Time Is of The Essence: Experimental Evidence on The Effects of Job Application Timing*

Sai Zhang†

March 26, 2026

[Click for the latest version]

Abstract

Using a large-scale resume correspondence experiment (Kline, Rose and Walters, 2022), I provide novel causal evidence on the impact of job application timing. Holding applicant attributes fixed within vacancies, submissions in the first half are 1.16 percentage points more likely to receive contact, equivalent to 57% of the racial contact gap in the same setting. Submitting one day earlier raises contact probability by 0.40 percentage points. Using within-vacancy contact histories, I show that recruiters begin evaluating before the full pool arrives: initiating contacts sharply reduce the contact probability of comparable later applications, consistent with sequential screening and early stopping. These dynamics are group-specific and consistent with the interaction of quota stopping and learning. The early application advantage is larger in tighter labor markets. I formalize these patterns in a model of sequential screening with endogenous contact initiation, estimate it by simulated moments, and quantify economically meaningful speed-quality tradeoffs, with rushing losses of 6.55% in proxied match quality, equivalent to $3,088 per vacancy-year. Together, the results link within-vacancy recruiting behavior to macro labor market conditions and highlight application timing as an important margin in job searching.

*I am indebted to Vittorio Bassi, Simon Quach and Pablo Kurlat for their invaluable guidance and support. I am also grateful to Patrick Kline, Thomas Chaney, Juan D. Carrillo, Jeff Weaver, Erin Kaplan, Robert Garlick, Cailin Slattery, Isaac Sorkin, David Deming, Patrick Button, Michelle Jiang and John Klopfer for the advice and suggestions. I thank Patrick Kline for sharing the experiment data. I also thank participants at the 17th All-California Labor Economics Conference, PolMeth Europe 2025, WEAI 100th Annual Conference, 6th World Labor Conference, 20th Economics Graduate Student Conference, 3rd Annual AEA CSQIEP Mentoring Conference and 90th Annual MEA Conference for valuable discussion. All errors are my own.

†University of Southern California, Department of Economics. Email: saizhang@usc.edu.

1 Introduction

This paper asks whether application timing affects hiring outcomes. I leverage a large-scale resume correspondence experiment (Kline, Rose and Walters, 2022) that submits multiple applications to the same vacancy at pre-scheduled times while randomizing resume content, creating within-vacancy variation in timing that is orthogonal to applicant characteristics. I document a sizable *early application advantage*: earlier submissions are substantially more likely to receive an employer contact. The advantage reflects employer-side dynamics: recruiters begin evaluating before the full pool arrives, and once a vacancy enters its contact stage, later comparable applications are far less likely to receive a response. The early application advantage systematically correlates with labor market conditions such as tightness. These reduced-form patterns motivate a firm-side partial equilibrium model that formalizes sequential screening and endogenous contact initiation. The model incorporates market conditions to discipline how firms balance speed against selectivity, and quantifies firms' loss from rushing.

Timing is central to both sides of the labor market. Job seekers face a choice between applying immediately and spending additional time to tailor materials and improve application quality; recruiters, on the other hand, face a choice between beginning evaluation early to increase the chance of filling the vacancy and waiting to assemble a larger pool with more qualified applicants. These choices shape match quality, the distribution of opportunities across job seekers, and the speed at which vacancies are filled. Yet, the timing effect in job application is mostly anecdotal and not casually identified.

The empirical gap is not for lack of interest. A growing literature documents that job search is dynamic and time allocation matters for job-finding and employment trajectories (Bandiera, Jalal and Roussille, 2025), and related work measures the flow of applications and recruiting activity in modern hiring markets (Davis and de la Parra, 2024). However, pinning down the effect of *when* an application arrives, separately from *who* applies and *how* they apply, has been difficult. There are three core empirical challenges. First, in observational data, early applications may differ in motivation, information, or ability, and application quality itself can be a function of timing. Second, timing is often measured imperfectly and offers limited within-vacancy variation. Third, employer behavior is endoge-

nous and heterogeneous, creating selection on the employer side that confounds standard comparisons. Hence, identifying a causal "return to applying earlier" requires holding applicant characteristics fixed while generating meaningful within-vacancy variation in arrival time.

The correspondence experiment I use is well-suited to overcoming these challenges. It submits eight applications to each of more than 11,000 entry-level vacancies posted by large U.S. firms, with resume characteristics randomized by design and contacts observed consistently across jobs. By construction, the design provides clean within-vacancy comparisons of employer responses across application arrival times. The resulting time variation is large enough to support both flexible estimation and tests of employer-side dynamics that would be difficult to implement with observational application data. This setting allows me to ask a simple question with unusually clean identification: *holding applicant attributes fixed, does submitting earlier increase the probability of receiving employer contact?* I call this causal effect the *early application advantage*.

I organize the reduced-form evidence into three findings. First, there is a quantitatively meaningful early application advantage. Applications submitted among the first half of a vacancy's experimental submissions are 1.16 percentage points more likely to receive contact within 30 days. This estimate is equivalent to 57% of the racial contact gap in the same setting. The daily early application advantage relative to the vacancy's last experimental submission is 0.40 percentage points. The timing gradient is approximately linear over the design window and is robust to alternative measures of relative timing.

Second, the early application advantage reflects employer-side sequential evaluation and premature stopping. Employers contact quickly: conditional on being contacted within 30 days, the average time from submission to contact is about six days. Using the observed timing of contacts, I construct measures of prior contact history within a vacancy at the time each application arrives. Prior contacts substantially reduces the contact probabilility of later comparable applications (by 50% to 70%). The observed early application advantage is attenuated sharply by the constructed contact history measures. I also use group-specific contact histories to separate two forces that can generate sequential patterns, quota-style stopping versus learning, and show that sequential contact dynamics can differ across applicant groups.

Third, the early application advantage varies systematically with labor market conditions in ways that map naturally into firms' incentives to move quickly versus wait. In tighter local labor markets, measured by vacancies per unemployed worker, the daily early application advantage is larger, consistent with greater vacancy pressure inducing earlier screening and contact initiation. The strength of such relationship varies over time, suggesting that other aggregate market conditions shift recruiting incentives in ways that are not fully summarized by tightness.

To interpret these patterns and quantify their implications, I formalize them in a model of sequential screening and endogenous contact initiation. A vacancy receives a stochastic inflow of applicants over time, and reviewing an application yields a noisy signal of match quality. The firm's contact-initiation decision is an optimal-stopping problem: delaying contact increases the option value of sampling additional applicants, but waiting is costly. In the model, local tightness shapes the expected inflow of applicants and the effective cost of delay, while aggregate conditions shift the importance of tightness. This framework delivers a disciplined link between the reduced-form early application advantage and economically interpretable primitives governing hiring pressure, market structure, and screening incentives.

The estimated model fits the empirical timing distribution well and reproduces the central comparative statics: environments with higher vacancy pressure generate earlier contact initiation and a steeper early application advantage. A disciplined specification that ties the model's macro state to observed labor-market flow series attributes time variation in recruiting conditions primarily to hiring and separation dynamics, providing a tractable way to bring aggregate shifters into the micro contact-initiation problem.

To quantify the cost of initiating contact early, I develop a novel method that recovers relative matching quality from detailed resume information and revealed contact decisions. Since resume correspondence experiments usually do not observe signals of realized matching quality (e.g., wages), I construct a data-driven proxy for initial screening value using high dimensional machine-learning predictions of contact based on detailed resume and posting attributes. Combining this screening index with the model-implied initiation weights yields two complementary measures of timing-induced loss. In the baseline environment, conditional on initiation occurring on an experimental application, the probability that the best

experimental resume has not yet arrived by the initiation date is roughly 50%, and the expected loss from missing the best experimental resume is about 6.55% of the contacted resume's screening score. Using an external wage benchmark to translate this scale-free loss into dollars, the back-of-envelope wage-equivalent match-value loss is $3,088 annually across experimental regimes per vacancy. Macro counterfactual exercises show that these losses rise in tight, high churning markets, mirroring the reduced-form heterogeneity. Micro policy experiments show that A one-day mechanical delay in contact initiation meaningfully reduces the implied losses ($1,011), and improvement in within-day screening can recover a large share of mis-selection conditional on timing, but it leaves the timing component largely intact, underscoring that the main friction operates through *when* firms initiate contact relative to the evolution of the applicant pool.

The model decomposes two channels attributing to the early application advantages idenfied and produces testable predictions: application arrival intensity and shadow cost of waiting. I return to reduced-form heterogeneity and provide model-guided validation using independent measures of firms' hiring prominence, market structure, offer competitiveness, screening complexity, and recruiting organization. Across these dimensions, environments consistent with thinner pools, higher urgency, and higher screening cost exhibit a steeper early application advantage, while environments consistent with lower recruiting pressure or greater screening capacity exhibit a flatter timing gradient.

Taken together, this paper provides, to the best of my knowledge, the first causal evidence on the role of application timing in large-scale online hiring. Exploiting randomized application timing within vacancies, I document a sizable early application advantage and trace it to an employer-side decision margin that is typically unobserved: recruiters begin screening before the full applicant pool arrives, and initiating contact sharply reduces the likelihood that later, observationally identical applications are contacted. I then formalize these facts in a model of sequential screening with endogenous contact initiation that links the magnitude of the early application advantage to labor market tightness, while allowing aggregate hires and separateions to shift the recruiting environment over time. This approach bridges micro recruiting behavior and macro labor market dynamics, and it highlights a channel through which hiring practices can shape match efficiency: not only *who* applies and *where*, but also *when* evaluation begins and *how*

quickly firms stop waiting.

Contribution to Literature This paper contributes to several distinct literatures.

First, it provides causal evidence on a central but under-identified margin of job search: the return to applying earlier. There is a recent and growing literature in understanding timing in labor markets. Bandiera, Jalal and Roussille (2025) validate the concept of "critical window" with apzplicant side experiments, and show that job search timing is a key predictor of female employment, and incentiving earlier search reduces gender gap by a third. This paper aims to provide evidence on the firm side, answering the question why applying earlier is linked to better chance of employment. It also echoes (Cortés et al., 2023), who focus on the last step of job search and show that offer acceptance timing is a key element in explaining college graduates' employment. Davis and de la Parra (2024) uses observational online application data to show that job seekers act relatively fast, which validates the design of this study.

Second, the paper provides direct evidence on employer recruiting dynamics. The hiring process is often summarized in macro and empirical work by reduced-form job-filling hazards (Mortensen and Pissarides, 1994; Shimer, 2005; Davis, Faberman and Haltiwanger, 2013; Mueller et al., 2024). This paper shows how contact decisions evolve within a vacancy, linking these micro recruiting dynamics to broader labor-market forces. It echoes recent studies in understanding firm behaviors and labor market structure (Dodini and Willén, 2025; Jarosch, Nimczik and Sorkin, 2024; Cederlöf et al., 2025; Kline, Rose and Walters, 2024), and microfounding labor market frictions through the lens of dyanmic searching (Birinci, See and Wee, 2025; Birinci, Mercan and See, 2025; Bradley, 2022; Engbom, 2021; Cai, Gautier and Wolthoff, 2025; Tekleselassie et al., 2025). By mapping firms' screening strategies to aggregate market conditions, the paper connects experimental evidence to the economic forces emphasized in modern search models and to empirical work on market structure and recruiting power.

Third, this paper contributes to the literature on belief-based discrimination and firm hiring practices. A growing body of work studies how potentially biased belief drives discrimination (Bordalo et al., 2016; Bohren, Imas and Rosenberg, 2019; Bohren et al., 2023; Coffman, Exley and Niederle, 2021), and how information can help correct such bias (Vuletich and Payne, 2019; Alesina et al., 2024).

Some recent work focuses on the endogenous acquisition of information. Cavounidis, Lang and Weinstein (2024) find that employers are more likely to monitor and then punish Black employees; Eyting (2022) shows in a series of online experiments that individuals incline to acquire information matched their motivation to justify discriminatory behavior; Coffman, Kostyshak and Saygin (2024) shows that endogenous information acquisition leads to under-acquisition and misevaluation in a stylized hiring experiment. Using field data in the financial advisory industry, Egan, Matvos and Seru (2022) show female advisers are more severely punished for misconduct. Leveraging the large-scale lab-in-the-field resume correspondence experiment, this paper provides novel evidence on the belief updating of recruiters and its interaction with the competing quota mechanism in the real-world hiring decision-making process.

Last, this paper speaks to the audit study design literature (Gaddis, 2018; Neumark, 2018; Kessler, Low and Sullivan, 2019). I identify timing as a dimension that could potentially bias the estimation. Correspondence experiments in similar settings where an optimal stopping structure is applicable for the receiving side should consider controlling for timing in the design.

Structure of the Paper. The remainder of the paper is organized as follows. Section 2 describes the experiment and complementary datasets. Section 3 defines timing measures and estimates the early application advantage. Section 4 constructs contact history measures and documents the sequential mechanism. Section 5 relates the timing gradient to tightness and motivates the model. Section 6 presents the sequential screening model, and Section 7 estimates it and conducts counterfactual exercises. Section 8 presents model-guided heterogeneity patterns. Section 9 concludes.

2 Empirical Context

The main analysis of this paper draws on the large-scale resume correspondence experiment conducted by Kline, Rose and Walters (2022). A key feature of this setting is that each sampled vacancy received multiple applications submitted over a nontrivial window of time, generating within-vacancy variation in application timing that is central to my empirical strategy.

This section summarizes the experiment and the complementary labor market data used to study heterogeneity in timing effects. Appendix Section E provides a detailed account of data sources, matching, and variable construction.

2.1 Experiment

The sampling frame begins with the 2018 Fortune 500 list. To ensure the feasibility of audit methods, the experiment selected 108 companies of distinct brands with proprietary hiring portals that posted a large number of entry-level vacancies with sufficient geographic dispersion. These firms spanned a sufficiently wide variety of industries and collectively employed around 15 million workers in 2020.

From each employer, 125 entry-level vacancies were sampled over five waves, with 25 jobs per wave, from October 2019 to April 2021. Each sampled vacancy corresponds to an establishment in a distinct U.S. county. A majority of firms (72) participated in all waves; some were excluded from the first wave because of COVID-19 disruptions and others from the last wave because of technological barriers. The order in which firms entered the sampling queue was randomized at the beginning of each wave.

For each sampled firm, vacancies were scraped daily from the firm's hiring portal. The study focused on entry-level jobs that did not require a bachelor's or advanced degree, substantial prior experience, or a specialized license. The *RandRes* platform monitored the scraped vacancies and added 5-10 jobs per firm at a time to match the daily submission capacity.

For each sampled vacancy, up to eight applications with randomized characteristics were submitted in four pairs, with each pair consisting of one Black applicant and one White applicant. Some vacancies stopped accepting applications before all eight submissions were completed. The final sample includes approximately 84,000 applications for about 11,000 vacancies. For a given vacancy, the submission procedure typically concluded within eight days (94% of vacancies), with a gap of one to two days between consecutive pairs. This feature of the design generates rich within-vacancy variation in application timing.

Following the resume correspondence experiment literature (Bertrand and Duflo, 2017), the experiment signaled race and gender using distinctive names. In addition, three other potential discriminatory dimensions were experimentally var-

ied: applicants were assigned an age drawn uniformly from 22 to 58; 10% of applicants listed LGBTQ-related high school clubs to signal sexual orientation; and another 10% listed pronouns, split evenly between gender-neutral (they/them) and gender-typical pronouns. A larger set of resume attributes was also randomized, including whether the applicant had an associate degree (50%) instead of a high school diploma, and employment histories comprising two or three jobs with relevant experiences.

As in most correspondence experiments, the main outcome is whether an employer attempted to contact the applicant. I focus on contacts occurring within 30 days after submission. Each fictitious applicant was assigned a unique phone number and email address, both monitored to record employer callbacks and interview requests. A key strength of the experimental design is its comparatively high contact rate: employers contacted 24% of submitted applications. Additional details on the experimental design are provided in Kline, Rose and Walters (2022).

2.2 Complementary Data

A central goal of this paper is to interpret application timing effects through the lens of market conditions. To do so, I assemble a set of complementary covariates measured at the state, industry, and firm levels.

First, I compile local labor market conditions at the state-by-month level. These measures include monthly indicators of job openings, hires, separations, and labor market tightness from Job Openings and Labor Turnover Survey (JOLTS), as well as pre-experiment income levels, mobility measures and time-varying minimum wage regulations.

Second, I collect industry-level covariates, including task-based measures of skill requirements and job content constructed from O*NET, baseline race and gender employment composition, and monthly measures of industry employment and compensation conditions.

Third, I assemble firm-level covariates capturing local labor market power and organizational characteristics, including human resource management practices, employee satisfaction, and financial performance. The primary data source is Revelio Labs (Revelio Labs, 2026). Revelio Labs aggregates publicly available workforce information from online professional profiles, job postings, and employee

reviews into longitudinal firm-level datasets. Revelio Labs' coverage spans a large share of U.S. online job postings and professional profiles, allowing consistent measurement of firm-level recruiting prominence across states and industries over time.[1]

I use the job postings data to construct measures of local labor demand and firm-level market power. Specifically, I aggregate posting records to the local labor market (state-month or state-industry-month) levels to compute posting shares of firms and concentration measures within local labor markets.[2] These measures proxy for a firm's relative position in the local hiring market. I also use employee review and sentiment data to proxy for workplace characteristics and organizational quality.[3] For each firm, I construct 6-month rolling averages of review-based ratings prior to the experimental submission month.

Additional firm-level covariates, including financial performance and employment size, are drawn from Compustat.

These datasets support the heterogeneity analysis of timing effects across firms, places, and labor market environments. Appendix Section E provides full details on sources, matching, and construction of all covariates.

3 Estimating Early Application Advantages

This section estimates the causal effect of applying earlier within a vacancy on employer contact. Applying earlier may increase the probability of employer contact if firms screen applications sequentially rather than waiting for a complete pool. In observational data, however, application timing is confounded by applicant selection, unobserved quality and preparation. In the experiment, multiple experimental applications with randomized resume characteristics are submitted over several days to each vacancy. This design allows me to isolate the effect of relative submission timing using within-vacancy comparisons.

I begin by defining the timing measures used throughout the paper and de-

[1]Revelio Labs, Data Dictionary documentation (accessed February 24, 2026): https://www.data-dictionary.reveliolabs.com/data.html.

[2]Revelio Labs, *Job Posting Dynamics* and *Individual Job Postings* datasets (accessed February 24, 2026).

[3]Revelio Labs, *Sentiment* datasets, including *Individual Reviews* and *Sentiment Scores* (accessed February 24, 2026).

scribe the identification strategies. I then document the magnitude and shape of the early application advantage. Finally, I examine whether the timing premium differs by resume characteristics.

3.1 Measuring Application Timing

For each vacancy, I observe the submission date and time of every experimental application. I use this information to construct three measures of application timing.

First, I define $Order_i \in \{1, \ldots, 8\}$, the realized within-vacancy order of submission. The sample contains 83,643 applications submitted to 11,114 vacancies. Most vacancies (88.4%) received the full set of eight experimental submissions (*Complete* hereafter). The distribution of realized submission orders and the prevalence of vacancies with less than 8 submissions (*incomplete* hereafter) are summarized in Table A.1.

Second, my primary timing indicator classifies an application as early if it arrives among the first four submissions to the vacancy,

$$early_i^1 \equiv \mathbf{1}\left(Order_i \leq 4\right),$$

with all submissions coded as late for vacancies that received at most four experimental applications (836 vacancies, 7.52%; 2321 submissions, 2.77%).[4] In the full sample, 41112 submissions (49.15%) are coded as early submissions.

Finally, I construct a continuous timing measure based on calendar distance to the last experimental submission to the vacancy. Let $D2L_i$ denote the number of days between application i and the last experimental submission to the same vacancy. As shown in Figure A.1, the distribution of $D2L_i$ is right-skewed, with 99.3% of all submissions having no more than 10 days of separation from the last submission for each job. I therefore use a trimmed version $D2L_i^t$ that top-codes observations beyond the intended design window in the main analysis.[5]

[4] Vacancies with at most four experimental applications closed during the experiment, it is therefore natural to classify such applications as late. I also consider two alternative early indicators as robustness checks: $early^2$ defines early submissions as those not among the last four realized submissions, and $early^3$ defines early submissions using the designed pair order. All alternative timing definitions and related diagnostics are discussed in Appendix Section F.1.

[5] By design, the days to the last submission should be at most 6, as the 8 applications for each job were planned to be submitted in 4 pairs either every other day (Wave 1) or daily (Wave 2-5).

Measuring application timing is central in the analysis, I therefore consider a range of alternative constructions and sample restrictions as robustness checks. Appendix Section F.2 reports the full set of robustness results.

3.2 Empirical Strategy

The outcome Y_i is an indicator for whether application i receives employer contact within 30 days of submission. I estimate within-vacancy timing effects by including job fixed effects $\lambda_{j(i)}$ throughout. Identification comes from random assignment of resume characteristics to submission slots within a vacancy, so timing measures are orthogonal to the randomized resume vector $\mathbf{X}_i$ by design. Table A.2 confirms balance.

I begin with the early indicator specification,

$$Y_i = \alpha_0 + \beta \cdot early_i^1 + \boldsymbol{\gamma}\mathbf{X}_i + \lambda_{j(i)} + \epsilon_i, \tag{3.1}$$

where $early_i^1$ is the constructed early application indicators.

To trace the timing profile more flexibly, I estimate an order-of-submission specification,

$$Y_i = \alpha_0 + \sum_{k=2}^{8} \beta_k \cdot \mathbf{1}\left(Order_i = k\right) + \boldsymbol{\gamma}\mathbf{X}_i + \lambda_{j(i)} + \epsilon_i, \tag{3.2}$$

where β_k can be interpreted as the contact gap between the kth submission and the first submission within the same vacancy, conditional on controlled characteristics.

Finally, I translate timing into days using the continuous measure $D2L^t$ as the regressor, and estimate

$$Y_i = \alpha_0 + \delta \cdot D2L_i^t + \boldsymbol{\gamma}\mathbf{X}_i + \lambda_{j(i)} + \epsilon_i. \tag{3.3}$$

where the coefficient δ measures how the contact probability changes with one additional day relative to the last submission. Model (3.3) assumes a linear effect of $D2L^t$ on contact probability. To test this assumption, I also estimate a nonparamet-

Following this time frame, I regroup all observations with $D2L$ larger than 6 and code them as 7, creating the top-coded measure $D2L^t$. Figure A.2 plots its distribution. In Appendix Section F.2, I discuss the design differences between Wave 1 and Wave 2-5 and the robustness of the relevent results.

ric analogue,

$$Y_i = \alpha_0 + \sum_{d=1}^{7} \delta_d \cdot \mathbf{1}\left(D2L_i^t = d\right) + \boldsymbol{\gamma}\mathbf{X}_i + \lambda_{j(i)} + \epsilon_i, \tag{3.4}$$

where coefficients $\{\delta_d\}_{d=1,\cdots,7}$ flexibly capture the potential functional form.

In all specifications, I cluster standard errors at the vacancy level to allow for within-vacancy dependence and heterogeneous treatment effects (Abadie et al., 2023). Appendix Section F.2 reports additional specification checks.

3.3 Evidence of Early Application Advantages

Column (1) of Table 1 reports estimates of Equation 3.1 on the full sample. Early submissions are 1.16 percentage points more likely to be contacted. Column (2) of Table 1 shows that the estimate is stable on the subset of vacancies with fewer than eight experimental submissions.[6] Relative to the average contact rate (24%) in the setting, early submissions gain an early application advantage of approximately 5%. This contact gap is economically significant, equivalent to around 57% of the racial contact gap (2.05 p.p.) in the same setting. Regarding resume characteristics, consistent with Kline, Rose and Walters (2022), applicants over 40 experience significant contact penalties. Listing pronouns, whether gender-typical or gender-neutral, leads to a marginally significant lower contact rate. Other characteristics, including gender (female versus male), do not have a significant impact in this context.

The results indicate a significant early application advantage within vacancies in this experimental setting conditional on comparable resume characteristics. I next show how the contact probability evolves over the submission sequence.

Figure 1 plots the estimated coefficients of interest $\hat{\beta}_k$ from Model (3.2). Within a vacancy, the first submission is the most likely to be contacted. The order effect is weakly monotonic: later submissions are not more likely to be contacted than earlier ones. A second feature is a pronounced decline (around 1 percentage point) in the contact rate between the fourth and fifth submissions. After this drop, the

[6] An important observation is that the constant for the *incomplete* job subsample is significantly lower than that of the full sample, supporting the existence of the early application advantage since by design, submissions to the *incomplete* jobs can be considered inherently late, as they occur near the closing of these vacancies.

Table 1: Estimation Results of Model (3.1) and (3.3)

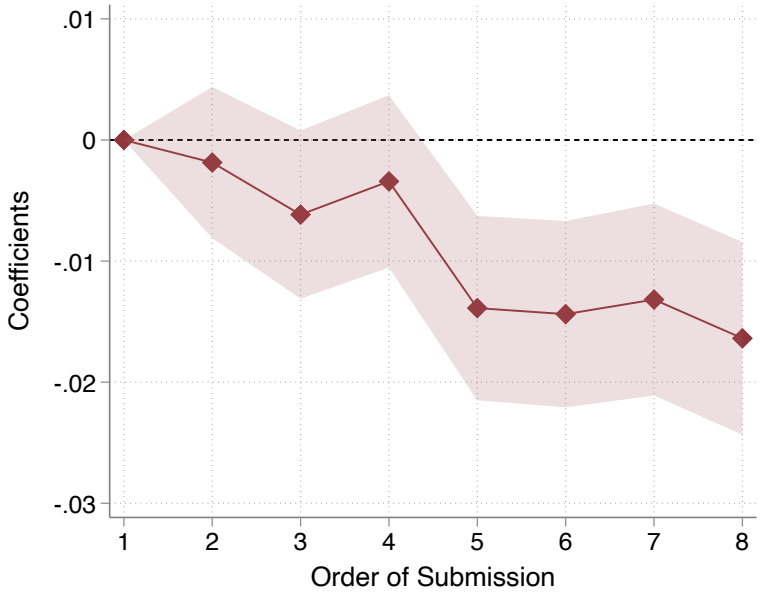
	(1) Full	(2) < 8	(3) Full	(4) < 8
$early_i^1$	0.0116	0.0240		
	(0.0024)	(0.0127)		
$D2L_i^t$			0.0040	0.0087
			(0.0008)	(0.0035)
Black	-0.0205	-0.0125	-0.0205	-0.0125
	(0.0018)	(0.0066)	(0.0018)	(0.0066)
Female	-0.0021	-0.0006	-0.0021	-0.0007
	(0.0022)	(0.0080)	(0.0022)	(0.0080)
Over 40	-0.0043	0.0041	-0.0043	0.0034
	(0.0020)	(0.0083)	(0.0020)	(0.0083)
Pol. club	0.0059	0.0037	0.0060	0.0026
	(0.0051)	(0.0251)	(0.0051)	(0.0253)
Acad. club	0.0049	0.0136	0.0050	0.0136
	(0.0051)	(0.0175)	(0.0051)	(0.0174)
LGBTQ club	-0.0003	0.0152	-0.0002	0.0154
	(0.0036)	(0.0171)	(0.0036)	(0.0171)
Typical pron.	-0.0077	0.0056	-0.0078	0.0040
	(0.0051)	(0.0203)	(0.0051)	(0.0202)
Neutral pron.	-0.0086	0.0065	-0.0085	0.0061
	(0.0051)	(0.0217)	(0.0051)	(0.0216)
Assoc. degree	0.0021	0.0082	0.0021	0.0083
	(0.0020)	(0.0085)	(0.0020)	(0.0085)
Constant	0.2475	0.1680	0.2453	0.1668
	(0.0025)	(0.0097)	(0.0027)	(0.0093)
Job FEs	✓	✓	✓	✓
N	83643	5011	83643	5011

Note: This table presents the effects of early submission and other characteristics on the probability of employer contact within 30 days. Column (1) and (2) estimate Model (3.1), Column (3) and (4) estimate Model (3.3). Columns titled "Full" include all jobs, while columns titled "<8" only include jobs with less than 8 submissions. The construction of $early_i^1$ and $D2L_i^t$ is described in 3.1. All regressions include job fixed effects. Standard errors in parentheses are clustered at the job level.

contact rate remains relatively stable for subsequent submissions.

So far, I have provided causal evidence of the existence and magnitude of the contact reward for early submissions using measures of relative application timing. As discussed earlier, the submission process of the experiment differs by waves. Expressing the early application advantage in days therefore makes the

Figure 1: Contact Gap and Order of Submission



Notes: The figure plots the estimated coefficients $\hat{\beta}_k, k \in \{2, \cdots, 8\}$ of Model 3.2 with the order of submission on the x-axis, controlling for resume characteristics. The regression includes job fixed effects, the standard errors are clustered at the job level. The area depicts the 95% confidence intervals. Sample: full sample.

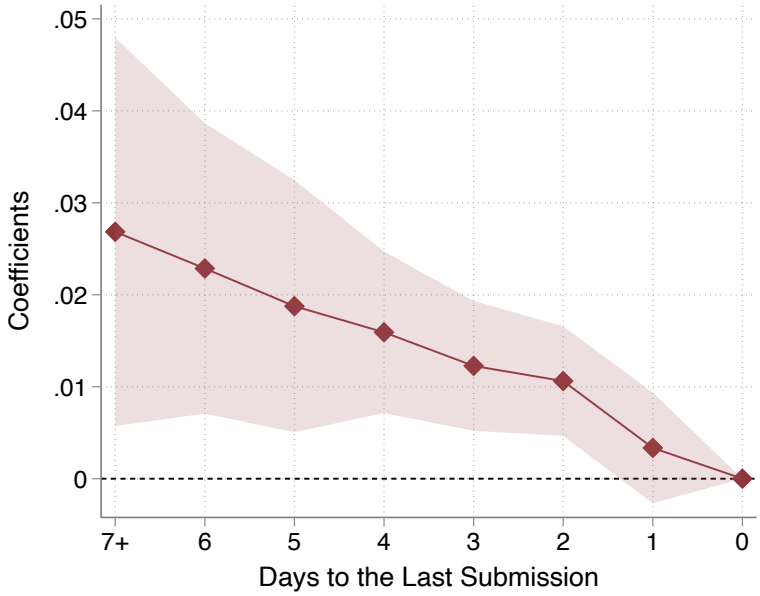
interpretation more consistent. Column (3) of Table 1 reports the estimation results with the timing measure $D2L^t$ as a continuous regressor on the full sample: compared to the last application submitted, submitting one day earlier to the same vacancy increases the contact rate by 0.40 percentage points, conditional on resume characteristics. Column (4) shows the robustness of this result for *incomplete* jobs.

As shown in Figure A.2, Wave 1 and Wave 2-5 differ in submission timing design. To address this, I re-run the estimation on the subsamples of Wave 1 and Wave 2-5 separately. As summarized in Column (3) to (6) of Table A.3, the full-sample results hold robustly in both subsamples. The estimated effect of submitting one day earlier than the last submission to a job is 0.36 percentage points in Wave 1 and 0.43 percentage points in Wave 2-5.

To test the linearity assumption in Model (3.3), I estimate the nonparametric Model (3.4) and plot the estimated coefficients $\hat{\delta}_d \in \{1, \cdots, 7\}$.

Figure 2 confirms that the early application advantage rises with $D2L_i^t$ approximately linearly on the full sample. When examining Wave 1 and Waves 2 to 5 separately in Figure A.4, the early application advantage remains pronounced and

Figure 2: Contact Gap and Days to the Last Submission



Notes: The figure plots the estimated coefficients $\hat{\delta}_d, d \in \{1, \cdots, 7\}$ of Model 3.4 with days to the last submission $D2L^t$ on the x-axis, controlling for resume characteristics. The regression includes job fixed effects, the standard errors are clustered at the job level. The area depicts the 95% confidence intervals. Sample: full sample.

the linear trend is generally robust. Since $D2L^t$ is constructed as the days elapsed backwards with the last submission to each job as the benchmark, *incomplete* jobs that were closed in the middle of the experiment process of submitting all 8 applications might exhibit different trends. Figure A.5 shows that, in addition to the robust significance of the early application advantage, the weak monotonicity still holds for the most part, but the estimations are noisier due to a much smaller sample size.

3.4 Timing and Resume Characteristics

Because timing varies within vacancies and is orthogonal to resume content, the early application advantage is naturally interpreted as employer-side behavior. Still, timing could matter differently across demographic groups if recruiters apply screening rules asymmetrically over the course of evaluation.

In Table 2, I examine whether the effects of various resume characteristics differ for early ($early_i^1 = 1$) and late applications. A joint test cannot reject the null

Table 2: Early Application Interacted with Resume characteristics

	(1) Late $early^1 = 0$	(2) Early $early^1 = 1$	(3) Diff. $(2) - (1)$
Black	-0.0148	-0.0157	-0.0009
	(0.0037)	(0.0036)	(0.0051)
Female	0.0012	0.0052	0.0040
	(0.0040)	(0.0041)	(0.0056)
Black × Female	-0.0089	-0.0123	-0.0034
	(0.0055)	(0.0055)	(0.0078)
Over 40	-0.0028	-0.0059	-0.0031
	(0.0028)	(0.0029)	(0.0040)
Pol. club	0.0046	0.0072	0.0026
	(0.0070)	(0.0074)	(0.0101)
Acad. club	0.0088	0.0012	-0.0076
	(0.0069)	(0.0072)	(0.0099)
LGBTQ club	0.0039	-0.0046	-0.0085
	(0.0051)	(0.0050)	(0.0071)
Typical pron.	-0.0043	-0.0112	-0.0069
	(0.0071)	(0.0073)	(0.0101)
Neutral pron.	-0.0106	-0.0067	0.0039
	(0.0072)	(0.0072)	(0.0102)
Assoc. degree	0.0044	-0.0001	-0.0045
	(0.0028)	(0.0029)	(0.0040)
Constant	0.2428	0.2585	0.0157
	(0.0035)	(0.0035)	(0.0051)
Job FEs	✓	✓	
N	42531	41112	83643
$F-$ stats for joint significance			0.58
$p-$value			0.8344

Note: This table presents the effects of early submission interacted with resume characteristics on the probability of employer contact within 30 days. Column (1) displays the results for late applicants (not the first 4 submissions to a job), Column (2) displays the results for early applicants (the first 4 submissions to a job), Column (3) shows the difference between early and late applicants. All regressions include job fixed effects. Standard errors in parentheses are clustered at the job level. $F-$ statistics and joint $p-$values come from tests that all differences in reported coefficients other than the constant term are zero. Sample: full sample.

hypothesis of no interaction effects ($F = 0.58$).

I then focus on the two key discriminatory factors in the literature, race and gender, and test how they interact with various application timing measures. Table A.4 shows that black applicants experience a smaller early application advan-

tage, while female applicants experience a larger early application advantage comparing to their male counterparts. The triple interaction design results show that black female applicants face a significant contact penalty compared to all other demographic groups and also experience the smallest early application advantage, making them the most vulnerable group of job seekers in this setting. The results are robust to the choice of application timing measures but statistically insignificant. In Figure A.6 and A.7, I visualize how the race and gender contact gaps evolve over time within vacancies, measured by submission order $\mathbf{1}\left(Order_i = j\right)$, $j = 2, \cdots, 8$ and days to last submission $\mathbf{1}\left(D2L_i^t = d\right)$, $d = 1, \cdots, 7$, respectively. The patterns largely align with the results reported in Table A.4 but the variation in neither racial nor gender contact gap over time is statistically significant.

Taken together, I find limited evidence that the early application advantage differs systematically by race or gender. I return to group-specific dynamics in Section 4.3, where I take advantage of the within-vacancy sequential structure of evaluation and provide sharper tests of underlying mechanisms.

4 Mechanism: Early Evaluation and Premature Stopping

Section 3 documents a robust early application advantage: within the same vacancy and holding resume characteristics fixed, earlier submissions are consistently more likely to receive employer contact. A natural hypothesis is that employers begin evaluating applications soon after they arrive and that early evaluation changes the environment faced by later applicants. If recruiters contact candidates as they screen and become less attentive once enough candidates have been identified, then later submissions can arrive after key contact decisions have already been made.

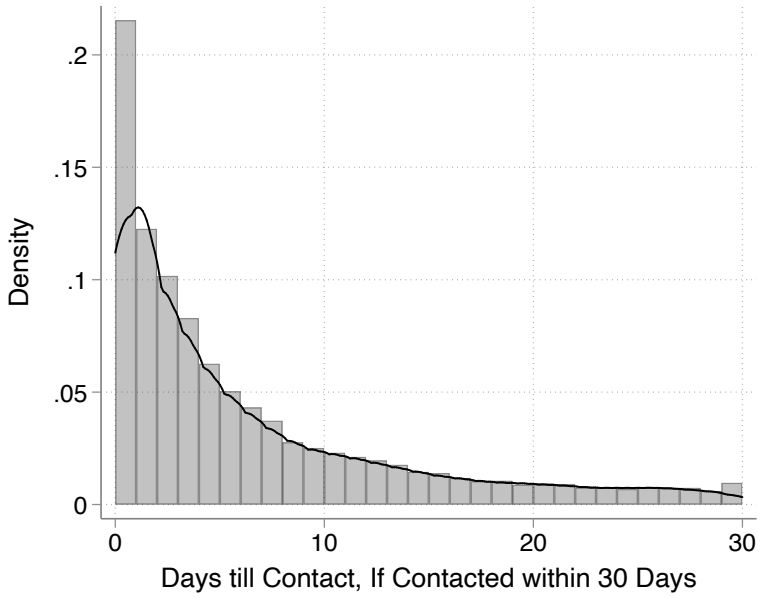
This section evaluates that hypothesis with the timing of observed employer contacts. I first show that contacts occur quickly relative to the experimental submission window. I then construct measures of *prior contact history* at the time each application is submitted and show that this history strongly predicts subsequent contact and largely absorbs the early application advantage identified ealier. Finally, I use race- and gender-specific contact histories to test whether the sequential

pattern is more consistent with quota stopping or learning-style belief updating.

