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

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

April 4, 2023

Abstract

Jobseekers face multiple barriers with potentially different implications for the level of search and returns to increasing search. An experiment on a job search platform in Pakistan shows that lowering users' psychological cost of initiating job applications increases applications by 600%. Returns to the marginal applications induced by treatment are approximately constant rather than decreasing, in contrast with intuitive job search models. This pattern is consistent with a model in which heterogeneous psychological costs of initiating applications, potentially due to heterogeneous present bias, lead some jobseekers to miss applying to even high-return vacancies. Additional experiments and measurement reject alternative behavioral and non-behavioral explanations. Our finding of constant returns to marginal search effort, combined with limited spillovers onto other jobseekers, raises the possibility of suboptimally low search effort due to psychological costs of initiating applications.

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

†Duke University, BREAD, NBER, field.erica@duke.edu

‡Duke University, robert.garlick@duke.edu

§Bates College, nsubrama@bates.edu

¶Duke University, katherine.vyborny@duke.edu

1 Introduction

Job search is a central feature of labor markets, and job search frictions can have important economic consequences. Macroeconomic models of frictional search can explain both employment levels and productivity of firm-worker matches (Pissarides, 2000). 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 job search can also be sensitive to behavioral factors such as present bias, reference-dependence, and motivated reasoning, which can reduce or delay search effort (e.g. Cooper and Kuhn 2020; Dellavigna et al. 2022; Mueller and Spinnewijn 2022).

In this paper, we study the search effort of jobseekers on a search and matching platform who receive monthly text messages listing vacancies relevant to their skills and interests. We show that adding follow-up calls that invite jobseekers to immediately start the application process substantially increases the number of applications, relative to leaving it to the jobseeker to take the initiative to phone and apply after getting text messages. Moreover, returns to the additional search effort are approximately constant rather than decreasing, in contrast with many job search models. This pattern is consistent with a model in which follow-up calls reduce the psychological cost of initiating applications, which is high enough for some jobseekers at some times that it deters applications to even high-return vacancies. This finding of constant returns, combined with the finding that additional search effort has limited spillovers onto other jobseekers, suggests the possibility of sub-optimally low search effort due to psychological costs of initiating applications.

To show this, we work with a novel job search platform in Lahore, Pakistan.[1] We 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. They have a wide range of education levels, ranging from incomplete primary school to graduate degrees, and a wide range of baseline labor force attachment, ranging from employed and searching to non-employed and non-searching. This sample breadth is unusual in experimental job search studies (Poverty Action Lab, 2022). 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 shifts how jobseekers communicate with the platform, lowering the psychological cost of initiating job applications. All users receive a monthly text message listing new vacancies that match their education, experience, and occupational preferences. These matches are determined by information jobseekers report at sign-up, before treatment assignment. Control group users must initiate job applications by calling the platform or asking the platform to call them. Treatment group users also receive a follow-up phone call after the text message that invites them to begin the application process, reducing the psychological cost of actively initiating a job application. The experimental design holds constant many other features of the economic environment: the phone call treatment has a negligible effect on the pecuniary and time costs of applying, provides no direct encouragement or pressure to apply, and does

[1] Such platforms are becoming an increasingly common feature of many labor markets. In Pakistan, Rozee, LinkedIn, and Bayt had respectively 9.5, 7.5, and 3 million users in 2021. Bayt reported 39 million users in 2021 across the Middle East, North Africa, and South Asia. LinkedIn reported over 10 million users in 2022 in at least 18 countries, 10 of them low- or middle-income.

not provide more information about the vacancies.

Our two main findings are that the phone call initiation treatment dramatically increases the job application rate, and that the average return to the additional applications is approximately constant rather than decreasing. Treatment increases the share of jobseeker $\times$ vacancy matches getting applications from 0.2 to 1.5%. Using treatment as an instrument for applications shows that marginal treatment-induced applications have a 5.9% probability of yielding interviews. This is neither substantively nor statistically significantly different from the 6.3% probability for inframarginal applications from the control group, implying that returns to job search are approximately constant over this large increase in search effort. Returns are also approximately constant for 'value-weighted' interviews, weighted by their desirability in terms of salary, hours, commute times, and non-salary benefits.[2] This finding is not explained by differences in the 'quality' of jobseekers who submit marginal vs inframarginal applications: we reject this type of selection using observed quality proxies and we replicate our main finding using an additional within-jobseeker through-time randomization.[3]

The finding of approximately constant returns is surprising. It clashes with an intuitively plausible model in which jobseekers prioritize applying to vacancies with the highest combination of interview probabilities and desirable attributes, so any additional applications would have decreasing returns. Variations on this intuitively plausible model form the basis for the literature on 'directed' job search, reviewed by Wright et al. (2021).[4] Constant returns are consistent with models of 'random' job search, where vacancies are homogeneous and jobseekers randomly choose where to apply, including the canonical model of Pissarides (2000). However, random search models do not match other results or features of our setting. We find substantial variation in both vacancy value and proxies for jobseeker-vacancy match value such as salary, commute times, and alignment between jobseekers' work experience and vacancies' experience preferences. Furthermore, control group jobseekers are more likely to apply when this value is higher. And we run an additional experiment deliberately designed to encourage random search, which generates sharply decreasing returns to marginal search. So, if jobseekers can and do direct applications to higher-value vacancies, why do additional applications induced by treatment have roughly constant returns?

To explain the constant returns finding, we present a simple model of job search with heterogeneous psychological application costs. Our goal is to provide a compelling explanation for our main findings, and several additional results, without claiming this is the only possible explanation. In this model, the platform matches jobseekers to vacancies each month and each jobseeker applies to every match for which the expected present value of applying exceeds the cost of applying. The key to this model is a *psychological cost of initiating applications* that *varies across jobseekers and/or within jobseeker through time*. This heterogeneity might be due to variation in the psychological cost itself (following Carroll et al. 2009) or variation

[2]Throughout the paper, we use 'returns to search' to refer to interview invitations or value-weighted interview invitations, acknowledging that these are gross returns and not net of application costs.

[3]Abebe et al. (2019) show that constant or even increasing returns to additional search are possible when search costs and jobseeker quality are positively correlated, as marginal applications can then come from higher-quality jobseekers. We do not see evidence of this pattern in our data.

[4]Some models of directed job search show that constant returns to marginal applications are possible if marginal and inframarginal returns are directed to different types of jobs (Belot et al., 2018). We find no evidence of this pattern in our data.

in jobseekers' degree of present bias that changes the ratio of benefits to psychological costs (following DellaVigna and Paserman 2005; Paserman 2008).[5] In either case, this heterogeneity means that some jobseekers in some months apply to at least one vacancy, while in other months they apply to no vacancies. This model also explains another pattern that we observe: many control group jobseekers miss applying to high-quality matches, potentially because they face particularly high psychological costs.

The phone call treatment reduces the psychological cost of initiating applications because the jobseeker passively receives the call rather than initiating it. Jobseekers who would have applied in that month without treatment may now submit additional applications, which will have on average lower returns because the lower cost lowers the 'bar' for applying. On the other hand, jobseekers who would not have applied in that month without treatment may now submit applications, which will go to the highest-return vacancies available to them and hence might have higher average returns than control group applications. The return to treatment-induced marginal applications can equal the return to inframarginal applications when averaged across these two groups of jobseekers. This prediction echoes research showing that eliminating the need for initiating decisions can increase financial and health investments (DellaVigna, 2009; DellaVigna and Malmendier, 2006; Madrian and Shea, 2001; Thaler and Benartzi, 2004).

We can test and reject many alternative explanations for our main findings. Additional experiments show that reducing the pecuniary or time costs of applying has minimal effects, implying an important role for the psychological cost of initiating applications. Alternative behavioral explanations – encouragement, pressure, or reminders – are not consistent with the platform design or results from additional experiments. Information- or belief-based explanations – additional information about matches or higher perceived returns to applications – are not consistent with the platform design, results from additional experiments, or survey measures of beliefs. Econometrically, we develop tests to show that the constant returns finding is robust to potential violations of the exclusion and monotonicity conditions in our instrumental variables analysis.

Importantly, we do not find evidence that this additional search has negative spillovers on other jobseekers. We treat 50% of jobseekers on the platform, which increases total search by enough that quantitatively large spillovers are possible. Instead, we find that individual jobseekers' interview probabilities are unaffected by competing against more treatment-induced applications from other users. However, we do not have data on job offers to test spillovers at that stage. We also use survey data to show that the treatment effect on off-platform search is close to zero, consistent with no crowd-out across search methods.

We also estimate treatment effects on employment. However, we are underpowered to study employment effects at the scale of this experiment. This occurs because treatment leads to dramatically more applications and interviews, but from a very low control group mean, so even treated jobseekers have too few interviews

[5] We do not attempt to test between heterogeneous psychological costs and heterogeneous present bias, as they produce equivalent predictions for both job applications and returns. This echoes a more general literature highlighting the difficulty of separating heterogeneous present bias from heterogeneous costs (Bernheim et al., 2015; Dean and Sautmann, 2021). DellaVigna et al. (2017) and Dellavigna et al. (2022) also document the role of present bias in job search, building on work by O'Donoghue and Rabin (1999) and O'Donoghue and Rabin (2001). They do not directly focus on heterogeneity in present bias but Duflo et al. (2011) show that heterogeneous present bias in particular can explain low adoption of highly profitable technology. The existing literature suggests several factors that could make applying psychologically costly including fear of rejection (reviewed by Bénabou et al. 2022) and cognitive costs of evaluating vacancies (Maćkowiak et al., 2023), particularly when this requires switching from another task (Rubinstein et al., 2001). We do not attempt to separate these factors.

to meaningfully shift the probability of employment. But treatments that lower the psychological cost of initiating applications on larger platforms might generate substantial job offer effects. For example, Rozee, Pakistan's largest job search platform, is almost one thousand times larger than our platform. And interviews do represent an important search outcome because they are a necessary condition for job offers and impose non-trivial costs on both job applicants and firms.

Our paper makes two main contributions. First, by studying psychological job search costs we add to a small literature on behavioral job search (reviewed by Cooper and Kuhn 2020). Existing work shows patterns of job search consistent with present bias, motivated reasoning, and reference dependence (DellaVigna and Paserman, 2005; DellaVigna et al., 2017; Dellavigna et al., 2022; Mueller and Spinnewijn, 2022; Paserman, 2008).[6] Babcock et al. (2012) suggest multiple possible policies to encourage and improve job search in the presence of behavioral biases. However, Abel et al. (2019) and Sanders et al. (2019) provide the only evaluations of policies designed to directly target behavioral factors, showing that job search action plans increase employment. We extend this work by running multiple field experiments to show how small, theory-informed policy changes to the job search environment can increase search without lowering returns, while other policy innovations that are only slightly different would fail to produce this result. Many other job search policies might have behavioral channels: motivated reasoning might affect how jobseekers process and use new information (Abebe et al., 2021a,b; Abel et al., 2020; Altmann et al., 2018; Bandiera et al., 2021; Bassi and Nansamba, 2020; Beam, 2016; Behaghel et al., 2020; Belot et al., 2018; Boudreau et al., 2022; Carranza et al., 2021; Dammert et al., 2015; Garlick et al., 2022; Spinnewijn, 2015; Subramanian, 2021), present bias and reference dependence might influence how jobseekers spend subsidies (Abebe et al., 2021a; Banerjee and Sequeira, 2020; Field and Vyborny, 2022; Franklin, 2017), and relationships between caseworkers and jobseekers might have behavioral components (Arni and Schiprowski, 2019; Bolhaar et al., 2020; Lechner and Smith, 2007; Schiprowski, 2020). However, research into these forms of job search assistance has not sought to pin down behavioral components.

Second, we provide a direct estimate of returns to additional search effort. Returns to search effort, typically interpreted as job applications, are a central feature of canonical job search models (Pissarides, 2000) and are important for evaluating policies such as search subsidies or search requirements for recipients of unemployment insurance. However, direct estimates are very rare, in part because identifying returns to search requires data on both search effort and outcomes, as well as exogenous search effort shifters. Many papers study the effect on employment of search subsidies or requirements for recipients of government benefits, but do not observe actual search effort (see reviews by Card et al. 2010, 2018; Filges et al. 2015; Heckman et al. 1999; Marinescu 2017b). A smaller, more recent literature studies the effect of search subsidies or requirements on online search effort, but without observing outcomes of search (Baker and Fradkin, 2017; Marinescu, 2017a; Marinescu and Skandalis, 2021). Without direct estimates of returns to search, it is difficult to understand why search shifters have different effects in different contexts (due to different effects on search or returns to search?) or to design search promotion policies (how many applications should be required or subsidised?). But only Arni and Schiprowski (2019) and Lichter and Schiprowski (2021)

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

show directly how additional policy-induced job applications affect labor market outcomes, specifically unemployment duration. We advance on this work by using application-level data that allow us to precisely describe how marginal and inframarginal search effort is directed and to compare the outcomes of marginal and inframarginal applications. However, we do not observe administrative data on employment, as they do. Most of the experimental work discussed in the preceding paragraph studies changes in search strategies, search technologies, or multiple dimensions of search, rather than isolating the role of search effort.[7]

Our findings of a positive and non-decreasing jobseeker-level return to search and a lack of spillovers on other jobseekers demonstrate the possibility of sub-optimally low search effort. This is relevant to debates about possible spillovers or congestion effects from rising search effort and what this implies for labor market policy.[8] Our results match those from recent studies using platform vacancy-level data to show that interviews and offers do not respond to application volumes (Fernando et al., 2021; Horton and Vasserman, 2021). Taken together, the three studies suggest that employer responses to application volumes on platforms are relatively inelastic. This might occur because vacancy fill rates are below 60% in all these studies, so more applications can increase the probability that any one applicant meets the firm's reservation quality.

Our findings about how jobseekers direct job applications to specific vacancies also relate to a growing literature on directed job search, a framework based on the idea that jobseekers send applications to jobs with higher wages or higher match quality (Wright et al., 2021). Recent microeconomic research finds mixed evidence for directed search (Alfonso Naya et al., 2020; Behaghel et al., 2020; Belot et al., 2018, 2022; Garlick et al., 2022; Gee., 2019; He et al., 2021; Marinescu and Wolthoff, 2020). We show that marginal and inframarginal applications are directed to similar types of vacancies and yield roughly equal returns in terms of interviews and quality-weighted interviews. This suggests that jobseekers miss applying to some high-value vacancies, which could be incorporated into future job search models.

Methodologically, we show the value of embedding multiple related experiments into a job search and matching platform to identify a specific barrier to search. This adds to a growing literature using these platforms as laboratories to study job search and hiring behavior (e.g. Horton 2017; Lyons 2017; Pallais 2014; Stanton and Thomas 2021), although rarely through multiple experiments (Pallais and Sands, 2016).

We describe the economic environment in Section 2: 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 develop and evaluate our preferred interpretation in Section 4 and show evidence against alternative interpretations in Section 5. We discuss spillover effects in Section 6 and treatment effects on off-platform outcomes in Section 7.

[7]In particular, our work differs from recent papers studying the effect of encouraging people to enroll on job search platforms (Afridi et al., 2022; Kelley et al., 2021; Jones and Sen, 2022; Wheeler et al., 2022). Joining a job search platform is a bundled experience that might shift factors ranging from wage expectations (Kelley et al., 2021) to information about specific vacancies (Wheeler et al., 2022). These have substantially different interpretations to our findings. The same is true of work studying the effect of access to (faster) online job search (Bhuller et al., 2019; Chiplunkar and Goldberg, 2022; Gurtzgen et al., 2020; Hjort and Poulsen, 2019; Kuhn and Skuterud, 2004; Kuhn and Mansour, 2014).

[8]See Blundell et al. (2004), Crepon et al. (2013), Ferracci et al. (2014), Gautier et al. (2018), Johnston and Mas (2018), LaLive et al. (2022), Lise et al. (2004), and Toohey (2014) for evaluations of search-encouragement policies in the presence of spillovers.

2 Economic Environment

2.1 Context

Our experiment takes place on Job Talash ("job search" in Urdu), a job search and matching platform in Lahore, Pakistan, created by our research partners at the Center for Economic Research in Pakistan. Lahore is a city of about 10 million people located in Pakistan's Punjab province. Pakistan's official Labor Force Survey for 2018 shows that 49% of adults in Lahore were in the labor force and 47% were employed, although the gender gap was large: women's labor force participation and employment were 11 and 10%, while men's were 85 and 83% (Table A.1). The job search rate was 50% for non-working adults who are 'available for work,' the only respondents who search status is recorded by the Labor Force Survey. Conditional on searching, direct applications were the most common method but are only fractionally more common than search through networks. The prevalence of informal, network-based search matches patterns in other developing economies (Government of Bangladesh, 2015; Government of South Africa, 2018; Government of Namibia, 2016). Job search and matching platforms are a growing feature of Pakistan's labor market, particularly in major urban areas such as Lahore, as we describe in footnote 1.

2.2 Samples of Jobseekers and Firms

We recruited participants by conducting a household listing from a random sample of 356 enumeration areas across Lahore between October 2016 and September 2017. This provides a representative listing of $49,506$ households and $182,585$ adult members across metropolitan Lahore, which included their age, education, and current work status. We invited each adult household member regardless of employment status to sign up for the Job Talash platform and $46,571$ expressed interest. The Job Talash call center then called each of these people to collect information on their education, work experience, job search, and occupational preferences. Platform staff used this information to populate CV templates for everyone who completed the sign-up process. These $9,838$ people 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: 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. Our sampling process is unique in covering a broader section of the population, which allows us to show that the search barrier we identify affects many different types of active and potential jobseekers.

Column 1 of Table 1 presents descriptive statistics for the control group in our study sample. At baseline, 20% of the sample are employed and searching through some channel other than Job Talash, 35% are searching but not employed, 14% are employed but not searching, and 31% are neither employed nor searching. Network search is the most common method, more than twice as common as applying directly and three times as common as visiting establishments to ask about vacancies. Later surveys of respondents show that only 4% used some other job search assistance program or online platform. The average respondent has 7.9 years of work experience with an interdecile range of 0-16. Respondents' education levels also vary widely: 15% have no education, 15% have completed secondary school, and 25% have a university degree. 31% are female and the average age is 30, with an interdecile range of 20-45. In Table A.1, we compare the study sample to the population of Lahore, captured by both the official Labor Force Survey and our household listing. Our sample is younger, more male, more educated, less likely to be employed, and more likely to be searching, suggesting greater openness to a novel search platform among these demographic groups.[9]

Firms are enrolled through a door-to-door listing in commercial areas of Lahore, described in more detail in Appendix A. Firms are invited to list any current vacancies during enrollment and periodically thereafter, providing each vacancy's job title, occupation, salary, benefits, and hours. Vacancies cover a range of education and experience levels and occupations, such as 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] The mean salary offer is roughly 60% of the mean salary in the Labor Force Survey data for Lahore (Figure A.2) 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.

Use of online job posting platforms was rare in this sample. At baseline, only 22% of firms had advertised a vacancy on a job search platform, while 67% had recruited through referrals, 35% from CVs dropped off directly 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 relatively easy scaling: 97% of urban households in Punjab have a mobile phone (MICS (2018)).

After signing up, jobseekers are matched to each listed vacancy using a very simple algorithm: the jobseeker must have at least the required years of education and experience, appropriate gender if specified by the employer, and must have indicated interest in the occupation category corresponding to the job.[11] Jobseekers can update education, experience, and occupation preferences in their profile whenever they choose, including adding missing information from their initial sign-up process; an active effort was made

[9] Our measures of search prevalence are not directly comparable to the Labor Force Survey. The Labor Force Survey reports that 50% of those who are 'available for work' but not employed are searching, but does not measure search for the employed or those not available for work. We do not observe a measure of 'available for work' in our sample. But 50% of the non-employed in our sample were searching at baseline, suggesting a higher rate if we were able to condition on availability for work.

[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.

[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 and Shen, 2013).

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

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

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

to update CVs by calling participants an average of 34 months after enrollment.

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 (2.4 per month since sign-up) and the interdecile range is 8-259.

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 their vacancies and how much jobseekers might value the vacancies to which they are matched. This heterogeneity creates the potential for non-constant returns to applications, which is important for interpreting our experimental results.

The platform sends jobseekers text message updates with matches approximately once per month if the jobseeker has matched to any vacancies in that month. See Figure A.1 for a sample text message. Jobseekers on average receive a text every 2.8 months. 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 search function. Conditional on receiving any matches in that month, the average jobseeker receives 3.3 matches per month and the interdecile range is 1-7. Participants can ask to pause or stop receiving matches at any time.

