. They are not given this information by the platform and are trained not to tell jobseekers about any beliefs they hold. Regular audits of call recordings confirm high compliance with these two aspects of the training. See Figure A.3 for an explanation of the call structure.

Third, call center agents seldom provide information about specific jobs. In approximately 80% of matching rounds covering 72% of matches, we gave the call center agents no additional information beyond the content of the text message scripts. All our results hold when restricting the sample to these rounds: treatment increases the application rate by more than 600% and we cannot reject equal returns for marginal and inframarginal applications (Table C.14).[32]

Fourth, treatment and control group jobseekers are equally likely to receive job matches from the platform. In principle, it might be possible that jobseekers are more likely to receive phone calls than text messages. For example, text messages may sometimes be blocked or unread.[33] To test this, we survey jobseekers to ask if they remember receiving a job match from the platform by either phone call or text message in the previous 14 or 30 days (recall period randomized). Treatment and control group respondents are equally likely to report receiving matches (Table C.15, column 1). This pattern also holds with the sample selection correction described in Appendix B.6 (Table C.15, column 2). The majority of jobseekers were not sent any matches by the platform in previous 14 or 30 days, because they did not match to any jobs in this period. 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. These treatment effects remain close to zero, with or without the sample selection correction (columns 3 & 4).[34]

[32] As an additional test, we record if the jobseeker asked for additional information on the call, irrespective of whether the call center agent could provide this information. Treatment effects on applications are positive on calls with and without requests for more information. This suggests that application effects are not driven by even requests for new information, let alone receiving new information. But this analysis should be viewed with caution because it involves splitting the sample by a post-treatment variable: the choice to ask for more information.

[33] We also test and find no difference in treatment effects between the 93% of respondents who indicated at registration that they were comfortable communicating with the platform by text message and the remaining 7%.

[34] These columns 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. 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: jobseekers may forget they received matches (and hence underreport receiving matches) or may forget the exact date they received them (which might lead to overreporting or underreporting receiving matches).

Table C.14: 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.15: 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)
# responses	2177	14069	2177	14069
# responses | T = 0	978	978	978	978
# responses | T = 1	1199	1199	1199	1199
Mean outcome | 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 Appendix 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.

C.6.4 Does Treatment Affect Jobseekers' Beliefs About The Value of Applications?

In principle, the phone call treatment might increase the job application rate by changing jobseekers' beliefs. For example, jobseekers might view a call from a professional recruiting service as a signal that platform firms are larger or wealthier and thus able to provide more benefits or opportunities for advancement (higher V), or as a signal that the firm sees her as a good fit for the job (higher P).

However, in this appendix we show two pieces of evidence that are not consistent with this mechanism. First, we directly test this explanation by collecting data on jobseekers' beliefs about P and V and estimating treatment effects on these two belief measures. Translated from Urdu, these questions ask: "Suppose Job Talash sends you one hundred job ads in the next year. Based on your past experience with our job matching service, how many of these jobs do you think would be desirable for you?" and "Suppose you apply for all of these jobs that you think are desirable. How many do you think would make you an offer?" Our main treatment assignment is time-invariant, so these questions are asking jobseekers about jobs sent by the mode of communication used in their treatment group. We therefore use a jobseeker-level version of equation (1) to estimate treatment effects on these two belief measures. Table C.16 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 decreases this by 1.1 percentage points (standard error 1.8). Jobseekers in the control group on average think that 31% of the vacancies on the platform would be desirable for them; the phone call treatment decreases this by 0.7 p.p. (standard error 1.6). The even-numbered columns show that results are similar when we adjust for survey non-response using the same method described in Appendix B.6. The survey data show that treatment does not increase respondents' perceptions of the average values of V or P on the platform, and hence cannot explain the treatment effect on applications.

Second, if phone calls influence job applications because a jobseeker views them as informative about the quality of a specific match, then phone calls should have larger effects on applications when the jobseeker views the phone call as unusual than when she views it as part of normal platform operations. We can test this idea using the "crossover" experiment described in Appendix C.6.2, which reassigns a random subset of control group respondents into the treatment group in a few rounds. The phone calls will seem more unusual for jobseekers receiving this temporary treatment. The treatment effect on applications is similar for the main and crossover treatments (Table C.12). This suggests that the phone call is unlikely to shift application decisions by signaling that these are unusually high-value matches.

Table C.16: Treatment Effects on 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.