4.1 Recruiters Contact Quickly

Figure 3 plots the distribution of the number of days between submission and contact, conditional on the application being contacted within 30 days. Contacts occur quickly in this setting: among the 20,102 applications that receive contact within 30 days, the mean time from submission to contact is 6.04 days. A substantial share of contacts occurs very early: 21.55% on the day of submission, 52.28% within three days, and 67.91% within six days.

Figure 3: Days from Submission till Contact, If Contacted



Notes: The figure plots the distribution of days from submission to contact given a submission was contacted within 30 days after submission. The epanechnikov kernel is used, the kernel half-width is set as 1.

These response times are short relative to the experimental cadence, under which the sequence of experimental submissions to a vacancy typically spans several days. As a result, for many vacancies, later experimental submissions plausibly arrive after the employer has already made at least some contact decisions. I therefore construct the contact history within vacancies prior to a submission based on this observed pattern.

4.2 Contact History and Early Application Advantages

Constructing contact history. The data report the day on which a contact occurred, but not the time of day. Because multiple contacts may occur on the same day,[7] and some applications are contacted on their day of submission, it is not always possible to order contacts precisely within a day. To accommodate this limitation, I construct upper- and lower-bound measures of the number of other experimental applications to the same vacancy that have already been contacted when application i is submitted.

Let $day(i)$ denote the calendar day of submission for i. The upper bound, NoC_i^{up}, counts all other applications to the vacancy that are contacted *no later than* $day(i)$, including contacts on the same day. The lower bound, NoC_i^{low}, counts all other applications to the vacancy that are contacted *at least one day* before $day(i)$. Table A.5 illustrates the construction of the measures with an example.

By construction, $NoC_i^{low} \leq NoC_i^{up}$. In Table A.6, I summarize the two measures by the submission order of i. On Average, there is a positive number of submissions contacted prior to the submission of i, and the number is monotonically bigger for later submissions. Table A.7 presents the pairwise correlations between the constructed contact history measures and various application timing measures.

Contact history drives early application advatnages. After constructing the contact history for a job prior to submission i, I reestimate the baseline timing specifications from Section 3.2, adding the two constructed contact history measures NoC_i^{up} and NoC_i^{low} as an additional control. The estimating equation takes the form

$$Y_i = \alpha_0 + \beta \cdot timing_i + \eta \cdot NoC_i + \boldsymbol{\gamma}\mathbf{X}_i + \lambda_{j(i)} + \epsilon_i, \tag{4.1}$$

where Y_i indicates contact within 30 days, $timing_i$ is one of the timing measures, $NoC_i \in \{NoC_i^{up}, NoC_i^{low}\}$. Resume characteristics $\mathbf{X}_i$ and job fixed effects $\lambda_{j(i)}$ are included. Standard errors are clustered at the job level.

Table 3 reports the results. Across specifications, prior contacts substantially reduce the probability that the current application is contacted. The implied magnitudes are large: an additional prior contact is associated with a 4.4-5.9 percentage

[7] Among all contacted submissions, 23.35% were the only one contacted on a specific day.

Table 3: Prior Contact History and Early Submission Advantages

	(1)	(2)	(3)	(4)	(5)	(6)
	Model (3.1)		Model (3.2)		Model (3.3)	
NoC^{up}	-0.0435		-0.0445		-0.0440	
	(0.0024)		(0.0024)		(0.0024)	
NoC^{low}		-0.0574		-0.0587		-0.0581
		(0.0028)		(0.0029)		(0.0029)
$early^1$	-0.0078	-0.0099				
	(0.0022)	(0.0022)				
$\mathbf{1}(Order = 2)$			0.0001	-0.0010		
			(0.0032)	(0.0032)		
$\mathbf{1}(Order = 3)$			0.0018	0.0004		
			(0.0035)	(0.0035)		
$\mathbf{1}(Order = 4)$			0.0060	0.0045		
			(0.0036)	(0.0036)		
$\mathbf{1}(Order = 5)$			0.0041	0.0041		
			(0.0038)	(0.0038)		
$\mathbf{1}(Order = 6)$			0.0057	0.0056		
			(0.0038)	(0.0038)		
$\mathbf{1}(Order = 7)$			0.0163	0.0182		
			(0.0038)	(0.0039)		
$\mathbf{1}(Order = 8)$			0.0155	0.0182		
			(0.0039)	(0.0039)		
$D2L^t$					-0.0030	-0.0038
					(0.0007)	(0.0007)
Controls $\mathbf{X}_i$	✓	✓	✓	✓	✓	✓
Job FEs	✓	✓	✓	✓	✓	✓
N	83643	83643	83643	83643	83643	83643

Note: This table presents the effects of prior contact histories and early submission, controlling for resume characteristics on the probability of employer contact within 30 days. NoC_i^{up} refers to the number of other submissions already contacted no later than the day when a submission was made; NoC_i^{low} refers to the number of other submissions already contacted at least one day before a submission was made. $early^1$ is the constructed early submission dummy; $\mathbf{1}(Order = j)$, $j = 2, \cdots, 8$ are the indicators of the submission orders; $D2L^t$ is the number of days before the last submission to a job. All regressions include resume characteristics $\mathbf{X}_i$ and job fixed effects. Standard errors in parentheses are clustered at the job level.

point lower contact probability, depending on the history measure. More importantly, adding prior contact history substantially attenuates the early application advantages, and in some specifications the coefficient on $early_i$ or $D2L_i^t$ changes sign. This pattern is consistent with sequential evaluation: later applications are

less likely to be contacted primarily because earlier applications have already generated contacts to move to the next step of the hiring process and, as a result, the marginal likelihood of contacting additional candidates declines even if the remaining applicants are equally qualified.

4.3 Implications for Discrimination Dynamics

The sequential structure also provides a way to study how discrimination evolves over the course of screening. In Table A.8, I show how the overall contact history affects subsequent applications of different races and genders differently, controlling for resume characteristics, job fixed effects and submission order fixed effects. The contact-history penalty is larger for Black applicants: the interaction $NoC \times Black$ is negative and statistically significant. In contrast, the interaction with gender is small and imprecise.

To furthur identify group-specific dynamics, I decompose contact history by demographic groups. For each application i of group g defined by race or gender, I construct $History_i^g$ as the number of previously contacted applications from group g up to the submission day of i, using NoC^{up} or NoC^{low}. This decomposition distinguishes two candidate mechanisms: *quota* versus *learning*. If employers apply a quota stopping rule, prior contacts of candidates from group g should reduce the probability of contacting subsequent candidates from the same group.[8] In contrast, if recruiters learn about group quality as they screen, holding resumes fixed by randomization, then prior contacts of group g is a positive signal that can non-negatively change the propensity to contact subsequent candidates from g.[9]

[8] As suggested by a large body of theoretical and empirical literature (Welch, 1976; Lundberg, 1991; Coate and Loury, 1993; Fang and Moro, 2011; Ahern and Dittmar, 2012; Bertrand et al., 2019; Eckbo, Nygaard and Thorburn, 2019; Bijkerk et al., 2021), hiring quotas of various forms can play a prominent role in labor market discrimination. The quota can be understood as a form of discrimination when employed only for the minority group, while it serves as a diversifying recruiting strategy when employed symmetrically across all demographic groups.

[9] Rooted in statistical discrimination and the ambiguity of the productivity of individual employees (Arrow, 1973; Bohren, Imas and Rosenberg, 2019; Lang and Spitzer, 2020), a growing strand of literature has focused on how employers update their beliefs about worker quality (Altonji and Pierret, 2001; Altonji, 2005; Lang and Lehmann, 2012; Campos-Mercade and Mengel, 2024). Due to the design of the experiment, no information on the evaluation procedure following the contact stage is observed. Therefore, the predictions focus only on the contact stage, under the assumption that applicants are contacted because they are deemed qualified. This ensures that, under the learning mechanism, more previous applicants being contacted always leads to a non-negative update in the perceived quality of applicants from the same group.

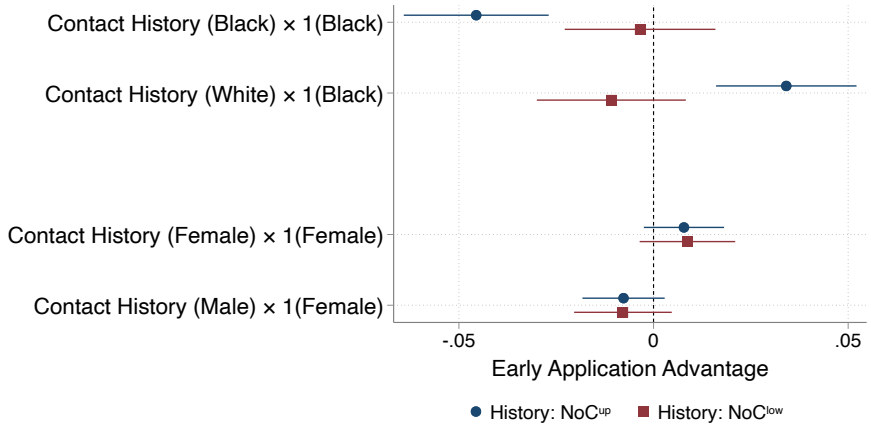
Specifically, I estimate

$$Y_i = \alpha_0 + \eta_1 \mathbf{1}(Group_i = g) + \eta_2 \text{History}_i^g + \eta_3 \text{History}_i^g \times \mathbf{1}(Group_i = g) \quad (4.2)$$
$$+ \iota_1 \text{History}_i^{g'} + \iota_2 \text{History}_i^{g'} \times \mathbf{1}(Group_i = g) + \boldsymbol{\gamma} \mathbf{X}_i + \lambda_{j(i)} + \pi_{Order_i} + \epsilon_i,$$

where g' denotes the complementary group, and I include job fixed effects $\lambda_{j(i)}$ and submission-order fixed effects π_{Order_i}. The parameter η_3 captures how the marginal effect of same-group prior contacts differs for group-g applicants; ι_2 captures the corresponding cross-group effect. Model (4.2) predicts that

- for application i's own group g, $\eta_3 < 0$ if the quota mechanism dominates, $\eta_3 > 0$ if the learning mechanism dominates;
- for the other group g', $\iota_2 > 0$ if quota mechanism also dominates, and $|\eta_3| = |\iota_3|$ if the quota mechanism applies symmetrically for both groups.

Figure 4: Group-Specific Contact Histories: Quota versus Learning



Notes: This figure plots selected interaction estimates from Model (4.2). Each point reports the coefficient on $\text{History}_i^h \times \mathbf{1}(\text{Group}_i = g)$, where g denotes the group of the focal applicant (Black or Female) and h denotes the group whose prior contacts define the history (Black/White or Female/Male). The outcome is an indicator for whether the employer contacts application i within 30 days. History_i^h is measured within vacancy as the number of other experimental applications from group h contacted prior to the submission of i. Circles use NoC^{up} (contacts occurring no later than the submission day); squares use NoC^{low} (contacts occurring at least one day before submission). All regressions include job fixed effects, submission-order fixed effects, and the full set of randomized resume controls. Horizontal bars denote 95% confidence intervals based on standard errors clustered at the job level.

Figure 4 summarizes the key interaction estimates from Model (4.2). For race, the upper-bound history measure (NoC^{up}) implies a pronounced asymmetry: additional previously contacted Black applicants substantially reduce the probability that a subsequent Black applicant is contacted relative to a White applicant, while

prior contacts of White applicants have a smaller and opposite-signed differential effect for Black applicants. This pattern suggests that the quota mechanism dominates for both racial groups, but asymmetrically tighter for Black applicants. Under the lower-bound history measure (NoC^{low}), these race-specific effects attenuate, suggesting that the quota domination is most salient for contacts within the submission day.

For gender, the corresponding interaction effects are smaller in scale, positive for within-group effects and negative for cross-group effects. That is, for both female and male applicants, point estimates suggest that prior contacts from the applicant's own group increase the probability of contact, consistent with learning mechanism dominating in contact decisions for both groups. These patterns are robust across the upper- and lower-bound contact history measures. Tables A.9 and A.10 report the underlying regression estimates, as well as results for within- and cross-group estimation separately.[10]

Finally, I construct two dummies, $\mathbf{1}\left(NoC_i^{up} > 0\right)$ and $\mathbf{1}\left(NoC_i^{low} > 0\right)$ to indicate any contacts prior to submission i within a vacancy, and repeat the analysis in Table 2 to explore how other resume characteristics interact with the contact history. Table A.11 reports the full set of interaction estimates. Joint tests reject the null hypothesis of no interaction effects when $\mathbf{1}\left(NoC_i^{up} > 0\right)$ is used, suggesting that the impact of resume characteristics varies marginally significantly across submissions with different traits of contact history.

Taken together, the micro evidence in this section supports the key intuition that the significant early application advantages observed in Section 3.3 are driven by sequential screening: employers contact quickly and contact decisions early in the submission window reduce the likelihood that later applicants are contacted. This sequential mechanism has asymmetric effects within and across demographic groups, which can be attributed to the interplay between two underlying forces, quota and learning.[11] The next section zooms out to ask how the strength of such

[10]One concern is that NoC_i^{up} and NoC_i^{low} rely on the observed submission-to-contact lag for contacted applications. Hence they may understate prior screening activity for applications that are never contacted. As a robustness check, I construct an alternative history measure that depends only on submission order, NoC_i^{alt}, which counts the number of earlier experimental submissions to the vacancy that eventually receive contact. Regressions using NoC_i^{alt} and the corresponding indicator $\mathbf{1}(NoC_i^{alt} > 0)$ yield similar qualitative conclusions. Results are reported in Appendix Section F.3.

[11]Even though the results are both economically and statistically significant, they should be in-

front-loaded screening strategy varies with local labor market conditions, setting the empirical targets for the model.

5 Market Conditions and Application Timing

Section 4 traces the early application advantage to employer-side decisions: recruiters begin evaluating before the full pool arrives, and once a vacancy starts generating contacts, the probability that later applications receive a contact falls sharply. This section asks when this front-loaded evaluation is most pronounced. I relate the *daily* early application advantage to one key market condition that maps into a firm's tradeoff between moving quickly and waiting for additional applicants: local labor market tightness.

I estimate heterogeneity in the early application advantage by interacting the top-coded days-to-last-submission measure $D2L_i^t$ defined in Section 3.1 with market covariates:

$$\begin{aligned} Y_i =& \alpha_0 + \delta\, D2L_i^t + \phi\left(D2L_i^t \times Z_{m(i),t(i)}\right) + \phi^{\text{pre}}\left(D2L_i^t \times \bar{Z}_{m(i),t(i)-3:t(i)-1}\right) \\ &+ \boldsymbol{\gamma}\mathbf{X}_i + \lambda_{j(i)} + \varepsilon_i, \end{aligned} \tag{5.1}$$

where Y_i indicates whether application i receives employer contact within 30 days, $Z_{m(i),t(i)}$ is a market covariate measured for the relevant market $m(i)$ and the month $t(i)$ of the vacancy's first experimental submission. $\bar{Z}_{m(i),t(i)-3:t(i)-1}$ is the three-month average of the same covariate prior to $t(i)$. All specifications include resume characteristics $\mathbf{X}_i$ and job fixed effects $\lambda_{j(i)}$. I report ϕ, which captures how the *daily* early application advantage varies with $Z_{m(i),t(i)}$. Standard errors are clustered at the job level. Because market conditions vary only across vacancies and are not experimentally assigned, the results in this section are descriptive and intended to discipline the model in Section 6.

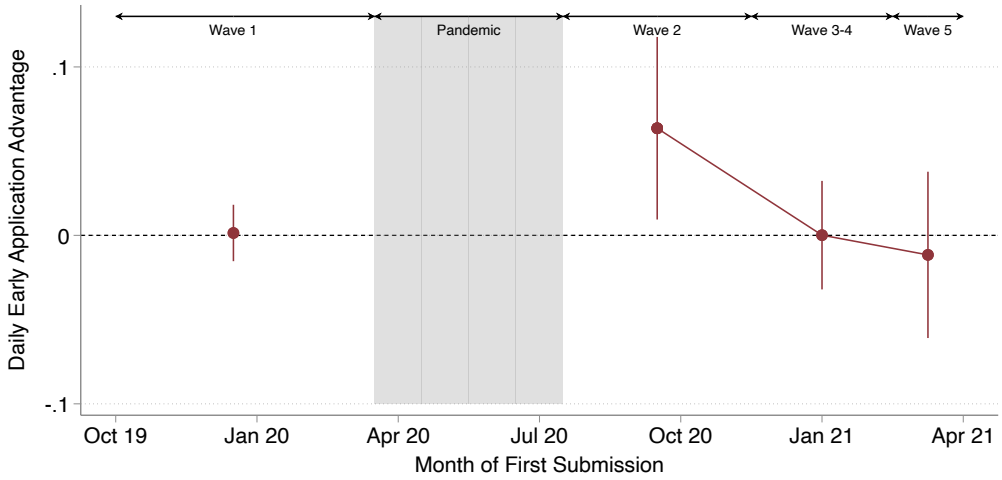
I measure labor market conditions using state-by-month series from JOLTS. The key object is labor market tightness, defined as vacancies per unemployed worker. Figure 5 visualizes how the tightness interaction evolves across waves.[12] The esti-

terpreted with caution. The applications submitted in the experiments are limited to a maximum of eight per job, likely accounting for only a part of the total application pool and the full contact history.

[12] I pool Waves 3 and 4 together for the analysis in this section. I do so for two reasons. First,

mated cofficient is close to 0 in Wave 1, rises sharply in Wave 2, the period immediately following the initial pandemic disruption, and then attenuates in subsequent waves. This pattern is consistent with a vacancy-pressure logic: when vacancies are plentiful relative to unemployment, the opportunity cost of delay is higher, and firms appear to shift evaluation earlier, increasing the daily early application advantage.

Figure 5: Labor Market Tightness and the Daily Early Application Advantage



Notes: This figure plots wave-specific estimates of ϕ in Equation (5.1) when $Z_{m(i),t(i)}$ is state-month tightness (vacancies per unemployed worker) measured in the month of the vacancy's first experimental submission. The shaded region marks the COVID-19 period in which the experiment was paused. Error bars denote 95% confidence intervals based on job-clustered standard errors.

To situate tightness relative to other labor market indicators, I also consider state-by-month job openings, hires, quits, layoffs/discharges, and total separations, each expressed as a rate.[13] Because labor market conditions exhibit sharp discontinuities around COVID-19, Figure A.9 reports estimated heterogeneity in the daily early application advantage separately for Waves 1 and 5 versus Waves 2-4. Across measures, tightness is the strongest predictor: the daily early application advantage is larger in tighter state-month labor markets, with the relationship

Waves 3 and 4 overlap closely in calendar time, both have substantial mass in January 2021 (7,520 of 17,982 submissions in Wave 3 and 5,164 of 17,140 in Wave 4). Therefore, pooling is necessary for the month-level analysis. Second, pooling increases statistical power. In unreported results, estimating the same specifications separately by Wave 3 and Wave 4 yields very similar patterns.

[13]Quit rates, layoff/discharge rates, total separation rates, hire rates are defined as the ratio of corresponding macro series over employment. Job opening rates is defined as the ratio of open positions over the sum of employment and open positions. See Appendix Section E for details.

most pronounced in Waves 2-4. In contrast, interactions with other JOLTS rates are comparatively small.

Taken together, these results connect the micro within-vacancy mechanism to the broader labor market. The early application advantage responds systematically to vacancy-filling pressure, and the tightness margin shifts over time with other aggregate market conditions. These reduced-form facts provide an empirical anchor for the model in the next section.

6 Modeling Timing in Labor Market

Sections 3 - 5 establish three facts. First, holding applicant attributes fixed within vacancies, earlier submissions are more likely to be contacted. Second, this early application advantage reflects history-dependent contact initiation: once a vacancy begins generating contacts, the probability that later, otherwise comparable applications receive a contact falls sharply. Third, this timing gradient varies systematically with local labor market tightness.

I formalize these facts in a model of sequential screening in which firms endogenously choose when to initiate contacts. The firm faces an optimal-stopping problem: waiting expands the applicant pool, but delay is costly. This is a partial equilibrium model where firms take the market condition as given and privately optimize, it establishes firms' shadow cost of keeping vacancies open that rationalizes the optimal contact initiation timing. Labor market tightness governs expected applicant inflows and the cost of delay, and aggregate conditions shift this environment over time. The model organizes the reduced-form evidence and delivers quantitative implications for the application timing while allowing the presence of organic applications that were not captured by the experiment.

6.1 Setup

6.1.1 Markets

The economy is indexed by a local labor market $m \in \mathcal{M}$ and a calendar period $t \in \{1, \ldots, T\}$. Each market-period (m, t) is summarized by unemployment U_{mt} and the vacancy stock V_{mt}. A market m is a geographically and occupationally segmented recruiting environment. In the data, m is a state-by-industry cell, but

the model only requires that m partitions environments relevant for applicant inflow and recruiting costs.

Each (m, t) is characterized by aggregate conditions $\mathbf{S}_{mt}$, taken as exogenous by individual vacancies. A key components of $\mathbf{S}_{mt}$ is market tightness:

$$\theta_{mt} \equiv \frac{V_{mt}}{U_{mt}}.$$

Higher θ_{mt} corresponds to a scarcer applicant environment per vacancy.

6.1.2 Firms and vacancies

I focus on recruiting decisions conditional on a vacancy being posted. The extensive margin of vacancy creation is taken as given. I model the within-vacancy recruiting problem that generates observed contact outcomes. In each (m, t), firms $\mathcal{F}_{mt}$ post vacancies. A firm may operate in multiple local labor markets; when needed, I interpret (f, m) as a local branch.

Vacancies. Firm f posts a set of vacancies $\mathcal{J}_{fmt}$. Each vacancy $j \in \mathcal{J}_{fmt}$ represents one position and can generate at most one hiring commitment. Vacancy j has observable characteristics $\mathbf{z}_j$ and two latent components observed by the firm but not by the econometrician:

- ν_j^λ captures the vacancy-level inflow shock that shifts the arrival intensity, conditional on (m, t, f) and $\mathbf{z}_j$. I model ν_j^λ as a vacancy-level random effect constant over the recruiting spell. I assume $\nu_j^\lambda \sim \mathcal{N}\left(-\frac{1}{2}\sigma^2_{\nu^\lambda},\, \sigma^2_{\nu^\lambda}\right)$, so that $\exp\left(\nu_j^\lambda\right)$ is log-normal with $\mathbb{E}\left[\exp(\nu_j^\lambda)\right] = 1$. $\sigma^2_{\nu^\lambda}$ governs the dispersion in unobserved vacancy-specific application pressure.
- ν_j^v is also a vacancy-level random effect constant over the recruiting spell that shifts the distribution of screened values faced by the firm, with the distribution $\nu_j^v \sim \mathcal{N}(0, \sigma^2_{\nu^v})$.[14]

[14] I allow correlation between the arrival intensity random effect ν_j^λ and the screened value random effect ν_j^v with a joint distribution

$$\begin{pmatrix} \nu_j^\lambda \\ \nu_j^v \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} -\frac{1}{2}\sigma^2_{\nu^\lambda} \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{\nu^\lambda} & \rho_v\, \sigma_{\nu^\lambda}\sigma_{\nu^v} \\ \rho_v\, \sigma_{\nu^\lambda}\sigma_{\nu^v} & \sigma^2_{\nu^v} \end{pmatrix} \right),$$

where $\rho_v \in [-1, 1]$ is the correlation between ν_j^λ and ν_j^v. ρ_v matters for the econometrician's joint distribution of queues and values, but does not affect the firm's optimal policy once it conditions

Throughout, I condition on the full vacancy state $(m, t, f, \mathbf{z}_j, \nu_j^\lambda, \nu_j^v)$. The inflow shock ν_j^λ enters only the arrival intensity λ_j, while the value shock ν_j^v enters only the distribution of match quality and screened values.

Within-vacancy time and vacancy resolution. Aggregate conditions vary at the calendar frequency t, while recruiting unfolds in within-vacancy time. For each vacancy j, let d denote days since the start of the recruiting spell. In the empirical implementation, the vacancy posting date is not observed; I therefore proxy d using the relative timing of the experimental submissions, normalizing $d = 0$ to the first experimental submission within each vacancy.

Resolution is the time when the firm initiates its contact stage and effectively stops active screening for the initial contact decision.[15] Let $\tau_j \in \mathbb{R}_+$ denote the (endogenous) contact stage beginning time for vacancy j. The recruiting window for the contact-stage decision is therefore $[0, \tau_j]$.

6.1.3 Applicants, match quality, and screening signals

Vacancy j receives a sequence of applications over its life. Each application (i, j) is characterized by a vector of observable resume attributes $\mathbf{x}_{ij} \in \mathcal{X}$. Employer decisions are made at the within-vacancy frequency and depend on the information contained in realized applications and the signals generated through screening.

Latent match quality. Applicant i has an unobserved, job-specific match quality $q_{ij} \in \mathbb{R}$, realized only if hired and not observed at the time of application. Conditional on $\left(\mathbf{x}_{ij}, \mathbf{z}_j, \nu_j^\lambda, \nu_j^v\right)$, match quality follows:

$$q_{ij} \sim F_q\left(\cdot \mid \mathbf{x}_{ij}, \mathbf{z}_j, \nu_j^v\right).$$

on (ν_j^λ, ν_j^v).

[15] I model the pre-contact screening-waiting problem, after initiating the first contact, the vacancy enters the interview-offer pipeline that I treat as a separate stage. If the expected value of entering the contact stage is an increasing function of the best-so-far screened value, e.g., because the firm can contact multiple candidates or contact sequentially until one accepts, then the optimal initiation rule for the contact stage remains a reservation value policy. Future extensions can allow for continued contacts after τ_j (e.g., backup candidates) without changing the core mechanism. See for more discussion.

For later use, it is convenient to express q_{ij} as

$$q_{ij} = \mu(\mathbf{x}_{ij}, \mathbf{z}_j) + \nu_j^v + u_{ij}, \qquad \mathbb{E}\big[u_{ij} \mid \mathbf{x}_{ij}, \mathbf{z}_j, \nu_j^v\big] = 0, \tag{6.1}$$

where $\mu(\cdot)$ captures systematic fit between applicant and job observables and $\{u_{ij}\}$ is i.i.d. across i within a vacancy conditional on (j, f, m, t).

Screening signal observed upon review. Reviewing an application produces a one-dimensional signal $a_{ij} \in \mathbb{R}$ that is informative about q_{ij}:

$$a_{ij} = q_{ij} + \varepsilon_{ij},$$

where ε_{ij} is i.i.d. across i and independent of u_{ij} conditional on $(\mathbf{x}_{ij}, \mathbf{z}_j, \nu_j^v)$. a_{ij} summarizes screening-relevant information available to the employer upon reviewing the application. Given $(\mathbf{x}_{ij}, \mathbf{z}_j, \nu_j^\lambda, \nu_j^v)$ and the realized signal a_{ij}, the firm forms a posterior over q_{ij}. Because firms' optimal stopping problem is driven by comparisons across applicants, it is without loss to treat a_{ij} as a sufficient statistic for ranking applicants by expected surplus.[16]

6.1.4 Application arrivals

Let $N_j(d)$ be the cumulative number of applications received by within-vacancy time $d \geq 0$. Let $d_{ij} \in \mathbb{R}_+$ denote applicant i's arrival time in days since posting, and define

$$A_j(d) \equiv \{i : d_{ij} \leq d\}$$

When comparing to daily data, the number of arrivals during day k is $N_j(k+1) - N_j(k)$ for $k \in \{0, 1, 2, \cdots\}$.

Exogenous inflow. The employer takes application inflow as exogenous over the recruiting spell, conditional on market conditions, firm identity, and vacancy characteristics. In the baseline, application arrivals follow a homogeneous Poisson process up to firm's stopping time τ_j:

[16] A sufficient condition is that $\mathbb{E}[q_{ij} \mid a_{ij}, \mathbf{x}_{ij}, \mathbf{z}_j, \nu_j^v]$ is weakly increasing in a_{ij}, e.g. under additive noise with a log-concave density.

Assumption 6.1 (Poisson application arrivals)**.** *Conditional on $(m, t, f, \mathbf{z}_j, \nu_j^\lambda, \nu_j^v)$, applications arrive to vacancy j according to a Poisson counting process $\{N_j(d)\}_{d\geq 0}$ with constant intensity $\lambda_j > 0$ (which depends only on ν_j^λ but not on ν_j^v). Equivalently, for any $0 \leq s < u$,*

$$N_j(u) - N_j(s) \sim \text{Poisson}\left(\lambda_j(u-s)\right),$$

with independent increments over disjoint intervals.

Assumption 6.1 implies $N_j(d+1) - N_j(d) \sim \text{Poisson}(\lambda_j)$ for integers $d \geq 0$. Equivalently, the integrated intensity up to time d is

$$\Lambda_j(d) \equiv \lambda_j d,$$

so $N_j(d) \sim \text{Poisson}(\Lambda_j(d))$ for each $d \geq 0$. This assumption is the within-vacancy analogue of the urn-ball application technology used in multi-application search models (Birinci, See and Wee, 2025). It provides a tractable way to model the stochastic evolution of the applicant pool.

In Appendix Section B.2, I consider a separable non-homogeneous Poisson specification with time-varying intensity $\lambda_j(d) = \lambda_j \rho(d)$, where $\rho(d) \geq 0$ is a common posting age profile normalized to have unit average over the empirical horizon $[0, \bar{T}]$: $\frac{1}{\bar{T}} \int_0^{\bar{T}} \rho(s)\mathrm{d}s = 1$.[17] The integrated intensity is then

$$\Lambda_j(d) = \int_0^d \lambda_j \rho(s)\mathrm{d}s = \lambda_j \int_0^d \rho(s)\mathrm{d}s.$$

The homogeneous baseline is nested as the special case $\rho(d) \equiv 1$ over the relevant horizon.

Factor structure for the arrival intensity I parameterize the arrival rate λ_j with a parsimonious factor structure that combines (i) aggregate labor market congestion, (ii) a time shifter that scales effective application pressure, and (iii) vacancy heterogeneity:

$$\lambda_j = \underbrace{\xi_{mt} c_{mt}^{\alpha_c}}_{\text{congestion}} \cdot \underbrace{\exp(\gamma' \mathbf{z}_j + \nu_j^\lambda)}_{\text{vacancy heterogeneity}}. \tag{6.2}$$

[17] This extension reflects the recent empirical evidence that postings typically receive many applications shortly after opening and taper off as job seeker attention decays (Davis and de la Parra, 2024), while preserving key tractability properties.

$c_{mt} \equiv 1/\theta_{mt}$ captures the market congestion. The shifter $\xi_{mt} > 0$ scales aggregate application intensity as in urn-ball application technologies (Birinci, See and Wee, 2025).[18] In the empirical implementation, I allow ξ_{mt} to be either wave fixed effects common across markets, or a function of state-month labor market statistics.

Marked arrivals and i.i.d. applicant draws. To link the arrival process to the stopping problem, I impose a stationarity restriction on the quality composition of arrivals:

Assumption 6.2 (Quality composition stationarity). $\{(\mathbf{x}_{ij}, q_{ij}, a_{ij})\}_{i \in A_j(\infty)}$ *is i.i.d. across arrivals and independent of* $\{N_j(d)\}_{d \geq 0}$*, conditional on* $(m, t, f, \mathbf{z}_j, \nu_j^\lambda, \nu_j^v)$.

This restriction implies that changes in λ_j shift the thickness of the applicant pool without mechanically shifting its quality composition. Any early-application advantage then arises from the firm's dynamic screening and stopping decisions rather than time-varying applicant selection. A central implication of i.i.d. applicant draws is diminishing returns of additional arrivals: the expected value of waiting is concave in the expected number of future arrivals. [19]

6.1.5 Firm payoffs and vacancy costs

I now specify the per-vacancy payoff of filling a position and the cost of delaying a hire.

Payoff from filling a vacancy. If vacancy j is filled with applicant i, match output is increasing in match quality:

$$y_{ij} = \chi_j + q_{ij},$$

where χ_j is a vacancy-specific productivity component absorbed into $\mathbf{z}_j$.

[18] In their framework, if a worker applies to a vacancy with probability ξ/v on average, the number of applicants a vacancy receives is Poisson with mean ξ/θ, where $\theta = v/u$ is labor market tightness. Birinci, See and Wee (2025) also interpret ξ as an application intensity parameter: raising ξ increases the expected number of applications a worker sends.

[19] The i.i.d. mark assumption is attractive here because the experiment holds resume quality fixed by construction. A useful extension would allow for perceived "quality drift" by letting the mark distribution depend on d, e.g., $q_{ij} = \mu(\mathbf{x}_{ij}, \mathbf{z}_j) + \phi d_{ij} + u_{ij}$, or more generally $F_q(\cdot \mid \mathbf{x}_{ij}, \mathbf{z}_j, \nu_j^\lambda, \nu_j^v, d_{ij})$.

Wages w_{ij} are set by Nash bargaining, with worker outside option b_{mt} and worker bargaining weight $\eta \in (0, 1)$:

$$w_{ij} = \arg\max_{w} (w - b_{mt})^{\eta} (y_{ij} - w)^{1-\eta},$$

which implies $w_{ij} = b_{mt} + \eta(y_{ij} - b_{mt})$. Therefore, the firm's flow profit is

$$\pi_{ij} \equiv y_{ij} - w_{ij} = (1 - \eta)\,(y_{ij} - b_{mt}) = (1 - \eta)\big(\chi_j - b_{mt}\big) + (1 - \eta) q_{ij},$$

that is, under standard bargaining, the firm's payoff is an affine transformation of q_{ij}. Because the optimal stopping rule below is invariant to positive affine transformations of payoffs that do not change the ranking of applicants, I normalize the firm's payoff from hiring to be linear in match quality:

$$S(q_{ij}) \equiv q_{ij}, \tag{6.3}$$

with any vacancy-level and market-level constants absorbed into $\mathbf{z}_j$ and the cost-of-delay term defined next.[20]

Vacancy cost of continued searching. While vacancy j remains unfilled, the firm pays a flow cost $\kappa_{fmt} > 0$ per day. It can be interpreted as the reduced-form shadow cost of waiting, which can include vacancy flow costs, recruiting urgency, candidate loss risk, and the burden of continued evaluation, and other costs of postponing the hire. κ_{fmt} depends parsimoniously on market tightness:

$$\kappa_{fmt} = \kappa_0 \cdot \theta_{mt}^{\kappa_\theta}, \tag{6.4}$$

The parameters κ_θ captures how urgency varies with tightness and market power. I later parameterize κ_θ as functions of labor market state index.

[20] Equivalently, one can carry $(1 - \eta)$ explicitly and treat b_{mt} as a shifter in the opportunity cost of delay. In my setting, wages are not observed for the matched hires and match quality is measured only up to a monotone transformation. As a result, the surplus-share factor $(1 - \eta)$ is not separately identified. The normalization (6.3) is therefore convenient and without loss for the timing and screening decisions studied here.

6.1.6 Sequential screening and the contact decision

For vacancy j, applications arrive stochastically as in Section 6.1.4. The firm trades off the option value of waiting for additional applicants against delay costs and decides (i) which arrived applications to screen and (ii) when to initiate contact.

Information. Upon arrival of application i, the employer observes $\mathbf{x}_{ij}$ and vacancy attributes $\left(\mathbf{z}_j, \nu_j^\lambda, \nu_j^v\right)$, but not the match quality q_{ij}. Screening reveals signal a_{ij}, and equivalently v_{ij}. Let $\mathcal{R}_j(d) \subseteq A_j(d)$ denote the set of screened applications by day $d \geq 0$. In the baseline, the firm reviews each application upon arrival and pays k_{fmt} per arrival to screen.[21] For screened applicant i, the posterior expected value

$$v_{ij} \equiv \mathbb{E}\left[q_{ij} \mid a_{ij}, \mathbf{x}_{ij}, \mathbf{z}_j, \nu_j^v\right]$$

is increasing in a_{ij} as discussed in Section 6.1.3. In the baseline solution, I work directly with the distribution of the posterior value v_{ij} revealed by screening and do not separately model (q_{ij}, a_{ij}).

State variable. Let

$$v_j^{\max}(d) \equiv \max\{v_{ij} : i \in \mathcal{R}_j(d)\}$$

be the best posterior value among screened applicants by time d, with $v_j^{\max}(d) = -\infty$ if no one has been screened. Let $T_{j,1} \equiv \inf\{d \geq 0 : N_j(d) \geq 1\}$ denote the first application arrival time. Since initiating contact requires screening at least one applicant, I restrict attention to stopping times $\tau_j \geq T_{j,1}$ so $v_j^{\max}(\tau_j) \in \mathbb{R}$ almost surely. Given Poisson arrivals and i.i.d. marks, the continuation value depends on history only through $v_j^{\max}(d)$.

Contact decision as optimal stopping. At any d while the vacancy is open, the firm chooses whether to continue recruiting or initiate contact. Stopping at τ_j yields $v_j^{\max}(\tau_j)$ and ends the recruiting spell.[22]

[21] It is referred to as immediate screening ($\mathcal{R}_j(d) = A_j(d)$) hereafter. Appendix B.2.2 relaxes immediate screening by allowing the firm to choose which applications to review.

[22] In the data, firms may contact multiple applicants and continue receiving applications after first contact. The model abstracts from subsequent contacts by treating the initiation of contact

Combining hiring payoffs and delay costs, the firm chooses a stopping time τ_j to maximize

$$\max_{\tau_j} \mathbb{E}\left[e^{-\delta\tau_j} v_j^{\max}(\tau_j) - \int_0^{\tau_j} e^{-\delta s}\kappa_{fmt}\mathrm{d}s\right], \tag{6.5}$$

where $\delta > 0$ is the daily discount rate. Expectations are taken over the arrival process and signal realizations.