If a jobseeker wants to apply to any of these vacancies, she is instructed to call the platform using a number listed in each month in the text message, before the deadline also stated in the text message. If the jobseeker reports that she wants to apply to a specific vacancy, the platform forwards her CV to the firm. The CVs are constructed by the platform by populating a template with respondent-specific information, so there is no variation in CV design. The platform sends all applications to the firm in a packet after the application deadline; thus timing of application 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 follows up with each firm a few weeks after the application packet is delivered to ask which applicants they interviewed. Column 2 of Table 1 shows that, within our sample, jobseekers who actively use the platform are slightly younger, better-educated, are more likely to be searching for jobs at baseline.

The platform design has two key advantages for our research, relative to standard job search platforms. First, we observe all information available to both sides of the market. We observe the same information about vacancies as jobseekers receive through the text messages, and the same information about jobseekers as firms receive through the CVs. We also gather a quality measure for the CVs of 1,470 jobseekers that

[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 if they do or do not have the preferred types of education and experience, e.g., their work experience might be in a non-preferred field. We use the alignment between jobseekers' education and experience and vacancies' preferred types as a measure of match quality.

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

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

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

applied to jobs on the platform that were willing to share their evaluation data with us. Second, respondents see only the vacancies to which they match. This generates a well-defined jobseeker-vacancy unit of analysis that we use throughout the paper, and refer to as a match. This is not possible on platforms that allow unrestricted search, as every jobseeker can apply to any vacancy on the platform and the researcher may not observe which vacancies the jobseeker has seen, making it difficult to distinguish between vacancies a jobseeker sees but decides not to apply to and vacancies she has not seen at all.[13]

[13]Jobseekers can influence which matches they receive by changing CV information, changing occupational preferences, or requesting to stop receiving matches temporarily or permanently. This potentially creates a sample selection problem for the match-level dataset but we show in Appendix B.2 that the amount of selection is small and correcting it does not affect our results.

2.4 Platform Use

We highlight four important patterns of platform use, focusing on the control group statistics in Tables 1 and 2. First, the application rate is low: the average jobseeker submits only 0.23 applications and applies to 0.2% of matches they receive, or an average of roughly 0.03 applications per month. While on the low end relative to other platforms, this is expected given that our sample deliberately includes people who were not actively searching at baseline (on or off the platform), and corresponds to rates of platform use in comparable samples.[14] The application count is unsurprisingly right-skewed: 80% of jobseekers submit zero applications and 2% submit more than 5 applications. 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 converting into an interview.[15]

Third, there is substantial variation in match value, and applications are directed to relatively high-value matches. For example, the standard deviation of monthly salary is roughly 9,200 Pakistani Rupees (275 USD PPP) and higher-salary vacancies get more applications (Table 2, column 2, row 1). This pattern persists within jobseeker: the average control group jobseeker faces a standard deviation of 4,900 Pakistani Rupees across the vacancies to which they match and is 445% more likely to apply to a match in the top than bottom quintile of their with-jobseeker across-match salary distribution (Figure C.1, panel A). At the match level, jobseekers are more likely to apply to vacancies where their work experience matches the firm's preferred as well as required level (Table 2, column 2, row 5). Combining our available proxies for vacancy and match value in a single index shows that applications are substantially more likely for high-value matches (row 8). This confirms that jobseekers can and do apply to higher-value matches, rather than randomly picking where to apply from relatively homogeneous matches, as random search models assume.

Fourth, however, control group jobseekers miss applying to many high-value matches. For example, jobseekers apply to only 0.46% of the matches in the top quintile of the within-jobseeker across-match salary distribution (Figure C.1, panel A). This pattern also persists for other measures of value such as hourly wage, commute-adjusted salary and an index of all vacancy- and match-level value measures we observe (Figure C.1, panel B).

These patterns naturally motivate our research. On the one hand, the facts that job applications are rare, including to high-value matches, and that applications have reasonably high interview probabilities suggest that further 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

[14] For example, in a sample of South African jobseekers who were encouraged to create LinkedIn accounts but were not already using the platform, jobseekers submit an average of 0.03 applications per month, almost identical to the rate we observe (Wheeler et al., 2022). In contrast, in samples restricted to *active* platform users studied in economics research, the number of job applications range from 0.2 - 3.6 applications per person per month (Banfi et al., 2019; Kudlyak et al., 2013).

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

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 to jobseekers in order to reduce the non-pecuniary costs 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 describe each match received by the jobseeker that month (described above) and tells jobseekers to call the Job Talash call center by a stated deadline if they wish to apply. The call center number is always included in text messages and stays the same for the entire experiment. The vacancy 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 within the same day to move forward with the application process. The financial cost of placing the call to initiate the application process is a maximum of PKR 5 (US 3 cents, or less than 1% of a day's earnings at minimum wage). In addition, mobile telephone service providers in Pakistan offer small "loan" packages allowing for customers to "borrow" 10-20 rupees of credit against a future top-up card, and the application period for each matching round stays open for at least a week, so a short-term zero balance is very unlikely to be a binding constraint.

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

Importantly, the text message and phone call scripts contain identical information: firm name, job title, location and salary for each matched vacancy. The phone call scripts are also identical for the treatment and control groups: the call center agent reads the information from the text messages to the jobseeker and then asks if they want to apply to any of the matched vacancies. The only difference between the two is that in the control group, the jobseeker must initiate the call, while the call center initiates the call for the treatment group. If the jobseeker requests more information about the job, in later matching rounds of the experiment the call center agent is permitted to provide a one-line job description, hours and flexibility, travel requirements, bonuses, and benefits. However, in these rounds we observe jobseekers requesting

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

additional information on fewer than 2% of phone calls, and we show in Section 5.3 that our results are robust to excluding these matching rounds. Additionally, call center agents are trained to not encourage or pressure jobseekers to apply at any moment during the call, and a supervisor audits the records of at least one call per call center agent per matching round to ensure agents are following the script.

We interpret treatment as a reduction in the cost of applying for jobs on the platform. The phone call allows jobseekers to apply at a specific point in time without any costs of initiating a call to 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. fear of rejection or cognitive costs of processing vacancy information). However, the platform is already designed to minimize the monetary and time costs jobseekers incur to initiate applications, and we show in Section 5 that additional experiments further reducing monetary and time costs do not replicate our main findings. 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 more formally in Section 4, showing what this implies for treatment effects on applications and the returns to treatment-induced applications. We show in Section 5 that we can rule out several other interpretations based on the platform design, additional experiments we run, and additional outcome measures we collect.

3 Search Effort and Returns to Search

In this section we first show that that the phone call initiation 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 the same probability as inframarginal applications submitted without treatment, and yield interviews for vacancies of similar quality. This implies roughly constant returns to additional search effort with respect to these outcomes.

3.1 Treatment Effects on Search Effort and Search Outcomes

We run all analyses 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. This means that these matches provide a well-defined unit of observation, unlike most job search and matching platforms where jobseekers can in principle observe and apply to any posted vacancy. We first estimate:

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

using the sample of all matches sent to all jobseekers over a four-year period. 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 treatment was randomized (see footnote 16). We estimate heteroskedasticity-robust standard errors clustered by jobseeker, the unit of treatment assignment.

The phone call initiation treatment leads to a large increase in job applications. Treated respondents apply to 1.3 percentage points more matches with standard error 0.08 p.p. (Table 3, column 1). This effect

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 and Paap (2006).

is seven times higher 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 mean interview invitation rate of 0.012%. At the jobseeker level, treatment also shifts the entire distribution of the number of interview invitations to the right (Figure B.1). In particular, treatment increases the proportion of jobseekers who ever receive a job interview on the platform from 1.3 to 6%. The interview data is collected from firms, not jobseeker surveys. Firms are unaware of respondent-level treatment assignments, so using firm reports of interview invitations minimizes measurement error from experimenter demand effects.[17]

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

The treatment effects on both applications and interview invitations are broad-based: treatment increases the job application and job interview rates for women and men, for people who were employed and not employed at baseline, for people who were searching and not searching at baseline, and for people with above- and below-median education and age (Table B.8). 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 present in Appendix B.2, including different ways of handling fixed effects, conditioning on baseline covariates, and reweighting the data to give equal weight to each jobseeker rather than each jobseeker $\times$ vacancy match. We highlight one particularly important robustness check here. Jobseekers can ask to pause or stop receiving matches at any point. Treatment might influence this decisions, which in turn could influence the composition of the sample of matches and potentially create a sample selection problem. In the main analysis reported here, we ignore these pauses and stops: we include in the sample the set of matches these jobseekers would have received during pauses/stops, coding their applications and interviews as zeros. This has an intention-to-treat spirit, in the sense that the sample size and composition depend entirely on treatment assignment and pre-treatment characteristics. In Table B.6, we show that our results are robust to instead dropping matches during pauses/stops.

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:

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

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

β recovers the local average effect on interviews of job applications induced by treatment under four conditions: treatment should be independent of all other factors influencing applications and interviews (independence), should influence interviews only through applications (exclusion), should influence applications (strength), and should 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.[18] We discuss potential complications with the monotonicity assumption and the exclusion restriction and how we address them at the end of this subsection.

Marginal applications submitted due to treatment have roughly the same return, measured in terms of interview invitations, as inframarginal applications. To see this, note that each treatment-induced application increases the probability of an interview invitation by 5.9 percentage points with standard error 0.5 p.p. (Table 3, column 3). This is very similar to the control group's mean interview probability conditional on

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

[18]More formally, Table 3 shows that the Kleibergen and Paap (2006) F-statistic for their test of weak identification is 312.8, substantially higher than conventional thresholds for instrument strength.

applying of 6.3% and we fail to reject equality of these two estimates (p = 0.647).

Marginal and inframarginal applications also have equal returns measured in 'value-weighted' interviews. This finding is important, as the return to an application, and the decision to apply, reflects both the probability of an interview P and the value of an interview V. To explore returns in terms of the overall expected return to an application, $P \cdot V$, we construct a proxy index 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 almost identical: 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.4).

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 with varying education and experience preferences, creating scope for heterogeneous returns. Furthermore, Table B.8 shows that the constant returns finding also holds for the subsamples of 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 to high-value vacancies, making the constant returns finding more surprising. We do not claim that returns are constant over all possible levels of search effort and acknowledge that these are likely to be near zero at sufficiently high levels. But our results show that returns are roughly constant over a large increase in search effort: applications to the first 0.18% of matches have a mean interview probability of 6.3% and the applications to the next 1.32% of matches have a mean interview probability of 5.9%. [19]

Before proceeding, we briefly discuss an extensive battery of robustness checks on the constant returns finding, shown in detail in Appendix B. First, we address the possibility that treatment increases applications from some jobseekers and decreases applications from others, which would violate the monotonicity condition used in our IV analysis. To do this, we derive a bound on the bias from violations of monotonicity in these data, which in turn 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, and is standardized by the platform's protocols; we show that treatment effects on measures of application quality that jobseekers can change (by updating their CVs on

[19] As a very speculative back-of-the-envelope calculation, we can estimate a linear returns curve using the inframarginal and marginal application rates and returns and use this to extrapolate the marginal interview probability at even higher application rates. The estimated curve is relatively flat: for example, if jobseekers applied to 4.625% of matches, the marginal application would have a extrapolated interview probability of 3.26%. This would represent a 25-fold increase in the application rate with a less than one half drop in the mean interview probability, relative to the control group.

file with the platform) 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 these characteristics, and that our key results are unchanged when we estimate them using a 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 slightly different assumptions, which also generates similar 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.

3.3 Role of Jobseeker Selection

Random assignment means that there are no systematic differences between jobseekers in the treatment and control groups. But job application is an endogenous decision, so there may be systematic differences between the treated and control jobseekers who submit applications. This would not bias any of the treatment effects that we estimate, but it would change the interpretation of the estimates. In particular, equal average returns to marginal and inframarginal applications might arise if each individual jobseeker experiences decreasing returns to additional search effort but treated jobseekers who are induced to apply are more positively selected *on gross returns to search*. In this subsection, we show that jobseeker selection into applications does not explain the constant returns finding, using four approaches.

First, and perhaps most importantly, we reduce the scope for jobseeker selection by running a within-jobseeker version of our experiment. This uses a "crossover" design in which we randomly assign some respondents from the control group to receive a phone call in some randomly selected matching rounds, in exactly the same way as the main treatment group. In total, 0.65% of matches in our sample are affected by this treatment, so it has minimal impact on the overall design. However, it allows us to replicate our main estimates on applications and returns with jobseeker fixed effects. This uses only within-jobseeker variation to identify effects, which reduces any role for jobseeker selection. Table B.2 shows that applications submitted due to this treatment have a 8.4% probability of yielding interviews (standard error 3.4 p.p.) and we cannot reject equality of the interview rate for marginal and inframarginal applications (p = 0.503).

Second, we control for jobseeker selection on observed characteristics. We repeat our analysis of the main experiment with controls using a post-double selection LASSO, following Belloni et al. (2014). Table B.1 shows that the point estimates and standard errors are almost identical.

Third, we show that jobseeker selection on observed characteristics does not differ between marginal and inframarginal applications. We do this using a complier or latent type analysis in a similar spirit to Abadie (2003), which we describe in detail in Appendix C.1. This estimates the mean characteristics of marginal applications submitted due to treatment and inframarginal applications submitted without treatment. Comparing these means provides a test for differential selection into applying. Table 4 shows that

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

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

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

mean education and and CV quality scores (provided by firms, as discussed in Section 2.3) are almost identical for the marginal and inframarginal applications. Marginal applications come from jobseekers with on average 1.1 years more work experience and this difference is close to statistically significant. But, as we note above, our main findings are almost identical when we control for these variables.[20]

Fourth, we show that jobseeker selection on latent interview probabilities does not differ between marginal and inframarginal characteristics. We do this by estimating latent interview probability using a data-driven approach and showing that this does not differ between marginal and inframarginal applications. Specifically, we first restrict the sample to the set of applications from control group jobseekers, i.e.

[20]This approach is conceptually different to a heterogeneous treatment effects approach, which compares the magnitude of treatment effects by values of observed characteristics. For interested readers, we also estimate heterogeneous treatment effects on applications by CV quality. The estimated effects are slightly larger for lower-quality CVs, although the differences are not statistically significant (Table B.3). This also suggests that marginal applications do not come from observably stronger jobseekers than inframarginal applications, and hence cannot explain the equal returns to marginal and inframarginal applications.

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 $\hat{P}|X_j = \hat{Pr}(INTERVIEW \mid APPLY = 1, X_j)$ for each jobseeker j. This is the probability the jobseeker will get an interview if she applies, given her observed characteristics.[21] The final row of Table 4 shows that the mean value of this measure does not differ between marginal and inframarginal applications.

Taken together, these results show that the constant return to treatment-induced job search is not explained by treatment changing patterns of jobseeker selection into applications. We next propose a simple framework that can explain the constant return finding.

4 Explaining Marginal Returns to Search

In our main experiment, a small reduction in search costs dramatically increases search, and treatment-induced marginal job search and inframarginal search have roughly the same average return, measured in terms of interviews and quality-weighted interviews. This presents a puzzle: why do jobseekers not apply to more jobs in the absence of treatment, given the non-decreasing returns to additional applications?

In this section, we develop a simple conceptual framework that can explain both the large treatment effect on applications and the constant returns to treatment-induced applications, show that this framework is also consistent with additional patterns in the treatment and control group data, and show why our results are unlikely to be explained by treatment effects on the pecuniary or time costs of applying. The framework is deliberately simple and stylized, as the paper's contribution is empirical, not theoretical. In Section 5, we consider a number of alternative frameworks and demonstrate that they are inconsistent with our results.

4.1 Conceptual Framework

We develop a simple framework in which each jobseeker receives a monthly batch of matches, ranks these vacancies based on the expected present gross return to applying, and applies to all vacancies with expected present gross return above the cost of applying. Because the platform already has very low financial and time cost of applying, we focus on the psychological cost of initiating applications. Our key assumption is that *this cost varies through time* within jobseeker and *can be high enough that some jobseekers apply to no vacancies at some times*. Intuitively, this assumption means that when the phone call treatment lowers the psychological cost of initiating applications, the marginal treatment-induced applications can come from two sources. Some are from jobseekers who would not apply to any vacancies in that month without treatment, so they can have high expected present gross returns. Some are from jobseekers would apply to at least one vacancy in that month without treatment, so they must have lower expected present gross returns than their inframarginal applications. When the population includes both jobseekers who would and would not apply without treatment, then the treatment-induced marginal applications, averaged over these two types of

[21] 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.

jobseekers, can have constant returns.

In this subsection, we write about months when jobseekers do and do not apply. *We could instead write this a between-jobseeker argument.* Our key assumption then becomes that the psychological cost of initiating applications varies across jobseekers, is high enough to deter all applications from some untreated jobseekers, and is lowered by treatment.

More formally, 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 net 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 gross return to applying is $PV\delta\beta$, where the quasi-hyperbolic discounting term $\delta\beta$ with $\beta, \delta \leq 1$ (following Laibson 1997) reflects the fact that interviews occurs 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: $PV\delta\beta - C > 0$, which we rewrite as

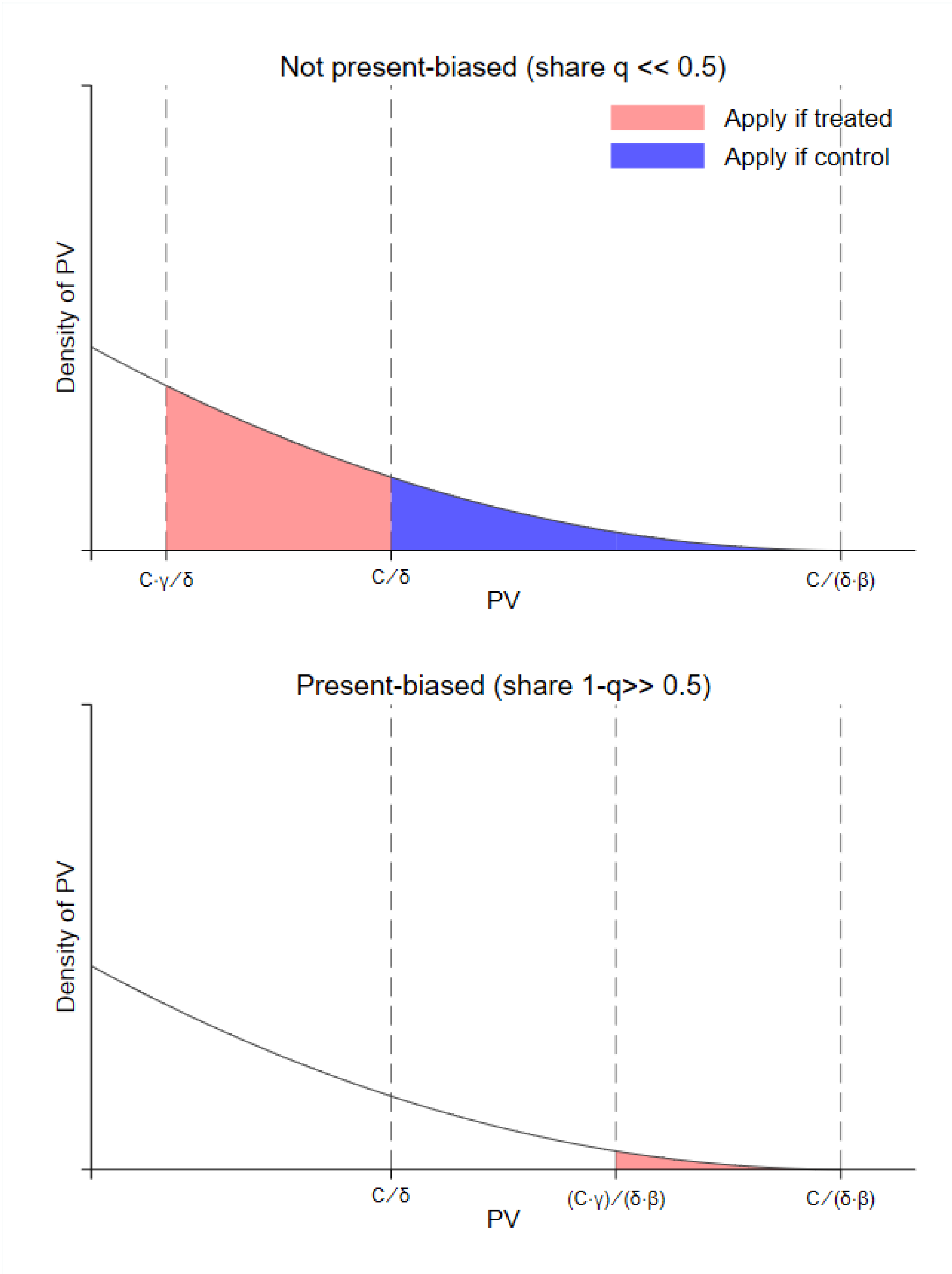
$$PV > \frac{C}{\delta\beta} \tag{4}$$

We can introduce heterogeneous application costs into this framework in multiple ways. We begin by considering heterogeneity in β, motivated by the association between present bias and job search effort documented by DellaVigna and Paserman (2005) and closely following the modeling approach that Duflo et al. (2011) use to study agricultural technology investments. We discuss other approaches at the end of this subsection. We show a simple form of heterogeneit, with only two types of jobseekers. But all predictions of the framework hold with continuously distributed heterogeneity, provided this leads some jobseekers to apply for no vacancies in some periods.