C.6.5 Random Search

In principle, our finding of roughly constant returns to treatment-induced search might be explained by a random search framework. If jobseekers were to apply to vacancies at random and the phone call treatment were to reduce the cost of applying, then treatment would 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 (e.g. Pissarides 2000). It may also be a reasonable approximation given empirical evidence that jobseekers have limited information about labor market conditions and match quality (e.g. Behaghel et al. 2023; Belot et al. 2018; Kiss ⓡ al. 2023).

However, in this appendix we note two pieces of evidence that are not consistent with this interpretation. First, recall from Section 2.4 and Appendix C.4 that applications are sent more often to vacancies with higher salaries and other positive attributes, 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. 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.

We find that listing vacancies first produces more applications with decreasing, rather than constant, returns. Table C.17 shows that the probability of application is 0.4 p.p. higher for vacancies listed first 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). This is substantially and statistically significantly lower than the 6.3% average interview probability for jobs listed second or later, and less than half the 5.9% interview probability for 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.[35]

The result of this experiment emphasizes that the return to marginal search depends on what 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 treatment causes marginal search that is directed in similar ways to inframarginal search and has

[35]Results 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 characteristics conditional on these fixed effects. Results are similar if 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.

roughly constant returns. This highlights that our constant returns finding is a consequence of the type of search induced by the phone calls, not inherent to this labor market or to these jobseekers.

Table C.17: Treatment Effects of Lowering Cost of Applying to Randomly Chosen Vacancies

	(1) Apply	(2) Interview	(3) Int. × V_{vm}	(4) Interview	(5) Int. × 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 ≥ 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 Spillover Effects

Increased search effort by some jobseekers may affect firms and other jobseekers. The sign of this effect is theoretically ambiguous. For firms, getting 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, 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 get 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. On such platforms, it is 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.

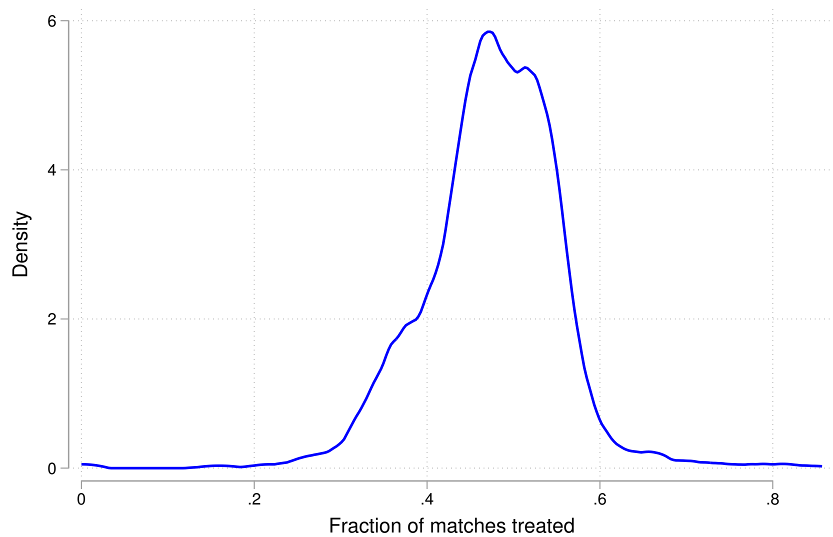
We estimate spillover effects using two methods. Our first method tests whether jobseeker-level outcomes are sensitive to the fraction of competing jobseekers who are treated, closely following Crepon et al. (2013). We define TR_{jv} as the fraction of jobseekers matched to vacancy v who are treated, 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:

$$\text{Interview}_{jv} = T_j \cdot \beta_1 + TR_{jv} \cdot \beta_2 + T_j \cdot TR_{jv} \cdot \beta_3 + \mathbf{X}_v \cdot \Lambda + \mu_b + \epsilon_{jv}, \qquad (10)$$