Under immediate screening, the within-vacancy recruiting problem reduces to the optimal stopping problem (6.5) with state variable given by the best screened value to date. Section 6.2 shows that the optimal policy is a reservation rule and provides a closed-form characterization under normal screened values.

6.2 Solution

This section characterizes the firm's optimal recruiting policy for a single vacancy under immediate screening. Fix a vacancy $j \in \mathcal{J}_{fmt}$ and condition on its state $(m, t, f, \mathbf{z}_j, \nu_j^\lambda, \nu_j^v)$, so that the arrival intensity λ_j and the flow vacancy cost κ_{fmt} are taken as given. For simplicity, I suppress subscripts when no confusion arises.

6.2.1 Baseline: immediate screening and stationary arrivals

Arrivals and screened values. Applications arrive in continuous within-vacancy time $d \in \mathbb{R}_+$ according to a Poisson process $\{N(d)\}_{d\geq 0}$ with constant intensity $\lambda > 0$ (Assumption 6.1). Let $\{T_n\}_{n\geq 1}$ denote the arrival times, with $T_1 < T_2 < \cdots$ almost surely. Under immediate screening, each arrival generates a scalar posterior value $v_n \in \mathbb{R}$.[23] Conditional on $(m, t, f, \mathbf{z}_j, \nu_j^\lambda, \nu_j^v)$, $\{v_n\}_{n\geq 1}$ are i.i.d. with CDF F_v. For tractability, assume:

Assumption 6.3 (Conditional normality of screened values)**.** *Screened values satisfy*

$$v_n \sim \mathcal{N}(\mu_j, \sigma_v^2), \qquad \mu_j \equiv \bar{\mu}(\mathbf{z}_j) + \nu_j^v,$$

as an absorbing decision for the purpose of studying early application advantage; Allowing for multiple contacts after initiation only changes the mapping from the best pre-contact candidate to the value of entering the post-contact pipeline. Appendix B.1.3 discuss the robustness of treating contact initiation as absorbing in details.

[23] In the setup, $v_{ij} \equiv \mathbb{E}[q_{ij} \mid a_{ij}, \mathbf{x}_{ij}, \mathbf{z}_j, \nu_j^v]$. For the baseline solution, it is without loss to work directly with the distribution of v_{ij} revealed by screening; a Gaussian signal microfoundation is discussed later.

conditional on $(m, t, f, \mathbf{z}_j, \nu_j^\lambda, \nu_j^v)$*, where* $\sigma_v^2 > 0$ *is the within-vacancy effective dispersion of posterior values and* ν_j^v *is the vacancy random effect defined in Section 6.1.2.*

Let $v^{\max}(d) \equiv \max\{v_n : T_n \leq d\}$ denote the best screened value by time d, with $v^{\max}(d) = -\infty$ if $N(d) = 0$.

Firm problem under immediate screening. Immediate screening implies that $\forall d < \tau_j$, $\mathcal{R}_j(d) = A_j(d)$. The only within-vacancy choice is when to initiate contact. The firm chooses a stopping time $\tau \in \mathcal{T}$ to maximize

$$\max_{\tau \in \mathcal{T}} \mathbb{E}\left[e^{-\delta\tau} v^{\max}(\tau) - \int_0^\tau e^{-\delta s}\kappa \mathrm{d}s\right]. \tag{6.6}$$

The firm only updates its information set when a new application arrives and is screened. That is, $v^{\max}(d)$ changes only at arrival times. Therefore, it is without loss to restrict stopping to arrival times.

Lemma 6.1 (No loss from stopping at arrival times)**.** *Under (6.6), there exists an optimal stopping time* $\tau^\star$ *such that* $\tau^\star \in \{T_1, T_2, \ldots\}$ *almost surely.*

Proof. Let τ be any admissible stopping time and define $\tilde{\tau} \equiv \sup\{T_n : T_n \leq \tau\}$ (with $\tilde{\tau} = 0$ if no arrival has occurred by τ). Between $\tilde{\tau}$ and τ there are no arrivals, so $v^{\max}(\tau) = v^{\max}(\tilde{\tau})$ and $N(\tau) = N(\tilde{\tau})$. Stopping at $\tilde{\tau}$ weakly increases the discounted terminal payoff (since $\tilde{\tau} \leq \tau$) and weakly reduces the discounted flow cost $\int e^{-\delta s}\kappa \mathrm{d}s$. Hence the objective under $\tilde{\tau}$ is weakly higher than under τ. ■

By Lemma 6.1, decisions can be analyzed at the sequence of arrival times. At each arrival n, after observing v_n and updating $v_n^{\max} \equiv \max\{v_{n-1}^{\max}, v_n\}$ for $n \geq 1$ and $v_0^{\max} \equiv -\infty$, the firm chooses whether to stop and realize $v_n^{\max}$ or continue.

Dynamic program at arrivals. Let $W(v)$ denote the maximal continuation value at an arrival when the current best screened value is v. If the firm continues, the next inter-arrival time Δ is exponential with rate λ and independent of the next value draw $v' \sim F_v$. Therefore,

$$W(v) = \max\left\{v, \underbrace{\mathbb{E}\left[e^{-\delta\Delta} W(\max\{v, v'\})\right]}_{\text{option value}} - \underbrace{\mathbb{E}\left[\int_0^\Delta e^{-\delta s}\kappa \mathrm{d}s\right]}_{\text{delay cost}}\right\}. \tag{6.7}$$

Since $\mathbb{E}[e^{-\delta\Delta}] = \frac{\lambda}{\lambda+\delta}$ and $\mathbb{E}\left[\int_0^\Delta e^{-\delta s}\kappa ds\right] = \frac{\kappa}{\lambda+\delta}$, (6.7) simplifies to

$$W(v) = \max\left\{v, -\frac{\kappa}{\lambda+\delta} + \frac{\lambda}{\lambda+\delta}\left(\mathbb{E}[W(\max\{v, v'\})]\right)\right\}. \tag{6.8}$$

Threshold policy. Define the continuation value

$$C(v) \equiv -\frac{\kappa}{\lambda+\delta} + \frac{\lambda}{\lambda+\delta}\left(\mathbb{E}[W(\max\{v, v'\})]\right),$$

so that $W(v) = \max\{v, C(v)\}$. The next lemma establishes a key monotonicity property that implies a cutoff rule.

Lemma 6.2 (Monotonicity and 1-Lipschitz property)**.** *$W(\cdot)$ is weakly increasing and 1-Lipschitz: for any $v_2 \geq v_1$,*

$$0 \leq W(v_2) - W(v_1) \leq v_2 - v_1.$$

Proof. Monotonicity follows because the firm can replicate any stopping rule feasible from v_1 when starting from $v_2 \geq v_1$, and the terminal payoff $v^{\max}(\tau)$ is weakly higher path-by-path. For the Lipschitz bound, couple two economies that share the same future arrivals and value draws. Starting from $v_2 = v_1 + \epsilon$, along any realization the best-so-far process in the v_2 economy is at most ϵ higher than in the v_1 economy at all times, so any feasible stopping rule yields a payoff difference bounded above by ϵ. Taking expectations and optimizing over stopping rules gives $W(v_2) - W(v_1) \leq \epsilon$. ■

Intuitively, the monotonicity of $W(v)$ means that having a better current candidate can never reduce the firm's continuation value, while the 1-Lipschitz property means that the best-so-far can only shift the eventual hire payoff one-for-one, and never amplifies it. Because stopping yields v immediately, while continuing postpones the payoff and only affects outcomes through future arrivals, discounting implies that the incremental value of continuing rises strictly less than one-for-one with v.

Proposition 6.1 (Reservation value)**.** *In the stationary baseline with Poisson arrivals of intensity λ, discount rate $\delta > 0$, flow cost $\kappa \geq 0$, i.i.d. screened value $v' \sim F_v$ with $\mathbb{E}[|v'|] < \infty$, there exists a unique reservation value $r \in \mathbb{R}$ such that it is optimal to stop*

at an arrival if and only if the current best value satisfies $v \geq r$. Equivalently, the optimal stopping time is the first arrival time at which $v^{\max}(d) \geq r$.

Proof sketch. Consider $H(v) \equiv C(v) - v$, the net gain from continuing rather than stopping at value v. Take $v_2 = v_1 + \epsilon$. Because $\max\{v_2, v'\} \leq \max\{v_1, v'\} + \epsilon$ for all v', Lemma 6.2 implies

$$\mathbb{E}[W(\max\{v_2, v'\})] - \mathbb{E}[W(\max\{v_1, v'\})] \leq \epsilon.$$

Hence,

$$H(v_2) - H(v_1) \leq \frac{\lambda}{\lambda + \delta}\epsilon - \epsilon = -\frac{\delta}{\lambda + \delta}\epsilon < 0,$$

so $H(\cdot)$ is strictly decreasing. As $v \to -\infty$, $H(v) \to +\infty$ (the firm surely prefers to wait for a better draw), while as $v \to +\infty$, $H(v) \to -\infty$ (the firm prefers to stop). Therefore there exists a unique r with $H(r) = 0$, and the optimal policy is to stop if and only if $v \geq r$. A full proof is provided in Appendix B.1.2. ■

Characterization of the reservation value. At the cutoff r, the firm is indifferent between stopping and continuing. Using $W(v) = v$ for $v \geq r$ and $W(r) = r$, the Bellman equation implies the scalar reservation condition

$$\delta r + \kappa = \lambda \mathbb{E}\left[(v - r)^+\right], \quad v \sim F_v, \quad (x)^+ \equiv \max\{x, 0\}. \tag{6.9}$$

This equation is intuitive: the left-hand side is the marginal cost of waiting, and the right-hand side is the marginal option value of drawing additional applicants. Under Assumption 6.3, if $v \sim \mathcal{N}(\mu_j, \sigma_v^2)$ and $z \equiv (r - \mu_j)/\sigma_v$, then

$$\mathbb{E}\left[(v - r)^+\right] = (\mu_j - r)\left(1 - \Phi(z)\right) + \sigma_v \phi(z), \tag{6.10}$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ are the standard normal CDF and PDF. r is therefore pinned down by a one-dimensional equation.

Implications for contact timing and the number of screened applications. Let $p \equiv \Pr(v \geq r) = 1 - F_v(r)$ denote the probability that a screened arrival meets the reservation value. Because arrivals are Poisson and marks are i.i.d., the process of "acceptable" arrivals is a thinned Poisson process with intensity λp. Therefore, the stopping time τ has an exponential distribution with hazard λp.

Corollary 6.1 (Duration and number screened)**.** *Under the cutoff rule in Proposition 6.1,*

$$\tau \sim \text{Exp}(\lambda p), \qquad \mathbb{E}[\tau] = \frac{1}{\lambda p}.$$

Moreover, the number of screened applications up to and including the contacted one is geometric with success probability p, so $\mathbb{E}[N(\tau)] = 1/p$.

Comparative statics. The cutoff r is increasing in applicant inflow and decreasing in the cost of waiting and screening.

Proposition 6.2 (Comparative statics of the reservation value)**.** *Let $p = 1 - F_v(r)$ and assume F_v is continuous. The unique solution r to (6.9) satisfies*

$$\frac{\partial r}{\partial \kappa} = -\frac{1}{\delta + \lambda p} < 0, \qquad \frac{\partial r}{\partial \lambda} = \frac{\delta r + \kappa}{\lambda(\delta + \lambda p)} > 0.$$

Proof. Differentiate (6.9) and use $\frac{d}{dr}\mathbb{E}[(v-r)^+] = -\Pr(v \geq r) = -p$. Then solve for the derivatives by the implicit function theorem. ■

Intuitively, Proposition 6.2 states that higher delay costs lower r and induce earlier stopping, while thicker application inflow raises r and makes the firm more selective.

Extensions. Appendix B.2.1 allows a posting-age profile $\rho(d)$ in arrivals. Appendix B.2.2 introduces selective and batch screening with a Pandora's box logic (Weitzman, 1979; Beyhaghi and Cai, 2024). In both extensions, I show that the main mechanism and comparative statics remain qualitatively robust.

7 Model Estimation and Counterfactuals

This section describes how I take the screening-and-wait model in Section 6 to the data and estimate its key parameters. I proceed in four steps. First, I map the model's objects to the experimental outcomes. Second, I define the moment vector and its sampling uncertainty. Third, I parameterize the model primitives as functions of observed job characteristics and labor-market conditions. Finally,

I describe the method of simulated moments (MSM) estimator, numerical implementation, and model-fit diagnostics.

Throughout, I use j to index vacancies (jobs), $m(j)$ to index the local labor market of vacancy j, and $t(j)$ to denote the calendar month in which the first experimental application to j is submitted.[24] To avoid confusion with wages (w_{ij} in Section 6.1.5), I denote the model-implied initiation weights by ω rather than w. The estimation is conducted on the *complete* jobs with 8 applications submitted in the experiment.

7.1 Mapping the model to experimental outcomes

7.1.1 First contact as the empirical analogue of stopping

The model focuses on the decision to initiate contact: the firm screens arrivals and stops waiting once the best screened applicant crosses the reservation threshold (Section 6.1.6). In the data, I observe contact-initiation events for the experimental applications. I therefore map the model stopping time τ_j to the *first observed contact initiation* among the experimental applications to vacancy j.

Let $C_{ij}^{\text{first}} \in \{0, 1\}$ indicate whether applicant i is the first contacted experimental applicant for vacancy j. The empirical analogue of the stopping decision is the event $\sum_i C_{ij}^{\text{first}} = 1$ (a first contact among the experimental set) versus $\sum_i C_{ij}^{\text{first}} = 0$ (no such contact observed within the experimental set). As discussed in Appendix B.1.3, allowing for subsequent contacts can be summarized by an increasing continuation payoff in the best screened value at τ_j; the optimal contact initiation rule remains a reservation policy, so this mapping targets the margin of interest.

7.1.2 Within-vacancy time in the experiment

The model's within-vacancy time d is defined from the start of the recruiting spell. Because the vacancy posting date is not observed, I proxy within-vacancy time using *relative submission timing* of the experimental applications.

Define the relative timing index

$$d_{ij} \equiv (\text{submission date of } i \text{ to } j) - \min_{i'} (\text{submission date of } i' \text{ to } j).$$

[24] The experiment does not observe the vacancy posting date. I therefore treat the month of the first experimental submission as the relevant calendar-time index for local conditions.

and discretize it into integer day bins $d_{ij} \in \{0, 1, \ldots, 7+\}$, where $7+$ is the design-based top-coded last bin for $d_{ij} \geq 7$.[25] The estimation sample is constructed so that each job has exactly eight experimental applications, and I summarize their timing by the within-job counts

$$n_{jd} \equiv \sum_{i:d_{ij}=d} 1, \quad d = 0, 1, \ldots, 7+.$$

In the moment vector, I keep seven day-bin level moments ($d = 0, \ldots, 6$) and a residual "none" category defined below. One additional day-bin level moment is omitted because the weights must sum to one within job.

7.1.3 From arrival intensity to initiation weights

For each vacancy j, the model delivers (i) an organic arrival intensity λ_j and (ii) a job-specific flow cost κ_j governing the contact reservation rule. Section 6.2 shows that (λ_j, κ_j) imply an acceptance probability $p(\lambda_j, \kappa_j)$.

The experiment injects eight *tracer* applications at known relative days $\{d_{ij}\}$. To connect the model to this design, I treat non-experimental inflow as an independent Poisson process with intensity λ_j and acceptance probability $p_j \equiv p(\lambda_j, \kappa_j)$. Under thinning, the arrival process of *acceptable* organic applicants has intensity $\lambda_j p_j$.

For a given day bin d, let $N_{j,<d} \equiv \sum_{\tau<d} n_{j\tau}$ be the number of tracer applications that arrive strictly before d. Then the probability that the first contacted applicant is a tracer who arrives in bin d is

$$P_{jd}(\lambda_j, p_j) = \underbrace{\exp(-\lambda_j p_j\, t_d)}_{\substack{\text{no acceptable organic} \\ \text{arrival before } t_d}} \cdot \underbrace{(1-p_j)^{N_{j,<d}}}_{\substack{\text{no acceptable tracer} \\ \text{before bin } d}} \cdot \underbrace{(1-(1-p_j)^{n_{jd}})}_{\substack{\text{at least one acceptable} \\ \text{tracer in bin } d}}, \tag{7.1}$$

where t_d is the bin boundary in days associated with d.[26] Conditional on $n_{jd} > 0$,

[25] This proxy is almost perfectly negatively correlated with the top-coded days-to-last-submission measure $D2L^t$ used in the reduced-form analysis (Model (3.3)). Controlling for job fixed effects and clustering standard errors at the job level, the regression coefficient of $D2L_i^t$ on d_{ij} is -0.9729 (s.e. 0.0018).

[26] In the code implementation, $t_d = d$ and the last bin is interpreted as $7+$.

each tracer application arriving in bin d receives equal expected initiation weight

$$\omega_{ij}(\lambda_j, p_j) = \frac{P_{j,d_{ij}}(\lambda_j, p_j)}{n_{j,d_{ij}}}. \tag{7.2}$$

Finally, the residual mass

$$\omega_j^{\text{none}}(\lambda_j, p_j) \equiv 1 - \sum_{i=1}^{8} \omega_{ij}(\lambda_j, p_j)$$

corresponds to the event that the first contacted applicant is *not* among the eight experimental applications (e.g., an organic applicant is contacted first, or contact occurs outside the experimental window). This "none" category is empirically important because many vacancies have no observed contact among the eight experimental applications.

7.2 Target moments and sampling uncertainty

Let $g^{\text{data}} \in \mathbb{R}^M$ denote the vector of empirical moments computed from the experimental job-application panel, and let $g(\Theta)$ denote the corresponding simulated moments from the model. The moment vector stacks three families of moments that discipline (i) the timing distribution of first contact among tracers and (ii) how that timing gradient varies with local tightness and firm prominence.

Wave groups. The experiment is fielded in multiple calendar waves. In estimation, I work with a wave-group indicator, denoted as $h(j)$. The wave grouping pools overlapping periods: $h(j) = 3$ combines Wave 3 and 4 of the experiment, and $h(j) = 4$ corresponds to Wave 5.[27] All moments are computed separately by wave group unless stated otherwise.

[27] I pool Waves 3 and 4 because the experimental periods overlap heavily and the reduced-form patterns are similar.

Level moments. For each wave group h, I compute the mean initiation-weight mass in each day bin $d = 0, \ldots, 6$, along with the residual "none" mass:

$$\mathbb{E}\left[\omega_{ij} \mid h(j) = h, d_{ij} = d\right], \quad d = 0, \ldots, 6,$$
$$\mathbb{E}\left[\omega_j^{\text{none}} \mid h(j) = h\right].$$

The $7+$ bin is omitted from the level-moment stack because the tracer-day weights plus the residual mass sum to one within each job by construction.

Tightness moments. To discipline how market tightness shapes the relationship between application timing and the probability of contact initiation, I target the coefficient ϕ from the auxiliary regression (5.1), using the state-month log tightness measure from JOLTS as the main market-level covariate and the initiation-weight mass as the dependent variable. As robustness checks, I construct two additional sets of moments based on local tightness terciles; see Section D.1 for details. I estimate the model under each alternative moment specification, generate simulated data, and recompute the benchmark regression moment from the simulations. This procedure verifies that the main results are robust to alternative tightness-based moment specifications.

Bootstrapping. To estimate sampling uncertainty in the moments, I implement a stratified job-cluster bootstrap with 500 draws. In each replication, I resample jobs with replacement within each wave group, reconstruct the job panel using the resampled job identifiers, and recompute the full moment vector. The bootstrap covariance matrix $\widehat{\Sigma}$ is the sample covariance of the bootstrap moment draws, and the diagonal standard errors $\widehat{se}$ are the square roots of its diagonal.

7.3 Parameterization

This subsection specifies how the model primitives (λ_j, κ_j) are mapped to observable job characteristics and labor-market conditions.

7.3.1 Arrival intensity

For each job j, I write the organic arrival intensity as:

$$\lambda_{0j}^{\text{cong}} = \exp\left(\log \xi_j + \alpha_c \widetilde{\log c_j}\right), \tag{7.3}$$

where $\log c_j$ is the local congestion measure (e.g., $-\log(\theta_j)$) and $\widetilde{\log c_j}$ is centered by wave group so that $\log \xi_j$ absorbs wave-level level shifts in arrival conditions. To allow flexible heterogeneity in arrival conditions across jobs, I introduce lognormal multiplicative shocks and integrate them out by simulation:

$$\lambda_{jk} = \lambda_{0j}^{\text{cong}} \exp(\nu_{jk}), \qquad \nu_{jk} \sim \mathcal{N}\left(-\tfrac{1}{2}\sigma_\nu^2, \sigma_\nu^2\right), \tag{7.4}$$

with common random numbers across objective evaluations.

7.3.2 Delay flow cost

I parameterize the job-specific delay flow cost κ_j in logs as

$$\log \kappa_j = \log \kappa_0 + \exp\left(\kappa_{\theta,0} + \kappa_{\theta,1} \cdot \widetilde{\log \xi_j}\right) \cdot z\left(\log \theta_j\right), \tag{7.5}$$

where $z(\cdot)$ denotes global standardization and $\widetilde{\log \xi_j}$ is a centered macro-state index. I impose theory guided sign restrictions via exponential parameterizations.

7.3.3 Macro state shifter

The macro state ξ_j shifts arrival conditions and interacts with the κ channel. I perform the estimation in two steps:

(i) Wave-group fixed effects: In the first step, I treat the macro state a wave-group fixed effect:

$$\log \xi_j = \log \xi_{h(j)}, \quad h(j) \in \{1, 2, 3, 4\}. \tag{7.6}$$

This flexible specification provides an anchor for disciplined specifications.

(ii) Disciplined macro-series parameterization In the second step, I warm-start from the wave-group fixed effect specification and parameterize $\log \xi_j$ as a linear function of observed labor-market flow series measured at month $t(j)$

and state-by-month deviations:

$$\log \xi_j = \xi_0 + \sum_{k \in \mathcal{K}} \xi_k z \left(\widetilde{M}_{t(j),k} \right) + \sum_{k \in \mathcal{K}} \xi_k^{\Delta} z \left(\widetilde{M}_{s(j),t(j),k} - \widetilde{M}_{t(j),k} \right), \tag{7.7}$$

where $\widetilde{M}_{t,k}$ is the transformed month-level macro series and $\widetilde{M}_{s,t,k}$ is the corresponding state-by-month series. I consider several candidate series and evaluate a small set of pre-specified combinations. The final model include hire rates to capture the job search outcomes, both quit rates and discharge rates to capture the market churning.[28]

7.4 MSM estimator and computation

Let Θ denote the vector of parameters to be estimated. In the benchmark wave-group specification,

$$\Theta = \left(\{\log \xi_h\}_{h=1}^{4}, \alpha_c, \log \sigma_\nu, \log \kappa_0, \kappa_{\theta,0}, \kappa_{\theta,1} \right).$$

In the macro-series specification, $\left(\xi_0, \{\xi_k\}_{k \in \mathcal{K}}, \{\xi_k^{\Delta}\}_{k \in \mathcal{K}}\right)$ replaces $\{\log \xi_h\}$ as in (7.7), with $\mathcal{K}$ determined by the specification under consideration.

Given Θ, I simulate model-implied initiation weights as follows. For each job j, I compute λ_{jk} from (7.3)-(7.4) and κ_j from (7.5). For each simulation draw k, I obtain $(\lambda_{jk}, \kappa_{jk})$ and compute $p_{jk} = p(\lambda_{jk}, \kappa_{jk})$ using the precomputed $(\log \lambda, \log \kappa)$ table and a two-dimensional interpolator. I then compute expected tracer initiation weights $\omega_{ij}(\lambda_{jk}, p_{jk})$ using (7.2) and average over draws.

Let $g(\Theta)$ denote the simulated moment vector computed from these simulated initiation weights. The baseline objective uses a diagonal weight matrix based on bootstrap standard errors $\widehat{se}_m$:

$$\hat{\Theta}_{\text{diag}} = \arg\min_{\Theta} Q_{\text{diag}}(\Theta), \quad Q_{\text{diag}}(\Theta) = \sum_{m=1}^{M} \left(\frac{g_m(\Theta) - g_m^{\text{data}}}{\widehat{se}_m} \right)^2. \tag{7.8}$$

I implement estimation by minimizing the corresponding residual vector via nonlinear least squares, and use common random numbers so that the objective is smooth in Θ. I also report a full-weight estimator with the bootstrap covariance

[28] The mapping is motivated by standard matching-function logic linking job-filling rates and congestion to an aggregate matching-efficiency shifter. Appendix D.2 provides details.

matrix $\widehat{\Sigma}$:

$$\hat{\Theta}_{\text{full}} = \arg\min_{\Theta} Q_{\text{full}}(\Theta), \quad Q_{\text{full}}(\Theta) = \big(g(\Theta) - g^{\text{data}}\big)' \widehat{\Sigma}^{-1} \big(g(\Theta) - g^{\text{data}}\big), \tag{7.9}$$

implemented with a ridge-stabilized Cholesky whitener rather than explicit matrix inversion. Because full-weight criteria can be numerically sensitive in macro-series specifications, I consider a regularized full-weight variant that augments the residual vector with weak quadratic penalties anchored at the benchmark wave-group estimates for selected structural parameters and centered at zero for macro slopes.

Table 4: Identification Map: MSM Moments and Parameter Blocks

Moments	Model objects	Parameters
Level moments (32)	Day-mass profile: hazard of acceptable outside arrivals λp and selectivity p	α_c, σ_ν
	Stopping incentive κ	$\log \kappa_0, \kappa_{\theta,0}, \kappa_{\theta,1}$
Tightness moment (1)	Tightness loading in timing profile	$\alpha_c, \kappa_{\theta,0}, \kappa_{\theta,1}$
Both	Macro index shifting $\log \xi_j$	$\{\log \xi_h\}_{h=1}^4$ $\xi_0, \{\xi_k\}_{k\in\mathcal{K}}, \{\xi_k^\Delta\}_{k\in\mathcal{K}}$

Note: Each wave contributes 8 level moments (day bins plus "none," dropping the "7+" day bin by adding-up). σ_ν governs unobserved *arrival* heterogeneity in λ. No estimation moments target match-quality dispersion; the value-shift random effect ν_j^ν is therefore normalized (set to zero) in estimation and is used only for post-estimation cost calculations via the ML screening index.

Table 4 summarizes the identification map. In the main specifications, I use 33 moments to estimate 9 parameters in the benchmark specification with wave-group fixed effects and 12 parameters in the macro-disciplined specification. Table A.12 reports the empirical moments and their bootstrap standard errors.

Because the experiment identifies contact timing rather than the full cardinal distribution of applicant values, The estimation focuses firm's stopping rule conditional on normalizing the latent quality distribution. Appendix Section B.3 shows that first-contact timing behavior is invariant to affine transformations of latent applicant value, so the estimated cost terms are normalized composites.

7.5 Estimation results

Table 5 reports parameter estimates under four specifications: the wave-group benchmark (Column 1 and 2) and the disciplined macro-series specification (Column 3 and 4), each estimated under diagonal and full bootstrap covariance weighting.[29] Panel A reports the structural parameters governing the arrival and selection objects that map into experimental outcomes; Panel B reports the estimated macro-state parameters, either wave-group shifters $\{\log \xi_h\}$ in the benchmark or macro-series coefficients in the disciplined specification. In all estimations, only jobs with 8 submissions are used. Two high-level patterns stand out.

First, the estimates are broadly robust to the weighting matrix. Moving from diagonal weighting to the full optimal weighting shifts levels of individual parameters, which is as expected when moments are correlated and the full weighting matrix reallocates emphasis across moment families. But the qualitative structure of the estimates is stable. In particular, the wave-group benchmark continues to load strongly on the wave shifters $\log \xi_h$, and the macro specification continues to attribute variation in $\log \xi$ primarily to the hiring and separation series and their within-month local deviations. Likewise, the key state dependence embedded in the κ parameterization remains present under both weighting methods.

Second, the macro-disciplined specifications allow for greater flexibility over time and deliver a better fit to the data. The benchmark specification with wave-group fixed effects allows $\log \xi_h$ to vary freely across the four experimental wave groups, but it cannot fully capture the month-to-month variation that is better reflected in the macro-disciplined specification. This additional flexibility along the time dimension improves overall fit.

I break down model fit by moment family in Panel C of Table 5 and Figure A.11. Across specifications, the model fits the level moments describing the distribution of first-contact timing and the no-contact mass well. It also continues to match the tightness-gradient moments closely. These moments play an important role in disciplining the state dependence embedded in the parameterization of κ. In Figure 6, I plot the empirical day-bin mass of contact initiation probabilities together with their simulated counterparts under the disciplined macro-series

[29] The full-weight macro estimates are reported with weak regularization to stabilize optimization in weakly identified directions. The objective comparisons reported in the appendix always separate the data component from any penalty terms.

Table 5: Parameter Estimation Results and Model Fitting

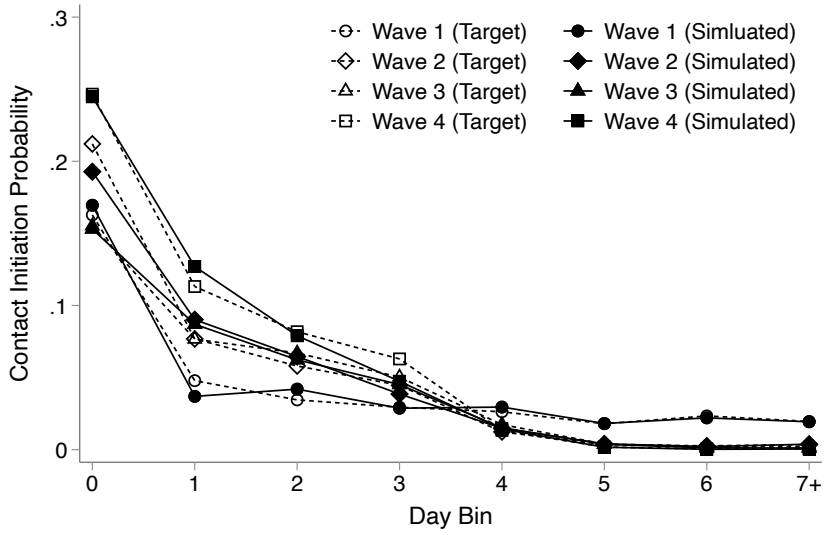
	(1) Wave (diag)	(2) Wave (full)	(3) Macro (diag)	(4) Macro (full)
Panel A: Structural parameters				
α_c	0.844	0.665	1.029	0.898
$\log \sigma_\nu$	0.711	0.509	0.713	0.512
$\log \kappa_0$	-4.524	-4.341	-4.600	-4.430
$\kappa_{\theta,0}$	-3.950	-4.000	-4.000	-4.000
$\kappa_{\theta,1}$	1.936	1.828	2.000	1.923
Panel B: ξ parameters				
$\log \xi_1$	2.464	2.482		
$\log \xi_2$	2.175	2.150		
$\log \xi_3$	2.474	2.506		
$\log \xi_4$	1.535	1.546		
ξ_0			1.654	1.656
ξ_{HIR}			-0.529	-0.580
ξ_{LDR}			-1.066	-1.144
ξ_{QUR}			-0.441	-0.439
$\xi^{\Delta}_{\text{HIR}}$			-0.208	-0.211
$\xi^{\Delta}_{\text{LDR}}$			0.688	0.663
$\xi^{\Delta}_{\text{QUR}}$			-0.495	-0.452
Panel C: Model fitting summary				
Q_{diag}	28.096	29.186	24.795	25.663
Q_{full}	30.723	29.821	27.123	26.689
Q_{level}	28.093	28.924	24.778	25.524
Q_θ	0.003	0.262	0.017	0.139
N(submissions)		78632		
N(jobs)		9829		

Note: This table reports parameter estimates and MSM objectives for four specifications: wave-group $\log \xi$ with diagonal weighting in Column (1), wave-group $\log \xi$ with full covariance weighting in Column (2), macro-series $\log \xi$ with diagonal weighting in Column (3), and macro-series $\log \xi$ with full covariance weighting in Column (4). Panel A reports structural parameters governing the arrival rate λ and the vacancy cost shifter κ. Panel B reports the macro-state parameters: in the wave-group specification these are wave-group shifters $\{\log \xi_h\}$; in the macro specification these are coefficients on standardized macro series and within-month deviations. All macro inputs are standardized prior to estimation. Bounds are imposed for numerical stability. Panel C reports the MSM objective and its decomposition by moment blocks. The diagonal criterion uses inverse bootstrap variances. The full-weight criterion uzses the inverse bootstrap covariance matrix with ridge adjustment and is implemented via a Cholesky whitener. Block contributions are sums of squared standardized residuals within each moment family.

specification with diagonal weighting. The simulated series captures the downward trend across wave groups well. Taken together, these comparisons suggest

that the model matches the empirical patterns effectively, and that the substantive conclusions do not hinge on a particular weighting choice, even though the full-weight estimator mechanically changes the relative importance assigned to different moments.

Figure 6: Level Moments, Simulation versus Empirical Targets



Notes: The figure plots the empirical and simulated level moments, that is, the contact-initiation probability mass by day bin and wave group, under the disciplined macro-series specification with diagonal weighting. Empirical targets are marked with hollow markers and dashed lines, while simulated moments are marked with solid markers and solid lines. Both the empirical calculations and the model simulations are based on jobs that received eight experimental applications.

Interpreting the macro coefficients in ξ. In the macro series specification, ξ_j functions as a vacancy-side shifter of effective arrival or contact intensity conditional on congestion. In matching language, it is best interpreted as a reduced-form index of recruiting intensity, the composition of vacancies and matching efficiency beyond what is accounted for by labor market tightness (Petrongolo and Pissarides, 2001). The estimated signs are consistent with this interpretation.

First, the negative loading on the hiring rate (HIR) and its local deviations (ΔHIR) are consistent with evidence that vacancy yields and job-filling rates move countercyclically, so that high-hiring periods can coincide with low per-vacancy success and arrival intensity once congestion is controlled for.[30]

[30]Using JOLTS data, Davis, Faberman and Haltiwanger (2013) document that the job filling rate

Second, the negative loading on the quit rate (QUR and ΔQUR) accords with job-to-job reallocation mechanisms: Quits proxy for employer-to-employer transitions that expands vacancy creation through replacement and poaching faster than it expands effective per-vacancy filling or contact, so vacancy yield falls and firms face more recruiting difficulty per posting, conditional on tightness (Fallick and Fleischman, 2004).[31]

Finally, the negative loading on layoffs/discharges (LDR) is naturally read as a matching efficiency channel: displacement and sectoral reallocation shift unemployment composition and amplify dispersion across labor market segments, lowering matching efficiency conditional on tightness (Barnichon and Figura, 2011; Birinci, Mercan and See, 2025). The positive loading on its local deviations (ΔLDR) is consistent with local displacement temporarily thickening the pool of active job seekers.[32] The estimation results also highlight the distinction between how quits and layoffs reflect underlying micro relations over the cycle (Davis, Faberman and Haltiwanger, 2006).

Importantly, coefficient magnitudes are not directly comparable across macro series because the volatility of $\log(\text{rate})$ differs sharply across LDR, QUR, and HIR. I therefore emphasize one-standard-deviation effects when discussing relative importance. Because the macro series are correlated and ξ is identified jointly with other structural parameters, these loadings should be interpreted as conditional correlations in a disciplined index rather than as causal elasticities.

To furthur validate the model and estimation, I conduct three more diagnostics: First, I plot the implied macro state $\log \xi$ over time in Figure 7. The wave-group

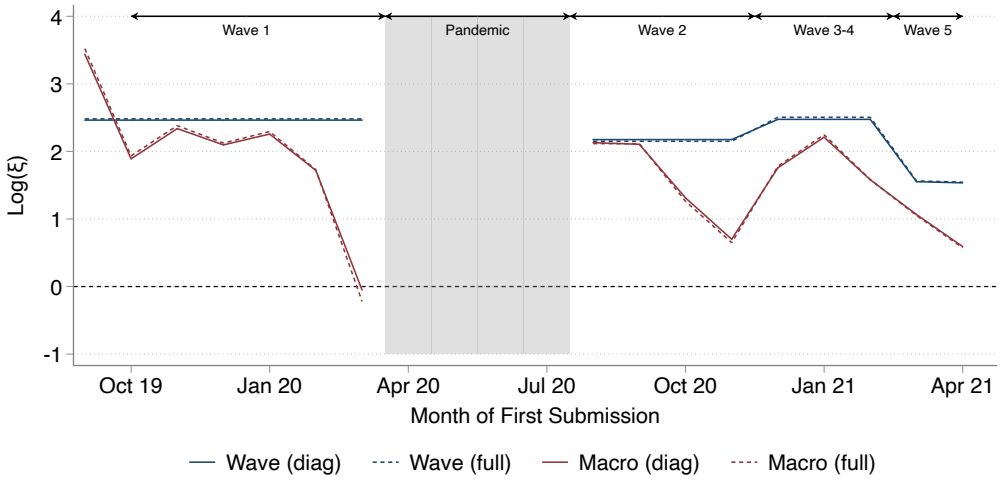
for vacant positions move counter to aggregate employment in the cross section. Recent empirical evidence shows that firms respond to tight conditions by increasing recruiting intensity per vacancy (Davis, Faberman and Haltiwanger, 2012; Forsythe and Weinstein, 2025), raising the cost and time of filling each position rather than mechanically raising the job-filling probability. See Faberman (2020) for a synopsis. Recent flow-based discussion explicitly highlights declines in job-filling rates and matching efficiency post-pandemic (Eusepi and Şahin, 2025).

[31] A dual-vacancy interpretation of recent Beveridge-curve dynamics formalizes this mechanism (Cheremukhin and Restrepo-Echavarria, 2022). Related work also links job-to-job mobility and quits to inflationary pressures and labor-market tightness (Autor, Dube and McGrew, 2023; Mercan, Schoefer and Sedláček, 2024). Bagga et al. (2025) show that in the 2020-2023 recovery, industries with higher quits posted more vacancies but had lower job-filling rates. Afrouzi et al. (2026) emphasize a related mechanism in which higher employer-to-employer (E-E) flows increase E-E vacancies but reduce the firm's E-E vacancy filling rate and raise vacancy duration.