Specifically, we assume 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 $\beta < 1$. We make two further assumptions. First, that $PV > \frac{C}{\delta}$ for some matches, so non-present-biased jobseekers apply to some jobs, matching the empirical pattern that some jobseekers submit applications. Second, we assume that $PV < \frac{C}{\delta\beta}$ for all matches, so present-biased jobseekers apply to no jobs, matching the empirical pattern that many control group jobseekers never apply or apply in only some periods. Figure 1 shows application behavior under these assumptions. In the top panel, non-present-biased jobseekers apply to the blue-shaded section of the density of PV over their matches. In the bottom panel, present-biased 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 psychological cost of initiating applications, reducing C to γC. Treated non-present-biased jobseekers apply if $PV > \frac{\gamma C}{\delta}$. Because $\gamma < 1$, these applications must have lower expected returns than control group non-present-biased jobseekers. These applications go to matches in the red-shaded section in the top panel of Figure 1. Treated present-biased jobseekers apply if $PV > \frac{\gamma C}{\delta\beta}$, shown in the red-shaded section in the bottom panel. If $\gamma > \beta$, i.e. if the treatment-induced drop in application costs

Figure 1: Application Decisions for Treated and Control Jobseekers with Different Time Preferences



Notes: This figure shows the application decisions for non-present-biased jobseekers (top panel) and present-biased jobseekers (bottom panel). The blue-shaded sections show the matches that control group jobseekers apply to under the assumptions stated in the text. The red-shaded sections show the additional matches that treatment group jobseekers apply to under the same assumptions. For simplicity, we show only the right tail of the density of PV.

is small relative to hyperbolic discounting, then these treated present-biased jobseekers' bar for applying is higher than $\frac{C}{\delta}$, the control non-present-biased jobseekers' bar for applying. This shows the core intuition of the model: marginal applications induced by treatment come from a mix of non-present-biased jobseekers, whose applications have returns lower than the inframarginal applications, and present-biased jobseekers, whose applications have returns higher than the inframarginal applications if $\gamma > \beta$. Averaged over these two types of jobseekers, marginal applications can have equal returns to inframarginal applications.[22]

This framework can also explain the large treatment effect on the application rate. The very low application rate in the control group suggests that the share of non-present-biased jobseekers in each month, q, is very low. When q is low, most marginal applications come from present-biased jobseekers, so the treatment effect on the application rate will be large relative to the control group application rate.[23] Low q is consistent with multiple studies finding relatively high rates of present bias, reviewed by Kremer et al. (2019).

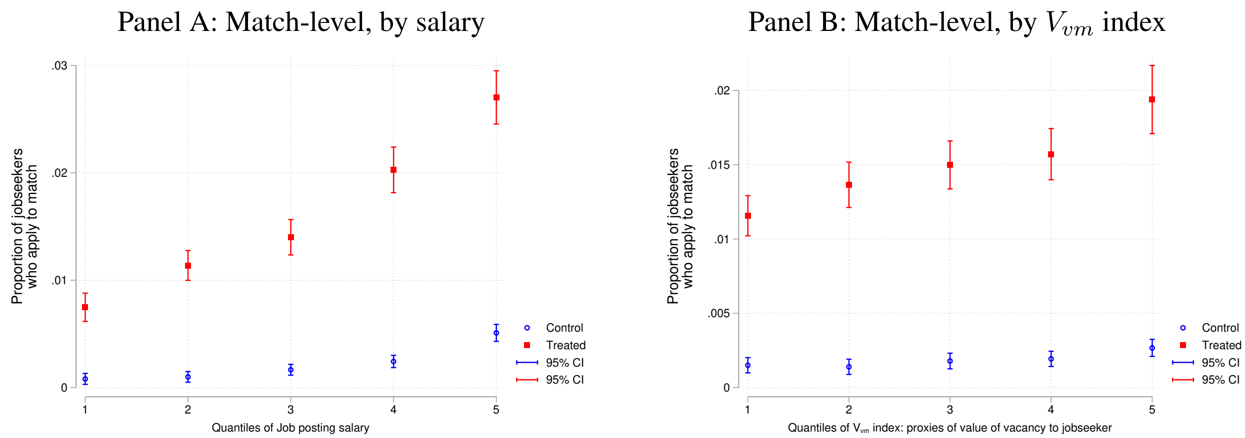
We could derive exactly the same conclusions with an assumption of heterogeneous psychological cost of initiating applications instead of heterogeneous present bias. To do this, we assume all jobseekers have the same $\beta \in (0, 1]$ and that share q face a lower application cost αC with $\alpha \in (0, 1)$. Then the previously non-present-biased agents with application threshold $\frac{C}{\delta} < \frac{C}{\delta\beta}$ become low-cost jobseekers with application threshold $\frac{\alpha C}{\delta\beta} < \frac{C}{\delta\beta}$, with the rest of the argument proceeding unchanged. This approach follows the modeling approach that Carroll et al. (2009) use to study financial investment decisions. Here again the heterogeneous psychological cost of initiating applications could be across jobseekers or within jobseeker through time. And the assumption of two types could be replaced by a continuously distributed α.

We do not test the framework's prediction that the returns to treatment-induced job applications should differ for jobseekers when they have different levels of present bias or different psychological costs of applying. The reason is sample: our data, like many datasets in labor economics, do not provide convincing proxies for these concepts, particularly not convincing proxies that vary through time. These are difficult to measure directly in short surveys. We might try to use proxies like education and work experience, as more educated and experienced jobseekers might face lower cognitive costs of processing the information in job listings and less fear of rejection. However, education and experience also determine the matches that jobseekers receive, which in turn might influence the returns they receive, confounding their interpretation as application cost proxies. Hence, we simply note that the framework can explain the low control

[22]Formally, the mean average return in the control group is $\mathbb{E}\left[PV|PV > \frac{C}{\delta}\right]$, while the average return in the treatment group is a weighted average of $\mathbb{E}\left[PV|PV > \frac{\gamma C}{\delta}\right]$ for non-present-biased jobseekers and $\mathbb{E}\left[PV|PV > \frac{\gamma C}{\delta\beta}\right]$ for present-biased jobseekers. Under our assumption that $\gamma \in (\beta, 1)$, 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 $q \cdot Pr\left(PV > \frac{\gamma C}{\delta}\right)$ and $(1-q) \cdot Pr\left(PV > \frac{\gamma C}{\delta\beta}\right)$ respectively. If the density of PV is strictly continuous, there exists a share q of non-present-biased jobseekers that equalizes the average return to control and treated applications.

[23]Formally, the control group application rate is $q \cdot Pr\left(PV > \frac{C}{\delta}\right)$. The treatment group application rate is $q \cdot Pr\left(PV > \frac{\gamma C}{\delta}\right) + (1-q) \cdot Pr\left(PV > \frac{\gamma C}{\delta\beta}\right)$. 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 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 $\gamma > \beta$. But the second term can still be substantially higher than the control group application for low values of q.

Figure 2: Heterogeneous treatment effects by value of vacancy



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

group application rate, both our key treatment effects, and several additional patterns we discuss in the next subsection.

4.2 Testing Additional Implications of the Framework

This framework can explain our two key findings – more applications with roughly constant returns – and generates three other predictions we can test in the data. First, *control group jobseekers should not apply to some high-value vacancies*, because some share of them are present biased. To test this, Figure 2 panel A shows the control group application rate by quintiles of the vacancy salary in blue. The application rate increases monotonically from the bottom to the top quintile, consistent with the idea that jobseekers value higher salaries. But under half of all applications are sent to top quintile matches, and under 0.1% of matches in the top quintile receive applications. This shows that control group jobseekers miss many high-value matches, consistent with the conceptual framework.

Second, in this framework, with $\gamma > \beta$, *treatment and control group applications should go to vacancies with similar average values*, as treatment should induce applications to a mix of higher- and lower-value applications that have similar average value to the control group. To test this, Figure 2 panel A shows the control and treatment group application rates by quintiles of the vacancy salary in respectively blue and red. Treatment effects increase monotonically from the bottom to the top quintile. But the share of total applications sent to each quintile does not differ between treatment and control groups. To show this, we test if 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). This shows that treatment does not systematically change the distribution of the value of vacancies that receive applications.

Salary is not a perfect proxy for match value. However, both of these patterns also hold for the index we introduced in Section 3.2 that combines all observed proxies of vacancy and match value (Figure 2 panel B) and for other proxies of vacancy value: quintiles of the within-jobseeker between-vacancy salary distribution (Figure C.1, panel A), quintiles of the within-jobseeker between-vacancy index (Figure C.1 panel B), and proxies that take into account the fact that some jobseekers receive multiple matches simultaneously (Figures C.3 - Figure C.4). This pattern also holds for interviews (Figure C.2), although the estimates are noisier.

The other observed value proxies also have relatively similar values across marginal and inframarginal job applications. To show this, we repeat the complier or latent types analysis introduced in Section 3.3. In that section, we examined the characteristics of jobseekers submitting marginal versus inframarginal applications; here we examine the characteristics of firms, vacancies, and jobseeker $\times$ vacancy matches receiving marginal versus inframarginal applications. There are modest differences between the mean values of some observed characteristics for marginal and inframarginal applications but these differences do not show consistently higher values for marginal or for inframarginal applications (Table C.1).

Third, in this framework, *treatment group applications should go to vacancies with more dispersed values*. This occurs because treatment lowers the bar for non-present-biased jobseekers to apply, as the top panel of Figure 1 shows, leading some lower-value matches to get applications. To test this, we estimate treatment effects on the variance, 10th percentile, and 25th percentile of log salary for matches that receive applications. Table C.4 shows that the salary variance is 3.1 log points in the control group and rises by 2.1 log points in the treatment group. Treatment also lowers the 10th percentile from 9.2 to 9 log points and the 25 percentile from 9.6 to 9.4 lot points, consistent with treatment encouraging some applications to lower-value vacancies. The same pattern holds for the proxy index V_{vm} that combines multiple measures of vacancy and match value, although the treatment effects are only statistically significant for log salary, not the proxy index. This shows that, consistent with the model, marginal treatment-induced applications go to vacancies with the same average value as inframarginal applications but more dispersed values.

The results on returns to marginal applications in Section 3.2, the heterogeneous treatment effects by proxies for vacancy value, the complier means analysis, and the treatment effects on the dispersion of application values are all consistent with treatment helping jobseekers overcome costs of initiating applications. We next examine which types of costs deter initiating applications.

4.3 Pecuniary, Time, and Psychological Costs of Job Applications

We run two additional mechanism experiments to help to understand *which* application costs are reduced by treatment, whose results suggest a central role for *reducing the psychological cost of initiating job applications*.[24] First, we evaluate an alternative treatment that lowers the pecuniary cost of applying. As described in Sections 2.3 and 2.5, the platform already provides a low pecuniary cost of job search. The platform

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

identifies potential job matches, presents them to the jobseeker, and then prepares, prints and delivers application materials to the firm. The only possible difference in financial costs is the cost of placing the call to initiate the application process. This costs a maximum of 5 Pakistani rupees (0.03 USD PPP, or less than 1% of a day's earnings at minimum wage) because call center agents always offer to call jobseekers back at the platform's expense to continue the call. In addition, mobile telephone service providers in Pakistan offer small "loan" packages allowing for customers to borrow 10-20 rupees of credit against a future top-up card, and the application period for each matching round stays open for at least a week, so a short-term zero balance is unlikely to be a binding constraint. To further lower pecuniary costs, we randomize a text message reminder that the jobseeker can ask the platform to call them back about a job posting, saving the cost of their calling the platform. Column 1 of Table C.5 compares the effect of our main phone call initiation treatment to the effect of providing this reminder to the control group. The results show that this free callback reminder treatment has an effect one hundredth of the size of the effect of the main phone call treatment, and is statistically significantly different (p = .017).

Second, we evaluate an alternative treatment that lowers the time cost of applying. Time costs are slightly lower for the main treatment group because control group jobseekers wait for their call to be answered and for the call center operator to find their record in the system. But this time is very small: call records show that this takes approximately 4 minutes on average, compared to 10-24 minutes for completing the application process itself over the phone. To further lower time costs, we randomly offer some control group jobseekers the option to text the platform and ask for a callback at a specific time. This eliminates the differential wait time between the main treatment and control groups. Column 2 of Table C.5 shows that this callback request treatment has an effect one quarter of the size of the effect of the main phone call treatment, and is statistically significantly different (p = .002).

The results of these two mechanism experiments suggest that pecuniary and time costs of applications do not explain why the phone call initiation treatment substantially increases applications. We view psychological costs of initiating applications as the most plausible remaining explanation, in line with work showing that eliminating the need for initiating decisions can increase financial and health investments (DellaVigna, 2009; DellaVigna and Malmendier, 2006; Madrian and Shea, 2001; Thaler and Benartzi, 2004).

5 Evaluating Alternative Explanations

5.1 Reminder Effects

Our main findings could be consistent with an alternative behavioral framework based on forgetfulness. In this framework, jobseekers are impatient, sometimes forget to apply, and underestimate their risk of forgetting. This framework predicts that jobseekers will postpone applications until near the deadline, forget to submit some applications, and hence not apply to some high-value vacancies. The distinction between forgetfulness and present bias (one possible source of the heterogeneity in our conceptual framework) matters for designing job search policies and systems. For example, reminders will partly offset forgetfulness but not offset present bias.

The forgetfulness framework is not consistent with the results of three mechanism tests that we run. First, for a subset of matching rounds, we send a second reminder text message to a random subsample of control group jobseekers at the same time that the treatment group jobseekers receive calls. If reminder effects explain our results, this should have a similar effect to that of the phone call treatment. Table C.6 shows the results. The effect of the reminder message is one-fourteenth as large as the effect of the phone call treatment in the same matching rounds and statistically significantly different ($p < 0.001$).

Second, we randomize the order in which we call jobseekers for the phone call treatment, within the application window between the text message job alert and the deadline for job applications. If reminder effects explain our results, we expect that the treatment should have a stronger effect for jobseekers called later within this window, as they will have had more time to forget to apply. Table C.7, Column 1, shows the results. The later the phone call made to the jobseeker, the smaller the treatment effect on applications. This suggests that reminder effects do not explain our results.

Third, we use non-experimental variation in the length of the window between the job alert text message and the deadline. A longer window gives more time for jobseekers to forget to apply. Thus, if reminder effects explain our results, control group application rates should be lower and treatment effects larger when the window is longer. The length of the window is not randomly assigned but it does vary 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 interacted with treatment to address variation over time in these logistical factors. Table C.7, Column 2 shows that treatment has a smaller effect when the window is longer, again suggesting that reminder effects do not explain our results. Instead, the smaller treatment effect on applications when the window is longer is more consistent with an explanation of time-varying psychological costs as outlined in our preferred conceptual framework: if the window is longer, there is a greater chance that at some point during the window a jobseeker will face a low psychological cost of applying even in the absence of treatment, increasing the probability that control group jobseekers apply and hence reducing the magnitude of the treatment effect.

5.2 Encouragement or Pressure to Apply

If call center agents encouraged or pressured respondents to apply on the call, this could lead to more search, although not necessarily to constant returns to the additional search. This could be viewed as a reduction in the cost of applying, with encouragement lowering the cost of applying or pressure imposing a price on not applying. However, the distinction between the general case of psychological application costs and the specific case of encouragement or pressure matters for designing job search policies and systems. For example, job search assistance caseworkers might provide encouragement or pressure to apply, while simplifying the application initiation process on search platforms facilitates applications in a different way.

The encouragement/pressure explanation is not consistent with two features of our platform design and results. First, platform staff are trained not to encourage or pressure jobseekers to apply and their payment and performance evaluations provide no incentive for them to offer encouragement or apply pressure. Regular audits of at least one call per call cycle per call center agent are used to help ensure that agents follow the script provided and do not pressure jobseekers. The second feature uses the fact that most jobseekers receive 'batches' of multiple matches at the same time. The lowest-cost way for pressured jobseekers to alleviate pressure is to apply to the first job listed within their batches. Thus, if pressure were responsible for the main treatment effect on applications, we would expect treated jobseekers to apply to the first job listed on the call at a higher rate than control jobseekers. Instead, the share of applications sent to the first job listed on the call is the same for the treatment and control groups (Figure C.6). We would also expect to see treatment group jobseekers applying to systematically lower-value or worse-matched vacancies than control group jobseekers. Instead, we showed in Section 4 that the two groups do not differ on these dimensions.

5.3 Differential Access to Information

Our main findings could be consistent with an alternative framework in which the phone call treatment *provides more information about specific jobs* to jobseekers, which both increases their job application rate and allows them to direct applications to vacancies that yield constant returns. The text message and phone call script contain identical information about each vacancy, to reduce the scope for differences in information. However, once the call has started, jobseekers might also ask the call center agent for more information about the position, which is available to call center agents in some matching rounds. Since treatment jobseekers are more likely to participate in phone calls, might treated jobseekers receive more information about jobs than control jobseekers?.

This information provision framework is not consistent with three additional results. First, our key findings hold when restricting the sample to the roughly 80% of matching rounds in which no additional information beyond the contents of the text message was available to the call center agents. Second, we record when jobseekers ask for more information and use this to identify the subset of calls without requests for more information. Our key findings hold when restricting to this sample, although we interpret the result with caution because this exercise requires restricting the sample based on a post-treatment variable. Third, we conduct a mechanism experiment in a subset of matching rounds in which randomly chosen subsets of treatment and control group jobseekers' text message alerts include information about benefits offered by specific jobs, in addition to the salary and location information included in all text messages.[25] If the phone call treatment influences application decisions by providing additional vacancy-specific information, then directly providing additional information might have the same effect. Instead, providing information about benefits has a near-zero impact on application rates in both the phone call treatment and control groups (Table C.8). This suggests that information about specific vacancies does not drive our main findings.

Our main findings might also arise if *jobseekers are more likely to receive phone calls than text messages*.

[25] Jobs listed on the platform provide a range of benefits: meals (51% of matches), health benefits (26%), travel allowance (15%), parental leave (14%), career development training (30%), and bonuses/commission (25%). We acknowledge that this intervention does not capture all vacancy-level benefits.

For example, the phone call could deliver the information in case the text message was blocked as a bulk delivery, or simply ensure the jobseeker picks up the phone and sees the message. This would explain both the phone call treatment's positive effect on application rates and, if the probability of missing a text message is uncorrelated with the value of the vacancy, the constant returns to additional applications.

This information receipt framework is not consistent with two additional results. First, we directly measure the probabilities that jobseekers receive text messages and phone calls. We phone jobseekers who recently received at least one match from the platform and ask them two questions: if they received a match by text message and if they received a match by phone. Treated jobseekers are more likely to report receiving a match by phone call (Table C.9, column 3), providing a validation of the survey measures. However, treated and control jobseekers are equally likely to report receiving a match at all, including text messages and phone calls, suggesting that differential information receipt does not account for our main findings (column 5).[26] The survey response rate differs between treatment and control groups, which might create a sample selection problem. To address this, we randomize some features of the survey data collection, e.g., number of call attempts, and use this to create instruments for a sample selection correction term, following (DiNardo et al., 2021). Treatment effects on information receipt are very similar with and without selection correction (Table C.9, even versus odd columns). We describe the selection correction method and how the randomized survey features influence response rates in detail in Appendix C.2.

Second, the effects of the phone call initiation treatment do not vary by two baseline proxies for the probability of receiving text messages from the platform. Treatment is not more effective for respondents who share a phone number with other members of the household; if information delivery failure in the control group explained our results, we would expect a larger treatment effect on applications among this group. In addition, 93% of respondents indicated at registration that they are comfortable communicating with the platform by text message, and treatment is no more effective for this group than for the 7% of jobseekers who indicated they might face difficulties with receiving or reading the job alerts by text message.