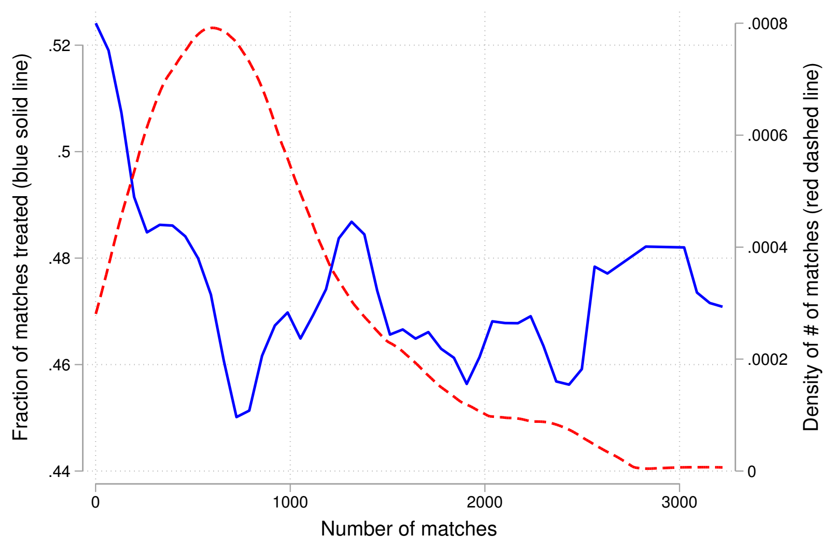
where $\mathbf{X}_v$ contains the number of jobseekers matched to vacancy v and vacancy-level factors that determine matches (e.g. occupation) and μ_b is a stratification block fixed effect. We cluster standard errors by both jobseeker and vacancy because treatment is assigned at the jobseeker level 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

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 5. 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.

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 β_2 and β_3 are both small and not statistically significant (Table D.1, 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 = 0.589) and -0.011 p.p. for a treatment group jobseeker (standard error 0.017, 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 (10) 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 TR_v. These effects are again close to zero for control or treatment group jobseekers (Table D.1, column 2).

We also estimate spillovers using the number of jobseekers matched to vacancy v who are treated, rather than the share of jobseekers matched to vacancy v who are treated. This measure has substantial variation, with an interdecile range of 226 to 1048. In these regressions we also replace the control for the total number of jobseekers matched to the vacancy with the number of control group jobseekers matched to the vacancy, to avoid including treated jobseekers in two regressors. We again find no evidence of a negative relationship: a vacancy exposed to the 75th percentile of the number of matched jobseekers who are treated rather than the 25th percentile would have a 0.120 percentage point higher treatment effect on interviews (standard error 0.058, p = 0.039).

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

$$\Delta\text{Interview}_v = TR_v \cdot \alpha + \mathbf{X}_v \cdot \Lambda + \varepsilon_v, \tag{11}$$

conditional on the same vacancy-level covariates $\mathbf{X}_v$ as the previous analysis. Finding $\alpha < 0$ 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 α (Table D.1, 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

Table D.1: 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

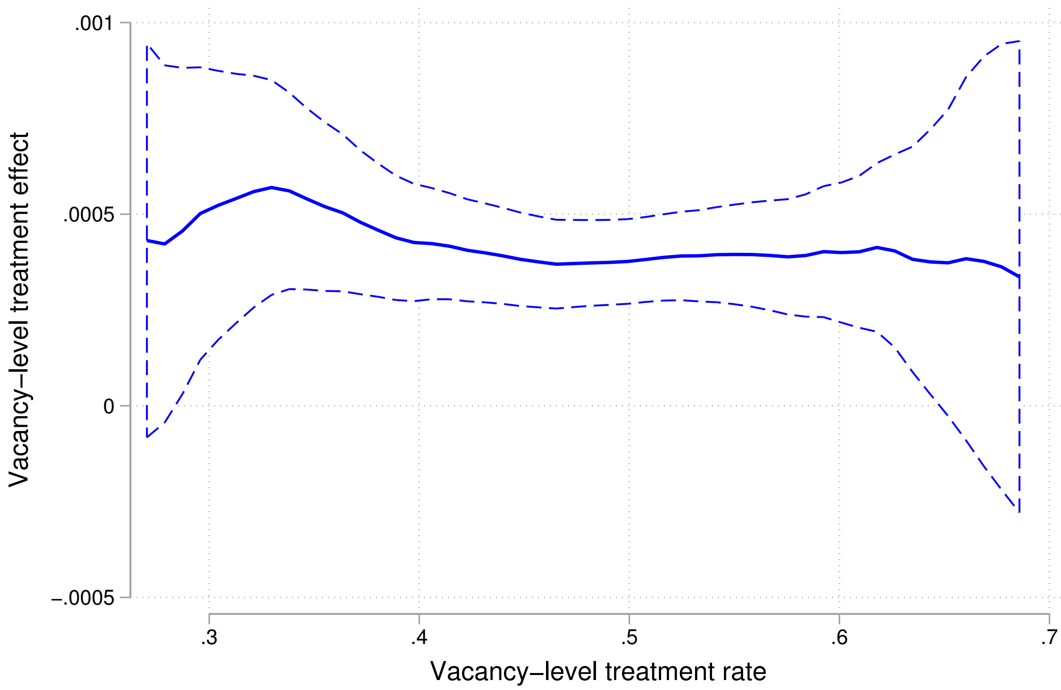
Notes: This table shows the results of tests for spillovers between jobseekers on interview invitations. Column (1) shows results from regressing match-level interview invitations on own treatment status, the fraction of other jobseekers matched to the same vacancy who are treated, and their interaction. 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.