[32] Policies that preserve search while employed around layoffs also highlight that the timing and nature of separations matter for efficiency (Cederlöf et al., 2025).

specification flexibly estimates discrete shifts across experimental regimes, while the macro-series specification produces a smoother and more interpretable path that tracks the broad dynamics of the benchmark state process. Both specifications produce similar trend of $\log \xi$, since the levels of $\log \xi$ are not separately identified from scale normalizations, the trend similarity shows that the macro series parameterization does capture the state-shifting nature of the experiment period.

Figure 7: Estimated Macro State $\log \xi$



Notes: The figure plots the implied macro state $\log \xi$ under the benchmark wave-group specification and the disciplined macro-series specification. For comparability, the figure reports month means of the implied $\log \xi$ across vacancies observed in each month. The wave-group specification allows discrete shifts across experimental regimes defined by pooled waves, whereas the macro specification generates a smoother process driven by observed macro series. The levels of $\log \xi$ are not separately identified from scale normalizations, so the figure emphasizes relative movements over time and differences across regimes.

Second, I provide an external validity check for the estimated time shifter ξ using national monthly vacancy-yield data from JOLTS, which are not targeted moments in the estimation. Motivated by the standard Cobb-Douglas matching identity, I construct a tightness-adjusted matching-efficiency proxy $\widehat{\log \Lambda_t}$ using the model-implied congestion elasticity $\widehat{\alpha}_c$, and examine whether it comoves with the estimated $\log \xi_t$.[33] Figure A.12 plots estimated $\log \xi_t$ under macro-disciplined specification with diagonal weighting against $\widehat{\log \Lambda_t}$ and reports the slope from a

[33]Under Cobb-Douglas matching, $H_t = \Lambda_t U_t^{\alpha} V_t^{1-\alpha}$ implies the vacancy yield $q_t \equiv H_t/V_t$ satisfies $\log q_t = \log \Lambda_t + \alpha \log(U_t/V_t)$. Accordingly, I construct the proxy

$$\widehat{\log \Lambda_t} \equiv \log\left(\frac{H_t}{V_t}\right) - \widehat{\alpha}_c \log\left(\frac{U_t}{V_t}\right),$$

bivariate regression of $\log \xi_t$ on $\widehat{\log \Lambda_t}$ (coefficient 0.805). With only 16 monthly observations, this check is primarily directional. The positive comovement supports interpreting ξ as a vacancy-side efficiency/intensity shifter rather than a mechanical artifact of the structural fit.

Lastly, Table A.13 decomposes the fitted variation in the macro state $\log \xi_j$ into the additive components implied by the macro-series parameterization in macro specification. Each row reports the mean and cross-job standard deviation of the contribution from a given macro term, evaluated at the estimated coefficients and taken over the job-level sample. Because the mapping into $\log \xi_j$ is linear in the macro covariates, this decomposition provides a transparent summary of which observed macro series are shaping the fitted macro state. The monthly layoff component is the largest driver of $\log \xi_j$ variation (S.D. 0.62 under diagonal weighting and 0.67 under full weighting). Deviation terms remain quantitatively meaningful, especially for both separation series. These patterns are consistent with the idea that ξ is disciplined by match formation intensity and reallocative dynamics that affect effective matching conditions.

In Table D.1 of Appendix D, I show the 10 moments with the largest standardized discrepancies in both wave and macro specification with diagonal weighting. It makes clear which specific wave-by-day cells and reduced-form gradients remain the primary sources of miss, and it shows that the pattern of miss is broadly stable across weighting choices. In the counterfactual exercises, I therefore focus on changes that are well identified by the moments the model matches tightly, and I treat the remaining deviations as a limitation of the parsimonious mapping from macro observables into the model's state.

7.6 Counterfactuals and Cost-Benefit Analysis

This subsection uses the estimated model to quantify the trade-off between *speed* and *match quality* in the firm's contact-initiation decision. The model rationalizes early contact as an optimal stopping rule in which the firm weighs the option value of waiting for better applicants against a flow cost of delay. A key empirical challenge is that match quality is not observed in the experiment. I therefore combine

where H_t/V_t is the national monthly hires-per-opening series from JOLTS and U_t/V_t is national monthly unemployed-per-opening; $\widehat{\alpha}_c$ is the estimated congestion elasticity from the baseline macro-disciplined specification (Column (3) of Table 5).

the model-implied initiation weights with a data-driven proxy for screening quality based on machine-learning predictions.

I first introduce the ML-based objects in Section 7.6.1, and conduct counterfactual exercises in Section 7.6.2 and 7.6.3, and cost-benefit analysis in Section 7.6.4. Throughout, the counterfactual exercises are implemented with the diagonal weighted macro-series specification (Column (3) of Table 5).

7.6.1 Machine learning screening index and cost-benefit measures

In the model (Section 6.1.3), screening reveals a signal a_{ij} and the employer forms a posterior value $v_{ij} = \mathbb{E}[q_{ij} \mid a_{ij}, \mathbf{x}_{ij}, \mathbf{z}_j, \nu_j^v]$. The contact-initiation problem depends on *rankings* of applicants by this posterior value, equivalently by a_{ij} under monotonicity. The empirical challenge is that neither a_{ij} nor v_{ij} is observed. I therefore construct an empirical one-dimensional index that is informative about the employer's latent screening value using machine-learning predictions.

Essence of the ML model: learning the employer's initial screening rule. Let $C_{ij} \in \{0, 1\}$ indicate whether an application i to vacancy j is contacted.[34] Conceptually, early in the recruiting spell the employer applies an unobservable screening rule that maps observables $(\mathbf{x}_{ij}, \mathbf{z}_j)$ into a latent evaluation index v_{ij}, and contacts applicants with high enough value. This implies that the conditional contact probability $\Pr(C_{ij} = 1 \mid \mathbf{x}_{ij}, \mathbf{z}_j)$ is monotone in the employer's evaluation index under weak regularity conditions. The ML model estimates this conditional probability flexibly, using rich posting and resume attributes, and uses it as a *revealed-preference* summary of how employers value observable resume content in this setting.

Why training on first applications removes timing. A central goal is to construct a screening index that is not mechanically confounded by application timing. I therefore train the prediction model using only the first experimental application within each vacancy. The normalized arrival time of all training observations is $d_{ij} = 0$. The fitted mapping therefore reflects the employer's initial-screen evaluation strategy. I then use the trained model to predict the contact probabilities of later experimental applications. The predicted probabilities can be interpreted as

[34] This contact indicator is distinct from C_{ij}^{first} used earlier to define the *first* contacted experimental applicant.

how likely a resume would be contacted if it arrived at the start of the recruiting spell, based only on resume and posting attributes. This interpretation effectively separate "quality" from "timing advantage" in the cost-benefit analysis.

Out-of-fold predictions and the empirical screening index. Let $\widehat{p}_{ij}$ denote the out-of-fold cross-fit predicted probability of contact from this model. Out-of-fold prediction ensures $\widehat{p}_{ij}$ is not mechanically inflated by in-sample overfitting. Because the structural model only requires a one-dimensional index that is monotone in the employer's posterior value, I define the empirical screened value as

$$\widehat{a}_{ij} \equiv h(\widehat{p}_{ij}),$$

where $h(\cdot)$ is any monotone transform. In the baseline I set $h(p) = p$.

Two measures of rushing loss within a vacancy. Fix a vacancy j and let I_j denote its eight experimental (tracer) applications. For each tracer (i, j), let $\widehat{a}_{ij}$ be the ML screening index defined above, and let $d_{ij} \in \{0, \ldots, 7+\}$ denote its within-vacancy arrival-day bin (with $7+$ top-coded). Define the best predicted screening value among all tracers and the best among those that have arrived by day d:

$$\text{BestAll}_j \equiv \max_{i \in I_j} \widehat{a}_{ij}, \qquad \text{BestPre}_j(d) \equiv \max_{i \in I_j:\ d_{ij} \le d} \widehat{a}_{ij}.$$

The rushing loss if contact is initiated by day d is

$$\text{Loss}_j(d) \equiv \text{BestAll}_j - \text{BestPre}_j(d) \ge 0,$$

and the corresponding "missed-best" indicator is

$$\text{MissedBest}_j(d) \equiv \mathbf{1}\{\text{BestPre}_j(d) < \text{BestAll}_j\} = \mathbf{1}\{\text{Loss}_j(d) > 0\}.$$

These are vacancy-level objects that capture the option value of waiting *within the experimental set*.

To aggregate these losses with the structural model, let $\omega_{ij}(\widehat{\Theta})$ denote the model implied initiation weight for tracer (i, j), and define the induced day-bin mass

$$P_{jd}(\widehat{\Theta}) \equiv \sum_{i \in I_j} \omega_{ij}(\widehat{\Theta}) \mathbf{1}\{d_{ij} = d\}, \qquad P_j^{\text{exp}}(\widehat{\Theta}) \equiv \sum_d P_{jd}(\widehat{\Theta}) = \sum_{i \in I_j} \omega_{ij}(\widehat{\Theta}).$$

Here $P_{jd}(\widehat{\Theta})$ is the probability that the initiating contact is a tracer arriving in bin d, and $P_j^{\text{exp}}(\widehat{\Theta})$ is the probability that the initiating contact is among tracers.[35]

Because ML scores are observed only for tracers, I report ML-based rushing costs conditional on an experimental initiator. **The first measure** is the probability that the ex post best tracer has not yet arrived when contact is initiated:

$$\Pr(\text{missed best} \mid \text{exp. contact}) = \frac{\sum_j \sum_d P_{jd}(\widehat{\Theta})\text{MissedBest}_j(d)}{\sum_j P_j^{\text{exp}}(\widehat{\Theta})}. \tag{7.10}$$

The second measure is a scale-free expected loss, defined as the ratio of the expected shortfall from the best tracer to the expected screening value of the contacted tracer:

$$\frac{\mathbb{E}[\text{loss} \mid \text{exp. contact}]}{\mathbb{E}[\text{score} \mid \text{exp. contact}]} \equiv \frac{\sum_j \sum_d P_{jd}(\widehat{\Theta})\text{Loss}_j(d)}{\sum_j \sum_{i \in I_j} \omega_{ij}(\widehat{\Theta})\widehat{a}_{ij}}. \tag{7.11}$$

In the implementation, I calculate $\mathbb{E}[\text{score} \mid \text{exp. contact}]$ as the initiation weighted mean of $\widehat{a}_{ij}$ across tracers; $\mathbb{E}[\text{loss} \mid \text{exp. contact}]$ and $\Pr(\text{missed best} \mid \text{exp. contact})$ are calculated by weighting $\text{Loss}_j(d)$ and $\text{MissedBest}_j(d)$ by the model-implied contact-day mass $P_{jd}(\widehat{\Theta})$, all normalized by the total experimental-initiation probability.

7.6.2 Counterfactuals exercises

Using the estimates in Column (4) of Table 5, I first calculate the baseline rushing loss implied by the model. Table 6 reports wave-specific averages of (i) the probability that the initiating contact falls on one of the eight experimental applications, and (ii) two ML-based measures of rushing loss conditional on an experimental initiator.

Conditional on an experimental initiator, the probability of having missed the best experimental resume by the initiation date is around 0.5 across waves. This indicates that the option value of waiting is quantitatively important even within the experimental set. The scale-free loss introduced in (7.11) is also economically meaningful: In the baseline, the expected shortfall from the ex post best experimental resume amounts to 6.55% of the expected screening score of the contacted

[35] The residual mass $1 - P_j^{\text{exp}}$ corresponds to the "none" category.

Table 6: Estimated Rushing Loss, Baseline

Sample	N(jobs)	$P_j^{\exp}(\widehat{\Theta})$	$1 - P_j^{\exp}(\widehat{\Theta})$	Prob. Loss	Scale Loss
Wave Group 1	1590	0.3662	0.6338	0.5391	0.0920
Wave Group 2	2064	0.4107	0.5893	0.4993	0.0621
Wave Group 3	4157	0.3686	0.6314	0.5121	0.0654
Wave Group 4	2018	0.5149	0.4851	0.5150	0.0597
Full	9829	0.4071	0.5929	0.5141	0.0655

Note: This table reports baseline rushing-loss measures implied by the estimated model (Column (4) of Table 5), computed on the sample of complete postings with eight experimental submissions. $P^{\exp}(\widehat{\Theta})$ is the wave-specific average probability that the initiating contact is one of the experimental applications. $1 - P^{\exp}(\widehat{\Theta})$ is the corresponding probability that the initiating contact falls outside the experimental set. "Prob. Loss" is calculated as in (7.10), "Scale Loss" is calculated as in (7.11). ML scores are out-of-fold predicted contact probabilities as discussed in Section 7.6.1 and C. Wave bins follow the paper's wave definition, with experiment Waves 3 and 4 pooled.

experimental resume, with the largest losses in Wave 1 (9.2%).

After establishing the baseline, Table 7 reports three counterfactual exercises conducted on all jobs, Table A.14 reports counterfactual results by wave groups. Each exercise perturbs a targeted element of the environment or the decision problem, recomputes the model-implied initiation weights, and then recomputes the same two loss measures. Both the value and percentage point change of counterfactual results from the baseline results are reported.

Counterfactual A: Market tightness θ_{mt}. I first shift local tightness θ_{mt} so that the mean tightness matches the $p10$, $p50$, or $p90$ of the baseline distribution while preserving cross-posting heterogeneity. Panel A of Table 7 shows a pronounced monotone response: moving from low (10th-percentile) to high (90th) tightness increases both the probability of missing the best and the scale loss. These patterns are consistent with the reduced-form evidence that early application advantages are larger in tighter labor markets.

Counterfactual B: Uniform macro-shifter shock. I shift the macro intercept by adding a constant $\Delta \in \{-2, +2\}$ to ξ_0, holding all other parameters fixed. This counterfactual changes the macro environment to mimick variations in match efficiency uniformly across postings and waves. Panel B of Table 7 shows that both loss measures vary significantly with changes in ξ. Moving from $\Delta = -2$ to

Table 7: Estimated Rushing Loss, Counterfactuals on Full Sample

	Prob. Loss	Scale Loss
Baseline	0.5141	0.0655
Counterfactual A: Change in market tightness θ_{mt}		
$\theta_{mt} = p10$	0.5020	0.0634
	[-1.203]	[-3.189]
$\theta_{mt} = p50$	0.5149	0.0656
	[0.082]	[0.219]
$\theta_{mt} = p90$	0.5251	0.0675
	[1.105]	[2.990]
Counterfactual B: Uniform change in macro shifter ξ_0		
$\Delta = -2$	0.5901	0.0792
	[7.605]	[20.958]
$\Delta = +2$	0.4595	0.0564
	[-5.458]	[-13.934]
Counterfactual C: Shadow cost of waiting κ_j		
$\times 0.5$	0.4925	0.0619
	[-2.156]	[-5.490]
$\times 2$	0.5461	0.0710
	[3.206]	[8.353]

Note: This table reports the two loss measures as in Table 6 under four counterfactual exercises. In each counterfactual, the model-implied initiation weights are recomputed holding fixed observed posting and resume attributes and the realized within-posting arrival times of the eight experimental applications. Panel A shifts market tightness θ_{mt} by shifting log tightness so that its mean matches the $p10$, $p50$, or $p90$ of the baseline distribution while preserving cross-posting heterogeneity. Panel B adds $\Delta \in \{-2, +2\}$ to the macro intercept ξ_0 holding all other parameters fixed. Panel C scales the vacancy flow cost κ_j by a multiplicative factor (0.5 or 2). All reported objects are conditional on an experimental initiator because ML scores are observed only for experimental resumes. The differences of loss in percentage point bewteen each counterfactual and the baseline are reported in brackets: For prob. loss, the numbers are calculated by subtracting the baseline results from the counterfactual results and times 100. For scale loss, the numbers are calculated as the difference between counterfactual and baseline results divided by baseline results, times 100. The sample is all jobs with 8 experimental applications (9829 vacancies).

$\Delta = +2$ changes the missed-best probability by 13%, and scale loss by around 35%. These results suggest that macro conditions other than tightness also contribute to the changes in application timing effects.

Counterfactual C: Vacancy flow cost κ_j. Lastly, I directly scale the shadow cost of waiting κ_j by a factor of 0.5 or 2, holding the arrival environment fixed. Panel C of Table 7 shows that scaling κ_j changes rushing losses in the expected direction:

halving κ_j lowers probability loss by 2.2% and scale loss by 5.4%, while doubling κ_j increases them by 3.2% and 8.3% respectively.

7.6.3 Channel Decomposition with Dose-Response Counterfactuals

The discrete counterfactuals quantify economically meaningful changes in rushing loss, but they bundle together two distinct primitives of the model. A given labor-market perturbation can (i) reshape the arrival environment by changing effective applicant inflow and congestion, summarized by λ_j, and (ii) reshape the stopping incentive by changing the flow cost of delay, summarized by κ_j. In the model, I allow tightness to move both λ_j and κ_j, the total effects in Table 7 do not directly reveal whether counterfactual changes come from shifts in the opportunity set, characterized by λ, or through shifts in the cost of waiting, characterized by κ. Hence, I decompose the key counterfactual effect into its λ and κ components.

Methodology: Shapley decomposition over λ and κ. Let $Y(\lambda, \kappa)$ denote any counterfactual target, For a counterfactual x that shifts primitives from (λ_0, κ_0) to (λ_x, κ_x), I evaluate Y in four environments:

$$Y_{00} \equiv Y(\lambda_0, \kappa_0), \qquad Y_{10} \equiv Y(\lambda_x, \kappa_0), \qquad Y_{01} \equiv Y(\lambda_0, \kappa_x), \qquad Y_{11} \equiv Y(\lambda_x, \kappa_x),$$

corresponding to baseline, λ-only, κ-only, and the full counterfactual. A naive one-at-a-time decomposition is order dependent when channels interact. I therefore use a two-player Shapley decomposition that averages marginal contributions across the two possible orders:

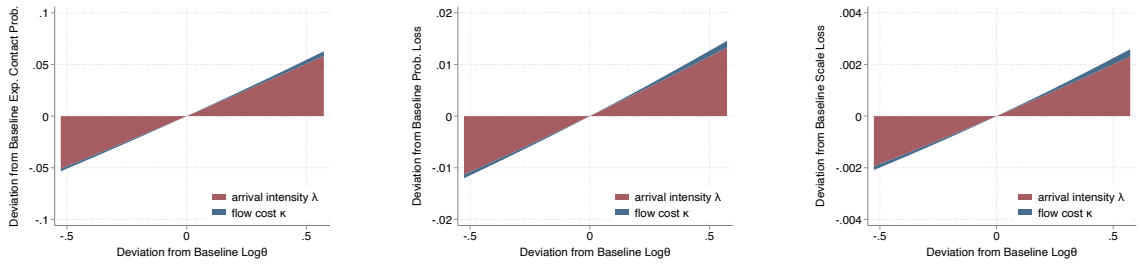
$$\phi_\lambda(Y) \equiv \frac{1}{2}\left[(Y_{10} - Y_{00}) + (Y_{11} - Y_{01})\right], \quad \phi_\kappa(Y) \equiv \frac{1}{2}\left[(Y_{01} - Y_{00}) + (Y_{11} - Y_{10})\right],$$

so that $\phi_\lambda(Y) + \phi_\kappa(Y) = Y_{11} - Y_{00}$.

To better characterize the effects and distinguish the two channels over the relevant range of the environment, I use continuous dose-response counterfactuals that trace $Y(x)$ as the counterfactual lever varies smoothly. I vary each lever x over a grid of deviations from baseline and recompute the total effect $Y_{11}(x) - Y_{00}$ and the channel contributions $\phi_\lambda(Y; x)$ and $\phi_\kappa(Y; x)$.

Decomposition of tightness counterfactuals. In Figure 8, I plot the Shapley decomposition on the full sample for the continuous tightness counterfactual. The results reveal that the effect of labor market tightness primarily operate through the arrival intensity λ channel. For both the extensive margin of whether contact is initiated in the experiment (Figure 8a) and the two intensive rushing loss objects (Figure 8b and 8c), The λ channel accounts for more than 90% of the counterfactual effects. The κ channel is slightly more prominent for intenstive margins since the shadow cost of waiting directly governs the stopping rule and tilts contact weight toward earlier versus later days, which is the central to rushing loss. In Figure A.13, I show these results are largely robust across waves but particularly pronounced in Wave 1. Next, I use the same dose-response counterfactual design to quantify the economic costs of rushing loss.

Figure 8: Tightness Counterfactual Shapley Decomposition, Full Sample



(a) $P_j^{\exp(\hat{\Theta})}$ (b) Prob. Loss (c) Scale Loss

Notes: The figures plot Shapley decomposition of the counterfactuals into the arrival-intensity and the flow cost channel, as in Section 7.6.3. The horizontal axis reports the counterfactual deviation from the baseline mean log tightness. In Figure 8a, the vertical axis reports Shapley contributions to the change in the probability that initiating contact is among experiment applications. In Figure 8b, the vertical axis reports Shapley contributions in probabiliy loss defined as in (7.10), while in Figure 8c, it reports Shapley contributions in scale loss defined as in (7.11), relative to baseline. Sample: all jobs with 8 experiment applications.

7.6.4 Cost-benefit analysis

This subsection translates the model-implied rushing costs into economically interpretable magnitudes. Because match quality is unobserved, all cost calculations are anchored in the ML-based screening index and are best interpreted as wage-equivalent back-of-envelope magnitudes rather than structural wage predictions.

Under the payoff normalization in Section 6.1.5, match surplus is linear in match quality. If wages are approximately proportional to the match-quality com-

ponent summarized by the screening index, then the scale loss in (7.11) can be interpreted as an approximate percentage loss in match value due to initiating contact before the best experimental resume has arrived. To translate this scale loss into dollars, I set $\bar{w}^{\text{entry}} = \$47{,}150$.[36] The resulting wage-equivalent rushing cost is

$$\Delta\hat{w}(x) \equiv \text{ScaleLoss}(x) \times \bar{w}^{\text{entry}},$$

where x denotes the deviation from the baseline environment.[37]

As a benchmark, I first calculate the baseline wage-equivalent rushing costs $\Delta\hat{w}(0)$ for the sample of all jobs with 8 experimental applications as \$3,088. Figure A.14 plots $\Delta\hat{w}(x)$ for continuous shifts in log mean of local market tightness and standardized shifts in the aggregate hire rate (HIR), quit rate (QUR) and layoff/discharge rate (LDR).

Two patterns are robust. First, wage-equivalent rushing costs rise with tightness. This aligns with the reduced-form early application advantage evidence and with the model's mechanism: in tighter markets, firms initiate contact early and increase the expected shortfall from the best available applicant. Second, increase in all three macro series is monotonically associated with higher ex-post rushing loss. This pattern is consistent with the estimated values of ξ_{HIR}, ξ_{LDR} and ξ_{QUR} and the interpretation that the three series are negatively loaded into the matching efficiency/vacancy yield index ξ.[38] Taken together, the counterfactual curves provide an interpretable picture of how much ex-post match value is at stake as market conditions vary over empirically relevant ranges.

[36] The salary benchmark is the 2024 median annual wage for occupations whose typical entry-level education is a high school diploma or equivalent (U.S. Bureau of Labor Statistics, 2025).

[37] This mapping is intended as a scale conversion rather than a claim about observed wages, which are not observed in the experiment. Importantly, $\Delta\hat{w}(x)$ is an *ex post* measure of match value forgone at the endogenous stopping time; it should not be interpreted as the firm's vacancy flow cost (captured by κ_j), which is the primitive that rationalizes why firms initiate contact early.

[38] All three counterfactuals are mechanically λ-driven in the implementation because a uniform month-level shift primarily rescales the inflow component $\log \xi_j$ while leaving the centered macro state that enters κ_j approximately unchanged. λ affects timing through two opposing forces. A higher λ increases the rate at which outside applicants arrive, but it also raises the reservation value and reduces the acceptance probability $p(\lambda, \kappa)$. Hence the relevant hazard of acceptable outside arrivals is $\lambda p(\lambda, \kappa)$, whose derivative $d(\lambda p)/d\lambda = p + \lambda \partial p/\partial \lambda$ is a priori ambiguous since $\partial p/\partial \lambda < 0$. The counterfactual exercises support that selectivity response donimates. Hence increase in arrival intensity reduces hazard and vacancies wait longer before initiating the first contact.

Micro counterfactuals. So far, I show that labor market conditions vary arrival intensity and shadow cost of waiting for vacancies, therefore change the stopping time of firms. Next, I implement two micro counterfactual exercises to directly explore firms' screening strategy in the sequential arrival environment.

First, I shift the model-implied initiation-day distribution one day earlier (or later) mechanically, holding fixed (i) the overall probability that initiation falls on an experimental application and (ii) the baseline within-day shares across experimental resumes. This exercise expresses the rushing costs in "per-day" terms. Full sample calculations imply that moving the initiation-day distribution one day earlier increases the ex-post rushing loss by \$1,011, while moving the distribution one day later decreases the loss by \$1,371. These wage-equivalent terms are robust across wave groups as shown in Table A.15.

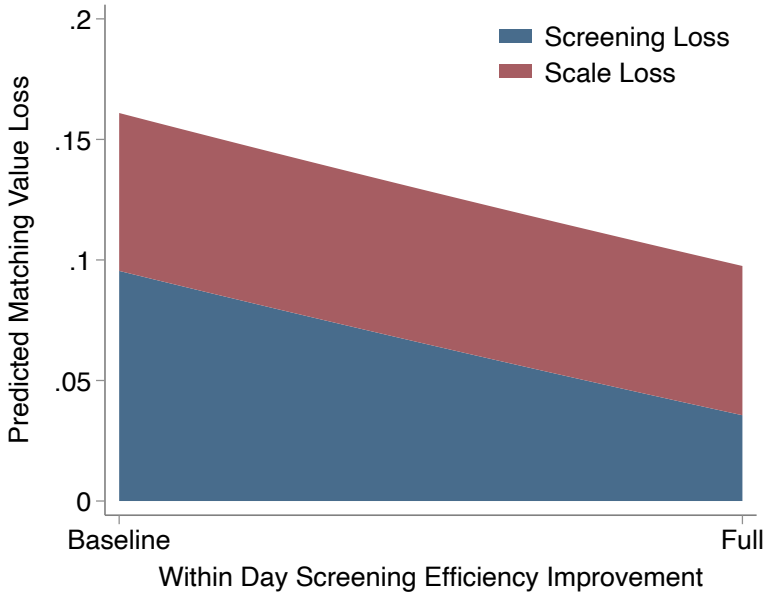
Second, I evaluate a stylized screening technology counterfactual that improves within-day ranking while holding fixed the contact-day distribution. In the estimation, screening cost is wrapped in the reduced-form shadow cost κ. Here, I break the cost of rushing into two parts: the scale loss that is generated from the sequential arrival, and a mis-selection cost conditional on timing, defined as

$$\text{Screening loss} \equiv \frac{\mathbb{E}[\text{BestPre} - \text{Score} \mid \text{exp. contact}]}{\mathbb{E}[\text{Score} \mid \text{exp. contact}]}, \tag{7.12}$$

This counterfactual exercise introduces a technology reallocates initiation weight to the experimental resume with the highest ML score within each day bin, leaving the timing margin unchanged. It therefore evaluates the scope for reducing rushing losses through better screening conditional on the same initiation speed.

In Figure 9, I vary the degree of screening technology adoption and plot the corresponding screening and scale loss. The adoption of such technology only reduce the scale loss by 5%, which reaffirms the conclusion that the rushing loss generated due to the sequential arrival nature of the environment can not be reduced by efficient screening alone. Meanwhile, the technology closes above 60% of the screening loss. The remaining 40% reflects cross-day queue-management frictions, e.g., imperfect recall or attention to earlier applicants, that cannot be eliminated by within-day ranking improvements alone and are distinct from the rushing loss due to not-yet-arrived applicants. Table A.16 shows the robustness of these results across waves. Overall, screening improvements can partially mitigate the cost of

Figure 9: Estimated Macro State $\log \xi$



Notes: The figure plots how the adoption of the within-day screening technology described in Section 7.6.4 changes the loss in predicted matching value. "Baseline" refers to zero adoption of the technology, "Full" refers to full adoption. Screening loss is defined as in (7.12), scale loss is defined as in (7.11). Sample: all jobs with 8 experimental applications.

rushing but cannot eliminate losses driven by initiating before high-quality applicants have arrived. This experiment suggests that the effects of timing may extend beyond not-yet-arrived applicants: as the pool accumulates over days, cross-day attention and recall constraints can generate additional time-related evaluation frictions. Quantifying these screening failures and the technologies that mitigate them is a promising direction for future research.

Final remark on model The estimated model is a *conditional equilibrium* (or *subgame equilibrium*) description of the firm's contact-initiation problem. Throughout, market conditions $\mathbf{S}_{mt}$ are treated as predetermined state variables from the perspective of an individual vacancy. Given $\mathbf{S}_{mt}$, the model pins down the firm's optimal reservation policy and the implied hazard of initiating contact, which jointly generate the observed early application advantage moments. In this sense, the paper estimates a firm-side policy rule for the contact-initiation stage, a margin typically left implicit, which can serve as a building block for future equilibrium work. In Appendix Section B.4, I sketch out how to embed the recruiting subgame

into the general equilibrium.

8 Model-guided heterogeneity in timing frictions

The preceding sections estimate a stopping-rule model of contact initiation and quantify the resulting speed-quality tradeoff. The counterfactual exercises show that environments in which firms place greater weight on early applicants also exhibit larger rushing losses in match quality. In the model, firms' behavior is governed by two channels: the application arrival rate, λ, and the shadow cost of delay, κ. This section returns to reduced-form heterogeneity in the daily early-application advantage (dEEA) and tests the model's external validity. Specifically, I examine how the early-application advantage varies with independent measures that may affect these two channels but are not directly targeted in the model estimation.

Interpretation and setup. Throughout, a larger dEEA indicates a steeper speed premium: applications arriving one day earlier are more likely to be contacted by 0.4%, and thus stronger timing frictions from the perspective of the model.[39] I use the speicification (8.1) to estimate the effect of observable market and firm level covariates on dEEA.

$$Y_i = \alpha_0 + \mu \cdot early_i^1 + \nu_m \cdot early_i^1 \times \text{Covar}_{m,i} + \boldsymbol{\gamma}\mathbf{X}_i + \epsilon_i \tag{8.1}$$

where $\text{Covar}_{m,i}$ is the value of covariate m for application i. I adopt the main specification with job fixed effects and standard errors clustered at the job level. To account for potential correlations between covariates, I also estimate a multivariate model

$$Y_i = \alpha_0 + \phi \cdot early_i^1 + \sum_m \psi_m \cdot early_i^1 \times \text{Covar}_{m,i} + \boldsymbol{\gamma}\mathbf{X}_i + \epsilon_i \tag{8.2}$$

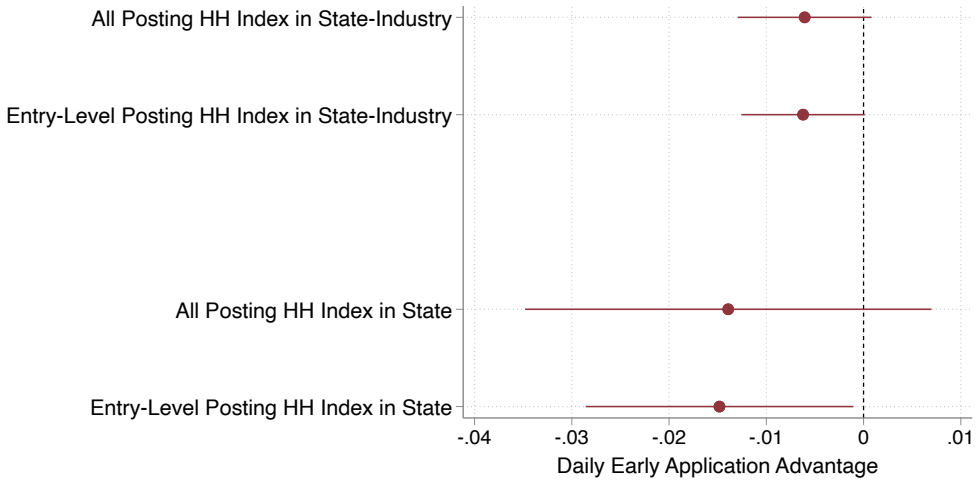
8.1 Market structure, competition and hiring pressure

Figure 10 relates dEEA to local concentration in posting markets, measured by posting Herfindahl-Hirschman indices (HHI) constructed at local labor market lev-

[39] A larger dEEA does not imply that firms make a mistake; it indicates a steeper speed-quality tradeoff. In the model, a steeper speed premium is consistent with stronger urgency (higher effective κ), a thinner arrival environment (lower effective λ), and time-related evaluation constraints.

els. Across specifications, higher concentration is associated with a smaller EEA, with the strongest and most precisely estimated relationship in the more granular state-industry measures. This pattern is consistent with reduced recruiting pressure in more concentrated markets: when employers face less competition for workers, they can be more patient in the initiation decision, and arriving early confers a smaller advantage.

Figure 10: Daily Early Application Advantage and Market Concentration



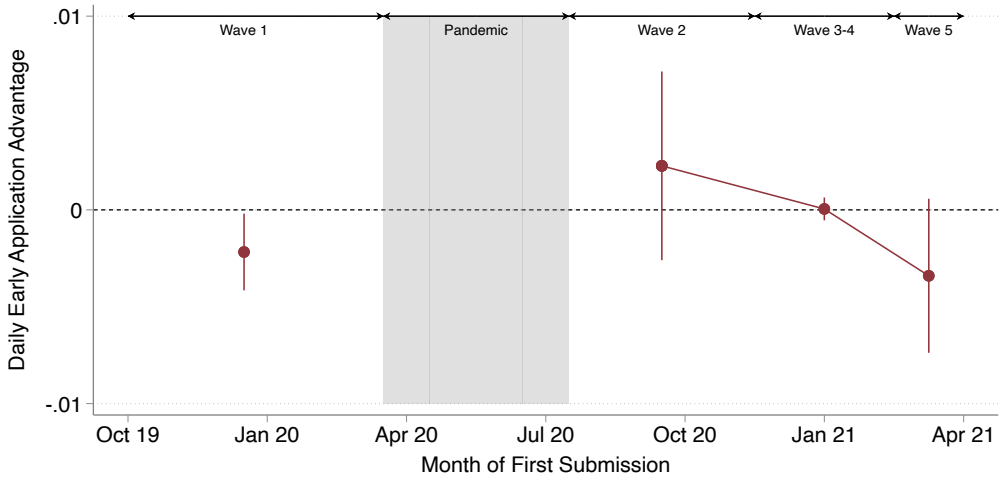
Notes: The figure plots how local labor market concentration, defined as the posting HH index in state or state-by-industry level affect the early application advantage. HH index is calculated using job posting data from Revelio Labs. All regressions include job fixed effects and resume characteristics, with standard errors clustered at the job level. The line depicts the 95% confidence intervals.

The second market force I consider is firm prominence in local recruiting. Using Revelio Labs job posting data, I construct measures of a firm's share of posting activity within a market-month. I aggregate posting records to the firm-state-month and firm-state-industry-month levels and compute the firm's posting share within the relevant market. The posting share measure summarizes the reduced-form relationship between local market structure and the per-vacancy applicant queue faced by the firm: Locally dominant employers attract a disproportionately large number of applications per vacancy, increasing the option value of waiting. This relationship is consistent with directed application behavior (Cai, Gautier and Wolthoff, 2025) and can be rationalized by granular search models (Jarosch, Nimczik and Sorkin, 2024).[40]

[40]Posting share measures are not observed for all vacancies. Shares are undefined when the Rev-

Figure 11 shows estimates of specification (5.1) using posting shares within state-industry markets by experiment waves. The interaction is negative outside the immediate post-pandemic recovery and close to zero during Waves 2 to 4. This pattern reflects the intuition that when firms have higher prominence in the local market, the arrival intensity λ for them is higher, they are rushing less to screen and initiate contacts, leading to a lower early appliation advantage. Appendix Figure A.10 shows that this negative relationship is robust across alternative market definitions (state and state-industry) and when restricting to entry-level postings only.

Figure 11: Market Share and the Daily Early Application Advantage by Waves



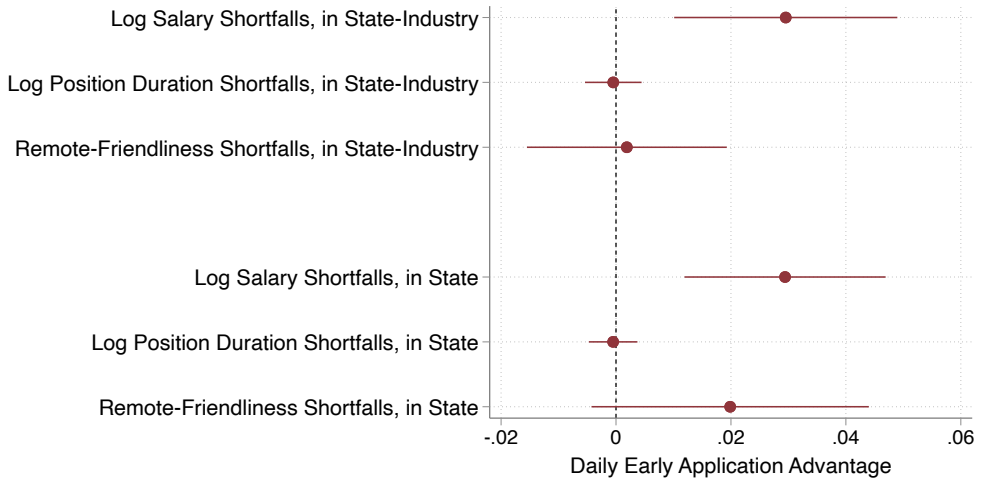
Notes: This figure plots wave-specific estimates of the market-share interaction in Equation (5.1). Each point reports the coefficient on $D2L_i^t \times Share_{m,t}$, where $Share_{m,t}$ is the firm's share of postings in the state-industry market of the vacancy at the month of first submission. Error bars denote 95% confidence intervals based on job-clustered standard errors.