5.4 Changes in Perceived Returns To Search

The phone call treatment might increase job applications by raising jobseekers' beliefs about the value of applying for jobs on the platform. For example, a call from a recruiting service with professional call center agents might signal to jobseekers that the firms recruiting on platform are larger or wealthier and able to offer more benefits or greater opportunities for advancement, suggesting higher V. Alternatively, it might signal to the jobseeker that the firm sees her as a good fit for the job, suggesting higher P. The distinction between psychological costs of initiating applications and belief updating matters for designing job search policies and systems. For example, platforms can make different communication choices to influence jobseekers'

[26] We expect some measurement error in these survey responses. Some jobseekers who received matches before the recall period will incorrectly report receiving matches during the recall period, a pattern called 'telescoping' in the survey methods literature. Some jobseekers who received matches during the recall period will incorrectly report not receiving matches due to inattention or forgetting. It is possible that measurement error could lead us to find no difference in self-reported match receipt when in fact treated jobseekers were more likely to receive matches. However, this is only possible if jobseekers are more likely to forget phone calls than text messages, which we view as unlikely.

beliefs about returns to applications versus simplify starting applications.

We first note that higher perceived P and V on the platform can increase the application rate but, by itself, is unlikely to generate constant returns to marginal search. Consider a simple decision rule, ignoring dynamics, in which jobseekers apply to all vacancies where $\tilde{P}{\cdot}\tilde{V} > C$, for perceived interview probability $\tilde{P}$ and perceived interview value $\tilde{V}$. If the phone call treatment raises $\tilde{P}$ or $\tilde{V}$ for all vacancies, then jobseekers will apply to vacancies with lower $P{\cdot}V$, leading to decreasing returns to additional search effort. To generate constant returns to additional applications, the phone call treatment would need to raise $\tilde{P}$ or $\tilde{V}$ specifically for vacancies that have high P or V (so they generate high returns) and low $\tilde{P}$ or $\tilde{V}$ (so they would not receive applications without treatment). This pattern is unlikely, given the results in Section 4 that treatment does not substantially change how jobseekers direct applications.

We also directly test if the phone call treatment shifts beliefs about P and V by surveying jobseekers. We ask: "Suppose Job Talash sends you one hundred job ads over a year. Based on your past experience with our job matching service, how many of these ads do you think would be desirable for you?" and "Suppose you apply for all the jobs you think are desirable jobs. How many of those do you think would make you an offer?"[27] We use a jobseeker-level version of equation (1) to estimate treatment effects on these two belief measures. Table C.10 shows that both results are close to zero. Jobseekers in the control group on average think that they will receive an offer from 43% of jobs they are interested in; the phone call treatment increases this by 1 percentage point (standard error 1.8. p.p.). Jobseekers in the control group on average think that 32% of the vacancies on the platform would be desirable for them; the phone call treatment decreases this by 0.5 p.p. (standard error 1.6 p.p.). Results are similar when we adjust for survey non-response using the same method introduced in Section 5.3 and described in detail in Appendix C.2 (Table C.10, odd versus even columns). The survey data indicate that the treatment does not increase respondents' perceptions of the average values of V or P on the platform, and hence cannot explain the large treatment effects on applications. It is possible that treatment induces shifts in beliefs we do not measure, such as about the relative ranking of vacancies by P or V. However, these changes would have to represent a large change in beliefs about average P and/or V to explain our treatment effects on applications.

In Appendix C, we show additional evidence that shifts in beliefs are not a mechanism for our treatment effects, using heterogeneous exposure to job characteristics in the baseline period, and jobseeker responses during the "crossover" matching rounds, in which a randomly selected sample of control group jobseekers receive phone calls. All the results further confirm that our main findings are not driven by a beliefs channel.

5.5 Random Search

If jobseekers apply to vacancies at random and the phone call treatment reduces the cost of applying, then treatment should increase the application rate and yield constant returns to marginal applications. Random job search may seem unintuitive. However, random search has been widely assumed in canonical search

[27] We measure beliefs about the probability of receiving an offer for the subset of jobs to which the respondent would consider applying, because shifting beliefs about P for jobs the jobseeker would not consider should not influence their application decisions

models, even if only as a simplifying benchmark (Pissarides, 2000). It may also be plausible given some empirical evidence that jobseekers have limited information about labor market conditions or their match quality with individual vacancies (Alfonso Naya et al., 2020; Behaghel et al., 2020; Belot et al., 2018).

However, the random search framework does not match three additional results. First, we showed in Sections 2.4 and 4 that applications are targeted towards vacancies with higher observable proxies for V. So applications are clearly not random. This is consistent with recent empirical evidence that jobseekers do direct applications based on some dimensions of vacancies, despite information frictions (Belot et al., 2022; He et al., 2021; Garlick et al., 2022; Marinescu and Wolthoff, 2020).

Second, we run an additional experiment designed to induce random search. This allows us to identify the marginal return to random search effort and compare that to the marginal return to search effort induced by the phone call treatment. To do this, we randomize the order in which vacancies are listed in all communications with jobseekers – text messages and phone calls – for some matching rounds. The goal is to encourage additional applications to the vacancies listed early and hence additional applications to randomly-chosen vacancies.[28] We analyze this experiment using exactly the same methods from Section 3 used to analyze the main experiment.[29]

Table 5 shows the results of this experiment. The match-level probability of application is 0.06 percentage points for vacancies listed second or later and 0.5 p.p. for vacancies listed first (column 1). The average interview probability is 6.3% for inframarginal applications and 2.5% for marginal applications submitted because the vacancy was listed first (column 4). The interview probability for marginal applications is both substantially and statistically significantly smaller than for inframarginal applications (p = 0.077), showing that random search produces decreasing, not constant, returns. Returns are also decreasing for quality-weighted interviews, although the difference between returns to marginal and inframarginal applications is not statistically significant (p = 0.109, column 5).

The marginal return to random applications (2.5% interview probability) is also substantially smaller than the marginal return to applications induced by the phone call treatment (5.9%). This suggests that the phone call treatment is not inducing random search. This is consistent with the fact that 69% of applications induced by the phone call treatment are sent to vacancies listed second or later in communication with jobseekers, showing that jobseekers are willing to apply to the slightly higher-cost later-listed vacancies to direct their applications non-randomly.

The result of this experiment emphasizes that "the" return to marginal search depends on which intervention causes the marginal search and how it is directed. The randomized order treatment causes marginal

[28] Vacancies listed earlier might attract more applications because applying to those takes less time or because jobseekers interpret the ordering as a signal of job quality or attainability. We do not attempt to separate these explanations.

[29] Order of job listings was randomised for 20% of the study period and determined by firm identifier for the rest of the study period. We include fixed effects for aggregated firm identifiers in all analysis; order of job listing is uncorrelated with job and jobseeker baseline characteristics conditional on these fixed effects. Results are similar but less precise when we use only the randomised period. Results are similar when we compare only the first job to all subsequent jobs or include indicators for first, second, etc. job. 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. This drops only 16% of matches from the sample.

Table 5: Treatment Effects of Lowering Cost of Applying to Randomly Chosen Vacancies

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

Notes: This table shows the effect of varying the relative marginal cost of applying to an individual vacancy within a round, by changing the order in which vacancies are listed on the application phone call. Column 1 shows the coefficient from regressing an indicator for job application on an indicator equal to 1 for a vacancy that is listed first in the call to the jobseeker during the round and 0 otherwise. Column 2 shows the coefficient from regressing an indicator for interview invitation on an indicator for vacancy listed first in the call. Column 3 shows the coefficient from regressing an indicator for interview invitation weighted by a proxy index for the value of the vacancy to the jobseeker, V_{vm}, on an indicator for vacancy listed first in the call. Column 4 shows the coefficient from regressing an indicator for interview invitation on job application, instrumented by vacancy listed first on the call. Column 5 shows the coefficient from regressing an indicator for interview invitation weighted by the proxy index for V_{vm} on job application, and instrumented by vacancy listed first on the call. See the note below Table 3 for a definition of V_{vm}. The p-value is for a test of equality between the IV treatment effect and the mean interview rate for control group applications. The first-stage F-statistic and p-value are for the test of weak identification from Kleibergen and 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.

search that is roughly randomly directed and has sharply decreasing returns. The phone call initiation treatment causes marginal search that is directed in similar ways to inframarginal search and has roughly equal returns. This highlights that the constant returns finding is a consequence of type of search, not inherent to this labor market or these jobseekers.

Third, we note that there is no case in which undirected search can lead to all our key results: that returns are constant in terms of both number of interviews and quality-weighted interviews (Table 3), and that jobseekers direct applications towards vacancies with high observable V (Table 2 and Figure 2). We lay out the details of each possible case in Table C.13.

6 Spillover Effects

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

We can identify spillover effects using variation in the treated share of respondents who are matched to each vacancy. 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; Crepon et al. 2013; Ferracci et al. 2014; Gautier et al. 2018; LaLive et al. 2022). This approach is only feasible because this platform's matching structure fully determines the set of respondents 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, making it difficult to define how much exposure each user has to other treated users without a full model of the job search process.

We first verify that the experiment generates sufficient variation across vacancies in the treatment rate to identify spillovers. The vacancy-level percentage of matches that are treated has interdecile range across job adverts 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 vacancies have standard deviation 0.054 in their treatment rates.

We estimate spillover effects using two methods. Our first method tests if 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 jobseeker-level interview invitations on the jobseeker's own treatment status, the treatment rate defined above and their interaction:

$$\text{Interview}_{jv} = T_j \cdot \beta_1 + TR_{jv} \cdot \beta_2 + T_j \cdot TR_{jv} \cdot \beta_3 + \mathbf{X}_v \cdot \Lambda + \mu_b + \epsilon_{jv}, \tag{5}$$

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 when more competing jobseekers are treated. $\beta_2 + \beta_3 < 0$ 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 (Table 6). Estimates of β_2 and β_3 are both small and not statistically significant (column 1). To interpret their magnitude, we consider the effect on a jobseeker's interview probability of moving from the 25th to 75th percentile of TR_{jv}, the treatment exposure rate. This effect is 0.006 percentage points for a control group jobseeker (standard error 0.011 p.p., p = 0.589) and -0.011 p.p. for a treatment group jobseeker (standard error 0.017 p.p., p = 0.511). As a benchmark, the effect of a jobseeker's own treatment status on interview invitations is substantially larger: 0.078 p.p. (from Table 3).

Equation (5) imposes a linear relationship. But negative spillover effects might be nonlinear and only substantial at high treatment rates. To test this idea, we repeat the analysis replacing the linear vacancy-level treatment rate TR_{jv} with indicators for the middle and top terciles of the treatment rate. These effects are again very small and not statistically significant for control or treatment group jobseekers (column 2).

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

$$\Delta\text{Interview}_v = TR_v \cdot \alpha + \mathbf{X}_v \cdot \Lambda + \varepsilon_v, \tag{6}$$

conditional on the same vacancy-level covariates $\mathbf{X}_v$ as the previous analysis. A negative value of α would be evidence of negative spillover effects, as this would show a smaller treatment effect on each individual jobseeker's interview probability at vacancies receiving more treatment-induced applications.

We do not find evidence of negative spillover effects using this second method (Table 6). Instead, we find a positive coefficient, although it is small (column 3). To interpret the magnitude, we note that this coefficient implies that a vacancy exposed to the 75th percentile of the treatment rate TR_v rather than the 25th percentile would have a 0.018 percentage point higher treatment effect on interviews (standard error 0.011 p.p., p = 0.096). To test for a nonlinear relationship, we repeat this analysis replacing the vacancy-level treatment rate with indicators for the middle and top terciles of the treatment rate. These coefficients are again positive (column 4). Vacancies with top-tercile rather than bottom-tercile treatment rates have

Table 6: 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.

0.05 p.p. higher treatment effects on interviews, although we cannot reject the null hypothesis that treatment effects are equal across all three terciles. A nonparametric regression of vacancy-level treatment effects on treatment rates also shows no evidence of negative spillover effects (Figure D.2).

The lack of negative spillovers is consistent with descriptive patterns in vacancy-level outcomes. If firms dislike congestion, then the relationship between application and interview numbers might be non-monotonic: a small increase in the number of applications might lead to more interview invitations and a large increase might lead to fewer interview invitations. At the extreme, a high number of applications might lead firms to ignore all applications and make no interview invitations. Instead, Table D.1 shows that vacancies get more applications and make more interview invitations if they are matched to more jobseekers and if more of these jobseekers are treated. Both the number of interviews and the probability of interviewing any jobseeker are monotonically increasing in terciles of the number of applications.

What might explain the negligible spillover effects we find? Our design cannot directly answer this question but we suggest three possible explanations. First we note that spillovers are not theoretically necessary, even if they are probable. In theory, firms may hire more when they face lower screening costs or receive higher-quality applications. Algan et al. (2020) and Le Barbanchon et al. (2023) show that reducing firms' hiring costs can increase vacancy-posting and hiring, while Carranza et al. (2021) and Fernando et al. (2021) show indirect evidence consistent with this mechanism. Second, firms in this context report filling only 60% of vacancies. Third, application volumes on this platform are relatively low: the average vacancy receives only 0.8 applications from control group applicants and another 6 applications from treated applicants (with pooled interdecile range 0-18). Firms report in surveys that they get on average 30% of their total applications through the platform. Taking these factors together, it is possible that firms in this labor market receive too few suitable applications in the absence of treatment for crowd-out to be relevant, at least at the interview stage that we observe.

7 Off-Platform Search and Employment

Treatment may shift off-platform search behavior and might increase job offers or employment, but our administrative platform data do not measure these outcomes. Hence we survey respondents on average 47 months after treatment about their off-platform search and employment. To the best of our knowledge, concurrent work by Ben Dhia et al. (2022) is the only other job search study that combines detailed data on on-platform search, off-platform search, and off-platform employment. Most other papers studying these job search platforms have rich administrative data, but only covering on-platform search and outcomes (e.g. Alfonso Naya et al. 2020; Behaghel et al. 2020; Belot et al. 2018). A related set of papers primarily use survey data on (prospective) platform users that covers off-platform search and employment and merge this with limited data on platform use (e.g. Jones and Sen 2022; Wheeler et al. 2022). Relative to the former literature, we add survey measures of off-platform search and employment; relative to the latter literature, we add detailed data on platform use, including match-level information about vacancy characteristics, applications, and interviews.[30]

The effect of the phone call treatment on off-platform search is theoretically ambiguous. A simple substitution effect could occur, resulting in lower off-platform search to compensate for the increased on-platform search. Alternatively, treatment might increase off-platform search effort if it shifts beliefs about the value of jobs available in this labor market or about off-platform labor market prospects. For example, the treatment could lead a respondent to believe that her skill set is in high demand overall, increasing her belief about P for off-platform vacancies and increasing off-platform search.

To evaluate these ideas, we survey jobseekers about their off-platform search. The phone call treatment has a near-zero effect on the extensive margin (Table 7, columns 1-2). 27% of control group jobseekers do any off-platform search in the past month or two weeks (recall period randomized); treatment decreases this by only 0.8 percentage points (standard error 1.6 p.p.). Treatment also has little impact on two measures of intensive margin search: the number of off-platform applications submitted and the number of off-platform search methods used (Table D.2). Both intensive-margin questions use the same recall period as the extensive-margin questions and are asked for random 50% subsamples of the jobseekers. All treatment effects on off-platform search are similar when we adjust for survey non-response using the same method introduced in Section 5.3 and described in detail in Appendix C.2.

Table 7 shows that treatment effects on employment are positive but very small and not close to statistically significant at conventional levels: 0.18 percentage points (standard error 1.6 p.p.) or 1.1 p.p. with the adjustment for survey non-response (standard error 2 p.p.). However, we are underpowered to detect treatment effects on employment at the scale of this experiment, like some other studies of platform-based job search (e.g. Belot et al. 2018).[31] Our minimum detectable effect on employment is roughly 5.7 p.p., assuming 80% test power and 5% test size. This would be an implausibly large effect of our treatment on employment, as the phone call treatment increases the share of matches resulting in interview invitations by 0.08 p.p and the share of jobseekers receiving any interview invitations by roughly 5 p.p. However, treatments like this may still lead to substantial increases in the number of employed users for larger platforms such as Rozee and LinkedIn, with respectively 9.5 and 7.5 million users in Pakistan alone. We also emphasize that interviews represent an important search outcome because they are a necessary condition for job offers and impose non-trivial costs on both job applicants and firms.

8 Conclusion

We show that job search effort can be substantially increased by reducing the psychological cost of initiating job applications. Returns to the additional search effort are constant rather than decreasing, in contrast with

[30]Online gig work platforms such as oDesk generate detailed data on both search and employment because employment occurs on the platform. The nature of this employment is, on average, quite different to the types of jobs posted on job search and matching platforms such as Job Talash.

[31]As Belot et al. (2018) note, this approach has a similar spirit to audit studies, which focus on interview invitations as an important early outcome in the hiring process. Lanning (2013) shows how using treatment effects on interviews in audit studies to forecast treatment effects on employment and earnings is sensitive to modeling assumptions.

Table 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. 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 C.2 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. Heteroskedasticity-robust standard errors shown in parentheses.

many intuitive job search models. This pattern is consistent with a model in which marginal applications are a mix of lower-return applications from jobseekers who would send some applications without treatment and higher-return applications from jobseekers who would not apply without treatment, at least at that period in time. This finding of constant returns, combined with limited spillovers on other jobseekers, suggests the possibility of suboptimally low search effort. This echoes findings that changing default options to avoid initiation costs can lead to economically significant increases in financial and health investments (DellaVigna, 2009; DellaVigna and Malmendier, 2006; Madrian and Shea, 2001; Thaler and Benartzi, 2004). Our findings are particularly striking because this is a platform directly 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 initiation 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 might help to motivate search encouragement 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 jobseekers' process of evaluating job listings or encouraging them to start applications, although the value of such design changes may be low on platforms with already-high application volumes (Horton and Vasserman, 2021). If, in particular, job search effort is sensitive to present-biased trade-offs between search costs and benefits, then future research might examine the impact of replacing monthly or weekly unemployment benefit payments conditional on job search with payments made immediately after applications. This relates to research into the optimal timing of unemployment insurance payments, which has not to date emphasized behavioral considerations (Bolhaar et al., 2019; Kolsrud et al., 2018). The exact design of these policies raises questions for future research; for example, results from O'Donoghue and Rabin (2008) suggest that the relative difficulty of initiating and finishing applications might be important for present-biased jobseekers.

References

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

ABADIE, A., S. ATHEY, G. IMBENS, AND J. WOOLDRIDGE (2017): "When Should You Adjust Standard Errors for Clustering?" Working Paper 24003, National Bureau of Economic Research.

ABEBE, G., S. CARIA, M. FAFCHAMPS, P. FALCO, S. FRANKLIN, AND S. QUINN (2021a): "Anonymity or Distance? Job Search and Labour Market Exclusion in a Growing African City," *Review of Economic Studies*, 88.

ABEBE, G., S. CARIA, M. FAFCHAMPS, P. FALCO, S. FRANKLIN, S. QUINN, AND F. SHILPI (2021b): "Matching Frictions and Distorted Beliefs: Evidence from a Job Fair Experiment," Working paper, University of Oxford.

ABEBE, G., S. CARIA, AND E. OSPINA-ORTIZ (2019): "The Selection of Talent: Experimental and Structural Evidence from Ethiopia," *American Economic Review*, 111.

ABEL, M., R. BURGER, E. CARRANZA, AND 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, 284–301.

ABEL, M., R. BURGER, AND P. PIRAINO (2020): "The Value of Reference Letters: Experimental Evidence from South Africa," *American Economic Journal: Applied Economics*, 12, 40–71.

AFRIDI, FARZANA AD DHILLON, A., S. ROY, AND N. SANGWAN (2022): "Social Networks, Gender Norms and Women's Labor Supply: Experimental Evidence using a Job Search Platform," Working paper, Indian Statistical Institute.

ALFONSO NAYA, V., G. BIED, P. CAILLOU, B. CREPON, C. GAILLAC, E. PERENNES, AND M. SEBAG (2020): "Designing labor market recommender systems: the importance of job seeker preferences and competition," Manuscript, LISN.

ALGAN, Y., B. CREPON, AND D. GLOVER (2020): "Are Active Labor Market Policies Directed at Firms Effective? Evidence from a Randomized Evaluation with Local Employment Agencies," Manuscript, Sciences Po.

ALTMANN, S., A. FALK, S. JÄGER, AND F. ZIMMERMANN (2018): "Learning about Job Search: A Field Experiment with Job Seekers in Germany," *Journal of Public Economics*, 164, 33–49.