on interviews (standard error 0.011, p = 0.096). To test for a nonlinear relationship, we repeat this analysis replacing the vacancy-level treatment rate TR_v with indicators for the middle and top terciles of TR_v. These coefficients are positive, although we cannot reject the null hypothesis that treatment effects are equal across all three terciles (Table D.1, column 4). A nonparametric regression of vacancy-level treatment effects on treatment rates also shows no evidence of negative spillovers (Figure D.2).

The lack of negative spillovers is consistent with descriptive patterns in vacancy-level outcomes. If firms did 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 but an extremely 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.2).

What might explain the negligible spillover effects we find? Our design cannot directly answer this question, but we suggest five possible explanations. First, spillover effects might be avoided 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, more offers need not mechanically lead to crowd-out because firms on this platforms report filling only 60% of vacancies, as on other platforms (Fernando et al., 2021). Third, firms may not get enough total applications from the platform to generate meaningful congestion: the average vacancy on this platform receives only 0.8 applications from control group applicants and another 6 applications from treated applicants (with pooled interdecile range 0-18). Fourth, the fact that only 30% of applications come through the platform attenuates the scope for search effort increases on the platform to generate spillover effects. The average firm receives 6.8 applications via the platform, of which 6 are due to the search encouragement treatment, but it also receives 15.9 applications from outside the platform. Taking these factors together, it is possible that firms in this labor market receive too few suitable applications via the platform in the absence of treatment for crowd-out to be relevant, at least at the interview stage that we are able to observe.

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.2: 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 5. 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.

Appendix 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, NBER.

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.

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.

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.

Behaghel, L., S. Dromundo, M. Gurgand, Y. Hazard & T. Zuber (2023) "The Potential of Recommender Systems for 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.

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.

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

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.

De Chaisemartin, C. (2017) "Tolerating defiance? Local average treatment effects without monotonicity," *Quantitative Economics*, 8, 367–396.

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.

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.

Faberman, J. & M. Kudlyak (2019) "The Intensity of Job Search and Search Duration," *American Economic Journal: Macroeconomics*, 3 (11), 327–357.

Garlick, R. & J. Hyman (2022) "Quasi-Experimental Evaluation of Alternative Sample Selection Corrections," *Journal of Business and Economic Statistics*, 40 (3), 101–125.

Gee, L. (2019) "The More You Know: Information Effects on Job Application Rates in a Large

Field Experiment," *Management Science*, 65 (5), 2077–2094.

He, H., M. Roel & Q. Weng (2023) "How Many Others Apply for the Jobs I Am Applying for? The Effect of Perceived Labor Market Competition on Job Search," Working Paper.

Heckman, J. (1974) "Shadow Prices, Market Wages, and Labor Supply," *Econometrica*, 42 (4), 679–694.

Imbens, G. & D. Rubin (2015) *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*: Cambridge University Press.

Kelley, E. M., C. Ksoll & J. Magruder (2021) "How do Online Job Portals Affect Employment and Search Outcomes? Evidence from India," Working paper.

Kleibergen, F. & R. Paap (2006) "Generalized Reduced Rank Tests Using the Singular Value Decomposition," *Journal of Econometrics*, 133 (1), 97–126.

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.

Laibson, D. (1997) "Golden Eggs and Hyperbolic Discounting," *The Quarterly Journal of Economics*, 112 (2), 443–478.

Lee, D. (2009) "Trimming, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects," *Review of Economic Studies*, 76, 1071–1102.

Marbach, M. & D. Hangartner (2020) "Profiling Compliers and Noncompliers for Instrumental-Variable Analysis," *Political Analysis*, 28 (3), 435–444.

Marinescu, I. & D. Skandalis (2021) "Unemployment Insurance and Job Search Behavior," *The Quarterly Journal of Economics*, 136, 887–931.

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.

Pakistan Bureau of Statistics (2018-2019) "Quarterly Labour Force Survey."

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.

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