Figure 12 provides a complementary lens based on offer competitiveness. It relates EEA to wage and amenity shortfall ("markdown") proxies constructed from posting data. Larger salary shortfalls are associated with a higher EEA, whereas other shortfall measures are weaker. Given that higher EEA reflects stronger timing frictions, this evidence suggests that less competitive offers are correlated with a steeper speed premium. A natural interpretation is that low offer competitive-

elio Labs universe contains no postings in the relevant market-month, and they are missing when firm-market posting coverage is insufficient. The main specifications therefore use the subsample with non-missing posting shares. Appendix Figure A.8 compares the distribution of submissions to vacancies with non-missing shares to the full sample across states and months of first submission.

ness coincides with thinner or less stable inflow of qualified candidates.

Figure 12: Daily Early Application Advantage and Offer Competitiveness



Notes: The figure plots how firms' offer competitiveness, defined as the log differences between the local market average and the firms' average in posting features, including salary, duration, and remote-friendliness, corresponds to different levels of daily early application advantages. All three measures are calculated with job posting data from Revelio Labs. All regressions include job fixed effects and resume characteristics, with standard errors clustered at the job level. The line depicts the 95% confidence intervals.

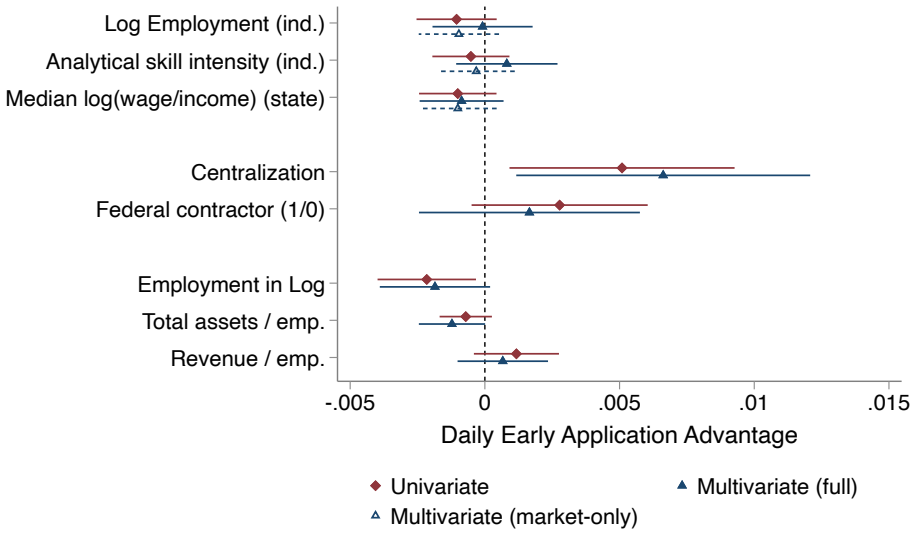
8.2 A compact summary of additional correlates

Figure 13 summarizes eight additional correlates of timing frictions motivated by the model. All continuous covariates are standardized (mean 0, s.d. 1), so coefficients can be interpreted as changes in daily EEA associated with a one-standard-deviation increase in the covariate; the federal-contractor indicator is unstandardized. To clarify the source of multivariate differences, the figure reports (i) univariate regressions, (ii) a "market-only" multivariate specification that jointly includes the three market covariates, and (iii) a full multivariate specification that jointly includes all eight covariates (all estimated on the same sample).

Market environment: thickness, screening complexity, and worker dependence. The first three coefficients capture market conditions that affect how quickly firms can accumulate and process suitable applicants. *Industry employment (log)* is negatively associated with EEA: timing frictions are smaller in thicker industry environments, consistent with a deeper pool attenuating the speed premium. *Industry*

analytical-skill intensity is also negatively associated with EEA: timing frictions are weaker in analytically intensive industries, consistent with more deliberative evaluation reducing the advantage of arriving early. Finally, *median log wage to income ratio at the state level* is negatively associated with EEA: timing frictions are smaller where households rely more on wage income, consistent with weaker speed premia in labor markets with stronger wage attachment. These market coefficients are stable between univariate and the market-only multivariate specification; in the full specification they attenuate and become less precise, consistent with part of cross-market variation operating through the composition and organization of recruiting firms rather than through independent market-level effects.

Figure 13: Daily Early Application Advantage and Model Motivated Factors



Notes: The figure plots how various factors that shift application inflow and firms' shadow cost of delay correspond to different levels of daily early application advantages. All regressions include job fixed effects and resume characteristics, with standard errors clustered at the job level. The line depicts the 95% confidence intervals.

Firm organization and firm characteristics. The remaining five coefficients capture within-market firm heterogeneity that may shape urgency and screening capacity. *Centralization* is strongly positively associated with EEA and remains so in the full specification, consistent with centralized recruiting pipelines creating bottlenecks and attention constraints that steepen the speed premium. The *federal contractor* indicator is also positively associated with EEA, though less pre-

cisely estimated in the full specification; one interpretation is that contractor recruiting processes and compliance environments are associated with greater time pressure or more constrained evaluation pipelines. Among firm scale and productivity proxies, *log firm employment* is negatively associated with EEA, consistent with larger firms having greater recruiting capacity and slack that mitigates timing frictions. *Assets per employee* is also negatively associated with EEA, suggestive of weaker speed premia in more capital-intensive firms. In contrast, *revenue per employee* is positively associated with EEA, consistent with higher vacancy stakes (value per worker) increasing urgency and amplifying the incentive to initiate early. Taken together, the firm covariates point to an important role for organizational constraints and vacancy stakes in shaping timing frictions, complementing the market-structure and offer-competitiveness evidence.

Summary of the model-motivated heterogeneity evidence. Taken together, the heterogeneity results reinforce the paper's central mechanism. Timing frictions are systematically smaller in environments consistent with thicker applicant pools and more deliberative evaluation (larger industry employment and higher analytical-skill intensity), and in markets where recruiting pressure is lower (higher concentration). They are systematically larger where offer competitiveness is weaker (larger salary shortfalls) and where recruiting is more centralized, consistent with time-related evaluation constraints. These patterns complement the structural estimates and counterfactuals: they suggest that the speed premium documented in the experiment reflects economically meaningful timing frictions that vary predictably with market structure, job attractiveness, and firm organization. The next section concludes by discussing the implications of these findings for recruiting design and for technologies that mitigate timing-induced losses.

9 Conclusion

This paper studies a simple but consequential margin in the hiring process: the timing of job applications. Using a large-scale resume correspondence experiment in which applications are randomized within vacancies, I isolate how the probability of employer contact changes as the same job receives applications over time, holding constant applicant characteristics by design. The central fact is an *early ap-*

plication advantage: earlier submissions are more likely to be contacted than equally qualified later submissions. The magnitude is economically meaningful and the estimation is robust.

The mechanism evidence points to employer-side decisions. Recruiters appear to begin evaluation before the full pool arrives, and the contact stage is front-loaded. Once a vacancy has already initiated contact, comparable later applications experience a sharp drop in contact probability. This pattern suggests a sequential pipeline in which early contact acts as an implicit stopping trigger for attention to later arrivals. The result is a timing gradient that is not a feature of applicant quality, but of how hiring unfolds as a dynamic process under time and attention constraints. I then relate the daily early application advantage to broader labor market conditions. One regularity stands out: The early application advantage is larger in tighter labor markets.This relationship varies across waves, indicating that other aggregate conditions also shift hiring incentives. These patterns connect the micro mechanism to the macro environment and motivate an organizing framework in which employers trade off speed against the option value of waiting for additional applicants.

To formalize that tradeoff, I develop a model of sequential screening and endogenous contact initiation. In the model, a vacancy receives a stochastic flow of applicants, screening is costly, and waiting incurs a flow cost of delay. The optimal policy takes the form of an initiation (stopping) rule: the firm initiates contact once the best screened applicant crosses a reservation threshold. Labor market tightness shapes expected inflows and the opportunity cost of delay, while aggregate conditions including hires and separations shift the recruiting environment over time. I estimate the model by simulated method of moments, mapping the model's contact-initiation decision to the first-contact timing observed among experimentally injected tracer applications. The estimated model matches the distribution of first contact timing and the way timing gradients vary with tightness closely.

A key advantage of the structural approach is that it allows a welfare-relevant accounting of what is *at stake* when contact is initiated early. Combining the estimated model with a data-driven screening proxy based on machine-learning predictions trained on first-arriving applications, I quantify firms' loss in matching quality as a 50% probability of missing the best experimental resume at initiation, and 6.55% of the contacted resume's predicted matching quality. Counterfactual

exercises show that environments associated with a steeper early application advantage also tend to exhibit larger implied rushing costs. A simple timing counterfactual provides a money-metric anchor: shifting initiation one day earlier raises implied rushing costs by roughly $1,011 annually per vacancy. Taken together, the estimates emphasize that early contact reflects a quantitatively important speed-quality tradeoff that varies with the state of the labor market.

These findings have important implications for job seekers. When the hiring pipeline is front-loaded, arriving early can materially change an applicant's probability of being considered, even when resume content is held fixed. This creates incentives to reallocate effort toward speed: monitoring postings closely, applying quickly, and compressing preparation time, potentially at the expense of match quality and thoughtful application targeting. Such speed arms race can be a meaningful friction since applicants' investment in speed does not translate into improvement in matching quality. constraints in such arms race also highlight a channel through which disparities can emerge or be amplified. If disadvantaged groups face higher constraints on timing, then even modest systematic differences in application timing can translate into meaningful differences in employer contact. Moreover, evidence that the effect of prior contacts can interact with demographic signals underscores that discrimination may operate dynamically across the recruiting sequence, not only through static differences in evaluation. In the case of gender discrimination, while recent correspondence experiments suggest that recruiters do not discriminate against female applicants at the contact stage (Kline, Rose and Walters, 2022), female applicants may still experience lower contact rates due to late application submissions. This delay could arise from spending too much time overpreparing applications due to underconfidence (Aksoy, Exley and Kessler, 2024; Tekleselassie et al., 2025), marriage (Bandiera, Jalal and Roussille, 2025), or facing tighter time constraints due to domestic responsibilities (Philippe and Skandalis, 2023).

The paper also speaks to recruiters and firms. The model-based quantification indicates that early initiation can carry nontrivial selection costs. This highlights the potential value of recruiting designs and technologies that relax cross-day attention constraints, for example, tools that improve recall and ranking across accumulated applications, or process changes that separate "collect" and "evaluate" stages more explicitly. A the same time, the evidence suggests that much of the

timing gradient is generated by the sequential arrival nature of applications and the organization of evaluation: screening begins early and attention to later applications drops sharply once contact begins. Therefore, the timing effects observed cannot be neutralized by increasing screening efficiency alone.

From a policy perspective, the results suggest that application timing is an understudied margin through which frictions can distort matching. Institutional rules that govern application windows, posting visibility, and the timing of recruiter actions can shape who is considered and when. Policies that increase transparency about recruiting timelines, standardize closing practices, or reduce the scope for de facto early closure may reduce purely timing-driven advantages. In practice, the same policy may have different consequences across labor market states, complicating the implementation. A key lesson is not that early screening is "wrong", but that it embeds a state-dependent tradeoff whose distributional and efficiency consequences depend on market conditions.

This paper opens several directions for future work. It provides one step toward opening the recruiting black box by showing that the timing of evaluation and contact initiation is a central determinant of who is considered, and by mapping that behavior into a tractable stopping-rule framework. A natural next step is to connect firm-side contact initiation more directly to downstream outcomes and to incorporate worker-side strategic timing decisions in general equilibrium. In addition, further work could study how attention and information acquisition evolve within recruiting pipelines, and whether new screening technologies change the nature of timing frictions or simply reallocate them across stages. This paper provides a framework in understanding within-vacancy timing. In the age of AI, application submission and initial screening might be compressed into a nearly instantaneous stage, the same sequential logic can help organize timing patterns in later stages of the hiring process. In that sense, the framework offers a way to study how timing frictions may persist, even as technology changes the stage of the hiring pipeline at which they operate. More broadly, bringing empirically disciplined employer-side dynamics into macro job-search models offers a promising route to understanding how market conditions shape not only the quantity of hiring, but the process by which hires are selected.

References

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey M Wooldridge. 2023. "When should you adjust standard errors for clustering?" *The Quarterly Journal of Economics*, 138(1): 1–35.

Afrouzi, Hassan, Andres Blanco, Andres Drenik, and Erik Hurst. 2026. "A theory of how workers keep up with inflation." *The Quarterly Journal of Economics*, qjag007.

Ahern, Kenneth R, and Amy K Dittmar. 2012. "The changing of the boards: The impact on firm valuation of mandated female board representation." *The Quarterly Journal of Economics*, 127(1): 137–197.

Aksoy, Billur, Christine L Exley, and Judd B Kessler. 2024. "The gender minority gaps in confidence and self-evaluation." National Bureau of Economic Research.

Alesina, Alberto, Michela Carlana, Eliana La Ferrara, and Paolo Pinotti. 2024. "Revealing stereotypes: Evidence from immigrants in schools." *American Economic Review*, 114(7): 1916–1948.

Altonji, Joseph G. 2005. "Employer learning, statistical discrimination and occupational attainment." *American Economic Review*, 95(2): 112–117.

Altonji, Joseph G, and Charles R Pierret. 2001. "Employer learning and statistical discrimination." *The quarterly journal of economics*, 116(1): 313–350.

Arrow, Kenneth J. 1973. "The Theory of Discrimination." In *Discrimination in Labor Markets*. . 1st ed., , ed. Orley Ashenfelter and Albert Rees, 3–33. Princeton, N.J.:Princeton University Press.

Autor, David, Arindrajit Dube, and Annie McGrew. 2023. "The unexpected compression: Competition at work in the low wage labor market." National Bureau of Economic Research.

Bagga, Sadhika, Lukas Friedrich Mann, Ayşegül Şahin, and Giovanni L Violante. 2025. "Job amenity shocks and labor reallocation." National Bureau of Economic Research.

Bandiera, Oriana, Amen Jalal, and Nina Roussille. 2025. "The Illusion of Time: Gender Gaps in Job Search and Employment." National Bureau of Economic Research Working Paper 34051.

Barnichon, Régis, and Andrew Figura. 2011. "What drives matching efficiency? a tale of composition and dispersion." Board of Governors of the Federal Reserve System (US).

Bertrand, Marianne, and Esther Duflo. 2017. "Field experiments on discrimination." *Handbook of economic field experiments*, 1: 309–393.

Bertrand, Marianne, Sandra E Black, Sissel Jensen, and Adriana Lleras-Muney. 2019. "Breaking the glass ceiling? The effect of board quotas on female labour market outcomes in Norway." *The Review of Economic Studies*, 86(1): 191–239.

Beyhaghi, Hedyeh, and Linda Cai. 2024. "Recent developments in pandora's box problem: Variants and applications." *ACM SIGecom Exchanges*, 21(1): 20–34.

Bijkerk, Suzanne H, Silvia Dominguez-Martinez, Jurjen Kamphorst, and Otto H Swank. 2021. "Labor market quotas when promotions are signals." *Journal of labor economics*, 39(2): 437–460.

Birinci, Serdar, Kurt See, and Shu Lin Wee. 2025. "Job applications and labour market flows." *Review of Economic Studies*, 92(3): 1438–1496.

Birinci, Serdar, Yusuf Mercan, and Kurt See. 2025. "Mismatch unemployment during COVID-19 and the post-pandemic labor shortages." *Journal of Economic Dynamics and Control*, 105142.

Bohren, J Aislinn, Alex Imas, and Michael Rosenberg. 2019. "The dynamics of discrimination: Theory and evidence." *American economic review*, 109(10): 3395–3436.

Bohren, J Aislinn, Kareem Haggag, Alex Imas, and Devin G Pope. 2023. "Inaccurate statistical discrimination: An identification problem." *Review of Economics and Statistics*, 1–45.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer. 2016. "Stereotypes." *The Quarterly Journal of Economics*, 131(4): 1753–1794.

Bradley, Jake. 2022. "Worker-firm screening and the business cycle." *IZA Discussion Paper*.

Cai, Xiaoming, Pieter Gautier, and Ronald Wolthoff. 2025. "Search, Screening, and Sorting." *American Economic Journal: Macroeconomics*, 17(3): 205–236.

Campos-Mercade, Pol, and Friederike Mengel. 2024. "Non-Bayesian statistical discrimination." *Management Science*, 70(4): 2549–2567.

Cavounidis, Costas, Kevin Lang, and Russell Weinstein. 2024. "The boss is watching: How monitoring decisions hurt black workers." *The Economic Journal*, 134(658): 485–514.

Cederlöf, Jonas, Peter Fredriksson, Arash Nekoei, and David Seim. 2025. "Mandatory notice of layoff, job search, and efficiency." *The Quarterly Journal of Economics*, 140(1): 585–633.

Cheremukhin, Anton, and Paulina Restrepo-Echavarria. 2022. "The dual Beveridge curve." FRB of Dallas Working Paper.

Coate, Stephen, and Glenn C Loury. 1993. "Will affirmative-action policies eliminate negative stereotypes?" *The American Economic Review*, 1220–1240.

Coffman, Katherine B, Christine L Exley, and Muriel Niederle. 2021. "The role of beliefs in driving gender discrimination." *Management Science*, 67(6): 3551–3569.

Coffman, Katherine, Scott Kostyshak, and Perihan Saygin. 2024. "Choosing and Using Information in Evaluation Decisions." *CESifo Working Paper*.

Cortés, Patricia, Jessica Pan, Laura Pilossoph, Ernesto Reuben, and Basit Zafar. 2023. "Gender differences in job search and the earnings gap: Evidence from the field and lab." *The Quarterly Journal of Economics*, 138(4): 2069–2126.

Davis, Steven J, and Brenda Samaniego de la Parra. 2024. "Application flows." National Bureau of Economic Research.

Davis, Steven J, R Jason Faberman, and John C Haltiwanger. 2012. "Recruiting intensity during and after the Great Recession: National and industry evidence." *American Economic Review*, 102(3): 584–588.

Davis, Steven J, R Jason Faberman, and John C Haltiwanger. 2013. "The Establishment-Level Behavior of Vacancies and Hiring." *The Quarterly Journal of Economics*, 128(2): 581–622.

Davis, Steven J, R Jason Faberman, and John Haltiwanger. 2006. "The flow approach to labor markets: new data sources and micro–macro links." *Journal of Economic perspectives*, 20(3): 3–26.

Dodini, Samuel, and Alexander Willén. 2025. "The Power to Discriminate." Norwegian School of Economics 10.

Doval, Laura. 2018. "Whether or not to open Pandora's box." *Journal of Economic Theory*, 175: 127–158.

Eckbo, B Espen, Knut Nygaard, and Karin S Thorburn. 2019. "Board gender-balancing and firm value." *Unpublished working paper. Tuck School of Business at Dartmouth*.

Egan, Mark, Gregor Matvos, and Amit Seru. 2022. "When Harry fired Sally: The double standard in punishing misconduct." *Journal of Political Economy*, 130(5): 1184–1248.

Engbom, Niklas. 2021. "Contagious unemployment." National Bureau of Economic Research.

Eusepi, Stefano, and Ayşegül Şahin. 2025. "Assessing Maximum Employment: A Flow-Based Approach." National Bureau of Economic Research.

Eyting, Markus. 2022. "Why do we discriminate? the role of motivated reasoning." *SAFE Working Paper*.

Faberman, R Jason. 2020. "Recruiting intensity." *IZA World of Labor*.

Fallick, Bruce, and Charles A Fleischman. 2004. "Employer-to-employer flows in the US labor market: the complete picture of gross worker flows." Board of Governors of the Federal Reserve System (US).

Fang, Hanming, and Andrea Moro. 2011. "Theories of statistical discrimination and affirmative action: A survey." *Handbook of social economics*, 1: 133–200.

Forsythe, Eliza, and Russell Weinstein. 2025. "Recruiting Intensity, Hires and Vacancies: Evidence from Firm-Level Data." *The Economic Journal*, 135(669): 1734–1748.

Gaddis, S. Michael, ed. 2018. *Audit Studies: Behind the Scenes with Theory, Method, and Nuance.* Cham:Springer International Publishing.

Jarosch, Gregor, Jan Sebastian Nimczik, and Isaac Sorkin. 2024. "Granular search, market structure, and wages." *Review of Economic Studies*, 91(6): 3569–3607.

Kessler, Judd B., Corinne Low, and Colin D. Sullivan. 2019. "Incentivized Resume Rating: Eliciting Employer Preferences without Deception." *American Economic Review*, 109(11): 3713–3744.

Kline, Patrick, Evan K Rose, and Christopher R Walters. 2022. "Systemic discrimination among large US employers." *The Quarterly Journal of Economics*, 137(4): 1963–2036.

Kline, Patrick, Evan K Rose, and Christopher R Walters. 2024. "A discrimination report card." *American Economic Review*, 114(8): 2472–2525.

Kondo, Illenin, Kevin Rinz, Brandon Hawkins, Natalie Gubbay, John Voorheis, and Abigail Wozniak. 2023. "Granular Income Inequality and Mobility using IDDA." U.S. Census Bureau Center for Economic Studies Working Paper CES-23-55. Associated dataset: version 1.0.

Lang, Kevin, and Ariella Kahn-Lang Spitzer. 2020. "Race discrimination: An economic perspective." *Journal of Economic Perspectives*, 34(2): 68–89.

Lang, Kevin, and Jee-Yeon K Lehmann. 2012. "Racial discrimination in the labor market: Theory and empirics." *Journal of Economic Literature*, 50(4): 959–1006.

Lundberg, Shelly J. 1991. "The enforcement of equal opportunity laws under imperfect information: affirmative action and alternatives." *The Quarterly Journal of Economics*, 106(1): 309–326.

Mercan, Yusuf, Benjamin Schoefer, and Petr Sedláček. 2024. "A congestion theory of unemployment fluctuations." *American Economic Journal: Macroeconomics*, 16(1): 238–285.

Mortensen, Dale T, and Christopher A Pissarides. 1994. "Job creation and job destruction in the theory of unemployment." *The review of economic studies*, 61(3): 397–415.

Mueller, Andreas I, Damian Osterwalder, Josef Zweimüller, and Andreas Kettemann. 2024. "Vacancy durations and entry wages: Evidence from linked vacancy–employer–employee data." *Review of Economic Studies*, 91(3): 1807–1841.

Neumark, David. 2018. "Experimental Research on Labor Market Discrimination." *Journal of Economic Literature*, 56(3): 799–866.

Olszewski, Wojciech, and Richard Weber. 2015. "A more general Pandora rule?" *Journal of Economic Theory*, 160: 429–437.

Petrongolo, Barbara, and Christopher A Pissarides. 2001. "Looking into the black box: A survey of the matching function." *Journal of Economic literature*, 39(2): 390–431.

Philippe, Arnaud, and Daphné Skandalis. 2023. "Motherhood and the Cost of Job Search." *IZA Discussion Paper*.

Revelio Labs. 2026. "Revelio Labs Workforce Intelligence Data." *Data Dictionary documentation*, Accessed February 24, 2026.

Shimer, Robert. 2005. "The cyclical behavior of equilibrium unemployment and vacancies." *American economic review*, 95(1): 25–49.

Tekleselassie, Tsegay, Marc Witte, Jonas Radbruch, Lukas Hensel, and Ingo E Isphording. 2025. "Feedback, Confidence and Job Search Behavior." IZA Institute of Labor Economics, IZA Discussion Paper.

U.S. Bureau of Labor Statistics. 2025. "Employment, wages, and projected change in employment by typical entry-level education." Employment Projections (EP) program, Table 5.2. Median annual wage (dollars), 2024. Last modified August 28, 2025. Accessed February 15, 2026. https://www.bls.gov/emp/tables/education-summary.htm.

Vuletich, Heidi A, and B Keith Payne. 2019. "Stability and change in implicit bias." *Psychological science*, 30(6): 854–862.

Weitzman, Martin L. 1979. "Optimal Search for the Best Alternative." *Econometrica*, 47(3): 641–654.

Welch, Finis. 1976. "Employment quotas for minorities." *Journal of Political Economy*, 84(4, Part 2): S105–S141.

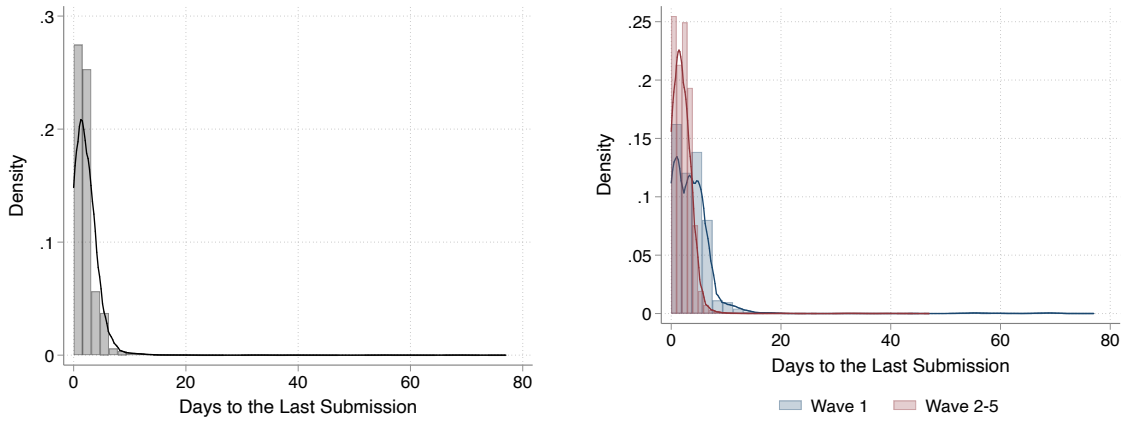
A Figures and Tables

Table A.1: Order of Submissions

Order	(1) Full	(2) Full (< 8)	(3) Wave 1	(4) Wave 2
1	11114	1285	2063	9051
2	11031	1202	2040	8991
3	10682	853	1883	8799
4	10606	777	1879	8727
5	10278	449	1748	8530
6	10204	375	1734	8470
7	9899	70	1604	8295
8	9829	0	1590	8239
N applications	83643	5011	14541	69102
N jobs	11114	1285	2063	9051
N firms	108		85	

Note: This table summarizes the number of applications ordered from 1 to 8. *Full* refers to full sample, *Full* (< 8) refers to the jobs with less than 8 submissions in the full sample, *Wave 1* refers to the subsample of Wave 1, *Wave 2-5* refers to the subsample of Wave 2-5.

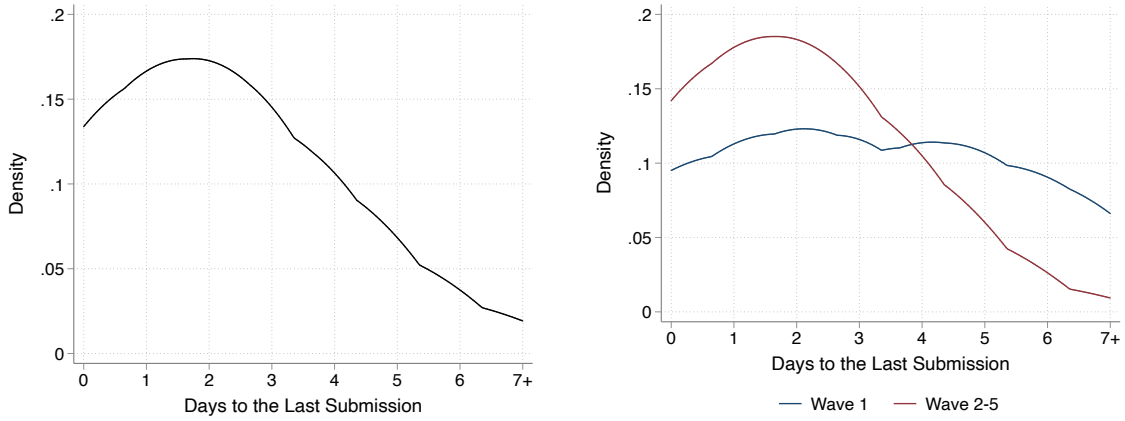
Figure A.1: Distribution of Days to the Last Submission $(D2L)$



(a) Full Sample (b) Wave 1 v.s. Wave 2-5

Notes: The variable $D2L$ is calculated as in Section 3.1. The distribution of all applications is depicted in the left panel; the distributions of applications in Wave 1 and Wave 2-5 are depicted separately in the right panel. In both panels, Epanechnikov kernel is used, the kernel half-width is set as 1.

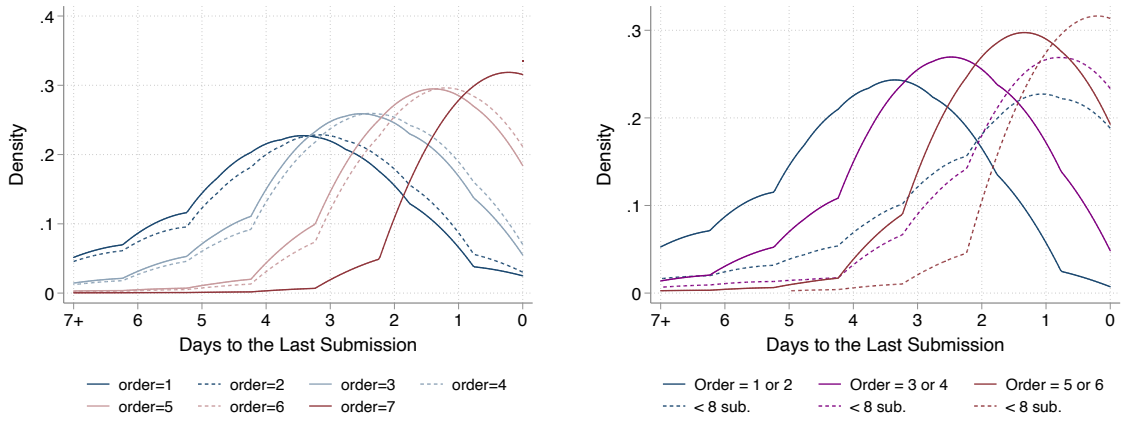
Figure A.2: Distribution of Days to the Last Submission, Top-Coded $(D2L^t)$



(a) Full Sample (b) Wave 1 v.s. Wave 2-5

Notes: The variable $D2L^t$ is calculated as in Section 3.1. The distribution of all applications is depicted in the left panel; the distributions of applications in Wave 1 and Wave 2-5 are depicted separately in the right panel. In both panels, Epanechnikov kernel is used, the kernel half-width is set as 1.5.

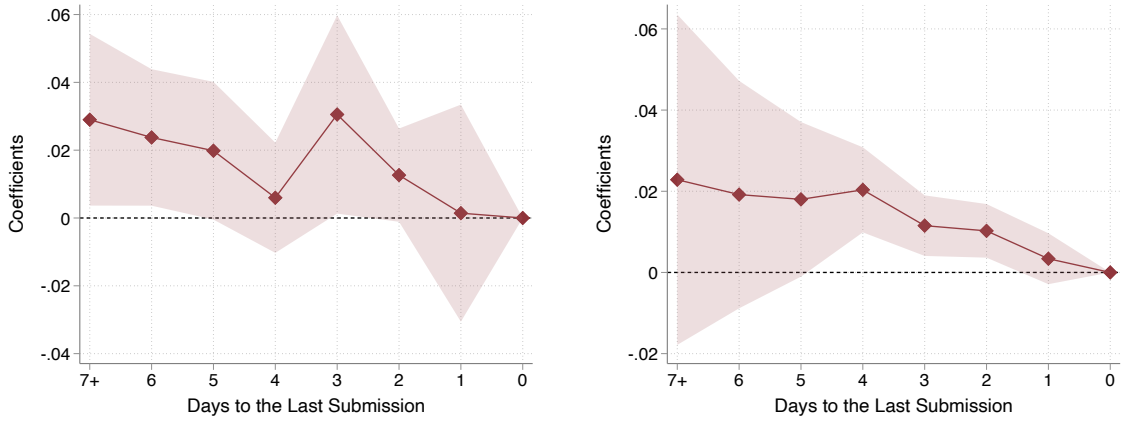
Figure A.3: Distribution of $D2L^t$, by Submission Order



(a) Full Sample (b) Complete v.s. Incomplete

Notes: The variable $D2L^t$ is calculated as in Section 3.1. In the left panel, applications are grouped by the within-job orders of submission. In the right panel, applications of order 1/2, 3/4, 5/6 are grouped together, the solid lines represent jobs with 8 submissions (*Complete*), the dashed line represent jobs with less than 8 submissions due to early closing (*Incomplete*). In both panels, Epanechnikov kernel is used, the kernel half-width is set as 1.

Figure A.4: Contact Gap and $D2L^t$, Wave 1 v.s. Wave 2-5



(a) Wave 1 ($N = 14541$) (b) Wave 2-5 ($N = 69102$)

Notes: The figures plot the estimated coefficients $\hat{\delta}_j, j \in \{1, \cdots, 7\}$ of Model 3.4 with days to the last submission $D2L^t$ on the x-axis, controlling for resume characteristics. The left panel shows Wave 1 results. The right panel, shows Wave 2-5 results. The regression includes job fixed effects, the standard errors are clustered at the job level. The area depicts the 95% confidence intervals.

Table A.2: Randomization Test of Constructed Timing Measures

Dep.Var.	(1) $early^1$	(2) $early^2$	(3) $early^3$	(4) $Order$	(5) $Order^{pair}$	(6) $D2L^t$
Specification A: with job FEs, clustering on job level						
Balanced	1.55	1.37	1.19	0.73	1.02	1.25
	[0.1228]	[0.1938]	[0.2967]	[0.6864]	[0.4212]	[0.2568]
Full	1.07	0.98	0.83	0.63	0.83	0.97
	[0.3831]	[0.4564]	[0.5857]	[0.7692]	[0.5876]	[0.4618]
Full$(=8)$	0.99	0.99	0.82	0.53	1.11	0.92
	[0.4481]	[0.4481]	[0.5988]	[0.8540]	[0.3519]	[0.5040]
Full(<8)	1.10	1.06	0.87	0.93	1.03	0.75
	[0.3566]	[0.3931]	[0.5478]	[0.5007]	[0.4142]	[0.6610]
Specification B: with job FEs, robust						
Balanced	1.24	1.25	1.21	0.81	0.86	0.62
	[0.2647]	[0.2582]	[0.2823]	[0.6096]	[0.5593]	[0.7791]
Full	1.03	1.01	0.87	0.68	0.59	0.61
	[0.4130]	[0.4252]	[0.5497]	[0.7313]	[0.8043]	[0.7879]
Full$(=8)$	0.93	0.93	0.83	0.59	0.56	0.51
	[0.4989]	[0.4989]	[0.5871]	[0.8066]	[0.8282]	[0.8676]
Full(<8)	1.26	1.16	0.91	0.94	1.15	0.95
	[0.2520]	[0.3192]	[0.5191]	[0.4856]	[0.3255]	[0.4792]
Specification C: with wave and region FEs, clustering on job level						
Balanced	1.97**	1.66*	1.16	1.13	2.03**	1.38
	[0.0388]	[0.0938]	[0.3181]	[0.3395]	[0.0319]	[0.1929]
Full	1.33	1.19	0.90	0.81	1.39	1.09
	[0.2129]	[0.2943]	[0.5221]	[0.6086]	[0.1864]	[0.3692]
Full$(=8)$	1.13	1.13	0.94	0.61	1.27	1.05
	[0.3383]	[0.3383]	[0.4911]	[0.7934]	[0.2454]	[0.4004]
Full(<8)	0.55	0.82	0.52	1.61	1.69*	0.32
	[0.8377]	[0.5972]	[0.8627]	[0.1065]	[0.0860]	[0.9691]

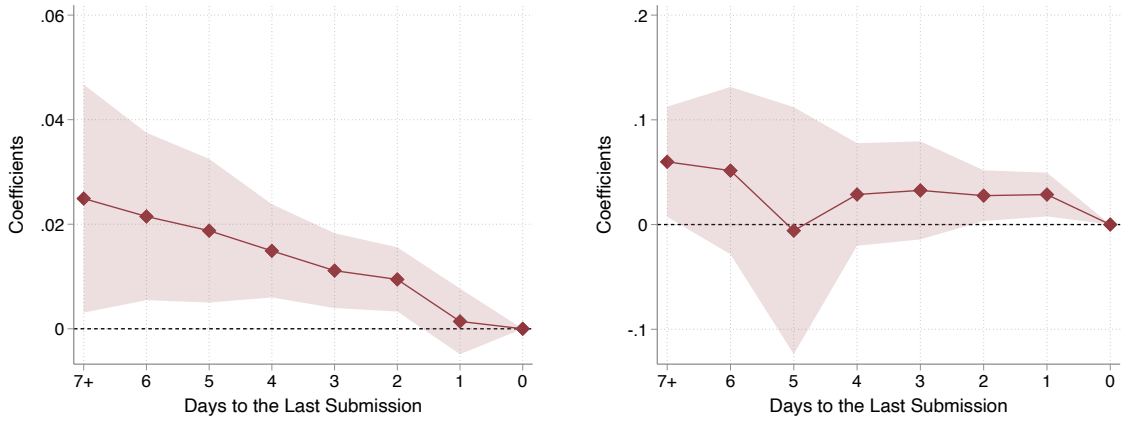
Note: All regressions include the list of covariates in Table 2 of Kline, Rose and Walters (2022). In each unit, joint $F-$ statistic (Degree of Freedom $= 9$) and $p-$values are reported in the square parenthesis. Each column refers to a constructed timing measure indicated by the column title. The four rows in each panel corresponds to balanced sample, full sample, full sample of jobs with 8 submissions, full sample of jobs with less than 8 submissions. Significant level: *** 0.01, ** 0.05, * 0.1.