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, 1481–1495.

ARNI, P. AND A. SCHIPROWSKI (2019): "Job search requirements, effort provision and labor market outcomes," *Journal of Public Economics*, 169, 65–88.

ATTANASIO, O., A. KUGLER, AND C. MEGHIR (2011): "Subsidizing Vocational Training for Disadvantaged Youth in Colombia: Evidence from a Randomized Trial," *American Economic Journal: Applied Economics*, 3, 188–220.

BABCOCK, L., W. J. CONGDON, L. F. KATZ, AND S. MULLAINATHAN (2012): "Notes on behavioral economics and labor market policy," *IZA Journal of Labor Policy*, 1, 2.

BAKER, S. R. AND 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, AND A. VITALI (2021): "The Search for Good Jobs: Evidence from a Six-year Field Experiment in Uganda," *Manuscript, London School of Economics*.

BANERJEE, A. AND S. SEQUEIRA (2020): "Spatial Mismatches and Imperfect Information in the Job Search," Discussion Paper 14414, Centre for Economic Policy Research.

BANFI, S., S. CHOI, AND B. VILLENA-ROLDAN (2019): "Deconstructing Job Search Behavior," *Social Science Research Network Working Paper*.

BASSI, V. AND 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, AND Y. HAZARD (2020): "Directing job search: a large scale experiment," Manuscript, Paris School of Economics.

BELLONI, A., V. CHERNOZHUKOV, AND C. HANSEN (2014): "Inference on treatment effects after selection among high-dimensional controls," *The Review of Economic Studies*, 81, 608–650.

BELOT, M., P. KIRCHER, AND P. MULLER (2018): "Providing Advice to Jobseekers at Low Cost: An Experimental Study on Online Advice," *Review of Economic Studies*, 86, 1411–1447.

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

BEN DHIA, A., B. CREPON, E. MBIH, L. PAUL-DELVAUX, B. PICARD, AND V. PONS (2022): "Can a Website Bring Unemployment Down? Experimental Evidence from France," *SSRN Electronic Journal*.

BERNHEIM, D., D. RAY, AND S. YELTEKIN (2015): "Povery and Self-Control," *Econometrica*, 83, 1877–1911.

BHULLER, M., A. KOSTØL, AND T. VIGTEL (2019): "How Broadband Internet Affects Labour Market Matching," .

BLUNDELL, R., M. C. DIAS, C. MEGHIR, AND J. VAN REENEN (2004): "Evaluating the Employment Impact of a Mandatory Job Search Program," *Journal of the European Economic Association*, 2, 569–606.

BOLHAAR, J., N. KETEL, AND B. VAN DER KLAAUW (2019): "Job Search Periods for Welfare Applicants: Evidence from a Randomized Experiment," *American Economic Journal: Applied Economics*, 11, 92–125.

BOLHAAR, J., N. KETEL, AND 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, AND T. MCCORMICK (2022): "Migrants, Experience, and Working Conditions in Bangladeshi Garment Factories," .

BROWN, M., C. J. FLINN, AND A. SCHOTTER (2011): "Real-Time Search in the Laboratory and the Market," *American Economic Review*, 101, 948–74.

BÉNABOU, R., A. JAROSZEWICZ, AND G. LOEWENSTEIN (2022): "It Hurts To Ask," Working Paper 30486, National Bureau of Economic Research.

CALIENDO, M., D. COBB-CLARK, AND A. UHLENDORFF (2015): "Locus of Control and Job Search Strategies," *The Reivew of Economics and Statistics*, 97, 88–103.

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

CARD, D., J. KLUVE, AND 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, 894–931.

CARRANZA, E., R. GARLICK, K. ORKIN, AND N. RANKIN (2021): "Job Search and Hiring with Limited Information about Workseekers' Skills," Manuscript, Duke University.

CARROLL, G., J. CHOI, D. LAIBSON, B. MADRIAN, AND A. METRICK (2009): "Optimal Defaults and Active Decisions," *Quarterly Journal of Economics*, 124, 1639–1674.

CHIPLUNKAR, G. AND P. K. GOLDBERG (2022): "The Employment Effects of Mobile Internet in Developing Countries," Working Paper 30741, National Bureau of Economic Research.

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

CREPON, B., E. DUFLO, M. GURGAND, R. RATHELOT, AND P. ZAMORA (2013): "Do Labor Market Policies Have Displacement Effects? Evidence from a Clustered Randomized Experiment," *Quarterly Journal of Economics*, 128, 531–580.

DAMMERT, A., J. GALDO, AND 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.

DEAN, M. AND A. SAUTMANN (2021): "Credit Constraints and the Measurement of Time Preferences," *The Review of Economics and Statistics*, 103, 119–135.

DELLAVIGNA, S. (2009): "Psychology and Economics: Evidence from the Field," *Journal of Economic Literature*, 47, 315–72.

DELLAVIGNA, S., J. HEINING, J. SCHMIEDER, AND 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, AND J. F. SCHMIEDER (2017): "Reference-Dependent Job Search: Evidence from Hungary," *The Quarterly Journal of Economics*, 132, 1969–2018.

DELLAVIGNA, S. AND U. MALMENDIER (2006): "Paying Not to Go to the Gym," *American Economic Review*, 96, 694–719.

DELLAVIGNA, S. AND M. D. PASERMAN (2005): "Job Search and Impatience," *Journal of Labor Economics*, 23, 527–588.

DINARDO, J., J. MATSUDAIRA, J. MCCRARY, AND 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, AND J. ROBINSON (2011): "Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya," *The American Economic Review*, 101, 2350–2390.

FALK, A., D. HUFFMAN, AND U. SUNDE (2006a): "Do I Have What It Takes? Equilibrium Search with Type Uncertainty and Non-Participation," Discussion paper 2531, Institute for the Study of Labor.

——— (2006b): "Self-Confidence and Search," IZA Discussion Paper 2525.

FERNANDO, N., N. SINGH, AND G. TOUREK (2021): "Hiring Frictions in Urban Labor Markets: Experimental Evidence from India," Working paper, Notre Dame.

FERRACCI, M., G. JOLIVET, AND G. J. VAN DEN BERG (2014): "Evidence of Treatment Spillovers Within Markets," *Review of Economics and Statistics*, 96, 812–823.

FIELD, E. AND K. VYBORNY (2022): "Women's Mobility and Labor Supply: Experimental Evidence from Pakistan," Report, Asian Development Bank.

FILGES, T., G. SMEDSLUND, A. D. KNUDSEN, AND 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, 2353–2379.

FU, J., M. SEFTON, AND R. UPWARD (2019): "Social Comparisons in Job Search," *Journal of Economic Behavior & Organization*, 168, 338–361.

GARLICK, R., L. HENSEL, A. KISS, AND K. ORKIN (2022): "Skills, Beliefs, and (Mis)Directed Job Search," Working paper, Duke University.

GARLICK, R. AND J. HYMAN (2022): "Quasi-Experimental Evaluation of Alternative Sample Selection Corrections," *Journal of Business and Economic Statistics*, 40, 101–125.

GAUTIER, P., P. MULLER, B. VAN DER KLAAUW, M. ROSHOLM, AND M. SVARER (2018): "Estimating Equilibrium Effects of Job Search Assistance," *Journal of Labor Economics*, 36, 1073–1125.

GEE. (2019): "The More You Know: Information Effects on Job Application Rates in a Large Field Experiment," *Management Science*, 65, 2077–2094.

GOVERNMENT OF BANGLADESH (2015): "Bangladesh Labor Force Survey 2015," Tech. rep.

GOVERNMENT OF NAMIBIA (2016): "Namibia Labor Force Survey 2016," Tech. rep.

GOVERNMENT OF SOUTH AFRICA (2018): "South Africa Quarterly Labour Force Survey 2013-2018," Tech. rep.

GURTZGEN, N., L. BENJAMIN, L. POHLAN, AND G. VAN DEN BERG (2020): "Does Online Search Improve the Match Quality of New Hires?" .

HE, H., D. NEUMARK, AND Q. WENG (2021): ""I Still Haven't Found What I'm Looking For": Evidence of Directed Search from a Field Experiment," Working Paper 28660, National Bureau of Economic Research.

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

HECKMAN, J., R. LALONDE, AND J. SMITH (1999): "The Economics and Econometrics of Active Labor Market Programs," in *Handbook of Labor Economics*, Elsevier, vol. 3, 1865–2097.

HJORT, J. AND J. POULSEN (2019): "The Arrival of Fast Internet and Employment in Africa," *American Economic Review*, 109, 1032–1079.

HORTON, J. (2017): "The Effects of Algorithmic Labor Market Recommendations: Evidence from a Field Experiment," *Journal of Labor Economics*, 35, 345–285.

HORTON, J. AND S. VASSERMAN (2021): "Job-Seekers Send Too Many Applications: Experimental Evidence and a Partial Solution," Manuscript, MIT.

IMBENS, G. AND D. RUBIN (2015): *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*, Cambridge University Press.

JOHNSTON, A. C. AND A. MAS (2018): "Potential Unemployment Insurance Duration and Labor Supply: The Individual and Market-Level Response to a Benefit Cut," *Journal of Political Economy*, 126, 2480–2522.

JONES, S. AND K. SEN (2022): "Labour Market Effects of Digital Matching Platforms: Experimental Evidence from Sub-Saharan Africa," .

KELLEY, E. M., C. KSOLL, AND J. MAGRUDER (2021): "How do Online Job Portals Affect Employment and Search Outcomes? Evidence from India," *Manuscript: University of California, Berkeley*.

KLEIBERGEN, F. AND R. PAAP (2006): "Generalized Reduced Rank Tests Using the Singular Value Decomposition," *Journal of Econometrics*, 133, 97–126.

KOLSRUD, J., C. LANDAIS, P. NILSSON, AND J. SPINNEWIJN (2018): "The Optimal Timing of Unemployment Benefits: Theory and Evidence from Sweden," *American Economic Review*, 108, 985–1033.

KREMER, M., G. RAO, AND F. SCHILBACH (2019): "Chapter 5 - Behavioral development economics," .

KUDLYAK, M., D. LKHAGVASUREN, AND R. SYSUYEV (2013): "Systematic Job Search: New Evidence from Individual Job Application Data," Working paper 12-03R, The Federal Reserve Bank of Richmond.

KUHN, P. AND H. MANSOUR (2014): "Is Internet Job Search Still Ineffective?" *The Economic Journal*, 124, 1213–1233.

KUHN, P. AND K. SHEN (2013): "Gender Discrimination in Job Ads: Evidence from China," *The Quarterly Journal of Economics*, 128, 287–336.

KUHN, P. AND M. SKUTERUD (2004): "Internet Job Search and Unemployment Durations," *American Economic Review*, 94, 218–232.

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

LALIVE, R., C. LANDAIS, AND J. ZWEIMULLER (2022): "Market Externalities of Large Unemployment Insurance Extension Programs," *American Economic Review*, 105, 3564–3596.

LANNING, J. (2013): "Opportunities Denied, Wages Diminished: Using Search Theory to Translate Audit-Pair Study Findings into Wage Differentials," *BE Journal of Economic Analysis and Policy*, 13, 921–958.

LE BARBANCHON, T., M. RONCHI, AND J. SAUVAGNAT (2023): "Hiring Frictions and Firm Growth," Discussion Paper 17891, CEPR.

LECHNER, M. AND J. SMITH (2007): "What is the value added by caseworkers?" *Labour Economics*, 14, 135–151.

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

LICHTER, A. AND A. SCHIPROWSKI (2021): "Benefit duration, job search behavior and re-employment," *Journal of Public Economics*, 193, 104–326.

LISE, J., S. SEITZ, AND J. SMITH (2004): "Equilibrium Policy Experiments and the Evaluation of Social Programs," Working paper 10283, National Bureau of Economic Research.

LYONS, E. (2017): "Team Production in International Labor Markets: Experimental Evidence from the Field," *American Economic Journal: Applied Economics*, 9, 70–104.

MADRIAN, B. C. AND D. F. SHEA (2001): "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior," *The Quarterly Journal of Economics*, 116, 1149–1187.

MARBACH, M. AND D. HANGARTNER (2020): "Profiling Compliers and Noncompliers for Instrumental-Variable Analysis," *Political Analysis*, 28, 435–444.

MARINESCU, I. (2017a): "The general equilibrium impacts of unemployment insurance: Evidence from a large online job board," *Journal of Public Economics*, 150, 14–29.

——— (2017b): "Job Search Monitoring and Assistance for the Unemployed," *IZA World of Labor*, 380.

MARINESCU, I. AND D. SKANDALIS (2021): "Unemployment Insurance and Job Search Behavior," *The Quarterly Journal of Economics*, 136, 887–931.

MARINESCU, I. AND R. WOLTHOFF (2020): "Opening the Black Box of the Matching Function: The Power of Words," *Journal of Labor Economics*, 38, 535–568.

MATSUDA, N., T. AHMED, AND S. NOMURA (2019): "Labor Market Analysis Using Big Data: The Case of a Pakistani Online Job Portal," Policy Research Working Paper 9063, World Bank.

MAĆKOWIAK, B., F. MATĚJKA, AND M. WIEDERHOLT (2023): "Rational Inattention: A Review," *Journal of Economic Literature*, 61, 226–73.

MCGEE, A. (2015): "How the Perception of Control Influences Unemployed Job Search," *Industrial and Labor Relations Review*, 68, 184–211.

MCGEE, A. AND 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. AND J. SPINNEWIJN (2022): *Expectations Data, Labor Market and Job Search*, Academic Press, Handbook of Economic Expectations.

O'DONOGHUE, T. AND M. RABIN (1999): "Doing It Now or Later," *American Economic Review*, 89,

103–124.
——— (2001): "Choice and Procrastination," *The Quarterly Journal of Economics*, 116, 121–160.
O'DONOGHUE, T. AND M. RABIN (2008): "Procrastination on long-term projects," *Journal of Economic Behavior and Organization*, 66, 161–175.
PAKISTAN BUREAU OF STATISTICS (2018-2019): "Quarterly Labour Force Survey," .
PALLAIS, A. (2014): "Inefficient Hiring in Entry-Level Labor Markets," *American Economic Review*, 104, 3565–3599.
PALLAIS, A. AND E. G. SANDS (2016): "Why the referential Treatment? Evidence from field experiments on referrals," *Journal of Political Economy*, 124, 1793–1828.
PASERMAN, M. D. (2008): "Job Search and Hyperbolic Discounting: Structural Estimation and Policy Evaluation," *The Economic Journal*, 118, 1418–1452.
PISSARIDES, C. A. (2000): *Equilibrium Unemployment Theory*, MIT Press, 2 ed.
POVERTY ACTION LAB (2022): "Reducing Search Barriers for Job Seekers," .
RUBINSTEIN, J., D. MEYER, AND J. EVANS (2001): "Executive Control of Cognitive Processes in Task Switching." *Journal of Experimental Psychology: Human Perception and Performance*, 27, 763–797.
SANDERS, M., G. BRISCESE, R. GALLAGHER, A. GYANI, S. HANES, E. KIRKMAN, AND 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, 1189–1225.
SPINNEWIJN, J. (2015): "Unemployed but Optimistic: Optimal Insurance Design with Biased Beliefs," *Journal of the European Economic Association*, 13, 130–167.
STANTON, C. T. AND C. THOMAS (2021): "Who Benefits from Online Gig Economy Platforms?" .
SUBRAMANIAN, N. (2021): "Workplace Attributes and Women's Labor Supply Decisions: Evidence from a Randomized Experiment," *Unpublished Manuscript*.
THALER, R. H. AND S. BENARTZI (2004): "Save More Tomorrow™: Using Behavioral Economics to Increase Employee Saving," *Journal of Political Economy*, 112, S164–S187.
TOOHEY, D. (2014): "Job Rationing in Recessions: Evidence from Work-Search Requirements," Manuscript, University of Michigan.
WHEELER, L., R. GARLICK, E. JOHNSON, P. SHAW, AND 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, AND V. GUERRIERI (2021): "Directed Search and Competitive Search Equilibrium: A Guided Tour," *Journal of Economic Litterature*, 59, 90–148.
ZHU, Y. (2021): "Phase transition of the monotonicity assumption in learning local average treatment effects," *arXiv:2103.13369 [econ, math, stat]*.

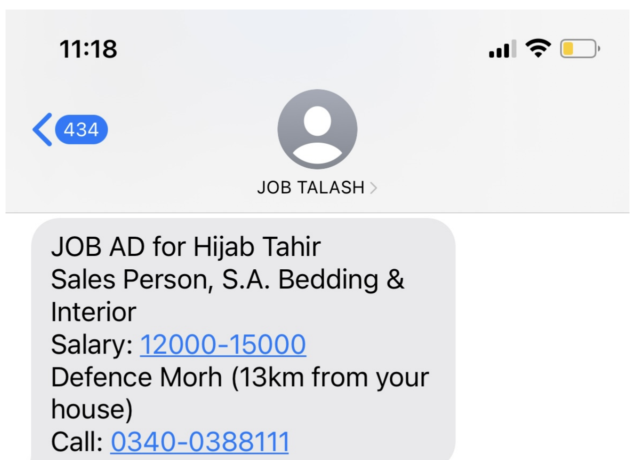
Appendices for Online Publication Only

A Additional Information about the Platform and Sample

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

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

Figure A.1: Sample Text Message



Notes: Sample text message shown in English; text messages sent on the platform are written in Urdu.

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.2 compares the distribution of salaries for vacancies posted on the platform to the distribution of salaries for the Lahore subsample of Pakistan's Labor Force Survey (Pakistan Bureau of Statistics, 2018-2019). These distributions should be compared with caution, as the former covers vacancies and the latter covers filled jobs, including jobs where incumbent workers have substantial experience with that firm.

Table A.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

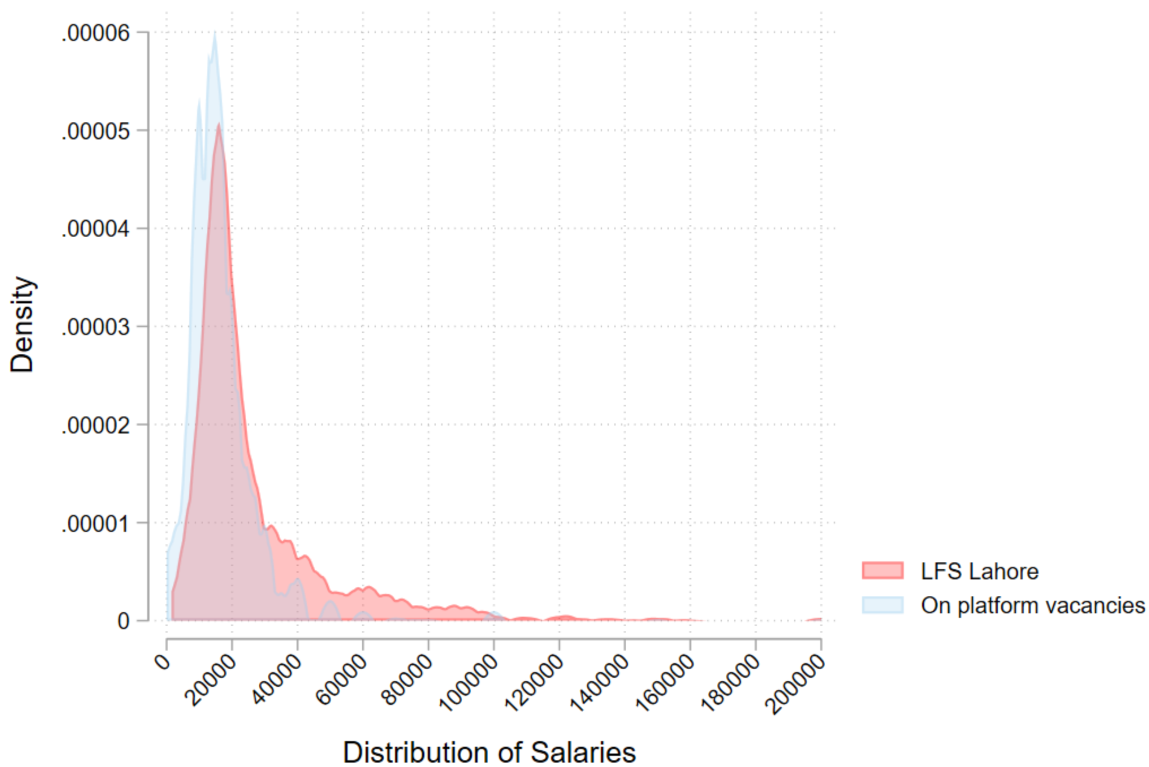
	LFS Pakistan	LFS Lahore	LFS Lahore Reweighted	HH listing Sample	Experimental Sample
	(1)	(2)	(3)	(4)	(5)
Age	33.9	33.8	32.7	32.6	30.7
	(11.6)	(11.6)	(11.0)	(11.1)	(9.5)
Highest Education Level					
Less than Intermediate/High School	0.853	0.679	0.700	0.706	0.491
Completed Intermediate/High School	0.073	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. The LFS search module is only asked for non-employed respondents. We ask the search module for all jobseekers.