Table A.3: Estimation of Model (3.3), Wave 1 vs. Wave 2-5

	(1)	(2)	(3)	(4)	(5)	(6)
	All		Wave 1		Wave 2-5	
	Full	< 8	Full	< 8	Full	< 8
$D2L^t$	0.0040	0.0087	0.0036	0.0083	0.0043	0.0093
	(0.0008)	(0.0035)	(0.0013)	(0.0042)	(0.0010)	(0.0064)
Black	-0.0205	-0.0125	-0.0205	-0.0112	-0.0205	-0.0132
	(0.0018)	(0.0066)	(0.0041)	(0.0113)	(0.0020)	(0.0080)
Female	-0.0021	-0.0007	0.0013	-0.0045	-0.0028	0.0014
	(0.0022)	(0.0080)	(0.0051)	(0.0131)	(0.0024)	(0.0102)
Over 40	-0.0043	0.0034	-0.0041	0.0035	-0.0043	0.0035
	(0.0020)	(0.0083)	(0.0048)	(0.0147)	(0.0022)	(0.0100)
Pol. club	0.0060	0.0026			0.0060	0.0026
	(0.0051)	(0.0253)			(0.0051)	(0.0252)
Acad. club	0.0050	0.0136			0.0050	0.0136
	(0.0051)	(0.0174)			(0.0051)	(0.0173)
LGBTQ club	-0.0002	0.0154			-0.0002	0.0153
	(0.0036)	(0.0171)			(0.0036)	(0.0171)
Typical pron.	-0.0078	0.0040			-0.0078	0.0040
	(0.0051)	(0.0202)			(0.0051)	(0.0202)
Neutral pron.	-0.0085	0.0061			-0.0085	0.0059
	(0.0051)	(0.0216)			(0.0050)	(0.0216)
Assoc. degree	0.0021	0.0083	-0.0013	0.0037	0.0028	0.0108
	(0.0020)	(0.0085)	(0.0050)	(0.0150)	(0.0022)	(0.0103)
Constant	0.2453	0.1668	0.1930	0.1744	0.2561	0.1623
	(0.0027)	(0.0093)	(0.0062)	(0.0150)	(0.0030)	(0.0121)
Job FEs	✓	✓	✓	✓	✓	✓
N	83643	5011	14541	1821	69102	3190

Note: This table presents the effects of early submission measured by days before the last submission and other characteristics on the probability of employer contact within 30 days. Columns titled "Full" include all jobs, while columns titled "<8" include jobs with less than 8 submissions. Column (1) and (2) are on full sample, same as Column (3) and (4) of Table 1, Column (3) and (4) are on Wave 1 where the variables "Pol. club", "Acad. club", "LGBTQ club", "Typical pron." and "Neutral pron." were not designed to vary, Column (5) and (6) are on Wave 2-5. All regressions include job fixed effects. Standard errors in parentheses are clustered at the job level.

Figure A.5: Contact Gap and $D2L^t$, *Complete* v.s. *Incomplete* Jobs

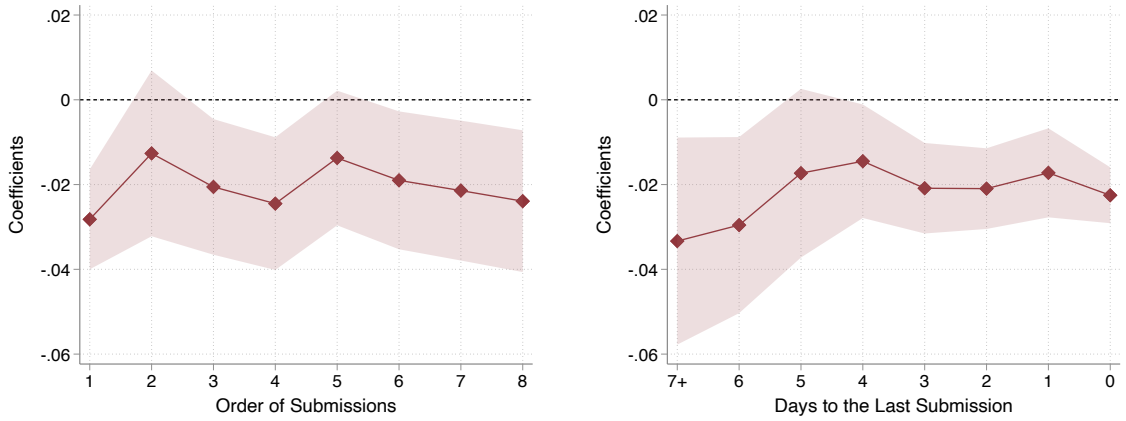


(a) *Complete* Jobs ($N = 78632$)

(b) *Incomplete* Jobs ($N = 5011$)

Notes: The figures plot the estimated coefficients $\hat{\delta}_j, j \in \{1, \cdots, 7\}$ of Model 3.4 with days to the last submission $D2L^t$ on the x-axis, controlling for resume characteristics. The left panel shows full sample results. The right panel, . The regression includes job fixed effects, the standard errors are clustered at the job level. The area depicts the 95% confidence intervals.

Figure A.6: Race Contact Gap Over Time



(a) $\mathbf{1}\left(Order_i = j\right) \times$ Black

(b) $\mathbf{1}\left(D2L_i^t = j\right) \times$ Black

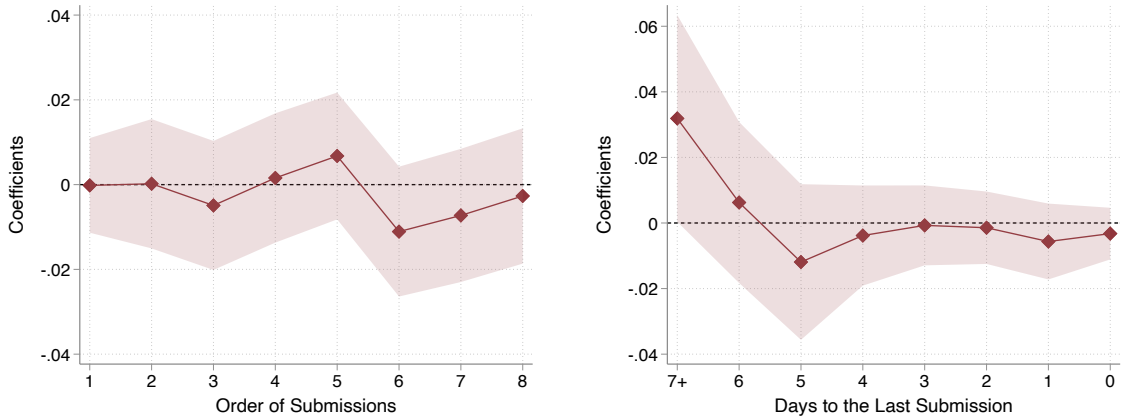
Notes: The figures plot how the race contact gap evolves over time. The left panel plots the estimated coefficients of $(Order_i = j) \times$ Black plus the estimated coefficient of Black for $j = 2, \cdots, 8$ and the coefficient of Black for $j = 1$; The right panel plots the estimated coefficients of $\left(D2L^t = d\right) \times$ Black plus the estimated coefficient of Black for $d = 1, \cdots, 7$ and the coefficient of Black for $d = 0$. The regression includes job fixed effects and all resume characteristics, the standard errors are clustered at the job level. The area depicts the 95% confidence intervals.

Table A.4: Early Application, Race and Gender

	(1)	(2)	(3)	(4)	(5)	(6)
		$early^1$			$D2L^t$	
E.S.I.	0.0129	0.0105	0.0109	0.0041	0.0032	0.0029
	(0.0029)	(0.0031)	(0.0041)	(0.0009)	(0.0010)	(0.0013)
Black	-0.0192	-0.0205	-0.0147	-0.0200	-0.0205	-0.0164
	(0.0024)	(0.0018)	(0.0037)	(0.0027)	(0.0018)	(0.0040)
E.S.I. × Black	-0.0027		-0.0009	-0.0002		0.0006
	(0.0033)		(0.0051)	(0.0010)		(0.0015)
Female	-0.0021	-0.0032	0.0012	-0.0021	-0.0054	-0.0017
	(0.0022)	(0.0029)	(0.0040)	(0.0022)	(0.0032)	(0.0044)
E.S.I. × Female		0.0022	0.0040		0.0017	0.0025
		(0.0039)	(0.0056)		(0.0012)	(0.0017)
Black × Female			-0.0089			-0.0074
			(0.0055)			(0.0059)
E.S.I. × Black × Fem.			-0.0034			-0.0016
			(0.0078)			(0.0023)
Constant	0.2468	0.2480	0.2451	0.2450	0.2469	0.2448
	(0.0026)	(0.0027)	(0.0031)	(0.0028)	(0.0029)	(0.0035)
Controls $\mathbf{X}_i$	✓	✓	✓	✓	✓	✓
Job FEs	✓	✓	✓	✓	✓	✓
N	83643	83643	83643	83643	83643	83643

Note: This table presents the effects of early submission interacted with race and gender on the probability of employer contact within 30 days, controlling for other characteristics. "E.S.I." stands for "Early Submission Indicator", referring to the dummy of first four submissions to each job $early^1$ in Column 1-3 and the days to the last submission $D2L^t$ in Column 4-6. All regressions are on the full sample, including job fixed effects and all resume characteristics. Standard errors in parentheses are clustered at the job level.

Figure A.7: Gender Contact Gap Over Time



(a) $\mathbf{1}\left(Order_i = j\right) \times$ Female (b) $\mathbf{1}\left(D2L_i^t = j\right) \times$ Female

Notes: The figures plot how the gender contact gap evolves over time. The left panel plots the estimated coefficients of $(Order_i = j) \times$ Female plus the estimated coefficient of Female for $j = 2, \cdots, 8$ and the coefficient of Female for $j = 1$; The right panel plots the estimated coefficients of $\left(D2L^t = d\right) \times$ Female plus the estimated coefficient of Female for $d = 1, \cdots, 7$ and the coefficient of Female for $d = 0$. The regression includes job fixed effects and all resume characteristics, the standard errors are clustered at the job level. The area depicts the 95% confidence intervals.

Table A.5: Construction of Contact History, Example

Order	Date: Submission	Date: Contact	Contacted	NoC^{up}	NoC^{low}	NoC^{alt}
1	10/8/19	10/8/19	1	1	0	0
2	10/8/19	10/8/19	1	1	0	1
3	10/12/19		0	3	2	2
4	10/12/19	10/12/19	1	2	2	2
5	10/14/19	10/15/19	1	3	3	3
6	10/14/19	10/15/19	1	3	3	4
7	10/16/19		0	5	5	5
8	10/16/19	10/20/19	1	5	5	5

Note: This table presents an example of the construction of the three contact history measures (job ID: 80). NoC^{up} is calculated including all other submissions contacted no later than the day of the current submission; NoC^{low} is calculated including other submissions contacted at least one day before the current submission; NoC^{alt} is calculated including all submissions ordered before the current submission. Order is constructed with the date and time of actual submissions.

Table A.6: Number of Submissions Contacted before A Submission

Order	N	NoC^{up}	NoC^{low}	NoC^{alt}
1	11114	0.0501	0	0
2	11031	0.0935	0.0154	0.2451
3	10682	0.2322	0.1127	0.4936
4	10606	0.2626	0.1355	0.7362
5	10278	0.4558	0.3070	0.9876
6	10204	0.5025	0.3410	1.2249
7	9899	0.7153	0.5359	1.4757
8	9829	0.7703	0.5913	1.7177

Note: This table summarize the number of submissions already contacted when a submission was made, by the order of that submission. In Column titled NoC^{up}, other submissions contacted no later than the day of the current submission; In Column titled NoC^{low}, only other submissions contacted at least one day before the current submission are included; In Column titled NoC^{alt}, all submissions before the current submission are included. Order is constructed with the date and time of actual submissions.

Table A.7: Number of Contacted Submissions and Timing Measures

	NoC_i^{up}	NoC_i^{low}	$early^1$	$early^2$	$early^3$	$Order$	$D2L^t$
NoC_i^{up}	1						
NoC_i^{low}	0.8843	1					
$early^1$	-0.1899	-0.2048	1				
$early^2$	-0.1859	-0.2006	0.9786	1			
$early^3$	-0.1850	-0.1993	0.9569	0.9710	1		
$Order$	0.2238	0.2385	-0.8164	-0.8076	-0.8006	1	
$D2L^t$	-0.1747	-0.1790	0.7169	0.7195	0.7146	-0.7406	1

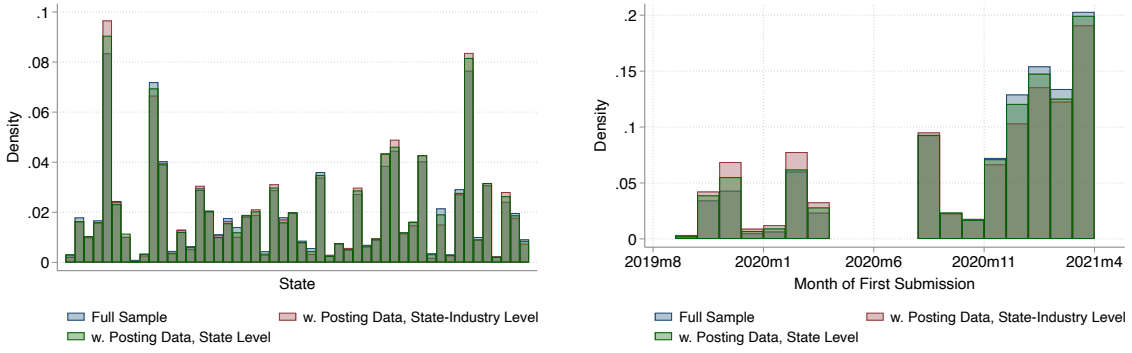
Note: This table summarize the pairwise correlation between the constructed measures. NoC_i^{up} refers to the number of other submissions already contacted no later than the day a submission was made; NoC_i^{low} refers to that while only including other submissions contacted at least one day before. $early^1$, $early^2$, $early^3$ are the 3 constructed early submission dummies; Order is constructed with the date and time of actual submissions; $D2L^t$ is the number of days before the last submission to a job. All coefficients are significant at 1% level.

Table A.8: Prior Contact History and Race/Gender

	(1)	(2) NoC^{up}	(3)	(4)	(5) NoC^{low}	(6)
NoC	-0.0445	-0.0415	-0.0447	-0.0587	-0.0550	-0.0589
	(0.0024)	(0.0024)	(0.0026)	(0.0029)	(0.0030)	(0.0032)
Black	-0.0203	-0.0181	-0.0203	-0.0205	-0.0186	-0.0205
	(0.0018)	(0.0018)	(0.0018)	(0.0018)	(0.0018)	(0.0018)
$NoC \times$ Black		-0.0060			-0.0074	
		(0.0018)			(0.0024)	
Female	-0.0022	-0.0022	-0.0024	-0.0024	-0.0024	-0.0025
	(0.0022)	(0.0022)	(0.0022)	(0.0022)	(0.0022)	(0.0022)
$NoC \times$ Female			0.0004			0.0005
			(0.0023)			(0.0031)
Controls $\mathbf{X}_i$	✓	✓	✓	✓	✓	✓
Job FEs	✓	✓	✓	✓	✓	✓
Order FEs	✓	✓	✓	✓	✓	✓
N	83643	83643	83643	83643	83643	83643

Note: This table presents the effects of prior contact histories, controlling for resume characteristics on the probability of employer contact within 30 days. In Column (1) - (3), NoC refers to NoC_i^{up}, the number of other submissions already contacted no later than the day when a submission was made; In Column (4) - (6), NoC refers to NoC_i^{low}, the number of other submissions already contacted at least one day before a submission was made. All regressions include submission order fixed effects and job fixed effects. Standard errors in parentheses are clustered at the job level.

Figure A.8: Distribution of Submissions



(a) across States (b) by First Submission Month

Notes: This figure compares the distribution of vacancies in the full experimental sample ("Full Sample") to the subsamples for which the Revelio-based posting-share covariates are observed. "w. Posting Data, State Level" restricts to vacancies for which the firm posting share is defined at the state-month level; "w. Posting Data, State-Industry Level" restricts to vacancies for which the posting share is defined at the state-industry-month level. Posting-share measures are missing when the relevant market-month posting universe is empty or when firm-market posting coverage is unavailable. The panels assess whether non-missing posting-share coverage is systematically concentrated in particular states or time periods.

Table A.9: Race-Specific Contact History

	(1)	(2)	(3)	(4)	(5)	(6)
	History : NoC^{up}			History : NoC^{low}		
Black	-0.0183	-0.0184	-0.0182	-0.0192	-0.0184	-0.0186
	(0.0019)	(0.0018)	(0.0018)	(0.0018)	(0.0018)	(0.0018)
History$^{\text{Black}}$	-0.0615		-0.0157	-0.0977		-0.0525
	(0.0041)		(0.0067)	(0.0058)		(0.0097)
History$^{\text{Black}}$ $\times$ Black	-0.0330		-0.0455	-0.0162		-0.0035
	(0.0039)		(0.0095)	(0.0047)		(0.0099)
History$^{\text{White}}$		-0.0827	-0.0697		-0.0974	-0.0574
		(0.0050)	(0.0078)		(0.0056)	(0.0096)
History$^{\text{White}}$ $\times$ Black		0.0093	0.0341		-0.0109	-0.0108
		(0.0037)	(0.0092)		(0.0047)	(0.0098)
Controls $\mathbf{X}_i$	✓	✓	✓	✓	✓	✓
Job FEs	✓	✓	✓	✓	✓	✓
Order FEs	✓	✓	✓	✓	✓	✓
N	83643	83643	83643	83643	83643	83643

Note: This table presents the effects of race-specific prior contact histories, controlling for resume characteristics on the probability of employer contact within 30 days. In Column (1) - (3), History refers to NoC_i^{up}, the number of other submissions already contacted no later than the day when a submission was made, by race (Black/White); In Column (4) - (6), History refers to NoC_i^{low}, the number of other submissions already contacted at least one day before a submission was made, by race (Black/White). All regressions include submission order fixed effects and job fixed effects. Standard errors in parentheses are clustered at the job level.

Table A.10: Gender-Specific Contact History

	(1)	(2)	(3)	(4)	(5)	(6)
	History : NoC^{up}			History : NoC^{low}		
Female	-0.0061	0.0011	-0.0029	-0.0061	0.0008	-0.0032
	(0.0022)	(0.0022)	(0.0022)	(0.0022)	(0.0022)	(0.0022)
History$^{\text{Female}}$	-0.0699		-0.0546	-0.0955		-0.0734
	(0.0046)		(0.0055)	(0.0057)		(0.0065)
History$^{\text{Female}}$ $\times$ Female	0.0001		0.0078	0.0038		0.0087
	(0.0042)		(0.0053)	(0.0054)		(0.0063)
History$^{\text{Male}}$		-0.0626	-0.0347		-0.0803	-0.0447
		(0.0048)	(0.0056)		(0.0057)	(0.0064)
History$^{\text{Male}}$ $\times$ Female		0.0010	-0.0077		-0.0021	-0.0079
		(0.0043)	(0.0054)		(0.0055)	(0.0064)
Controls $\mathbf{X}_i$	✓	✓	✓	✓	✓	✓
Job FEs	✓	✓	✓	✓	✓	✓
Order FEs	✓	✓	✓	✓	✓	✓
N	83643	83643	83643	83643	83643	83643

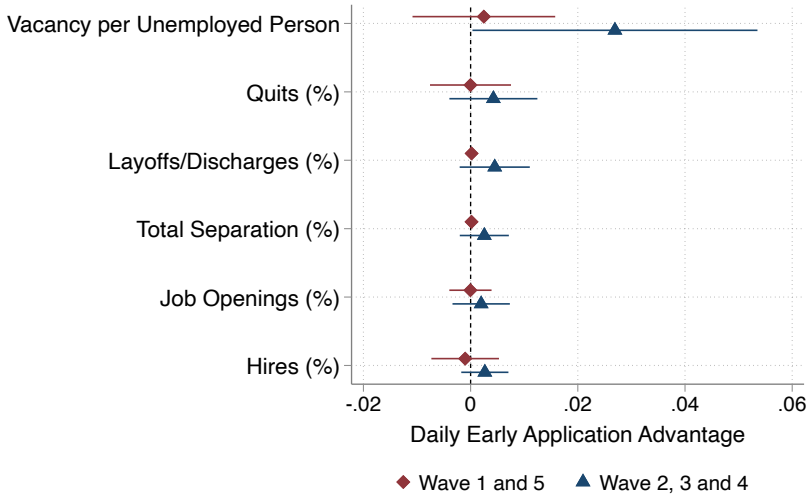
Note: This table presents the effects of gender-specific prior contact histories, controlling for resume characteristics on the probability of employer contact within 30 days. In Column (1) - (3), History refers to NoC_i^{up}, the number of other submissions already contacted no later than the day when a submission was made, by gender (female/male); In Column (4) - (6), History refers to NoC_i^{low}, the number of other submissions already contacted at least one day before a submission was made, by gender (female/male). All regressions include submission order fixed effects and job fixed effects. Standard errors in parentheses are clustered at the job level.

Table A.11: Prior Contact History and Resume Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
		NoC^{up}			NoC^{low}	
	$= 0$	$= 1$	$(2) - (1)$	$= 0$	$= 1$	$(5) - (4)$
Black	-0.0140	-0.0215	-0.0075	-0.0149	-0.0200	-0.0051
	(0.0026)	(0.0095)	(0.0099)	(0.0026)	(0.0111)	(0.0114)
Female	0.0019	0.0088	0.0069	0.0007	0.0179	0.0172
	(0.0029)	(0.0105)	(0.0108)	(0.0030)	(0.0120)	(0.0123)
Black $\times$ Female	-0.0075	-0.0289	-0.0214	-0.0071	-0.0359	-0.0288
	(0.0038)	(0.0148)	(0.0152)	(0.0038)	(0.0171)	(0.0175)
Over 40	-0.0027	-0.0122	-0.0095	-0.0037	-0.0066	-0.0029
	(0.0020)	(0.0079)	(0.0082)	(0.0020)	(0.0090)	(0.0092)
Pol. club	0.0080	-0.0049	-0.0129	0.0071	0.0006	-0.0066
	(0.0051)	(0.0189)	(0.0196)	(0.0051)	(0.0218)	(0.0223)
Acad. club	0.0044	0.0081	0.0037	0.0040	0.0143	0.0103
	(0.0050)	(0.0179)	(0.0186)	(0.0051)	(0.0201)	(0.0207)
LGBTQ club	-0.0045	0.0265	0.0310	-0.0034	0.0299	0.0333
	(0.0036)	(0.0132)	(0.0137)	(0.0036)	(0.0159)	(0.0163)
Typical pron.	-0.0059	-0.0178	-0.0119	-0.0082	0.0010	0.0092
	(0.0052)	(0.0177)	(0.0184)	(0.0051)	(0.0204)	(0.0210)
Neutral pron.	-0.0085	-0.0089	-0.0003	-0.0085	0.0002	0.0087
	(0.0050)	(0.0187)	(0.0193)	(0.0050)	(0.0221)	(0.0226)
Assoc. degree	0.0008	0.0082	0.0073	0.0023	-0.0044	-0.0067
	(0.0020)	(0.0074)	(0.0077)	(0.0020)	(0.0086)	(0.0089)
Constant	0.2638	0.1382	-0.1256	0.2591	0.0818	-0.1773
	(0.0034)	(0.0120)	(0.0127)	(0.0033)	(0.0131)	(0.0133)
Job FEs	✓	✓		✓	✓	
N	71662	11981	83643	75074	8569	83643
$F-$ stats, joint			1.77			1.44
$p-$value			0.0598			0.1563

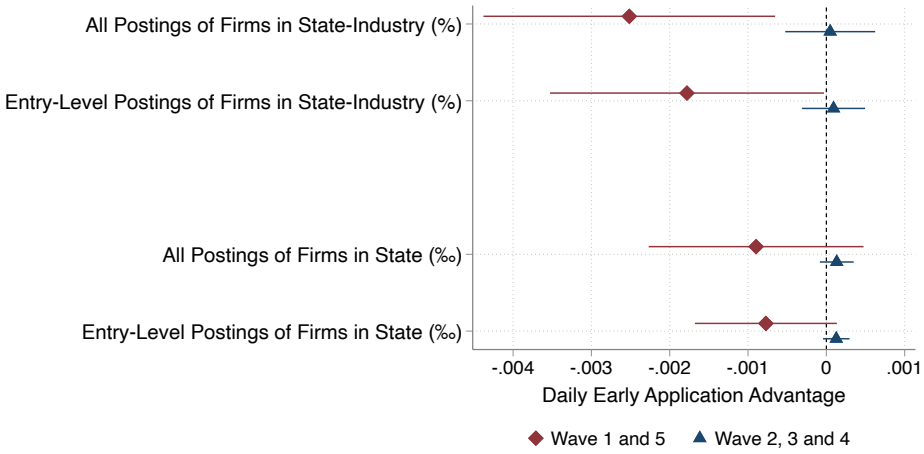
Note: This table presents the effects of the prior contact history interacted with resume characteristics on the probability of employer contact within 30 days. NoC_i^{up} refers to the number of other submissions already contacted no later than the day when a submission was made; NoC_i^{low} refers to the number of other submissions already contacted at least one day before a submission was made. Column (3) displays the differences of between Column (2) and Column (1). Column (6) displays the differences of between Column (5) and Column (4). All regressions include job fixed effects and submission order fixed effects. Standard errors in parentheses are clustered at the job level. χ^2 statistics and joint $p-$values come from tests that all differences in reported coefficients other than the constant term are zero.

Figure A.9: Labor Market Conditions and the Daily Early Application Advantage



Notes: This figure plots estimates of ϕ in Model (5.1) across alternative state-month labor market measures from JOLTS. Each marker reports the coefficient on $D2L_i^t \times Z_{m(i),t(i)}$, where $Z_{m(i),t(i)}$ is the indicated measure in the vacancy's state-month of first submission. Diamonds use the subsample of Waves 1 and 5; triangles use the subsample of Waves 2-4. All specifications include job fixed effects and the full set of randomized resume controls. Error bars denote 95% confidence intervals based on job-clustered standard errors.

Figure A.10: Tightness and the Daily Early Application Advantage by Waves



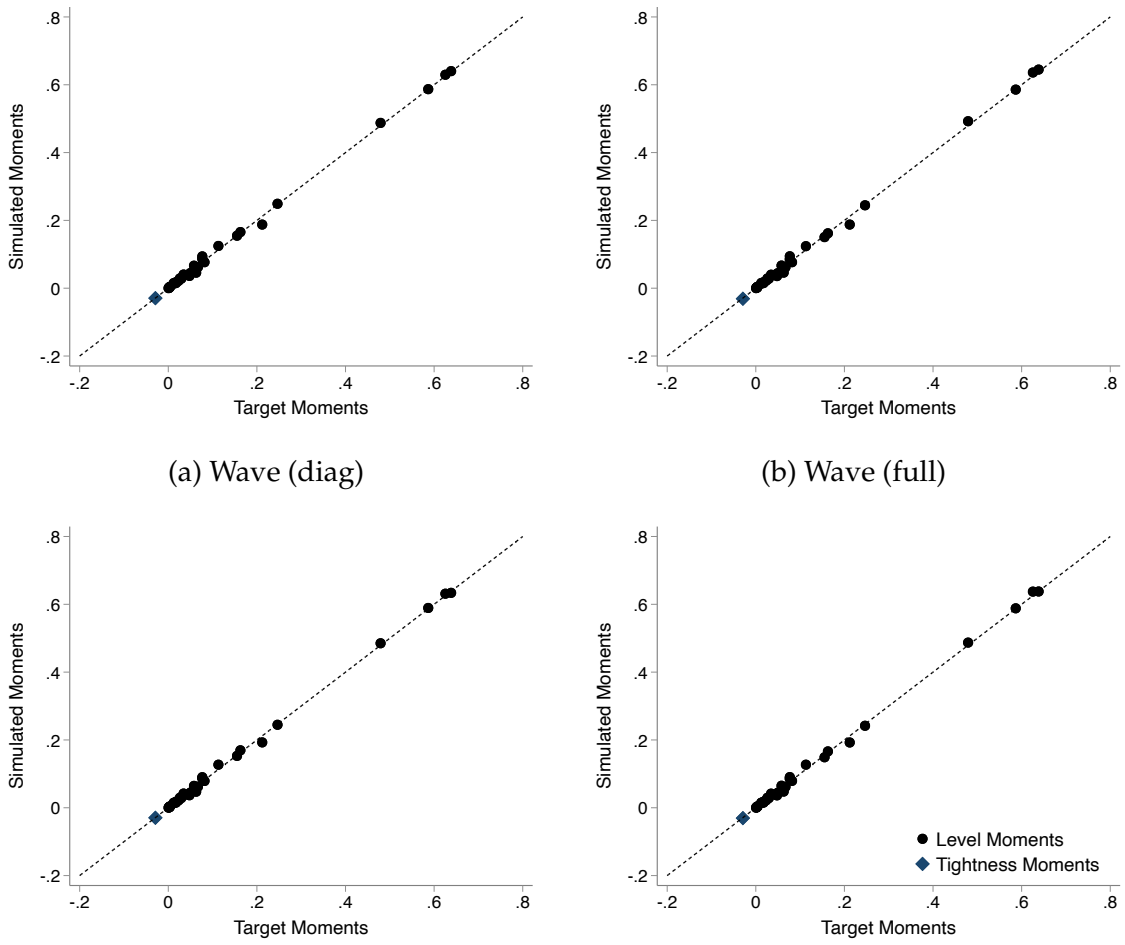
Notes: This figure plots estimates of how the daily early application advantage varies with firm posting shares constructed from Revelio Labs job postings. Each marker reports the coefficient on $D2L_i^t \times Share_{m,t}$ in Model (5.1), where $Share_{m,t}$ is the firm's share of postings in the indicated market (state or state-industry) and month of first submission. "Entry-level" restricts postings to O*NET Job Zones 1-2. Diamonds represent estimates on the subsample of Waves 1 and 5; triangles represent estimates on the subsmaple of Waves 2-4. All specifications include job fixed effects and the full set of randomized resume controls. Error bars denote 95% confidence intervals based on job-clustered standard errors.

Table A.12: Empirical Moments

Level moments: Contact initiation day mass $\mathbb{E}[\omega_{ij} \mid h(j), d_{ij}]$								
	$d_{ij}=0$	$d_{ij}=1$	$d_{ij}=2$	$d_{ij}=3$	$d_{ij}=4$	$d_{ij}=5$	$d_{ij}=6$	None
$h(j)=1$	0.1628	0.0478	0.0345	0.0294	0.0263	0.0179	0.0234	0.6384
	(0.0146)	(0.0078)	(0.0067)	(0.0061)	(0.0057)	(0.0046)	(0.0054)	(0.0201)
$h(j)=2$	0.2120	0.0768	0.0580	0.0446	0.0123	0.0035	0.0026	0.5867
	(0.0140)	(0.0078)	(0.0069)	(0.0063)	(0.0037)	(0.0017)	(0.0016)	(0.0190)
$h(j)=3$	0.1552	0.0766	0.0670	0.0503	0.0174	0.0039	0.0021	0.6255
	(0.0087)	(0.0057)	(0.0053)	(0.0048)	(0.0029)	(0.0014)	(0.0010)	(0.0131)
$h(j)=4$	0.2465	0.1131	0.0818	0.0629	0.0134	0.0018	0.0005	0.4792
	(0.0133)	(0.0099)	(0.0079)	(0.0073)	(0.0035)	(0.0014)	(0.0008)	(0.0198)
Tightness moments								
Main: regression moments						-0.0291		
						(0.0037)		
Tightness tercile based moments, basic								
$\mathbb{E}\left[\omega_{ij}^{\text{none}} \mid \text{high}\right] - \mathbb{E}\left[\omega_{ij}^{\text{none}} \mid \text{low}\right]$						-0.0639		
						(0.0207)		
$\mathbb{E}[\omega_{ij} \mid d_{ij} <= 1, \text{high}] - \mathbb{E}[\omega_{ij} \mid d_{ij} <= 1, \text{low}]$						0.0294		
						(0.0158)		
Tightness tercile based moments, shape								
$\mathbb{E}[\omega_{ij} \cdot d_{ij} \mid \text{high}] - \mathbb{E}[\omega_{ij} \cdot d_{ij} \mid \text{low}]$						-0.0767		
						(0.504)		

Note: This table reports the empirical moments targeted in the MSM procedure. The moments are constructed as described in Section 7.2. Level moments refer to the day-specific contact-initiation mass, $\mathbb{E}[\omega_{ij} \mid h(j), d_{ij}]$, where $h(j)$ denotes the wave group of vacancy j and d_{ij} denotes the day bin. The column labeled "None" reports the probability that contact initiation occurs outside the experiment. The main estimation results are based on the regression-based tightness moments. All estimates are based on the sample of jobs with eight experimental applications. Bootstrapped standard errors are reported in parentheses.

Figure A.11: Model Fitting: Target and Simulated Moments



(a) Wave (diag)

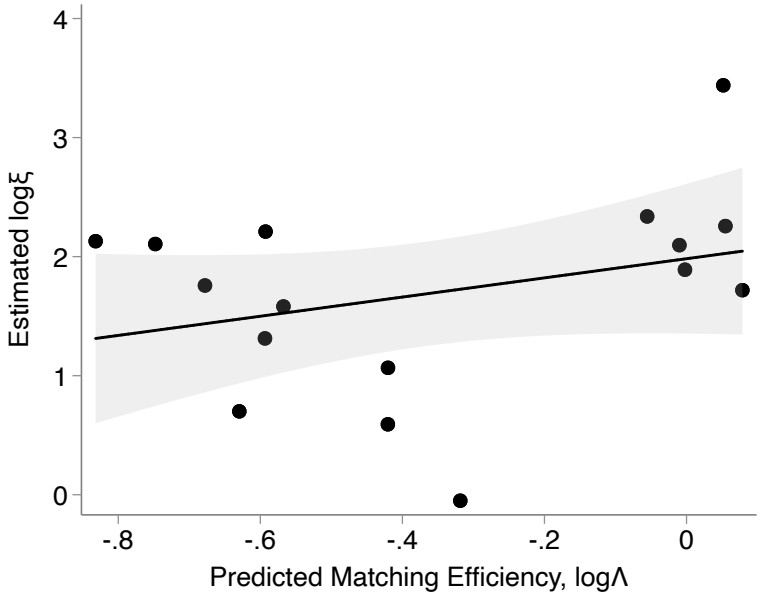
(b) Wave (full)

(c) Macro (diag)

(d) Macro (full)

Notes: The figures plot simulated moments against empirical targets for each specification, with points colored/symbolized by moment block. The 45-degree line indicates perfect fit. Moments are computed using identical definitions in the data and simulation. The three families of moments are as discussed in 7.2.

Figure A.12: External Validity of Estimated Macro State Shifter$\log \xi_t$



Notes: This figure plots the estimated monthly shifter $\log \xi_t$ against a tightness-adjusted matching-efficiency proxy $\widehat{\log \Lambda}_t \equiv \log q_t - \widehat{\alpha}_c \log c_t$. The vacancy yield is $q_t \equiv H_t/V_t$, computed from national monthly JOLTS hires H_t and job openings V_t; congestion is $c_t \equiv U_t/V_t$, where U_t is national monthly unemployment. The elasticity $\widehat{\alpha}_c$ is taken from the baseline macro-disciplined MSM estimate with diagonal weighting (Column (3) of Table 5). The line shows the fitted relationship from a bivariate regression of $\log \xi_t$ on $\widehat{\log \Lambda}_t$, the shaded band reports robust 95% confidence intervals.

Table A.13: Macro State $\log \xi$ SD Decomposition

	Macro (diag)		Macro (full)	
	S.D.	mean	S.D.	mean
Hires	0.3643	-0.1810	0.3997	-0.1987
ΔHIR	0.1914	0.0323	0.1946	0.0329
Layoffs/Discharges	0.6201	0.2429	0.6654	0.2607
ΔLDR	0.6605	-0.0564	0.6370	-0.0544
Quits	0.3626	-0.1551	0.3612	-0.1545
ΔQUR	0.5208	0.0527	0.4757	0.0481

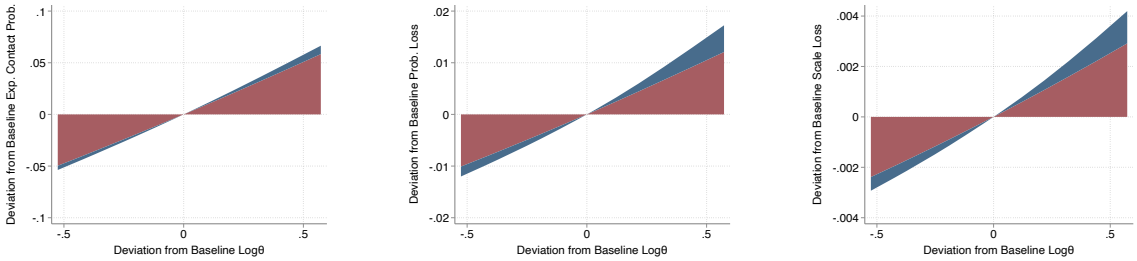
Note: This table reports the mean and cross-job standard deviation of each macro term's contribution to the fitted macro state $\log \xi_j$ under Specification A. Hires, Layoffs/Discharges and Quits denote the national month component of macro series constructed from the JOLTS-based rate series after the transformation used in estimation, and Δ denotes the state-level deviation from the national month component for that same month. Each reported contribution is the product of the estimated slope coefficient and the corresponding standardized macro regressor, evaluated at the job's state and first-submission month. Standard deviations are in $\log \xi$ units and summarize the typical magnitude by which each macro series shifts $\log \xi_j$ in the sample. This table should be interpreted as a descriptive decomposition of the fitted index rather than a variance decomposition. Column (2) uses the regularized inverse bootstrap covariance matrix as the weighting matrix.

Table A.14: Estimated Rushing Loss, Counterfactuals by Wave Groups

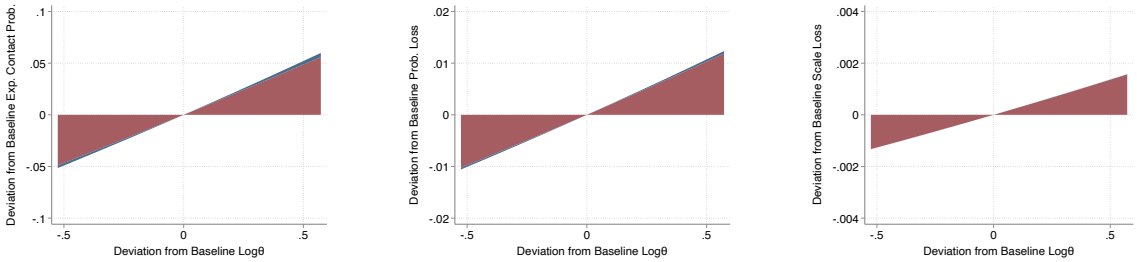
	Prob. Loss				Scale Loss			
Wave Group	1	2	3	4	1	2	3	4
Baseline	0.5391	0.4993	0.5121	0.5150	0.0920	0.0621	0.0654	0.0597
Panel A: Change in market tightness θ_{mt}								
$\theta_{mt} = p10$	0.5271	0.4888	0.5002	0.5018	0.0890	0.0608	0.0634	0.0576
	[-1.202]	[-1.058]	[-1.185]	[-1.319]	[-3.185]	[-2.142]	[-3.104]	[-3.443]
$\theta_{mt} = p50$	0.5400	0.5001	0.5129	0.5159	0.0922	0.0622	0.0655	0.0598
	[0.088]	[0.070]	[0.080]	[0.087]	[0.231]	[0.144]	[0.211]	[0.230]
$\theta_{mt} = p90$	0.5520	0.5087	0.5228	0.5266	0.0951	0.0633	0.0673	0.0615
	[1.292]	[0.938]	[1.070]	[1.161]	[3.389]	[1.922]	[2.858]	[3.078]
Counterfactual B: Uniform change in macro shifter ξ_0								
$\Delta = -2$	0.6112	0.5665	0.5893	0.5985	0.1097	0.0714	0.0792	0.0735
	[7.212]	[6.719]	[7.719]	[8.356]	[19.31]	[14.99]	[21.20]	[23.16]
$\Delta = +2$	0.4925	0.4486	0.4599	0.4519	0.0811	0.0557	0.0567	0.0501
	[-4.660]	[-5.076]	[-5.225]	[-6.310]	[-11.81]	[-10.27]	[-13.22]	[-15.97]
Panel C: Vacancy flow cost κ_j								
$\times 0.5$	0.5153	0.4783	0.4918	0.4925	0.0867	0.0590	0.0620	0.0562
	[-2.374]	[-2.105]	[-2.034]	[-2.245]	[-5.744]	[-5.030]	[-5.165]	[-5.886]
$\times 2$	0.5755	0.5295	0.5432	0.5474	0.1003	0.0666	0.0707	0.0649
	[3.643]	[3.014]	[3.114]	[3.246]	[9.030]	[7.276]	[8.118]	[8.814]

Note: This table reports the two loss measures as in Table 6 under four counterfactual exercises. In each counterfactual, the model-implied initiation weights are recomputed holding fixed observed posting and resume attributes and the realized within-posting arrival times of the eight experimental applications. Panel A shifts market tightness θ_{mt} by shifting log tightness so that its mean matches the $p10$, $p50$, or $p90$ of the baseline distribution while preserving cross-posting heterogeneity. Panel B adds $\Delta \in \{-2, +2\}$ to the macro intercept ξ_0 holding all other parameters fixed. Panel C scales the vacancy flow cost κ_j by a multiplicative factor (0.5 or 2). All reported objects are conditional on an experimental initiator because ML scores are observed only for experimental resumes. The differences of loss in percentage point bewteen each counterfactual and the baseline are reported in brackets: For prob. loss, the numbers are calculated by subtracting the baseline results from the counterfactual results and times 100. For scale loss, the numbers are calculated as the difference between counterfactual and baseline results divided by baseline results, times 100. Wave groups refer to the pooled wave groups, with Wave 3 being Wave 3 and 4 of the experiments.

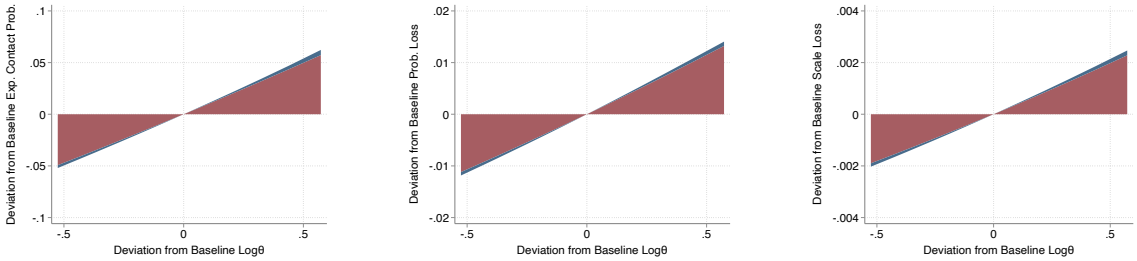
Figure A.13: Tightness Counterfactual Shapley Decomposition, by Wave Groups



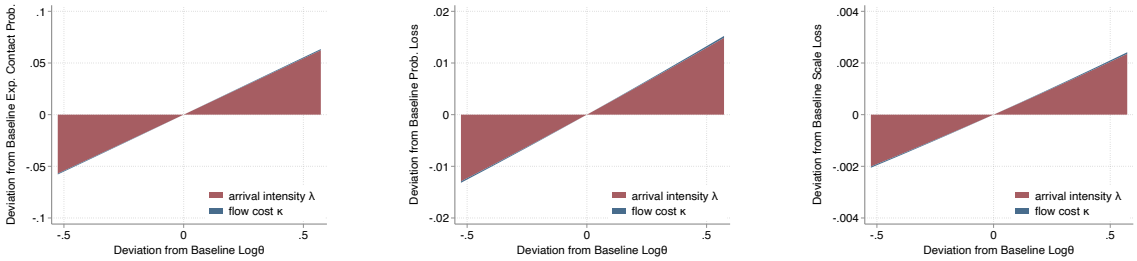
(a) $P_j^{\exp(\hat{\Theta})}$, Wave 1 (b) Prob. Loss, Wave 1 (c) Scale Loss, Wave 1



(d) $P_j^{\exp(\hat{\Theta})}$, Wave 2 (e) Prob. Loss, Wave 2 (f) Scale Loss, Wave 2



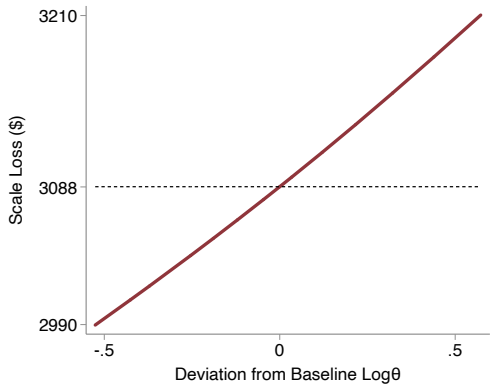
(g) $P_j^{\exp(\hat{\Theta})}$, Wave 3 (h) Prob. Loss, Wave 3 (i) Scale Loss, Wave 3



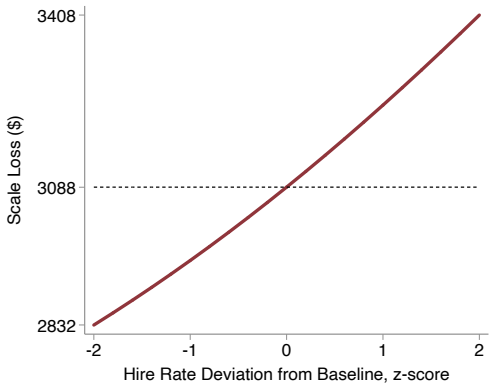
(j) $P_j^{\exp(\hat{\Theta})}$, Wave 4 (k) Prob. Loss, Wave 4 (l) Scale Loss, Wave 4

Notes: The figures plot Shapley decomposition of the counterfactuals into the arrival-intensity and the flow cost channel, as in Section 7.6.3. The horizontal axis reports the counterfactual deviation from the baseline mean log tightness. In Figures of the left column, the vertical axis reports Shapley contributions to the change in the probability that initiating contact is among experiment applications. In Figures of the middle column, it reports Shapley contributions in probabiliy loss defined as in (7.10), while in Figures of the right column, it reports Shapley contributions in scale loss defined as in (7.11), relative to baseline. Waves refer to the pooled wave groups, with Wave 3 being Wave 3 and 4 of the experiments.

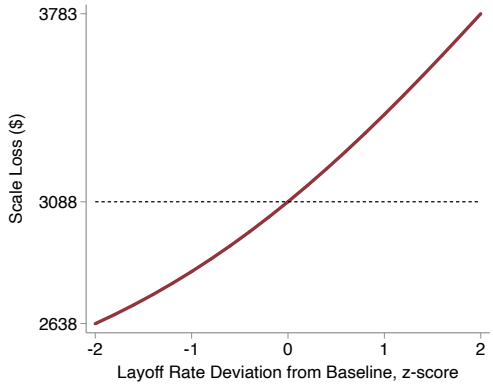
Figure A.14: Wage Equivalent Loss, by Market Condition Series



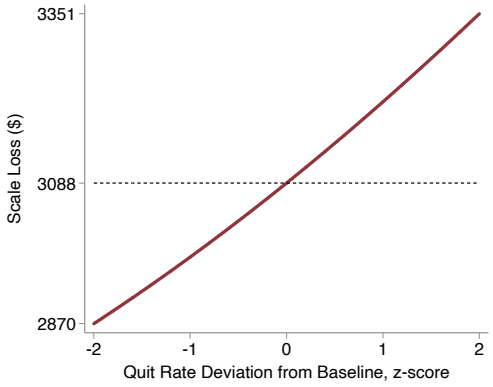
(a) Tightness



(b) Hire Rates



(c) Layoff/Discharge Rates



(d) Quit Rates

Notes: The figures plot the wage equivalent loss of the counterfactuals as discussed in Section 7.6.4. The horizontal axis reports the counterfactual deviation from the baseline environment. Tightness uses shifts in mean log value. Hire rates, layoff/discharge rates and quit rates use standardized z-score shifts The vertical axis reports the dollar amount of wage equivalent loss in match values. The dashed horizontal lines represent the baseline wage equivalent loss. Sample: all jobs with 8 experiment applications.

Table A.15: Estimated Rushing Loss, One-Day Shift Counterfactual

	Prob. Loss				Scale Loss			
Wave Group	1	2	3	4	1	2	3	4
Baseline	0.5391	0.4993	0.5121	0.5150	0.0920	0.0621	0.0654	0.0597
Panel A: Change in levels								
1 day early	0.5764	0.6196	0.6507	0.6412	0.1018	0.0828	0.0903	0.0818
	[3.729]	[12.07]	[13.86]	[12.62]	[10.66]	[33.32]	[38.18]	[37.13]
1 day late	0.4853	0.3189	0.3269	0.3112	0.0772	0.0331	0.0350	0.0275
	[-5.376]	[-18.04]	[-18.52]	[-20.38]	[-16.05]	[-46.80]	[-46.46]	[-53.94]
Panel B: Wage equivalent loss in dollar value								
Baseline					4336	2929	3083	2813
1 day early					4798	3905	4260	3858
Changes from baseline					462	976	1177	1044
1 day late					3640	1558	1651	1296
Changes from baseline					-696	-1371	-1432	-1518

Note: This table reports the two loss measures as in Table 6 under the one-day shifting counterfactual exercises described in Section 7.6.4. "1 day early" refers to the counterfactual where the initiation-day mass is shifted one day earlier, "1 day later" the opposite. In each counterfactual, the model-implied initiation weights are recomputed holding fixed observed posting and resume attributes and the realized within-posting arrival times of the eight experimental applications. Panel A reports the level results as in Table 7, the changes from baseline in percentage point are reported in brackets as in Table A.14 . Panel B reports the wage equivalent loss in match values as discussed in Section 7.6.4. All reported objects are conditional on an experimental initiator because ML scores are observed only for experimental resumes.

Table A.16: Timing Loss vs. Screening Loss, Screening Technology Counterfactual

	Timing Loss (Scale Loss)				Screening Loss			
Wave Group	1	2	3	4	1	2	3	4
Baseline	0.0920	0.0621	0.0654	0.0597	0.1126	0.1026	0.0962	0.0856
Technology	0.0861	0.0585	0.0619	0.0566	0.0421	0.0389	0.0370	0.0304
	[-6.340]	[-5.772]	[-5.403]	[-5.081]	[-62.64]	[-62.04]	[-61.57]	[-64.47]

Note: This table reports the two loss measures as in Table 6 under the within-day screening technology counterfactual exercise. The model-implied initiation weights are recomputed holding fixed observed posting and resume attributes and the realized within-posting arrival times of the eight experimental applications. The changes from baseline in percentage point are presented in brackets as in Table A.14. Timing loss (scale loss) and screening loss are constructed as discussed in Section 7.6.3. All reported objects are conditional on an experimental initiator because ML scores are observed only for experimental resumes.

B Theoretical Appendix

B.1 Proofs

This subsection provides proofs of the main theoretical proposition and discuss why post-contact pipeline does not change the main theoretical structure.

B.1.1 Auxiliary Results

Lemma B.1 (Finiteness of the value function). *Assume $\delta > 0$, $\lambda > 0$, $\kappa \geq 0$, and $\mathbb{E}[v^+] < \infty$ for $v \sim F_v$. Then the dynamic program (6.8) admits a well-defined value function $W : \mathbb{R} \to \mathbb{R}$ that is finite for all v. Moreover, for all $v \in \mathbb{R}$,*

$$v \leq W(v) \leq v^+ + B, \qquad\qquad B \equiv \frac{\kappa + \lambda\left(\mathbb{E}[v^+]\right)}{\delta}.$$

In particular, $W(v_0)$ is finite for any fixed v_0 (e.g. $v_0 = 0$), so $C(v_0)$ and $H(v_0)$ are finite as well.

Proof. Let $\beta \equiv \lambda/(\lambda + \delta) \in (0, 1)$ and define the Bellman operator $\mathcal{T}$ acting on measurable functions $f : \mathbb{R} \to \mathbb{R}$ by

$$(\mathcal{T}f)(v) \equiv \max\left\{v, -\frac{\kappa}{\lambda+\delta} + \beta\left(\mathbb{E}\left[f(\max\{v, v'\})\right]\right)\right\}, \qquad v' \sim F_v.$$

Note that $\mathcal{T}$ is monotone: if $f \leq g$ pointwise, then $\mathcal{T}f \leq \mathcal{T}g$.

Define the candidate upper bound $\bar{W}(v) \equiv v^+ + B$. Since $\max\{v, v'\}^+ \leq v^+ + (v')^+$, we have

$$\mathbb{E}[\bar{W}(\max\{v, v'\})] = \mathbb{E}\left[\max\{v, v'\}^+\right] + B \leq v^+ + \mathbb{E}[(v')^+] + B.$$

Therefore the continuation part of $(\mathcal{T}\bar{W})(v)$ satisfies

$$\begin{aligned}
-\frac{\kappa}{\lambda+\delta} + \beta\left(\mathbb{E}[\bar{W}(\max\{v, v'\})]\right) \leq & -\frac{\kappa}{\lambda+\delta} + \beta\left(v^+ + \mathbb{E}[v^+] + B\right) \\
= & v^+ + B \\
& + \frac{-\kappa + \lambda(\mathbb{E}[v^+]) - \delta B}{\lambda+\delta} - \frac{\delta}{\lambda+\delta} v^+.
\end{aligned}$$

By the definition of B,

$$-\kappa + \lambda(\mathbb{E}[v^+]) - \delta B \leq -\kappa + \lambda(\mathbb{E}[v^+]) - \big(\kappa + \lambda(\mathbb{E}[v^+])\big) = -2\kappa \leq 0,$$

and $-\frac{\delta}{\lambda+\delta} v^+ \leq 0$, so the continuation value is bounded above by $v^+ + B = \bar{W}(v)$. Also, the stopping value satisfies $v \leq v^+ + B = \bar{W}(v)$. Hence $(\mathcal{T}\bar{W})(v) \leq \bar{W}(v)$ for all v.

Now define a sequence $W_0(v) \equiv v$ and $W_{n+1} \equiv \mathcal{T}W_n$. By monotonicity of $\mathcal{T}$ and the fact that $\mathcal{T}\bar{W} \leq \bar{W}$, we have

$$W_0 \leq W_1 \leq \cdots \leq \bar{W} \quad \text{pointwise.}$$

Thus $\{W_n(v)\}_{n\geq 0}$ is increasing and bounded above for each fixed v, so $W(v) \equiv \lim_{n\to\infty} W_n(v)$ exists and is finite. Standard dynamic programming arguments imply $W = \mathcal{T}W$, and the bound $W(v) \leq \bar{W}(v)$ follows from $W_n \leq \bar{W}$ for all n. ■

B.1.2 Proof of Proposition 6.1

Proof. By Lemma 6.1, there is no loss from restricting attention to stopping at arrival times. Let v denote the current best screened value at a decision time, i.e. $v = v^{\max}(d)$ in the main text, and let $W(v)$ be the maximal expected discounted payoff from that point onward.

Let $\Delta \sim \text{Exp}(\lambda)$ denote the time to the next arrival and $v' \sim F_v$ the next screened value, independent of Δ. As shown in (6.8), W satisfies

$$W(v) = \max\{v, C(v)\}, \tag{B.1}$$

where the continuation value is

$$C(v) \equiv -\frac{\kappa}{\lambda+\delta} + \frac{\lambda}{\lambda+\delta}\left(\mathbb{E}[W(\max\{v, v'\})]\right). \tag{B.2}$$

Define the gain from continuing relative to stopping as $H(v) \equiv C(v) - v$. At state v, continuing is (weakly) optimal if and only if $H(v) \geq 0$.

Step 1: $H(\cdot)$ is strictly decreasing. Let $\beta \equiv \lambda/(\lambda+\delta) \in (0,1)$ and fix $\varepsilon > 0$. From (B.2),

$$C(v+\varepsilon) - C(v) = \beta\left(\mathbb{E}[W(\max\{v+\varepsilon, v'\})] - \mathbb{E}[W(\max\{v, v'\})]\right).$$

For every realization of v', we have

$$\max\{v+\varepsilon, v'\} \leq \max\{v, v'\} + \varepsilon.$$

By the 1-Lipschitz property of $W(\cdot)$ (Lemma 6.2),

$$W(\max\{v+\varepsilon, v'\}) - W(\max\{v, v'\}) \leq \varepsilon, \qquad \text{for all } v'.$$

Taking expectations yields $C(v+\varepsilon) - C(v) \leq \beta\varepsilon$. Therefore,

$$H(v+\varepsilon) - H(v) = (C(v+\varepsilon) - C(v)) - \varepsilon \leq (\beta - 1)\varepsilon = -\frac{\delta}{\lambda+\delta}\varepsilon < 0.$$

Hence $H(\cdot)$ is strictly decreasing.

Step 2: $H(v) \to +\infty$ as $v \to -\infty$. Consider the feasible continuation strategy that (i) continues until the next arrival and screens it, then (ii) stops immediately. Starting from state v, this strategy yields the continuation payoff

$$J(v) \equiv -\frac{\kappa}{\lambda+\delta} + \beta\left(\mathbb{E}[\max\{v, v'\}]\right).$$

Because $C(v)$ is the optimal continuation value, $C(v) \geq J(v)$ for all v. Moreover, $\mathbb{E}[\max\{v, v'\}] \geq \mathbb{E}[v']$ for all v, and $\mathbb{E}[v']$ is finite by assumption, so $J(v)$ is bounded below by a finite constant. Thus

$$H(v) = C(v) - v \geq J(v) - v \to +\infty, \quad \text{as } v \to -\infty$$

since $-v \to +\infty$ and $J(v)$ is bounded below.

Step 3: $H(v) \to -\infty$ as $v \to +\infty$. By Lemma B.1, $H(v_0)$ is finite for $v_0 = 0$. Since $H(\cdot)$ is strictly decreasing, for any $v \geq v_0$, apply Step 1 repeatedly (or directly with $\varepsilon = v - v_0$) to obtain

$$H(v) \leq H(v_0) - \frac{\delta}{\lambda+\delta}(v - v_0).$$

Letting $v \to +\infty$ implies $H(v) \to -\infty$.

Step 4: existence and uniqueness of the reservation value, and optimality of the cutoff rule. By Lemma 6.2, W is 1-Lipschitz and therefore continuous. Since $C(v)$ is an affine transformation of $\mathbb{E}\left[W\left(\max\left\{v, v'\right\}\right)\right]$, it is continuous in v, hence $H(v)$ is continuous. By Steps 2–3, $H(v)$ takes both positive and negative values, and by Step 1 it is strictly decreasing. Therefore, there exists a unique $r \in \mathbb{R}$ such that $H(r) = 0$. Since H is strictly decreasing, $H(v) \geq 0$ if and only if $v \leq r$, and $H(v) \leq 0$ if and only if $v \geq r$. Because continuing is optimal when $H(v) \geq 0$ and stopping is optimal when $H(v) \leq 0$, the optimal policy is: continue if $v < r$ and stop if $v \geq r$. Equivalently, the optimal stopping time is the first arrival time at which the best-so-far value crosses r. ■

B.1.3 Initiating contact versus completing hiring

The baseline model treats the initiation of contact as absorbing for the purpose of analyzing early application advantage. This appendix states the lemma that allowing for a generic post-contact pipeline does not change the threshold structure of the contact initiation decision.

Lemma B.2 (Robustness to post-contact processing)**.** *Consider the environment in Section 6.2.1. Suppose that when the firm initiates contact at an arrival, the expected present discounted value of entering the post-contact pipeline is $G(v)$, where v is the current best screened value and $G : \mathbb{R} \to \mathbb{R}$ is weakly increasing. Then the optimal contact initiation policy remains a reservation rule: there exists a cutoff $r \in \mathbb{R}$ such that it is optimal to initiate contact at an arrival if and only if $v \geq r$, equivalently, $G(v) \geq G(r)$.*

Proof. Under immediate screening and stationary Poisson arrivals, the arrival-time Bellman equation with a generic post-contact payoff $G(\cdot)$ is

$$W(v) = \max\left\{G(v),\ -\frac{\kappa}{\lambda+\delta} + \frac{\lambda}{\lambda+\delta}\left(\mathbb{E}\left[W(\max\{v, v'\})\right]\right)\right\}, \tag{B.3}$$

where v is the current best screened value at an arrival and $v' \sim F_v$ is the next screened draw.

Let $y \equiv G(v)$ and $y' \equiv G(v')$, and let F_y denote the induced distribution of y'.

Since G is weakly increasing, for all (v, v') we have

$$\max\{y, y'\} = \max\{G(v), G(v')\} = G(\max\{v, v'\}). \tag{B.4}$$

Define the generalized inverse $G^{-1}(y) \equiv \inf\{v : G(v) \geq y\}$ and the transformed value function $\widetilde{W}(y) \equiv W(G^{-1}(y))$.[41] Using (B.4), equation (B.3) is equivalent to

$$\widetilde{W}(y) = \max\left\{y, -\frac{\kappa}{\lambda+\delta} + \frac{\lambda}{\lambda+\delta}\left(\mathbb{E}\left[\widetilde{W}(\max\{y, y'\})\right]\right)\right\}. \tag{B.5}$$

Equation (B.5) has exactly the same form as the baseline Bellman equation in Equation (6.8). Therefore Proposition 6.1 applies to (B.5): there exists a cutoff $r_G \in \mathbb{R}$ such that it is optimal to stop (initiate contact) if and only if $y \geq r_G$.

Mapping back to the original index, the stopping rule is $G(v) \geq r_G$, which is equivalent to $v \geq r \equiv \inf\{v : G(v) \geq r_G\}$. This proves Lemma B.2. ■

Lemma B.2 shows that allowing for post-contact processing, including contacting multiple candidates after outreach begins, affects the pre-contact problem only through the monotone transformation $G(\cdot)$. It does not change the threshold structure of the contact initiation decision. For example, if the first contacted candidate ultimately converts into a hire with probability $\alpha \in (0, 1]$, and failure yields a normalized continuation value of zero, then $G(v) = \alpha v$ and the initiation policy remains a reservation rule in v with an appropriately rescaled cutoff. Accordingly, in the main text I normalize v to be the firm's expected payoff from entering the contact stage without loss for the timing and comparative statics here.

B.2 Extension

This subsection extends the baseline model in three directions: allowing nonstationary posting-age profile, including adjustment costs in screening technology, and non-immediate/batch review.

B.2.1 Extension A: nonstationary posting-age profile

This appendix studies the case in which the application arrival rate varies with posting age. Fix a vacancy state $(m, t, f, \mathbf{z}_j, \nu_j^\lambda, \nu_j^v)$ and suppress indices. Applica-

[41] If G is strictly increasing then G^{-1} is the usual inverse.

tions arrive according to a non-homogeneous Poisson process with deterministic intensity

$$\lambda(d) = \lambda\rho(d), \qquad d \geq 0,$$

where $\lambda > 0$ is a scale parameter and $\rho(d) \geq 0$ is a common posting-age profile as described in Section 6.1.4. Screened values are i.i.d. across arrivals with CDF F_v and $\mathbb{E}[v^+] < \infty$.[42]

Time-dependent value function and cutoff rule. Let $W(d, v)$ denote the maximal expected discounted payoff at posting age d when the current best screened value equals v, under immediate screening. In the continuation region, $W(d, v)$ satisfies the (time-inhomogeneous) HJB equation

$$0 = -\kappa - \delta W(d, v) + \partial_d W(d, v) + \lambda\rho(d) \left(\mathbb{E}\left[W(d, \max\{v, v'\}) \right] - W(d, v) \right), \tag{B.6}$$

where $v' \sim F_v$ and κ, k, δ are constant within the vacancy spell.

Reservation rule and boundary condition. As in the stationary case, the same Lipschitz argument applies pointwise in d and the gain from continuing is decreasing in v, so the optimal policy is a cutoff rule: there exists a reservation function $r(d)$ such that it is optimal to stop if and only if $v \geq r(d)$. In the stopping region, $W(d, v) = v$. I impose the standard regularity condition used in optimal stopping problems with a smooth payoff:

Assumption B.1 (Smooth pasting)**.** *The reservation function $r(d)$ is differentiable and $W(d, v)$ is continuously differentiable in v at the boundary, with*

$$\partial_v W(d, r(d)^-) = 1.$$

Since $W(d, r(d)) = r(d)$ for all d, differentiating yields

$$\partial_d W(d, r(d)^-) + \partial_v W(d, r(d)^-) r'(d) = r'(d),$$

and Assumption B.1 implies $\partial_d W(d, r(d)^-) = 0$. Evaluating (B.6) at $v = r(d)$ and using $W(d, r(d)) = r(d)$ and $W(d, \max\{r(d), v'\}) = \max\{r(d), v'\}$ gives the bound-

[42] I take $\rho(d)$ as defined on $[0, \bar{T}]$ (the empirical posting horizon) and extend by $\rho(d) = 0$ for $d > \bar{T}$.

ary condition

$$\delta r(d) + \kappa = \lambda \rho(d) \mathbb{E}\left[(v - r(d))^+\right], \qquad v \sim F_v. \tag{B.7}$$

If $v \sim \mathcal{N}(\mu_j, \sigma_v^2)$ and $z(d) \equiv \frac{r(d)-\mu_j}{\sigma_v}$, then (B.7) becomes

$$\delta r(d) + \kappa = \lambda \rho(d) \left[(\mu_j - r(d))\left(1 - \Phi(z(d))\right) + \sigma_v \phi(z(d))\right].$$

Sufficient condition for monotonicity of $r(d)$ Let $p(r) \equiv \Pr(v \geq r) = 1 - F_v(r)$ and $G(r) \equiv \mathbb{E}[(v-r)^+]$. If F_v is continuous, then $G'(r) = -p(r)$.

Proposition B.1 (Monotonicity of the reservation function)**.** *Suppose that: (i) $\rho(\cdot)$ is differentiable and weakly decreasing, i.e. $\rho'(d) \leq 0, \forall d$; (ii) F_v has a continuous density so that $G'(r) = -p(r)$; (iii) Assumption B.1 holds. Then the reservation function satisfies*

$$r'(d) = \frac{\lambda \rho'(d) G(r(d))}{\delta + \lambda \rho(d) p(r(d))} = \frac{\rho'(d)\left(\delta r(d) + \kappa\right)}{\left[\delta + \lambda \rho(d) p(r(d))\right] \rho(d)}. \tag{B.8}$$

In particular, if $\rho'(d) \leq 0$ and $\delta r(d) + \kappa \geq 0$ for all d, then $r'(d) \leq 0$ for all d. If $\rho'(d) < 0$ and $\delta r(d) + \kappa > 0$ on an interval, then $r'(d) < 0$ on that interval.

Proof. Differentiate the boundary condition (B.7) with respect to d:

$$\delta r'(d) = \lambda \rho'(d) G(r(d)) + \lambda \rho(d) G'(r(d)) r'(d).$$

Using $G'(r(d)) = -p(r(d))$ and rearranging yields

$$\left(\delta + \lambda \rho(d) p(r(d))\right) r'(d) = \lambda \rho'(d) G(r(d)),$$

which is the first equality in (B.8). The second equality follows from (B.7), which implies

$$\lambda \rho(d) G(r(d)) = \delta r(d) + \kappa.$$

The sign statements follow because $\delta + \lambda \rho(d) p(r(d)) > 0$. ■

Monotonicity is therefore governed by two forces: First, $\rho'(d) \leq 0$ means inflow thins out with posting age, reducing the option value of waiting so the firm becomes weakly less selective. Second, the sign of $\delta r(d) + \kappa$ determines whether additional arrivals are valuable at the margin. Since $G(r(d)) \geq 0$, $\delta r(d) + \kappa \geq 0$,

new arrivals raise the continuation value on net, thus a decline in inflow makes the firm less selective over time.

B.2.2 Extension B: Selective screening and batch review

In this section, I relax the baseline *immediate screening* assumption by allowing the firm to choose which applications to review. The purpose is not estimation, but to show that the main mechanism, a stopping problem with diminishing returns to waiting, is robust when screening frequency is endogenous.

Setup. Applications arrive according to the Poisson process described in Section 6.1.4, conditional on $(m, t, f, \mathbf{z}_j, \nu_j^\lambda, \nu_j^v)$. Upon arrival, the firm observes costless resume and posting attributes $(\mathbf{x}_{ij}, \mathbf{z}_j, \nu_j^\lambda, \nu_j^v)$ but not the realized screened value. If the firm reviews an application, it pays k_{fmt} and observes a posterior value $v \in \mathbb{R}$, with distribution

$$v \sim F_v(\cdot \mid \mathbf{x}_{ij}, \mathbf{z}_j, \nu_j^\lambda, \nu_j^v)$$

Let $v^{\max}(d)$ denote the best *revealed* value among reviewed applicants by within-vacancy time d, and let $\mathcal{U}(d)$ be the set of arrived but unreviewed applications.

Pandora index and review order. For an unreviewed application with attributes $(\mathbf{x}_{ij}, \mathbf{z}_j, \nu_j^\lambda, \nu_j^v)$, define its *Pandora index* $\zeta(\mathbf{x}_{ij}, \mathbf{z}_j, \nu_j^\lambda, \nu_j^v)$ as the unique solution to

$$\mathbb{E}\left[(v - \zeta)^+ \mid \mathbf{x}_{ij}, \mathbf{z}_j, \nu_j^\lambda, \nu_j^v\right] = 0. \tag{B.9}$$

Equation (B.9) says: if the firm already has an outside option equal to ζ (reservation value), then waiting this application yields zero net expected gain because it only matters when v exceeds ζ.

In a static environment with a fixed set of arrived applications, the classic Pan-

dora's box rule implies:[43]

$$u^\star(d) \in \arg\max_{u \in \mathcal{U}(d)} \zeta_u, \tag{B.10}$$

$$\text{stop if } v^{\max}(d) \geq \max_{u \in \mathcal{U}(d)} \zeta_u, \tag{B.11}$$

where Equation (B.10) governs review choices, and Equation (B.11) governs review stopping. Intuitively, the firm reviews the application with the greatest upside first, and stops reviewing once the best revealed applicant dominates the remaining upside of all unreviewed applications.

Dynamic arrivals and robustness of the baseline mechanism. With ongoing arrivals, endogenous review adds a *screening margin* (which applications to open) on top of the baseline *stopping margin* (when to initiate contact). However, when the arrival environment is stationary with constant λ_j and stable conditional mark distributions, the distribution of future indices ζ is time-invariant. This preserves the same single-crossing intuition as in the baseline.

Selective screening changes *which* applicants are processed before stopping: it prioritizes high-ζ applications and may leave some unreviewed. But it does not overturn the core comparative statics: thicker inflow raises the option value of waiting and increases selectivity, whereas higher delay costs push toward earlier stopping and lower selectivity.

Batch screening and review frequency. Suppose the firm reviews applications only at review times $\{h_n\}_{n\geq 1}$, with review interval $\Delta_n \equiv h_n - h_{n-1}$ (and $h_0 = 0$). During $(h_{n-1}, h_n]$ applications accumulate. At time h_n the firm pays a fixed setup cost $c_0 \geq 0$ to initiate a review session and then reviews all accumulated applications.Let v denote the current best revealed value at the end of the previous review session.

For a given interval length $\Delta > 0$, the number of arrivals is $N_\Delta \sim \text{Poisson}(\lambda\Delta)$. Let M_Δ denote the best screened value in that batch:

$$M_\Delta \equiv \max\{v_1, \ldots, v_{N_\Delta}\},$$

[43]The index characterization follows Weitzman's (1979) Pandora's box problem; see also Olszewski and Weber (2015) and Doval (2018) for extensions and Beyhaghi and Cai (2024) for a recent survey. The result applies under independent values across applications.

with $M_\Delta = -\infty$ if $N_\Delta = 0$. Under i.i.d. marks $v_i \sim F_v$, the CDF of M_Δ is

$$\Pr(M_\Delta \leq x) = \exp\left(-\lambda\Delta\left(1 - F_v(x)\right)\right). \tag{B.12}$$

Equivalently, the expected improvement from waiting Δ when the current best is v is

$$\mathbb{E}\left[(M_\Delta - v)^+\right] = \int_v^\infty \left(1 - \exp\left(-\lambda\Delta(1 - F_v(x))\right)\right) \mathrm{d}x, \tag{B.13}$$

which is increasing and concave in Δ.

A reduced-form dynamic program at review times is:

$$W_\Delta(v) = \max\{v,\ C_\Delta(v)\}, \tag{B.14}$$

with continuation value

$$C_\Delta(v) = -\int_0^\Delta e^{-\delta s}\kappa \mathrm{d}s - e^{-\delta\Delta}c_0 + e^{-\delta\Delta}\mathbb{E}\left[W_\Delta\left(\max\{v, M_\Delta\}\right)\right]. \tag{B.15}$$

This is an optimal adjustment frequency problem: the firm chooses the review interval Δ trading off fixed per-session costs and the concave option value of accumulating information. The fixed setup cost c_0 creates an additional frequency tradeoff. Smaller Δ yields more frequent review but incurs c_0 more often; larger Δ reduces setup costs but increases delay and produces diminishing marginal gains via (B.13). The baseline immediate-screening model is the limiting case $\Delta \to 0$ with $c_0 = 0$.

Discussion. The selective screening extension is intended only as a robustness check for two reasons. First, in the recruiting process, contact initiation is itself the first observable screening stage. Therefore, the contact decision can be interpreted as the outcome of an initial, low-cost review of arriving applications, and later stages of the hiring pipeline are the natural domain for richer selective-screening dynamics. Second, the data are suggestive that screening intensity does not decline mechanically over queue position: in auxiliary exercises, machine learning models trained separately by application order yield similar predictive power of detailed resume features across early and later arrivals. While not conclusive, this evidence supports the immediate screening approximation for the contact stage.

B.3 Normalization of latent applicant quality

This section clarifies why the location and scale of the latent applicant-value distribution are not separately identified from first-contact timing data, and why the estimation normalizes them.

Let v denote the latent value of a screened applicant, and the firm's reservation rule is characterized by

$$\delta r + \kappa = \lambda \mathbb{E}\left[(v - r)^+\right], \tag{B.16}$$

where r is the reservation value, $\delta > 0$ is the discount rate, κ is the flow cost of keeping the vacancy open, λ is the application arrival rate.

Lemma B.3 (Affine non-identification of latent value)**.** *Fix λ and δ, and suppose v has distribution F. For any $a \in \mathbb{R}$ and any $b > 0$, define*

$$v' = a + bv, \qquad r' = a + br, \qquad \kappa' = b\kappa - \delta a,$$

and let F' denote the distribution of v'. Then (r, κ, F) satisfies (B.16) if and only if (r', κ', F') satisfies

$$\delta r' + \kappa' = \lambda \mathbb{E}\big[(v' - r')^+\big]. \tag{B.17}$$

Moreover, $\Pr(v' \geq r') = \Pr(v \geq r)$, so the transformed economy implies the same acceptance probability and the same contact-initiation hazard.

Proof. Using the definitions above,

$$\delta r' + \kappa' = \delta(a + br) + (b\kappa - \delta a) = b(\delta r + \kappa).$$

Also, $(v' - r')^+ = (a + bv - a - br)^+ = b(v - r)^+$, therefore

$$\lambda \mathbb{E}\left[(v' - r')^+\right] = \lambda b \mathbb{E}\left[(v - r)^+\right].$$

Hence (B.17) holds if and only if (B.16) holds.