Figure A.2: 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 $\approx$ USD 0.03 in purchasing power parity terms during the study period.

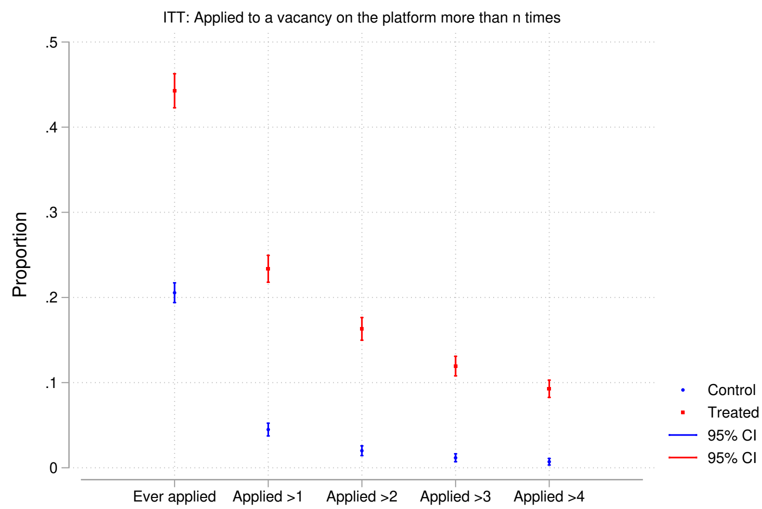
B Additional Analysis on Search Effects and Returns to Search

B.1 Treatment Effects on Additional Interview and Application Outcomes

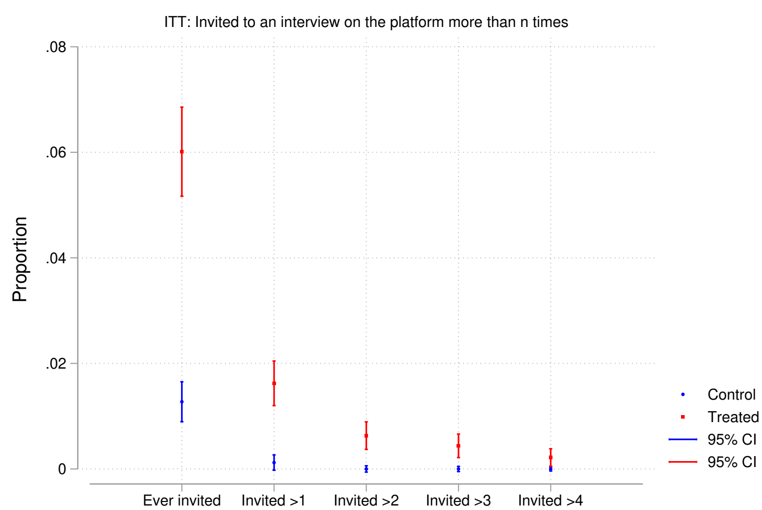
Figure B.1 shows treatment effects on the number of times each jobseeker applies to and is invited to an interview for a job. This figure shows that treatment raises the probability of submitting K applications and getting L interviews for all K and for L $\leq$ 4..

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.

B.2 Robustness Checks

Table B.1 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 and Rubin (2015); column 3 drops stratification block fixed effects. Results are similar across the three specifications: the effect on applications ranges from 1.28 to 1.34 percentage points and the marginal applications have a mean interview probability between 5 and 5.9%. We also show results conditioning on jobseeker-level covariates in column 4, vacancy- and match-level covariates in column 5, and all three sets of covariates in column 6. All covariates are selected using a post-double selection LASSO, following Belloni et al. (2014). The effect on applications ranges from 1.33 to 1.34 percentage points and the marginal applications have a mean interview probability between 5.9 and 6.8%. The findings in columns 4, 5, and 6 reinforce our argument in Sections 3.3 and 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.1: 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 and 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.

Table B.2 presents treatment effects on applications and interviews using the within-jobseeker through-time randomization. Table B.3 shows that treatment effects on applications do not differ substantially by the quality scores assigned by employers to CVs. Both tables support the argument in Section 3.3 that constant returns to search are not explained by different types of jobseekers applying for jobs in the treatment and control groups.

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

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

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

Notes: All specifications are identical to those in Table 3 except that the treatment indicator varies both through time and between jobseekers. Column 1 shows the coefficient from regressing an indicator for job application on treatment assignment. Column 2 shows the coefficient from regressing an indicator for interview invitation on treatment assignment. Column 3 shows the coefficient from regressing an indicator for interview invitation weighted by a proxy index for the value of the vacancy to the jobseeker, V_{vm}, on treatment assignment. Column 4 shows the coefficient from regressing an indicator for interview invitation on job application, instrumented by treatment assignment. Column 5 shows the coefficient from regressing an indicator for interview invitation weighted by the proxy index V_{vm} on job application, instrumented by treatment assignment. The proxy index is defined in the note to Table 3. All regressions use one observation per jobseeker $\times$ vacancy match, include jobseeker and round fixed effects, and use 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 and Paap (2006).

Table B.3: Heterogeneous Treatment Effects by Employer-Scored CV Quality

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

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

Table B.4: 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.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:

$$\beta_{2SLS} = \frac{\mathbf{P_C} \cdot \mathbf{\Delta I_C} - \mathbf{P_D} \cdot \mathbf{\Delta I_D}}{\mathbf{P_C} - \mathbf{P_D}}. \tag{7}$$

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 (7) as

$$\mathbf{\Delta I_C} = \frac{\beta_{2SLS} \cdot (\mathbf{P_C} - \mathbf{P_D}) + \mathbf{P_D} \cdot \mathbf{\Delta I_D}}{\mathbf{P_C}} = \frac{\beta_{2SLS} \cdot \beta_{S1} - \mathbf{P_D} \cdot \mathbf{\bar{I}_D}}{\beta_{S1} + \mathbf{P_D}}, \tag{8}$$

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 (8) 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

$$0 \leq \mathbf{P_D} \leq \bar{A}_0. \tag{9}$$

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

$$0 \leq \bar{\mathbf{I}}_{\mathbf{A}} \leq 1 \tag{10}$$

$$0 \leq \bar{\mathbf{I}}_{\mathbf{D}} \leq 1. \tag{11}$$

Evaluating equation (8) 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

$$\mathbf{P_D} \cdot \left(\bar{\mathbf{I}}_{\mathbf{D}} - \bar{\mathbf{I}}_{\mathbf{A}}\right) = \bar{A}_0 \cdot \left(\bar{I}_0 - \bar{\mathbf{I}}_{\mathbf{A}}\right). \tag{12}$$

Combining (8), (9), (10), (11), and (12) 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.

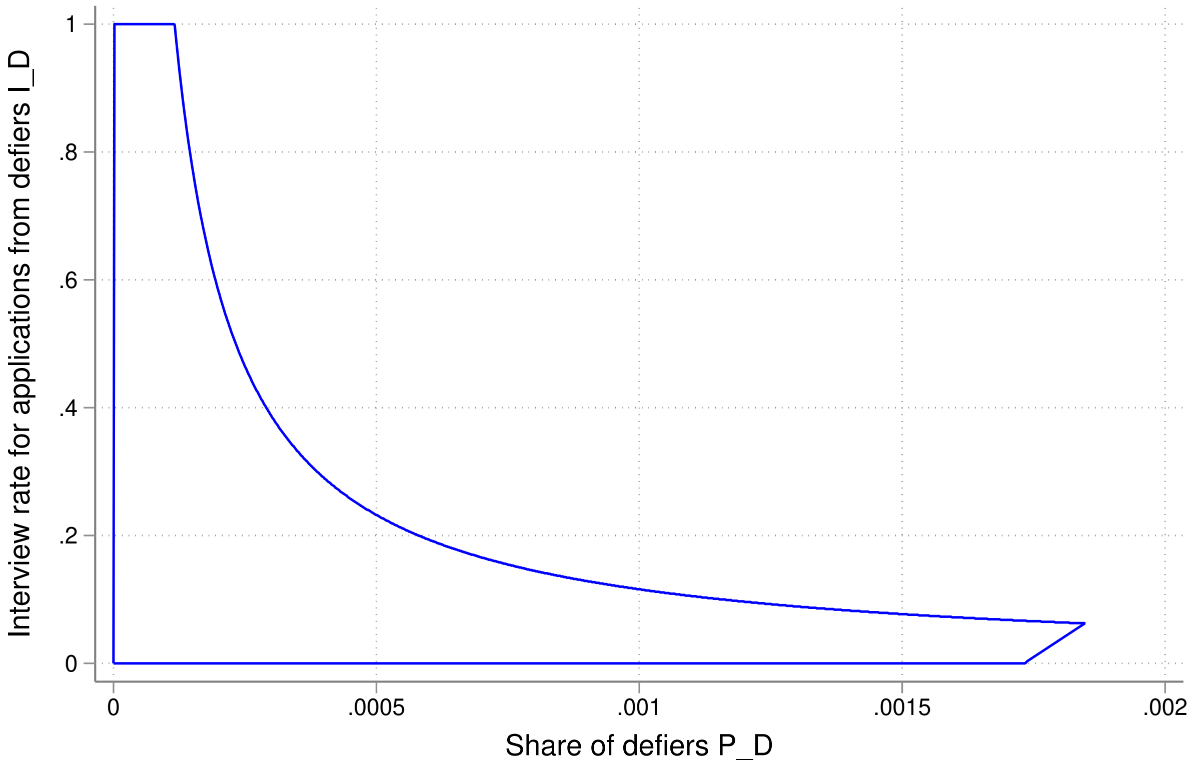
Fourth, we consider each value P_D satisfying (9), solve for the set of values of $\bar{\mathbf{I}}_{\mathbf{D}}$ and $\bar{\mathbf{I}}_{\mathbf{A}}$ consistent with all the restrictions, and then solve for the set of values of $\mathbf{\Delta I_C}$ consistent with all the restrictions. Let $\{\mathbf{\Delta I_C}\}_{\mathbf{P_D}}$ 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 (10) 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 (12) from becoming so large that it can only be satisfied by a value of $\bar{\mathbf{I}}_{\mathbf{A}}$ that violates condition (11). 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 (7): 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, 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.

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.

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.5, 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.5, 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

Table B.5: 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.

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.6, 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 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.5, column 3). But replicating our main analysis for respondents who completed their entire CV at registration replicates our main findings (Table B.6, 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

Table B.6: 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 × vacancy. Heteroskedasticity-robust standard errors are shown in parentheses, clustered by jobseeker.

text messages) to active disengagement (asking to stop calls and text messages). Our main analysis retains matches from jobseekers who request stops and codes applications and interviews as zeros for these matches. As a sensitivity check, we can instead drop observations from jobseekers during periods when they have requested stops. This mechanically slightly increases treatment effects on applications and interviews (Table B.6, 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.

B.5 An Alternative Approach to Testing Constant Returns to Search

In this subsection 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.7, 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.7, column 2, row 3).

Table B.7: Alternative Test for Constant Returns to Search

	(1) Interview	(2) Interview $\times$ V_{vm} index
Treatment effect	0.00078 (0.00009)	0.00281 (0.00036)
Constant-returns implied effect	0.00083 (0.00005)	0.00314 (0.00018)
Difference	-0.00006 (0.00007)	-0.00033 (0.00028)
# matches	1,116,952	1,116,952
# jobseekers	9831	9831
Mean outcome | 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.6 Heterogeneous Treatment Effects by Baseline Employment, Search, and Gender

Table B.8 shows heterogeneous treatment effects on applications and interviews by baseline employment status, search activity, gender, education, and age.

Table B.8: Heterogeneous Treatment Effects

Panel A: Applications

	(1)	(2)	(3)	(4)	(5)
Phone call treatment	0.01356	0.01174	0.01421	0.01161	0.01165
	(0.00091)	(0.00117)	(0.00088)	(0.00084)	(0.00091)
Phone call treatment × Group	-0.00080	0.00392	-0.00378	0.00309	0.00260
	(0.00113)	(0.00135)	(0.00119)	(0.00116)	(0.00110)
Group	-0.00047	0.00064	0.00097	0.00174	0.00036
	(0.00024)	(0.00029)	(0.00039)	(0.00035)	(0.00022)
# matches	1,116,160	921,011	1,116,952	1,116,952	1,116,952
Group	Employed	Searching	Female	High school or higher education	Age $<$ 30
Proportion in Group	0.41427	0.58115	0.22850	0.46970	0.58101
Outcome Control Mean	0.00185	0.00189	0.00185	0.00185	0.00185
Total Effect on HTE group	0.01276	0.01566	0.01044	0.01469	0.01425
	(0.00097)	(0.00109)	(0.00099)	(0.00103)	(0.00093)

Panel B: Interview Invitations

	(1)	(2)	(3)	(4)	(5)
Phone call treatment	0.00084	0.00085	0.00070	0.00085	0.00060
	(0.00011)	(0.00014)	(0.00009)	(0.00011)	(0.00010)
Phone call treatment × Group	-0.00014	0.00005	0.00027	-0.00014	0.00028
	(0.00013)	(0.00016)	(0.00017)	(0.00013)	(0.00012)
Group	-0.00003	0.00006	0.00002	0.00004	-0.00000
	(0.00003)	(0.00004)	(0.00005)	(0.00004)	(0.00003)
# matches	1,116,160	921,011	1,116,952	1,116,952	1,116,952
Group	Employed	Searching	Female	High school or higher education	Age $<$ 30
Proportion in Group	0.41427	0.58115	0.22850	0.46970	0.58101
Outcome Control Mean	0.00012	0.00012	0.00012	0.00012	0.00012
Total Effect on HTE group	0.00069	0.00090	0.00097	0.00071	0.00089
	(0.00011)	(0.00013)	(0.00016)	(0.00011)	(0.00011)

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

C Additional Analysis on Mechanisms

C.1 How are Marginal and Inframarginal Applications Directed?

This appendix provides additional results about how marginal and inframarginal applications are directed and explains the complier or latent type method used in sections 3.3 and 4.

Heterogeneous Treatment Effects by Vacancy Value: We show several figures confirming the pattern documented in Section 4 that marginal and inframarginal applications are sent to vacancies with similar values to jobseekers. Figure 2 showed heterogeneous treatment effects by two proxies for the value of a vacancy – salary and an index of multiple value proxies – to show that the share of applications sent to high-value vacancies does not differ between treatment and control groups. Figure C.1 replicates this using the within-jobseeker between-vacancy distributions of salary and the index, showing loosely the same pattern. Figure C.2 replicates Figure 2 using the value of interviews rather than vacancies, showing the same pattern. Recall that we test if the share of applications sent to high-value vacancies differs between treatment and control groups by testing if the ratio of the treatment group to control group application rate is equal over the five quintiles shown in each figure. Test results are reported in the note below each figure.

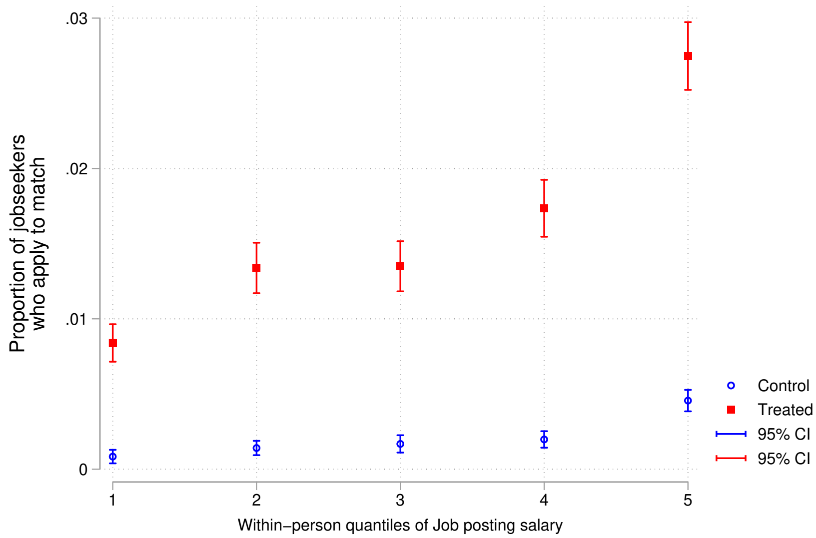
Matches are sent to jobseekers roughly every month, as part of a matching round. Any jobseeker who has received multiple matches in that round receives a *batch* of multiple matches. Roughly two thirds of matches are sent in batches and one third are sent individually. We use this structure to show in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

Figure C.3 shows how the phone call treatment shifts the number of applications that respondents make in each of these rounds. Panel A shows the full distribution, while Panel B shows the distribution conditional on a positive number of applications. The conditional distributions are similar between treatment and control group, with confidence intervals fully overlapping. This shows that the entire treatment effect on applications comes from the shift from applying to zero vacancies in a given round to a positive number of applications. This pattern is consistent with the conceptual framework: some jobseekers miss applying to some batches of matches due to temporarily high present bias or psychological application costs. If, instead, treatment shifted some jobseekers from making one to making two or more applications within a batch of matches, this would not be explained by a reduction in the psychological cost of initiating applications.

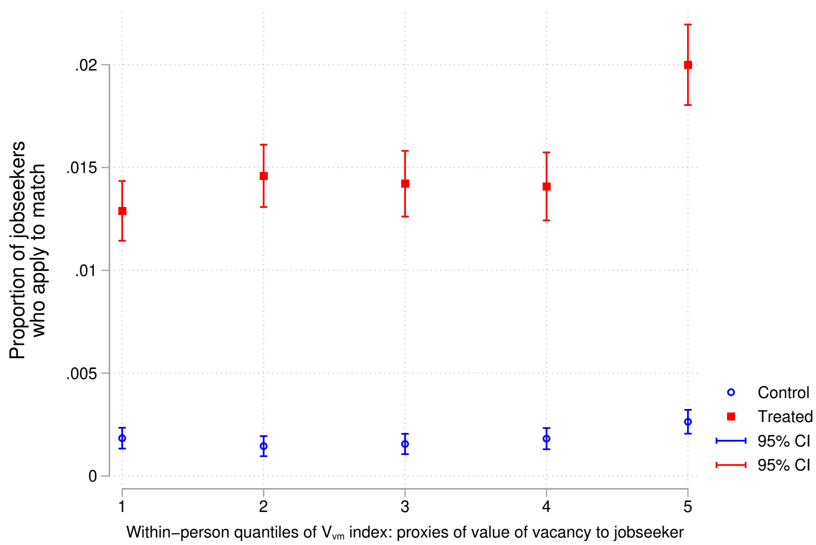
Figure C.4 shows heterogeneous treatment effects collapsing the data to the level of the matching round, replicating the results in Figure 2. Finally, Figure C.5 repeats this analysis measuring the value of a round based on the highest-value vacancy rather than average over the vacancies in the round. Results are similar across all approaches, showing that treatment and control group jobseekers apply to observationally similar batches of matches.

Figure C.1: Heterogeneous Treatment Effects on Applications by Vacancy Value Using Within-Jobseeker Variation in Value

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



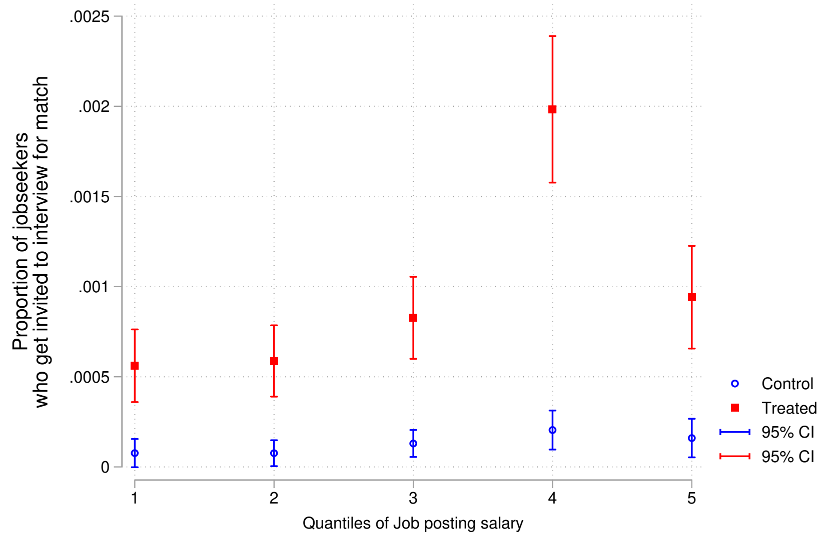
Panel B: Match level: Quintiles of V_{vm} index within-jobseeker



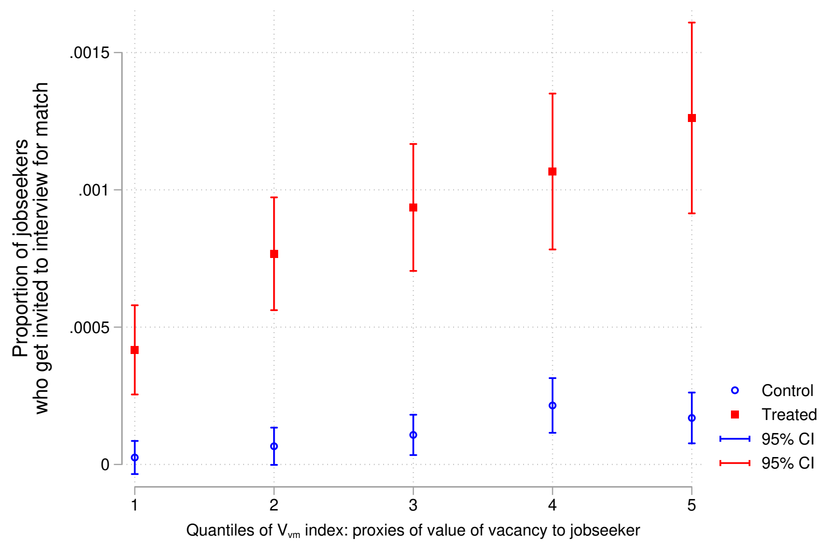
Notes: This figure shows heterogeneous treatment effects of the phone call treatment on applications by proxies for the value of the job posting. Panel A shows heterogeneity by job posting salary, defining quintiles based on the distribution of salary within-jobseeker. Panel B shows heterogeneity by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

Figure C.2: Heterogeneous Treatment Effects on Interviews by Vacancy Value

Panel A: Match level, by posted salary



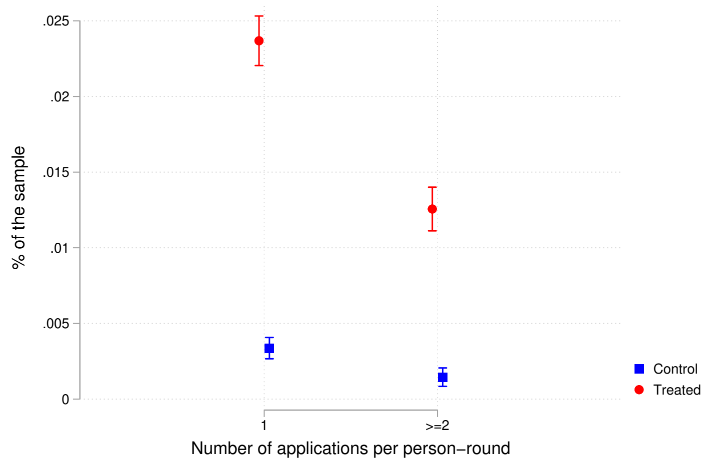
Panel B: Match level, by V_{vm} index



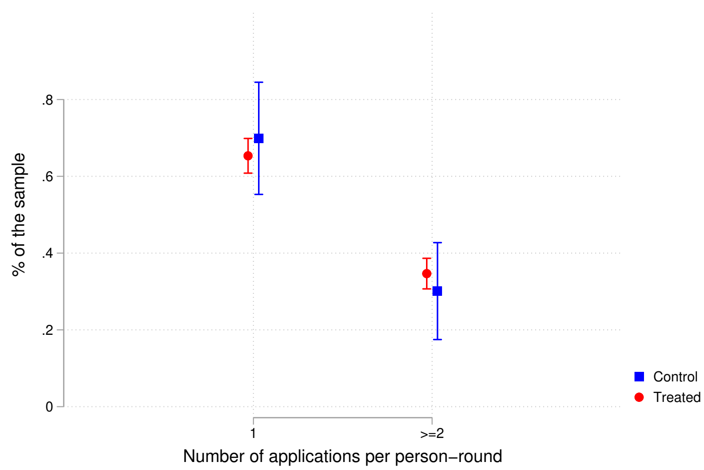
Notes: This figure shows heterogeneous treatment effects of the phone call treatment on interviews by proxies for the value of the job posting. Panels A and B show heterogeneity by job posting salary and V_{vm} index described in Section 3.2 using the within-jobseekers between-vacancy distribution. The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

Figure C.3: Treatment Effects on the Number of Applications per Jobseeker $\times$ Matching Round

Panel A: Treatment Effects on Each Positive Number of Applications

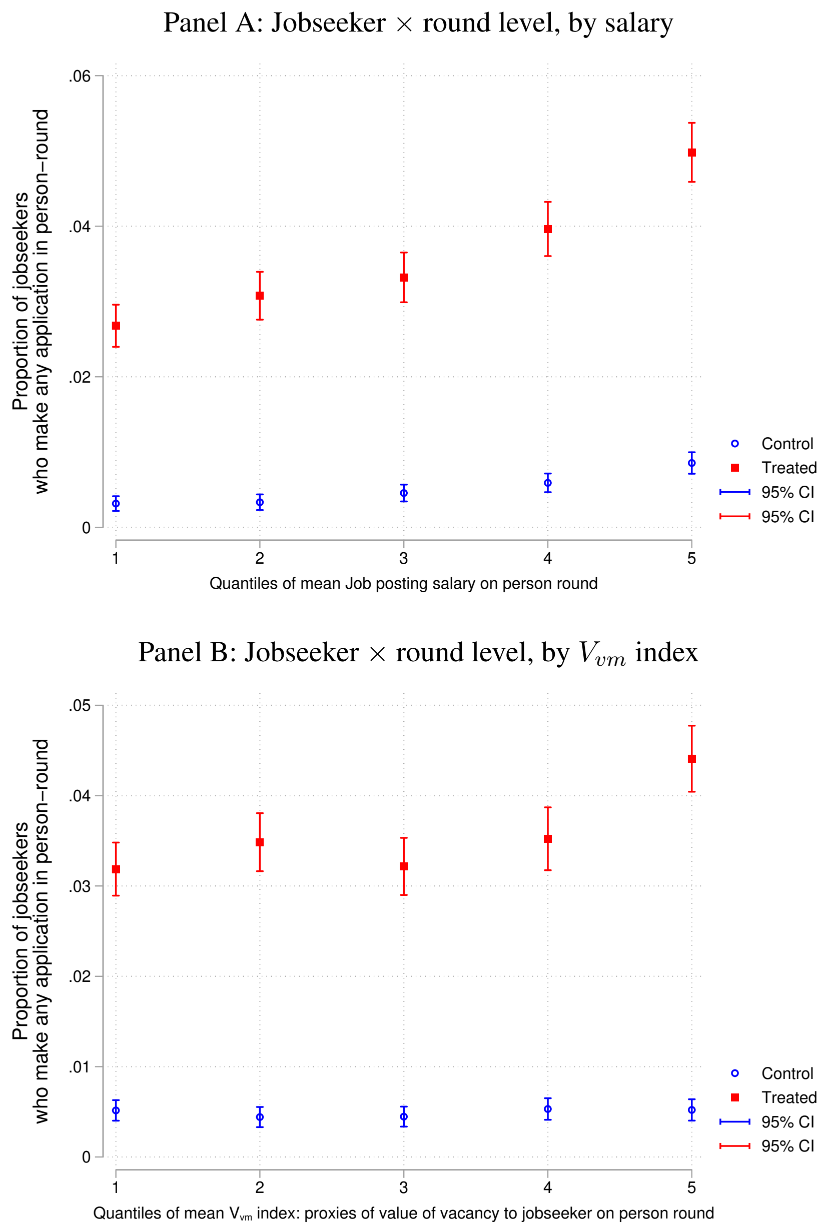


Panel B: Treatment Effects on Each Positive Number of Applications, Scaled by the Probability of > 0 Applications



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.

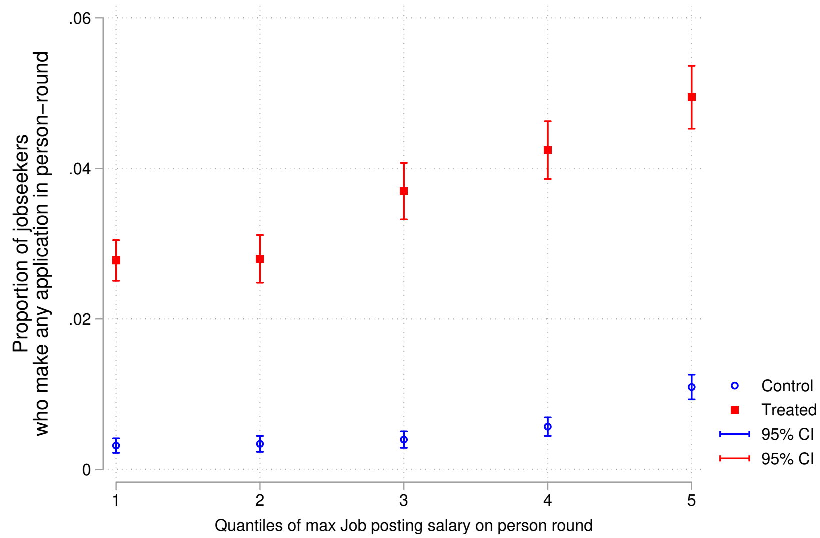
Figure C.4: Heterogeneous Treatment Effects on Applications by Vacancy Value Using Jobseeker × Matching Round Level Data



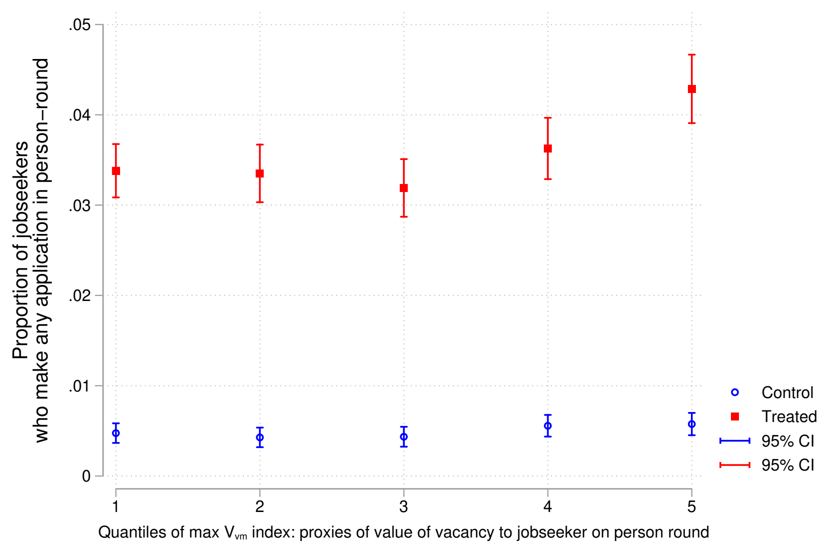
Notes: This figure shows heterogeneous treatment effects of the phone call treatment on applications by proxies for the value of the job posting. All analysis in this figure uses one jobseeker × round as a unit of observation, averaging over the values of the vacancies in that unit. Panels A and B show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

Figure C.5: Heterogeneous Treatment Effects on Applications by Vacancy Value Using Jobseeker × Matching Round Data

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



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



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

Method for Complier / Latent Type Analysis: We argue that the marginal job applications submitted due to treatment and inframarginal job applications submitted in the absence of treatment have similar characteristics, in terms of jobseeker characteristics (Section 3.3, Table 4) and vacancy/match characteristics (Section 5). In this appendix we describe the method used to support this argument, which is adapted from Marbach and Hangartner (2020).[32]

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

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

We can estimate 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.

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

Latent interview probabilities are another proxy for the value of each application; these also do not differ on average between marginal and inframarginal applications. To show this, we estimate the latent probability that each match would yield an interview if an application were submitted using the same method introduced in Section 3.3, but now incorporating vacancy and match level characteristics into the prediction model. The mean probability is similar between marginal and inframarginal applications when estimated using only these vacancy- and match-level characteristics or also including jobseeker characteristics. Finally,

[32] 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.

we interact each latent interview probability measure with the value index to create an omnibus proxy for PV. The means for marginal and inframarginal applications do not differ.

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

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

	(1) Inframarginal applications	(2) Marginal applications	(3) Difference (p-value)
Firm characteristics			
Leave one out ratio of firm interviews to applications (on platform)	0.061	0.058	-0.003
			(0.583)
Firm baseline ratio of interviews to applications (off-platform)	0.705	0.738	0.033
			(0.053)
Firm # employees	88.347	43.720	-44.627
			(0.001)
Firm # vacancies in last year	12.301	9.096	-3.205
			(0.002)
Vacancy characteristics			
Ln(posted salary)	9.848	9.704	-0.143
			(0.000)
$<$ median working hours	0.600	0.582	-0.018
			(0.489)
Allows employees to work flexible hours	0.697	0.803	0.106
			(0.000)
Offers any benefits	0.767	0.758	-0.010
			(0.624)
Match characteristics			
Exact education match $\mid$ 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.2: 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.

C.2 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 C.2).

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

Table C.3: Effect of Randomized Survey Features on Probability of Answering Survey Modules

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

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

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 C.3. 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).[33]

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 7), specific off-platform search activities (Table D.2), receipt of calls/text messages (Table C.9), and beliefs about jobs on the platform (Table C.10). Adjusting for selection has small effects on most estimated treatment effects.

Many researchers instead focus on bounding a different parameter: the average treatment effect in the subpopulation of respondents who respond irrespective of treatment status, following Lee (2009). This approach does not require instruments but is uninformative in our setting because the bounds are very wide.

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.

[33]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, the statistically strong relationship between response and the instruments is reassuring. As an alternative metric for evaluating instrument strength, following Garlick and 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.

C.3 Additional Mechanisms Results

We explain in Section 4.2 that the conceptual framework predicts that the value of matches receiving applications should be more dispersed in the treatment group. To test this, we estimate treatment effects on the variance, 10th percentile, and 25th percentile of log salary for matches that receive applications, using a nonparametric bootstrap clustered by jobseeker to obtain standard errors on these treatment effects. Table C.4 shows treatment raises the variance and lowers the 10th and 25th percentiles for both log salary and the proxy index V_{vm} that combines multiple proxies for match and vacancy value. This is consistent with the framework's prediction that marginal treatment-induced applications should go to vacancies with the same average value as inframarginal applications but more dispersed values.

Table C.4: Treatment Effects on Dispersion of Value of Matches Receiving Applications

	Ln(Salary)			V_{vm} index		
	Variance	10th pctile	25th pctile	Variance	10th pctile	25th pctile
Control	3.13	9.2	9.6	0.926	2.37	2.82
	(0.471)	(0.022)	(0.060)	(0.058)	(0.035)	(0.038)
Treatment	5.18	8.99	9.4	0.964	2.34	2.80
	(0.229)	(0.004)	(0.091)	(0.033)	(0.015)	(0.017)
Treatment effect	2.06	-0.223	-0.223	0.038	-0.025	-0.017
	(0.532)	(0.022)	(0.108)	(0.067)	(0.038)	(0.041)

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, 10th, and 25th percentiles – of two proxies for vacancy value – log monthly salary and the index V_{vm} of vacancy- and match-level proxies for vacancy value defined in the note to Table 3. The table rows show the levels of these dispersion statistics for the treatment and control groups and the treatment effect. Standard errors are estimated using 1000 iterations of a nonparametric bootstrap, clustering by jobseeker.

Pecuniary and time costs: Here we show results for the mechanism experiments described in Section 4.3. Column 1 of Table C.5 compares the effects of our main phone call initiation treatment to the effects of a randomized text message reminder that the jobseeker can ask the platform to call them back about a job posting. The free callback reminder treatment has an effect one hundredth of the size of the effect of the main phone call treatment, and the two effects are statistically significantly different (p = .017).

Column 2 of Table C.5 compares the effects of our main phone call initiation treatment to the effects of randomly offering some control group jobseekers the option to text the platform and ask for a callback at a specific time. This eliminates the differential wait time between the main treatment and control groups. This callback request treatment has an effect one quarter of the size of the effect of the main phone call treatment, and the two effects are statistically significantly different (p = .002).

Each column uses only the set of jobseeker $\times$ vacancy matches from rounds in which the relevant feature was randomized.

Table C.5: Mechanism Experiment: Treatment Effects on Applications of Reductions in Pecuniary and Time Costs

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

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

Reminder effects: Here we show results for the mechanism experiments and non-experimental analysis described in Section 5.1. Table C.6 shows the effect of a reminder text message sent to a random subsample of control group jobseekers at the same time that the treatment group jobseekers receive calls. If reminder effects explain our results, this should have a similar effect to that of the phone call treatment. The effect of the reminder message is one-fourteenth as large as the effect of the phone call treatment in the same matching rounds and statistically significantly smaller ($p < 0.001$).

Table C.6: Mechanism Experiment: 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.

Table C.7, Column 1, shows the effect of the timing of the phone call made to the treatment group. We randomize the order in which we call jobseekers for the phone call treatment, within the application window between the text message job alert and the deadline for job applications. If reminder effects explain our results, we expect that the treatment should have a stronger effect for jobseekers called later within this window, as they will have had more time to forget to apply. Instead, we find that the later the phone call made to the jobseeker, the smaller the treatment effect on applications. This suggests that reminder effects do not explain our results.

Table C.7, Column 2, tests for heterogeneous treatment effects by the duration between the job alert text message and the application deadline. This duration is not randomly assigned, but varies due to logistical factors such as the number of call center agents on staff at the time of the matching round. We interact the duration of this window with treatment, controlling for quarter fixed effects to address variation over time in these logistical factors. Table C.7, Column 2 shows the results. If reminder effects explained our results, we would expect treatment to have a larger effect when there is a longer application window, as jobseekers will have had more time to forget to apply. The results show that treatment has a smaller effect when the window is longer, again suggesting reminder effects do not explain our results.

Table C.7: Mechanism Analysis: Treatment Effects on Applications by Timing of Phone Call and Length of Application Window

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

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

Pressure to apply: Here we show results for the mechanism analysis described in Section 5.2. Figure C.6) shows the proportion of applications that are directed to the first vacancy listed on the phone call versus later vacancies listed. If pressure were responsible for the main treatment effects on applications, we would expect to see treatment group jobseekers applying to the first vacancy listed on the call at a higher rate than control group jobseekers. Instead, 31% of applications go to the first vacancy listed on the call in both the treatment and control groups. To help contextualize this result, we note that 22% of all vacancies are listed first on the call. So jobseekers are disproportionately likely to apply to first-listed vacancies, but this pattern does not differ between treated and control jobseekers.

Figure C.6: Proportion of Applications by Order in which Vacancies are Listed

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

Differential information receipt: Here we show results for the mechanism experiments described in Section 5.3. That section raised the possibility that the phone call treatment might influence application rates simply by providing additional information. To test this, in some matching rounds, we randomize some jobseekers to receive additional information about benefits offered by specific vacancies in their text messages. We then estimate

$$\begin{aligned} \text{Apply}_{jvt} = &\beta_0 + \beta_1 \cdot T_j + \beta_2 \cdot \text{Benefits}_v + \beta_3 \cdot \text{BenefitsInfo}_{jt} + \beta_4 \cdot T_j \cdot \text{Benefits}_v \\ &+ \beta_5 \cdot T_j \cdot \text{BenefitsInfo}_{jt} + \beta_6 \cdot \text{Benefits}_v \cdot \text{BenefitsInfo}_{jt} \\ &+ \beta_7 \cdot T_j \cdot \text{Benefits}_v \cdot \text{BenefitsInfo}_{jt} + \mu_b + \epsilon_{jvt} \end{aligned} \tag{13}$$

where Benefits_v indicates whether the vacancy offers benefits, and BenefitsInfo_{jt} indicates whether the jobseeker was randomly selected during round t to receive information about benefits for vacancies that offer them, and T_j indicates the main phone call treatment. We include the interactions with Benefits_v because not all jobs offer benefits, so this mechanism experiment is conducted only in this subset of vacancies. The coefficients of interest are β_6 and $\beta_6 + \beta_7$, the effects of getting information about benefits for jobseekers assigned to respectively the phone call control and treatment groups.

Table C.8 shows that the effect of providing information about benefits on application rates is one between one eighth and one quarter the size of the effect of the main phone call treatment, and is not statistically significant, either by itself or interacted with the phone call initiation treatment. This suggests that receipt of information about specific vacancies does not drive our main findings.

Table C.8: Mechanism Experiment: Treatment Effects of Providing Additional Information About Vacancy Benefits in Text Messages

	Apply	
	(1)	(2)
Phone call treatment$_j$	0.00217	
	(0.00059)	
Vacancy offers benefits$_v$	0.00006	
	(0.00018)	
Benefits information treatment$_{jt}$	-0.00003	0.00002
	(0.00033)	(0.00036)
Vacancy offers benefits$_v$ $\times$ Benefits information treatment$_{jt}$	0.00050	0.00041
	(0.00040)	(0.00043)
Phone call treatment$_j$ $\times$ Vacancy offers benefits$_v$ $\times$ Benefits information treatment$_{jt}$	-0.00076	-0.00027
	(0.00079)	(0.00089)
# matches	212301	212301
# vacancies	289	289
# jobseekers	8108	8108
Mean outcome | T = 0	0.00076	0.00076
Round FE	No	Yes
Round $\times$ vacancy with benefits FE	No	Yes
Round $\times$ Phone call treatment FE	No	Yes
Round $\times$ Phone call treatment $\times$ vacancy with benefits FE	No	Yes

Notes: Sample includes thirteen matching rounds; the mechanism experiment was rolled out in the last seven rounds of this sample. We restrict the sample to jobseekers who receive at least one ad for a job offering benefits over a period of thirteen rounds. Estimates are shown for Equation 13, where Benefits$_v$ is an indicator for a vacancy that offers any non-salary benefit; BenefitsInfo$_{jt}$ = 1 for jobseekers who are randomly selected to receive information about benefits in round t. Additional coefficients from Equation 13 are included in the estimation but suppressed from the table for readability. 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.

Section 5.3 also asked if the phone call treatment might increase application rates because jobseekers were more likely to receive phone calls than text messages. To test this, we survey respondents and ask if they have received matches by phone call and/or text message in the previous 14 or 30 days (recalled period randomized). Table C.9 shows treatment effects on respondents' reports of receiving information about job matches. Throughout the table, odd numbered columns show the uncorrected estimates, while even numbered columns show the estimates with selection correction discussed in Section C.2.

Columns 1-2 show that treatment group respondents are less likely to report receiving a vacancy update by text message, but this effect is not statistically significant. Such a decrease could occur because some treated respondents pay attention to the phone call and disregard the text message. Columns 3-4 show that treatment increases the probability that a respondent reports receiving a vacancy update by phone call. This acts as a validation check on the survey data, and confirms that the survey responses capture the treatment protocol implementation. The key result, in columns 5-6, shows that treatment does *not* change the probability of a jobseeker receiving information about the job match. This suggests that differential access to information does not explain our results.

It is possible that some jobseekers do not receive the text message, due to the text message being blocked as spam; or that treatment group jobseekers miss the attempted phone call. We also expect some measurement error in these survey responses. Some jobseekers who received matches before the recall period will incorrectly report receiving matches during the recall period, a pattern called 'telescoping' in the survey methods literature. Some jobseekers who received matches during the recall period will incorrectly report not receiving matches due to inattention or forgetting. It is possible that such sources of measurement error could lead us to find no difference in self-reported match receipt when in fact treated jobseekers were more likely to receive matches. However, this is only possible if jobseekers are more likely to forget phone calls than text messages, which we view as unlikely.

Table C.9: Mechanism Analysis: Treatment Effects on Recalling Receiving Matches

	Respondent reported receiving:					
	Text		Phone call		Either	
	(1)	(2)	(3)	(4)	(5)	(6)
Phone call treatment	-0.04406	-0.04124	0.09653	0.12486	-0.00091	0.02964
	(0.02946)	(0.04475)	(0.02274)	(0.03430)	(0.02964)	(0.04377)
N	1520	10582	1543	10582	1544	10582
# responses | T = 0	685	685	692	692	693	693
# responses | T = 1	835	835	851	851	851	851
Mean outcome | T = 0	0.40146	0.40146	0.12283	0.12283	0.44012	0.44012
Adjusted for non-response	No	Yes	No	Yes	No	Yes
IV strength test: F-stat		56.872		51.884		46.510
IV strength test: p-value		0.000		0.000		0.000

Notes: This table shows treatment effects on the probability that respondents report receiving matches from the platform through different modes of communication. The sample consists of respondents who did receive a match in the last two matching rounds; some respondents were asked this question even if they did not receive a match in the exact date of the recall period because of survey instrument randomization between 14 and 30 day recall periods. The outcome in columns (1) and (2) is an indicator for reporting that they received a match by text message, in columns (3) and (4) is an indicator for reporting that they received a match by phone, and in columns (5) and (6) an indicator for reporting that they received a match by text message or phone call. Each outcome is regressed on an indicator for treatment assignment, fixed effects for recall periods (which are randomly assigned), and stratification block fixed effects. Even-numbered columns include selection adjustment terms for survey non-response described in Section C.2, following DiNardo et al. (2021). The first-stage F-statistics jointly test the strength of the four excluded instruments. Heteroskedasticity-robust standard errors shown in parentheses, clustered by jobseeker. The unit of observation is a survey response, as jobseekers were surveyed up to twice. Only 0.6% of jobseekers complete two surveys.

Beliefs about returns to search on the platform: Section 5.4 introduced the possibility that treatment shifts application rates by changing jobseekers' beliefs about returns to search on the platform. Table C.10 shows that treatment does not shift jobseekers' self-reported beliefs about their own probability of being selected if they apply (columns 1-2) or the quality of job opportunities available on the platform (columns 3-4). Both results are close to zero, both with and without the selection correction approach described in Section C.2.

Table C.10: Mechanism Analysis: Beliefs About Potential Returns to Search on Job Talash Platform

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

Notes: This table shows treatment effects on beliefs collected as part of jobseeker followup surveys. Each outcome is regressed on an indicator for treatment assignment and stratification block fixed effects. Columns (2) and (4) include selection adjustment terms for survey non-response as described in Section C.2, 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. Heteroskedasticity-robust standard errors are shown in parentheses.

We also present additional mechanism tests beyond the evidence described in Section 5.4. During the randomized rollout of the baseline surveys, some jobseekers were surveyed and enrolled on the platform by the survey team before the start of the phone call experiment. During this baseline period, these jobseekers all received phone calls on each matching round in which they received any matches, following the same protocol applied during the experiment to the treatment group. We use this initial period in three ways.

First, we test for differences in response to treatment between those who did and did not receive this baseline period of phone calls initiated by the platform. If treatment works by shifting beliefs about vacancies on the platform, we might expect a differential response for jobseekers who receive phone calls for their initial rounds. For example, if the phone calls serve to create a more professional image of the platform and a signal of higher quality jobs available, this belief might persist after the baseline period. Column 1 of Table C.11 shows heterogeneous treatment effects by whether the jobseeker received any rounds with phone calls before the experiment began. There is no difference in effects between the two groups.

Second, we use variation in the characteristics of vacancies included in this baseline period. If treatment increases application by raising the perceived value of vacancies on the platform, then jobseekers who happen to receive particularly high-value vacancy announcements during this baseline may experience a smaller treatment effect, because both treatment and control group jobseekers who receive baseline phone calls with high-value matches might have already updated their beliefs. For each jobseeker who received baseline matching phone calls, we calculate what proportion of the vacancy ads she received in this baseline period offered salaries above the median for her distribution on the platform, i.e. were "better draws" from her own distribution at baseline. Column 2 of Table C.11 shows heterogeneous treatment effects by the proprotin of high-value vacancies the jobseeker received before the experiment began. There is no difference in treatment effects by this variable.

Table C.11: Mechanism Analysis: Heterogeneous Treatment Effects on Applications by Baseline Experiences with the Platform

	Apply	
	(1)	(2)
Phone call treatment	0.01327	0.01395
	(0.00091)	(0.00166)
Any baseline matches	-0.00131	
	(0.00359)	
Phone call treatment × any baseline matches	-0.00015	
	(0.00159)	
Proportion of baseline matches that are high salary		0.00197
		(0.00106)
Phone call treatment × proportion of baseline matches that are high salary		-0.00565
		(0.00507)
Vacancy v is high salary for jobseeker j		0.00413
		(0.00039)
# matches	1,116,952	375,010
# jobseekers	9831	2129
Mean outcome	0.00814	0.00622
Sample	All matches	Jobseekers with baseline matches

Notes: This table shows how treatment effects uses initial exposure on the platform to test for beliefs as a mechanism for treatment effects. Column 1 regresses an indicator for application on the phone call treatment, an indicator for whether the jobseeker received any matches at baseline before the experiment started, during which all jobseekers received phone call initiation, and the interaction between the two. The sample is all matches during the experiment. Column 2 regresses an indicator for application on phone call treatment, the proportion of baseline matches that were high salary for jobseeker j (above the median in that jobseeker's distribution), and the interaction between the two. An indicator for whether the observation represents a high-salary vacancy is included as a control. The sample consists of matches during the experiment for the subsample of jobseekers who received any baseline matches and therefore phone call initiation at baseline. All regressions include stratification block fixed effects. The unit of observation is the jobseeker × vacancy match. Heteroskedasticity-robust standard errors are shown in parentheses, clustered by jobseeker.

Third, we use the "crossover" design described in Section 3.3, in which we randomly assigned some respondents from the control group to receive a phone call in some randomly selected matching rounds. This phone call is identical to the phone call received by the treatment group. This provides us another avenue to test for changes in beliefs. If respondents interpret the call as a one-time signal of unusually high potential returns only for the vacancies advertised at that time, we should expect the treatment effect to wear off with repetition as respondents update their interpretation of the call; thus the treatment effect should be larger on the "crossover group" than on the treatment group. Table C.12 shows the results. Randomly treating a subset of the control group produces similar effects to the long-run effects on the treatment group. This suggests that beliefs are not a likely explanation for our treatment effects on applications. This result also suggest that dynamics such as jobseekers becoming accustomed to the phone call and anticipating it are not an important factor in the design.

Table C.12: Effects of Short-term and Long-Term Treatment are Similar

	(1) Apply
Main phone call treatment	0.00284
	(0.00059)
Short term crossover treatment	0.00245
	(0.00078)
# matches	90016
# jobseekers	8053
# vacancies	115
Mean outcome | Main T = 0	0.00054
Mean outcome | Main T = 0, Crossover T = 0	0.00029

Notes: This table compares our main phone call initiation treatment, which continued over four years, with the effects of a short term "crossover" assignment of a randomized subset of the control group to receive treatment phone calls for one round each. The short-term treatment was implemented over five rounds close to the end of the main experiment; the estimation sample includes data from these rounds. 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.

Inaccurate beliefs about returns: Section 5.5 asked if our finding of constant returns could arise due to jobseekers searching as-good-as-randomly, with treatment simply leading to a greater quantity of random search and hence constant returns. Search that is not directed towards high-return vacancies could occur if jobseekers have inaccurate beliefs about potential returns. Section 5.5 shows that deliberately inducing random search produces lower marginal returns than the phone call treatment, helping to reject this explanation.

In addition, we show here that if jobseekers direct applications toward higher-value matches (as we have shown in Table 2) then jobseekers searching as-good-as-randomly with respect to the probability of an interview, P, cannot explain the treatment effects we observe. We show this using an informal proof by contradiction. We assume that jobseekers do search as-good-as-randomly with respect to P and show that for each possible correlation between P and the value of a match, V, this assumption yields a prediction that contradicts one of the treatment effects that we estimate:

1. roughly constant returns to marginal search in terms of P, shown in column 4 of Table 3, and

2. roughly constant returns to marginal search in terms of $P \cdot V$, shown in column 5 of Table 3.

First, suppose that P is uncorrelated with V (Table C.13, row 1). We know that jobseekers target vacancies with higher value V. So marginal applications should go to lower-value vacancies. This would lead to constant returns in terms of number of interviews P (consistent with result 1), but decreasing returns in terms of quality-weighted interviews $P \cdot V$ (contradicting result 2).

Second, suppose that P is negatively correlated with V (Table C.13, row 2). This could happen for example if there is simply more competition for high-salary jobs. Again, we know that jobseekers target vacancies with higher value V. Thus marginal applications should go to lower V vacancies, which would have higher P in this case. Thus marginal applications could lead to constant returns in terms of quality-weighted interviews $P \cdot V$ (consistent with result 2), but *increasing* returns in terms of number of interviews P (contradicting result 1), as the treatment encourages jobseekers to apply for more "realistic" job opportunities with lower salaries that they have a better chance of obtaining.

Finally, suppose that P is positively correlated with V (Table C.13, row 3). This could occur due to high match quality leading to a jobseeker preferring a job and also being preferred by the employer. As before, marginal applications should go to lower V vacancies, which would have lower P in this case. Thus marginal search should lead to decreasing returns in terms of both the number and quality of interviews (contradicting both results 1 and 2).

Thus, we conclude there is no pattern of correlation between P and V in which an increase in the quantity of search that is not directed towards P can rationalize our results.

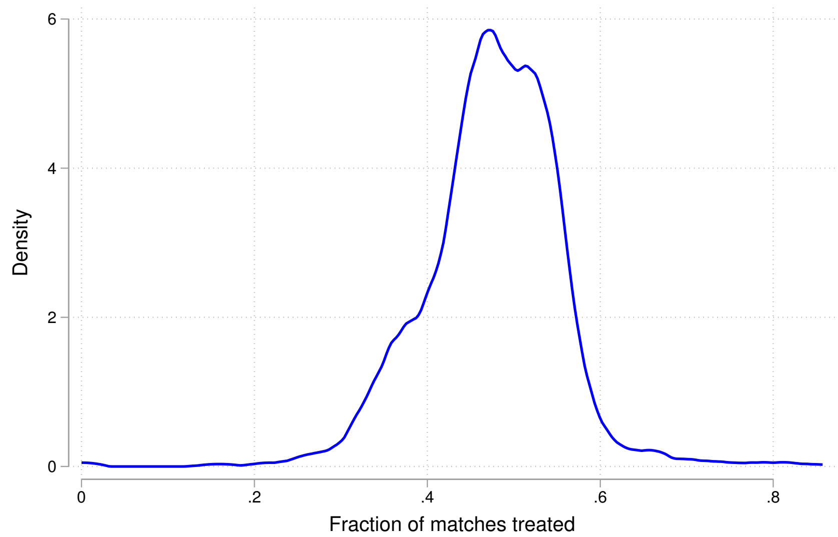
Table C.13: Returns to Marginal Applications Assuming Jobseekers Target V But Not P

Correlation of P and V	Returns to marginal applications in terms of:	
	number of interviews	value-weighted interviews
Uncorrelated	Constant	Decreasing
Negative	Increasing	Constant
Positive	Decreasing	Decreasing

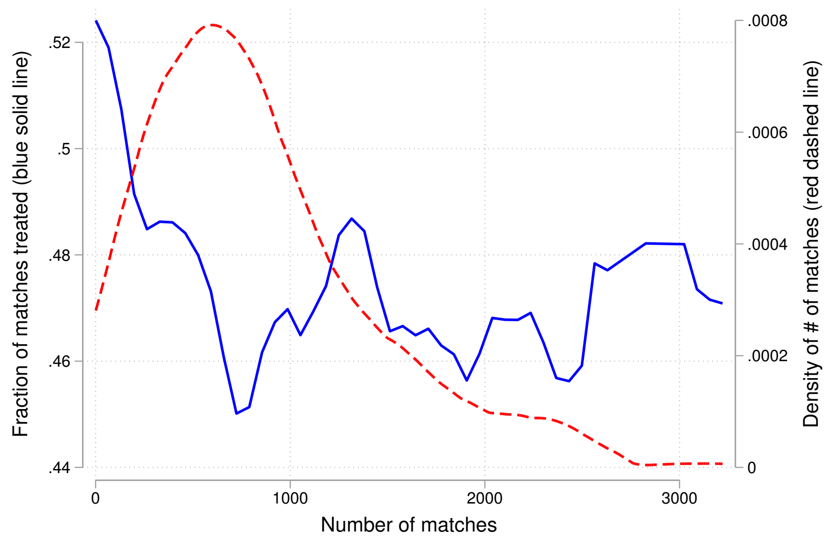
D Additional Analysis on Spillovers and Off-Platform Outcomes

Figure D.1: Variation in Treatment Rate Between Vacancies

Panel A: Density of Vacancy-Level Treatment Rate

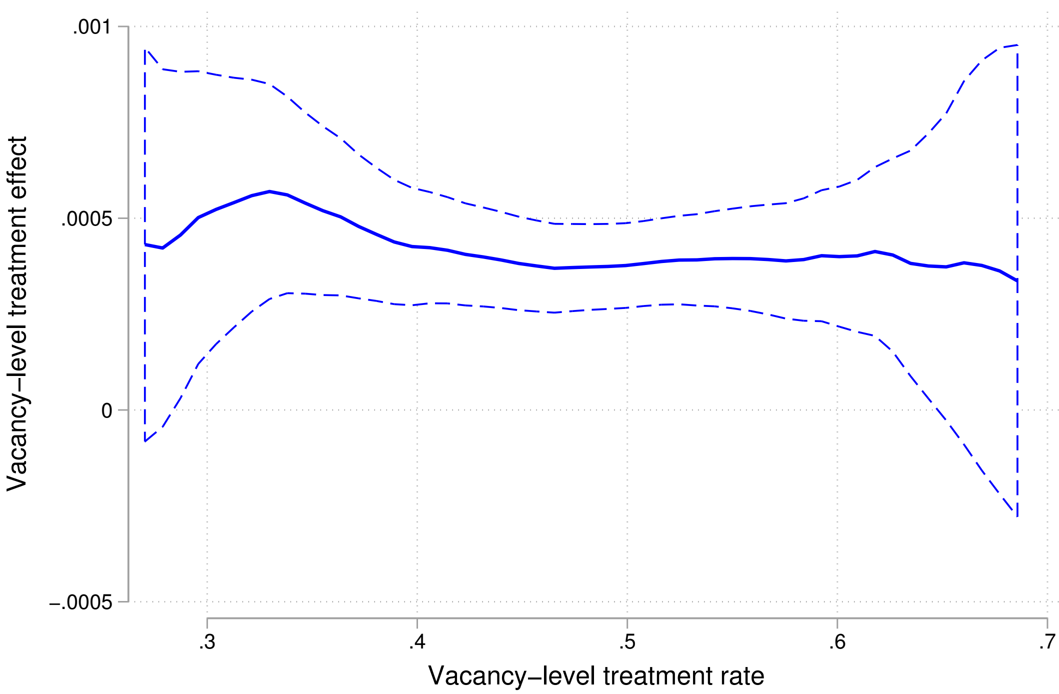


Panel B: Mean of Vacancy-Level Treatment Rate by Number of Matches



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

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



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

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

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

Notes: This table shows the relationship between the number of applications and interviews at the vacancy level, to contextualize the spillovers analysis in Section 6. Column (1) shows that vacancies get more applications if they are matched to more jobseekers and if more of these jobseekers are treated. Column (2) shows that vacancies that get more applications issue more interview invitations. Column (3) shows that the positive relationship between applications and interviews persists when we instrument the number of applications with the fraction of matched jobseekers who are treated, although the instrument is relatively weak and the second stage estimate is imprecise. Column (4) replicates column (2) but replaces the number of interviews with indicators for the middle and top terciles of the number of applications. Columns (5) and (6) replicate columns (2) and (4) but replace the number of interviews with an indicator for conducting any interviews as an outcome. Columns (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.

Table D.2: Treatment effects on Off-Platform Search (Intensive Margin)

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

Notes: This table shows treatment effects on specific off-platform search behaviors. The outcome in columns (1) and (2) is the number of applications submitted off-platform in the last 30 days and in columns (3) and (4) is the share of the following 7 search methods the respondent reported using: searching for clients, preparing CV or other related document, seeking assistance from friends or relatives, visiting employers, searching in newspaper/magazine/social media, contacting some organization, and other steps. Each outcome is regressed on an indicator for treatment assignment and stratification block fixed effects. Odd-numbered columns include selection adjustment terms for survey non-response as described in Section C.2, 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. Heteroskedasticity-robust standard errors shown in parentheses.