Finally,

$$\Pr(v' \geq r') = \Pr(a + bv \geq a + br) = \Pr(v \geq r),$$

since $b > 0$. Thus the transformed primitives imply the same acceptance probability and therefore the same stopping/contact hazard. ■

Lemma B.3 shows that first-contact timing behavior is invariant to affine transformations of latent value. As a result, the location and scale of F are not separately identified from the stopping threshold and cost parameters.

Corollary B.1 (Normalized reservation equation)**.** *Suppose $v = \mu + \sigma\tilde{v}, \sigma > 0$, where $\tilde{v}$ has normalized distribution F_0. Define the normalized cutoff*

$$\tilde{r} \equiv \frac{r - \mu}{\sigma}.$$

Then (B.16) *is equivalent to*

$$\delta\tilde{r} + \frac{\delta\mu + \kappa}{\sigma} = \lambda\mathbb{E}\left[(\tilde{v} - \tilde{r})^+\right]. \tag{B.18}$$

Therefore the data identify only the composite normalized objects

$$\tilde{\kappa} \equiv \frac{\delta\mu + \kappa}{\sigma},$$

together with the normalized cutoff $\tilde{r}$, but not μ, σ, and κ separately.

Proof. Write $r = \mu + \sigma\tilde{r}$ and substitute $v = \mu + \sigma\tilde{v}$ into (B.16):

$$\delta(\mu + \sigma\tilde{r}) + \kappa = \lambda\mathbb{E}\left[(\mu + \sigma\tilde{v} - \mu - \sigma\tilde{r})^+\right] = \lambda\sigma\mathbb{E}\left[(\tilde{v} - \tilde{r})^+\right].$$

Dividing both sides by σ yields (B.18). ■

Remarks. In the empirical implementation, the latent quality distribution is normalized to a reference family F_0 (e.g. mean zero and unit variance), and the estimated cost/intercept parameters should be interpreted as the normalized composites in (B.18), rather than as primitive cardinal costs.

The normalization argument above applies to constant location and scale shifts of latent value. By contrast, if applicant quality varies systematically with application timing, e.g.

$$v_d = \mu(d) + \sigma\tilde{v},$$

then no constant affine transformation can absorb $\mu(d)$ into the flow cost term. Such a specification would introduce a distinct mechanism: time-varying applicant composition, which generally renders the stopping problem non-stationary,

with a time-dependent reservation rule. Accordingly, the model in this paper normalizes the latent value distribution but leaves timing-dependent quality as a separate extension for future work.

B.4 Embedding the recruiting subgame into general equilibrium

This appendix outlines how the within-vacancy recruiting problem can be embedded into a standard equilibrium search environment. The purpose is interpretive: it clarifies the equilibrium meaning of the objects treated as exogenous in the main text and shows how the estimated firm-side policy rules would enter a fully closed model.

B.4.1 Environment and timing

Time is indexed by calendar periods $t = 1, 2, \ldots$ (e.g. months), and the economy is partitioned into local labor markets $m \in \mathcal{M}$. At the beginning of each (m, t), aggregate stocks (U_{mt}, V_{mt}) are realized and imply tightness $\theta_{mt} \equiv V_{mt}/U_{mt}$ and congestion $c_{mt} \equiv U_{mt}/V_{mt}$. Within the period, each posted vacancy runs a high-frequency recruiting subgame in continuous within-vacancy time $d \in \mathbb{R}_+$ as in Section 6.2.

A key modeling choice is a time-scale separation: (U_{mt}, V_{mt}) and other aggregates are treated as constant over the initial recruiting window relevant for contact initiation. This approximation is appropriate when the within-vacancy horizon of interest (days) is short relative to the calendar period (month).

B.4.2 Worker-side applications

Let each unemployed worker in market m choose an application intensity $\xi_{mt} \geq 0$ (applications per unit time), with per-period cost $\mathcal{C}(\xi_{mt})$. Workers route applications across vacancies; a minimal *undirected* benchmark assumes each application is sent uniformly at random across active vacancies in (m, t).

Under undirected routing, if the unemployment stock is U_{mt} and the vacancy stock is V_{mt}, then each vacancy receives applications according to a Poisson process with intensity

$$\lambda_{mt} = \xi_{mt} \frac{U_{mt}}{V_{mt}} \;=\; \xi_{mt} c_{mt}. \tag{B.19}$$

This delivers the congestion term used in the main text and gives ξ_t a structural interpretation as equilibrium application intensity.

More generally, *directed* application behavior can be accommodated by allowing routing weights. For example, let each application be sent to firm–market branches (f, m) with probability proportional to an "attractiveness" index A_{fmt}, and then uniformly across that branch's vacancies. This implies a per-vacancy inflow

$$\lambda_j \;=\; \xi_{mt} U_{mt} \cdot \frac{A_{fmt}}{\sum_{f'} A_{f'mt} V_{f'mt}} \cdot \frac{1}{V_{fmt}} \;=\; \xi_{mt} c_{mt} \cdot \underbrace{\frac{A_{fmt}/V_{fmt}}{\sum_{f'} A_{f'mt} V_{f'mt}/V_{mt}}}_{\text{directed-routing wedge}}. \tag{B.20}$$

B.4.3 Firms, vacancy creation, and the recruiting subgame

Firms post vacancies in market m at cost χ_{mt} per vacancy (posting and overhead costs). Let J_{fmt} denote the number of vacancies posted by firm f in (m, t), and $V_{mt} = \sum_f J_{fmt}$.

Given market aggregates and firm characteristics, each vacancy solves the recruiting subgame described in the main text. Let $\mathcal{V}(\mathbf{S}_{mt}, s_{fmt}, z_j, \nu_j)$ denote the ex-ante value of a newly posted vacancy at the beginning of its recruiting spell (before any applications arrive), induced by the optimal reservation policy and the associated distribution of stopping times.

A simple free-entry condition for vacancy creation is

$$\mathbb{E}\left[\mathcal{V}(\mathbf{S}_{mt}, s_{fmt}, z_j, \nu_j)\right] \;\leq\; \chi_{mt}, \qquad \text{with equality when } J_{fmt} > 0, \tag{B.21}$$

where the expectation averages over vacancy heterogeneity within the firm-market cell.

Equation (B.21) makes clear how the estimated recruiting policy rules would enter a GE closure: changes in ξ_{mt} or θ_{mt} shift inflows and urgency, which change $\mathcal{V}(\cdot)$ and thus equilibrium vacancy creation.

B.4.4 From contact initiation to hires

The experiment observes contact initiation rather than completed hires. To close the equilibrium model, one must specify how contact initiation maps into job-filling.

A parsimonious reduced-form closure is to assume that initiating contact generates a hire with probability $h(v^{\max}) \in [0,1]$, where $v^{\max}$ is the best screened value at stopping and $h(\cdot)$ is weakly increasing. Then the per-vacancy job-filling hazard equals the contact-initiation hazard times the conditional hire probability:

$$q_{mt} = \underbrace{\lambda_{mt} \cdot \Pr(v \geq r_{mt})}_{\text{contact-initiation hazard}} \times \underbrace{\mathbb{E}\left[h(v^{\max}) \mid \text{stop}\right]}_{\text{post-contact conversion}}. \tag{B.22}$$

In the knife-edge case $h(\cdot) \equiv 1$, contact initiation coincides with job-filling. Alternatively, one can model an explicit post-contact stage (offers, acceptance, continued screening), but that requires additional outcomes not observed in the experiment.

B.4.5 Market-level flow equations

Let s_u denote the separation rate into unemployment (or more generally, let separations be market-specific). Given job-filling rate q_{mt}, the unemployment and vacancy stocks evolve as

$$\begin{aligned} U_{m,t+1} &= U_{mt} + \text{separations}_{mt} - \text{hires}_{mt}, \\ \text{hires}_{mt} &= q_{mt} V_{mt}, \\ V_{m,t+1} &= V_{mt} + \text{new postings}_{mt} - \text{filled vacancies}_{mt}, \\ \text{filled vacancies}_{mt} &= q_{mt} V_{mt}. \end{aligned} \tag{B.23}$$

In steady state, these conditions pin down (U_m, V_m) jointly with posting and application choices.

B.4.6 Equilibrium definition

An equilibrium in market m and period t consists of: (i) worker application intensity and routing policies (summarized by ξ_{mt} and routing wedges), (ii) vacancy posting decisions $\{J_{fmt}\}_f$ (hence V_{mt}), (iii) firms' within-vacancy recruiting policy functions (reservation values) and induced contact hazards, and (iv) implied stocks (U_{mt}, V_{mt}), such that:

- Given $\mathbf{S}_{mt}$ and routing, each vacancy follows the optimal recruiting policy (main text).
- Given the recruiting subgame value $\mathcal{V}(\cdot)$, vacancy creation satisfies (B.21).

- Given posting and recruiting behavior, application intensity is optimal for workers (first-order condition from maximizing unemployment value net of $\mathcal{C}(\xi_{mt})$).
- Stocks evolve according to (B.23).

B.4.7 Why the present paper estimates the subgame only

The experiment disciplines the firm's within-vacancy timing and selection behavior holding resume content fixed, which is sufficient to identify the recruiting policy rule conditional on $(\theta_{mt}, \xi_t, s_{fmt})$. However, closing equilibrium requires additional objects that are not identified by the experimental data, including: worker application costs/preferences (to pin down ξ_{mt}), vacancy posting costs (to pin down entry and V_{mt}), and post-contact conversion outcomes (to map contact to hires). For this reason, the paper estimates the recruiting subgame and interprets counterfactuals as partial equilibrium changes holding aggregate states fixed.

C Details of Machine Learning Predictions

This section discusses the contact propensity prediction model in detail. The purpose of the machine-learning (ML) component is to construct a resume-job contact propensity score that can be interpreted as an estimate of contact likelihood under a state-0 screening environment. This score is then used to quantify the foregone match quality from contacting early. To preserve a clean interpretation, the ML models exclude variables that directly encode the submission order and other timing-related variation. Instead, I build the model using only posting characteristics and resume attributes as predictors.

C.1 State-0 Environment and Feature Set

The ML models are designed to approximate contact propensity under an early-stage screening environment ("state-0"), before the recruiting process becomes endogenously shaped by prior contacts, time pressure, or other recruiting-state constraints. Specifically, I train the models on the first application submitted within each posting. This design has two advantages: First, it anchors the estimated mapping from resume and job features to contact outcomes to a common early point in the recruiting timeline and avoids implicitly conditioning the model on later-stage dynamics. Second, the model is able to predict how later submissions would be evaluated under the early-stage screening environment within the experiment, thereby creating a reasonable counterfactual to quantify the potential quality loss from rushing to contact. This yields a training and evaluation sample of 11,114 submissions, one per posting.

I use the following feature set as predictors:

- **Posting characteristics**, including state, firm ID, experimental wave, submission month, industry SIC code, job SOC3 category, and a hiring intermediary platform indicator.
- **Resume headline attributes**, including race, gender, age, education, political leaning indicators, and pronoun indicators.
- **Resume work-history details**: number of prior positions listed, and for up to three prior positions: (i) job type (1,512 categories), (ii) job title (23 categories), and (iii) duration in months.

Binary indicators are treated as numerical regressors. For numerical covariates

(including binaries), missing values are imputed using the training-fold median; for binaries this coincides with imputing the modal value in the common case where the median is 0.

Categorical representation Since the job type of resume entries contains 1,512 categories, it can generate a large sparse matrix, which may lower the predictive performance of some algorithms. I therefore consider two encoding approaches to group rare categories and shrink the effective number of categories:

- **Frequency-based grouping.** I cap the number of distinct categories retained for high-cardinality type fields by keeping the K most frequent labels and mapping all other labels to an "Other" bucket. Specifically, for the latest resume entry (the most recent job prior to application submission), I one-hot encode the type variable with a maximum of 300 categories using infrequent-category handling, and the other resume entries are pooled and mapped similarly.
- **Text-based clustering.** I also use an unsupervised text-based approach for encoding. I construct a TF-IDF representation of the raw type labels and cluster the labels into 300 clusters using MiniBatch k-means. Each type label is replaced by its cluster ID, then encoded analogously to the frequency-based approach (the most recent entry separately, all other entries pooled). This provides a data-driven grouping based on lexical similarity rather than purely frequency.

C.2 Candidate Models and Evaluation Metrics

I considered three "workhorse" model classes as initial benchmarks:

- **Elastic-net logistic regression**: A linear model for binary outcomes (logistic regression) with combined L1/L2 regularization. It is fast and stable in sparse feature spaces, but relies on encoded inputs and can miss nonlinearities and complex interactions.
- **ExtraTrees classifier**: An ensemble of heavily randomized decision trees that captures nonlinearities and interactions with minimal feature engineering, but its probability outputs may be less well-calibrated and less interpretable.
- **CatBoost gradient boosting**: A gradient boosting algorithm designed for

tabular problems, especially with many categorical predictors. It uses specialized categorical handling, but is more complex, less transparent, and requires careful validation.

I compare the three models using sparse encoded features to assess overall performance. Since CatBoost uses specialized categorical handling and boosting schemes, it handles high dimensional categorical variables efficiently. I therefore also include a CatBoost model with the native, ungrouped categorical variables as predictors. It is attractive in our setting because it avoids high dimensional one-hot expansions for high cardinality categories, while still allowing flexible nonlinearities and interactions via gradient boosting. Models are evaluated with probability and ranking metrics computed on out-of-fold predictions for first applications.

Out-of-fold procedure A central design choice is to generate **out-of-fold (OOF)** predictions at the posting level. This avoids overstating predictive performance due to posting-level leakage. I implement 5-fold cross-fitting using GroupKFold:

1. Construct folds using only the first submission of each posting, with postings partitioned into 5 folds.
2. For each fold f, train the model on order-1 applications from postings in the training folds.
3. Score applications in held-out postings.
4. Concatenate predictions across folds to calculate model performance statistics.

C.3 Model Selection

I first evaluate which model performs the best in predicting contact outcomes for the first submissions of each posting. Table C.1 summarizes performance on first applications (order $= 1$) using OOF predictions. The CatBoost model with native handling of categorical variables delivers the strongest overall ranking and probability performance among the candidate models. Figure C.1 plots the ROC curve and the precision-recall curve of the four candidate models. It further shows that the native CatBoost model consistently outperforms the other models. I therefore select CatBoost with native handling of categorical variables as the main specification.

Table C.1: Candidate Models' Out-of-Fold Performance on First Submissions

Model	ROC-AUC	Log Loss	Brier	Precision
CatBoost (native categories)	0.758	0.474	0.154	0.525
ExtraTrees (encoded)	0.743	0.491	0.160	0.511
CatBoost (encoded)	0.740	0.484	0.158	0.504
Elastic-net logit (encoded)	0.734	0.491	0.161	0.488

Note: This table compares four predictive models used to construct a contact propensity score. Elastic-net is a penalized logistic regression estimated with an elastic-net penalty (L1/L2), ExtraTrees is an ensemble of extremely randomized decision trees, and CatBoost is a gradient-boosted tree model. The three models that use frequency-clustering encoded categorical variables are noted with "(encoded)". CatBoost (native) uses CatBoost's native handling of categorical predictors. Performance is measured on the state-0 sample consisting of the first application in each posting, and all reported statistics are computed from posting-level out-of-fold (OOF) predictions using 5-fold GroupKFold cross-fitting. The four reported metrics are: ROC-AUC (ranking quality), log loss (probability accuracy), Brier score (mean squared error of predicted probabilities), and **average precision** (area under the precision-recall curve; reported as "Precision" in the table). Lower values are better for log loss and Brier; higher values are better for ROC-AUC and precision.

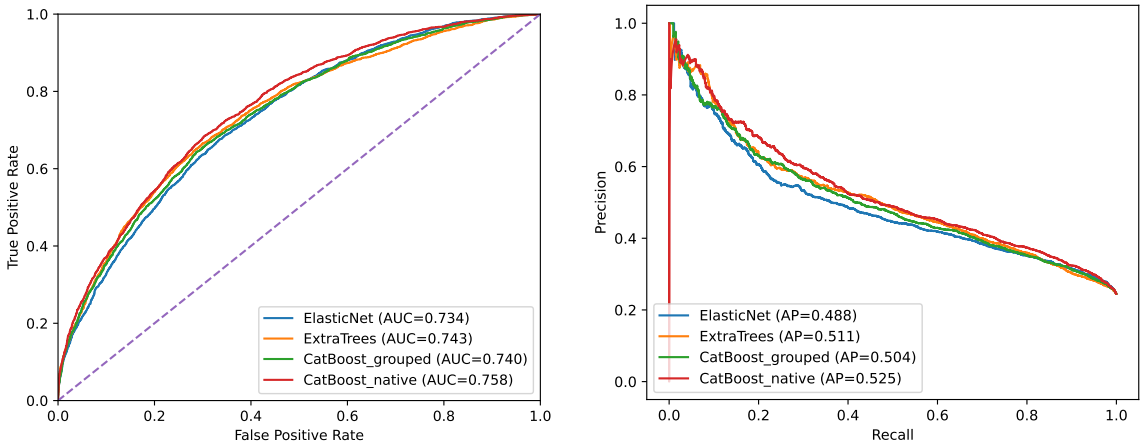
Since the model uses the first submissions of each posting as the training and evaluation sample, one concern is that the predictive power comes mainly from the job-side predictors. To assess whether the candidate models leverage resume information in prediction, I perform qualitative diagnostics of within-posting score dispersion. Specifically, I use the model trained on first submissions to generate out-of-fold predictions for later submissions and plot the distribution of within-posting score ranges and top-1 vs. top-2 gaps. Figure C.2 plots the empirical cumulative distribution function (ECDF) across postings for the two dispersion measures. It shows that the native CatBoost model generates prediction scores that vary meaningfully within postings. This feature of the model significantly benefits downstream comparisons of "best available" applicants within a posting.

C.4 Hyperparameter Tuning

I tune the native CatBoost model using the same posting-level OOF design described in Section C.3. Tuning targets ROC-AUC with log loss as a tie-breaker, favoring better-calibrated probabilities when AUC differences are small.

Due to computational limitations, I perform the hyperparameter tuning via a two-stage grid search design. The initial grid search is over a moderately sized set of parameter values: depth $\in \{4, 6, 8\}$, learning rate $\in \{0.03, 0.06\}$, ℓ_2 regularization $\in \{1, 3, 10\}$, and iteration count $\in \{800, 1200\}$. The best-performing configuration

Figure C.1: Candidate Models' Out-of-Fold Performance on First Submissions



(a) ROC-AUC Curve (b) Precision-Recall Curve

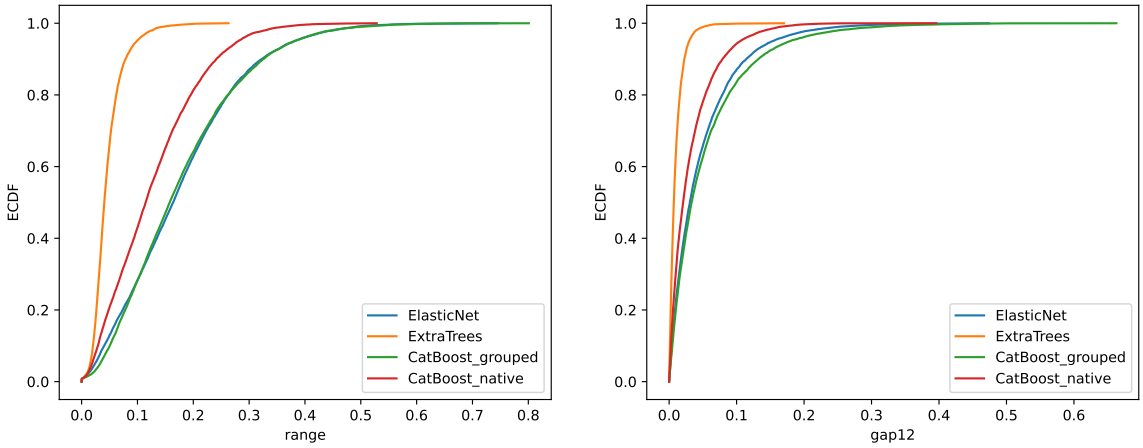
Notes: The figures plot the ROC curve in Panel C.1a and the precision-recall curve in Panel C.1b for the four candidate models as described in Section C.2. The numbers in the legends represent the area under the curve in both panels.

is concentrated in the "shallow + regularized" region, consistent with the high-cardinality categorical structure of the feature set. I then refine around the Stage 1 best-performing parameter region with a second grid that varies learning rate, ℓ_2 regularization, iterations, and minimum leaf size, with depth fixed at 4. The final tuned configuration is depth$=4$, learning rate$=0.04$, ℓ_2 regularization$=20$, iterations$=1200$, and minimum training sample in a leaf (min_data_in_leaf)$=50$.

This tuned configuration improves the OOF performance on first-application contact prediction consistently. The ROC-AUC increases to 0.763 from 0.758, the average precision increases to 0.534 from 0.525, and both log loss and the Brier score decrease moderately. I proceed to use this configuration as the finalized version.

Robustness As robustness checks on the feature representation, I re-estimate CatBoost using the grid-search optimal parameters but replacing the native categorical inputs with (i) the frequency-based grouped encoding and (ii) the TF-IDF clustering-based encoding described in Section C.1. Model comparisons are performed using the same out-of-fold evaluation protocol. Figure C.3 shows that the native CatBoost model outperforms the two encoded models in prediction but does not preserve as much within-posting dispersion. I use the predictions of the

Figure C.2: Candidate Models' within-Posting Dispersion Performance



(a) Predicted Score Range

(b) Top 2 Score Gap

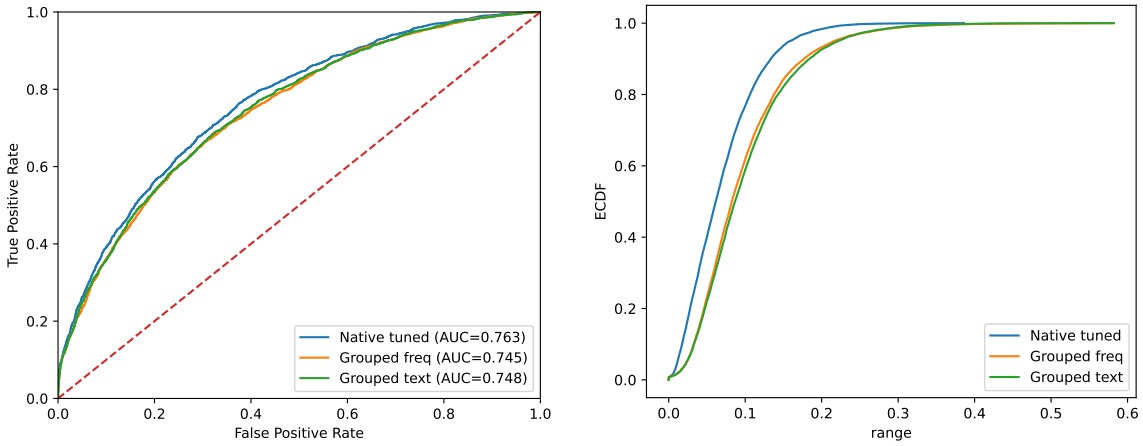
Notes: The figures plot the within-posting score ranges in Panel C.2a and the top-1 vs. top-2 gaps in Panel C.2b for the four candidate models as described in Section C.2. In both panels, the y-axis is the empirical cumulative distribution function (ECDF). At any x-value it shows the fraction of job postings whose within-posting dispersion statistic is less than or equal to x.

two encoded models in robustness checks.

C.5 Model Outputs and Reproducibility

I use the tuned native CatBoost model and the two encoded CatBoost models to predict the state-0 contact probability of later submissions. I use out-of-fold by posting predictions to ensure that each posting's applications (orders 1-8) are scored using a model trained only on other postings. This yields uniformly out-of-sample predictions at the posting level, mitigates posting-level leakage through job characteristics, and provides a cleaner object for inference and for constructing within-posting counterfactual comparisons. Algorithm 1 summarizes the procedure. All reported downstream analyses use the score generated through this algorithm to avoid in-sample artifacts and posting-level leakage.

Figure C.3: CatBoost Out-of-Fold Performance: Native vs. Encoded



(a) ROC-AUC Curve

(b) Predicted Score Range

Notes: Panel C.3a plots the ROC curve (the numbers in the legend report the area under the curve). Panel C.3b plots within-posting score ranges; the y-axis is the empirical cumulative distribution function (ECDF) as in Figure C.2. The blue line represents the tuned native CatBoost model, the orange line represents the tuned encoded CatBoost model with frequency-based grouping, and the green line represents the tuned encoded CatBoost model with text-based grouping as discussed in Section C.1.

Algorithm 1 OOF-by-posting prediction (GroupKFold cross-fitting)

Require: Dataset $\mathcal{D} = \{(X_{ij}, Y_{ij}, j)\}$ with job ID j, order i, number of folds F, outcome $Y_{ij} = \mathbf{1}(\text{contacted})$.

Ensure: Out-of-fold predictions $\widehat{p}_{ij}^{OOF}$ for all applications.

1: Define state-0 sample $\mathcal{D}_0 \leftarrow \{(X_{ij}, Y_{ij}, j) \in \mathcal{D} : i = 1\}$.
2: Partition postings in $\mathcal{D}_0$ into folds $\{\mathcal{J}_1, \ldots, \mathcal{J}_F\}$ using GroupKFold by job ID.
3: Initialize $\widehat{p}^{OOF}$ for all observations in $\mathcal{D}$.
4: **for** $f = 1, \ldots, F$ **do**
5: Train on $\mathcal{T}_f \leftarrow \{(X_{ij}, Y_{ij}, j) \in \mathcal{D}_0 : j \notin \mathcal{J}_f\}$.
6: Score $\mathcal{S}_f \leftarrow \{(X_{ij}, Y_{ij}, j) \in \mathcal{D} : j \in \mathcal{J}_f\}$ (all orders).
7: Compute preprocessing parameters on $\mathcal{T}_f$ (e.g., medians for numeric imputation; `MISSING` category for categorical NA).
8: Fit model $\widehat{m}_f$ on $\mathcal{T}_f$ and set $\widehat{p}_{ij}^{OOF} \leftarrow \widehat{m}_f(X_{ij})$ for all $(i, j) \in \mathcal{S}_f$.
9: **end for**
10: **return** $\widehat{p}^{OOF}$.

D Estimation Appendix

This appendix discuss implementation details for the MSM estimator, additional diagnostics, and robustness checks. It is written to make the estimation replicable and to clarify which modeling choices are required for identification versus which are convenience or numerical-stability choices.

D.1 Estimation details

Estimation Sample. The estimation sample is restricted to vacancies that receive exactly eight experimental applications. Within each vacancy, I observe each tracer application's submission timing (in coarse day bins) and whether that tracer is the first contacted applicant. I reshape the data into a job-major panel: each job contributes eight rows, one per tracer application, and all job-level variables are repeated within job.

Robustness tightness moments. I construct two sets of moments characterizing the contact initiation based on local tightness terciles for robustness. Specifically, within each wave I split jobs into the top and bottom terciles of local tightness and compute differences in model-implied initiation outcomes between the high- and low-tightness groups, then average these wave-specific differences using job counts as weights. The first set of moments contains two moments: (i) the difference in the probability that no experimental application receives the initiating contact, $\Delta P(\text{none})$, which captures the extensive margin, and (ii) the difference in the probability that initiation occurs by day 1 conditional on an experimental initiator, $\Delta P(d \leq 1 \mid \text{exp. contact})$, which captures the intensive timing margin. The second set of moments adds a third moment, the difference in the expected initiation day conditional on an experimental initiator, $\Delta E[d \mid \text{exp. contact}]$, so as to summarize more of the timing distribution beyond the early-mass measure alone.

I use these grouped moments in addition to the pooled regression moment because they relax the linear functional-form restriction implicit in the regression slope, are transparent to interpret in terms of extensive and intensive margins, and provide a robustness check that the estimated tightness channel is not driven by the particular auxiliary regression used in the baseline MSM.

Acceptance mapping, grids, interpolation, and clipping The acceptance probability $p(\lambda, \kappa)$ is obtained from the model's reservation rule. Evaluating this mapping repeatedly inside the MSM objective is computationally expensive, so I precompute p on a dense grid in $(\log \lambda, \log \kappa)$ and construct a two-dimensional interpolator. In simulation, I hard-clip (λ, κ) to the grid support before interpolation and clip the resulting probabilities to an interior interval $[10^{-12}, 1 - 10^{-12}]$ for numerical stability.

I monitor the fraction of simulated observations that are clipped and choose grid bounds so that clipping is negligible in the baseline estimates. I also report summary statistics of the implied (λ, κ) distributions to document that the parameter estimates are not driven by boundary artifacts.

D.2 Parameterization details

Centering and scaling of covariates Several covariates enter the arrival and κ parameterizations. To reduce collinearity between intercept-like terms and slopes, I use the following default normalizations (which match the implementation): (i) $\log c$ is centered within wave group before entering $\lambda_{0j}^{\text{cong}}$, so that $\log \xi$ captures group-level shifts in arrival conditions; (ii) the tightness entering κ_j is z-scored globally.

Macro-series parameterizations of $\log \xi$ The benchmark specification treats $\log \xi$ as a wave-group fixed effect. The disciplined specifications instead parameterize $\log \xi$ as a function of observed labor-market flow series. Candidate macro series include quit rates (QUR), hiring rates (HIR), layoff/discharge rates (LDR), job opening rates (JOR), and total separations rates (TSR) from JOLTS state-month estimation series (see E.2), each transformed into logs for rates expressed in percentage points or log-odds for openings, and then standardized.

In month-only specifications, the macro index is the national/month mean. In "with-deviation" specifications, I additionally include state-by-month deviations from the month mean, allowing local departures from aggregate conditions. I implement a three-tier estimation strategy: first, I test the five series individually and select HIR as the winner for model fitting and stability; second, I estimate multiple-series specifications with HIR as given, and select HIR + QUR + LDR as

the winner; Lastly, I estimation the with-deviation specification with HIR, QUR and TSR. The goal of this small set of pre-specified combinations is not to perform high-dimensional model selection, but to assess which economically interpretable shifters best replicate the evolution of the estimated macro state.

Here, I discuss the reason why I consider hires and separations as the natural macro shifters. The macro-series specification is guided by a simple matching-efficiency logic. In a standard Cobb-Douglas matching function, hires in market m and month t satisfy

$$H_{mt} = \Lambda_t U_{mt}^{\alpha} V_{mt}^{1-\alpha},$$

where Λ_t is the time-varying matching efficiency. Dividing by vacancies yields the job-filling rate per vacancy,

$$q_{mt} \equiv \frac{H_{mt}}{V_{mt}} = \Lambda_t \left(\frac{U_{mt}}{V_{mt}}\right)^{\alpha} = \Lambda_t c_{mt}^{\alpha},$$

with $c_{mt} \equiv U_{mt}/V_{mt}$ denoting congestion. This expression motivates using the hiring rate as a direct proxy for the flow of successful matches, and using separation flows as complementary indicators of the intensity and composition of labor market turnover. In particular, quits and layoffs/discharges are distinct margins of separations that respond differently to aggregate conditions and can shift effective search and matching frictions through channels not fully summarized by tightness alone (e.g., changes in worker outside options, churn, and vacancy chain dynamics). Accordingly, the disciplined macro parameterization allows $\log \xi$ to load on hiring and on both separation components, with the goal of capturing time variation in the arrival environment through observables that have a clear connection to the underlying matching technology.

To improve numerical stability in macro-series specifications, I initialize macro slopes using an OLS projection of the benchmark wave-group $\log \xi$ estimates onto the candidate macro series, and in the regularized estimator I add weak quadratic penalties that keep selected structural parameters near their benchmark values while leaving the data moments to discipline the macro slopes.

D.3 Computation details

I solve the MSM problem with a nonlinear least-squares routine with trust-region reflective, which is well-suited to residual-vector objectives and supports simple bounds. All simulations use common random numbers across objective evaluations to reduce simulation noise. I check robustness to local minima using multiple starting values and verify that the best solutions are stable across starts. I also report diagnostics that track implied (λ, κ) quantiles and clipping rates. In Algorithm 2, I summarize the MSM procedure.

D.4 Estimation Evaluation

In Table D.1, I summarize the 10 moments that are most mismatched in both the wave-specific and the macro-disciplined specification. Level moments are precisely estimated hence they contribute significantly to the model mismatch, but the differences are quite small.

Next, I show the robustness of the estimation using the simulated moments. First, I compare the counterfactuals based on the main (macro-disciplined) specification to the benchmark (wave-specific) specification. Figure D.1 shows that the two models generate virtually same moments, suggesting that the two-stage estimation procedure guarantees the consistency of the macro-disciplined specification and the flexible benchmark.

In Figure D.2, I plot the three loss measures defined in Section 7.6.4, scale loss, screening loss and probability loss, generated with two alternative machine learning algorithms (CatBoost with frequency-based categories and with text-based categories). The results show that the cost-benefit analysis is robust to the choice of machine learning algorithms. As discussed in Appendix Section C.3, the categorical CatBoost algorithms generate more within vacancy dispersion in predicted contact probability, which mechanically increases the magnitude of both the scale and screening loss without changing the relative positions of applications. This explains why scale loss and screening loss points are above the 45-degree line in both panels of Figure D.2. The loss presented in the main text is therefore the lower bound among the family of considered algorithms.

Algorithm 2 Method of simulated moments with precomputed $p(\lambda, \kappa)$

Require: Vacancy-level panel $\{j\}_{j=1}^{J}$ with wave-group $h(j)$, covariates (θ_j, s_j, M_j), and experimental application-day counts $\{n_{jd}\}_{d=0}^{7}$.

Require: Target moments g^{data} and bootstrap covariance $\widehat{\Omega}$.

Require: Grid bounds and size for $(\log \lambda, \log \kappa)$; simulation draws K and common random numbers $\{z_{jk}\}$.

Precompute $p(\lambda, \kappa)$ on a grid: for each $(\log \lambda_g, \log \kappa_g)$ solve the reservation cutoff $r(\lambda_g, \kappa_g)$ and set $p_g = 1 - \Phi(r)$; store a bivariate interpolator $\widehat{p}(\log \lambda, \log \kappa)$.

Compute g^{data} from the data and draw B vacancy-cluster bootstrap resamples to obtain $\widehat{\Omega}$ (or $\widehat{se}$).

Set weighting: (i) diagonal $\widehat{W}_{\text{diag}} = \text{diag}(\widehat{\Omega})^{-1}$, or (ii) full $\widehat{W}_{\text{full}} = (\widehat{\Omega} + \lambda I)^{-1}$ with ridge $\lambda > 0$.

Define the MSM residual map $\Theta \mapsto r(\Theta)$:

for $j = 1, \ldots, J$ **do**

    Construct baseline arrival rate $\lambda_{0j}(\Theta)$ from the congestion chanel, using $h(j)$ and covariates.

    Construct $\kappa_j(\Theta)$ from the κ parameterization (including sign-restricted loadings).

    for $k = 1, \ldots, K$ **do**

        Draw $\nu_{jk}(\Theta)$ from common random numbers and set $\lambda_{jk} = \lambda_{0j} \exp(\nu_{jk})$.

        Evaluate $p_{jk} = \widehat{p}(\log \lambda_{jk}, \log \kappa_j)$ with hard clipping to grid support.

        Compute implied initiation probabilities $\{P_{jd}(\lambda_{jk}, p_{jk})\}_{d=0}^{7}$ and weights $\{\omega_{jd}(\lambda_{jk}, p_{jk})\}_{d=0}^{7}$.

    end for

    Average over k to obtain $\{\bar{\omega}_{jd}(\Theta)\}_{d=0}^{7}$.

end for

Compute simulated moments $g^{\text{sim}}(\Theta)$ from $\{\bar{\omega}_{jd}(\Theta)\}$ using the same definitions as in the data (including wave-group pooling).

Form the moment difference $\Delta(\Theta) = g^{\text{sim}}(\Theta) - g^{\text{data}}$.

Return residuals $r(\Theta)$ such that $\|r(\Theta)\|^2 = \Delta(\Theta)^{\top} \widehat{W}\, \Delta(\Theta)$ (diagonal scaling or full whitening).

Estimate $\widehat{\Theta} \in \arg\min_{\Theta \in \mathcal{B}} \|r(\Theta)\|^2$ using a bound-constrained nonlinear least squares algorithm.

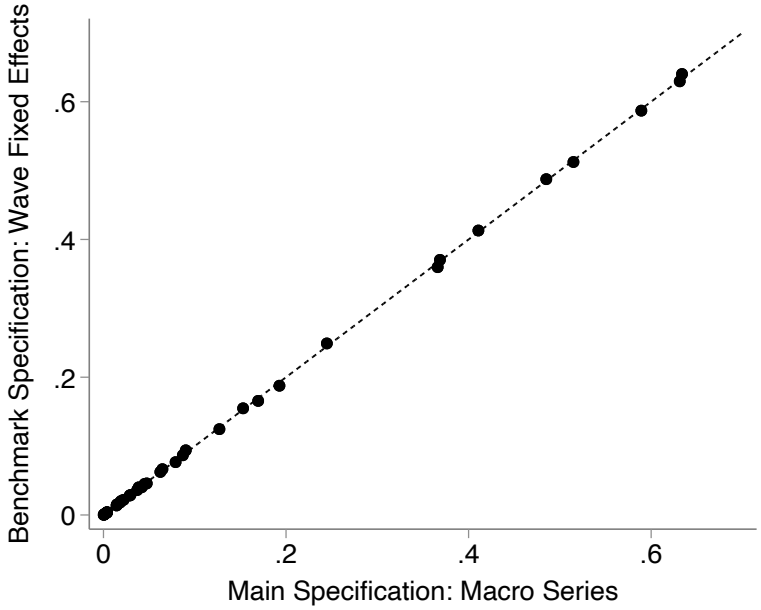
Report $\widehat{\Theta}$, block-specific fit diagnostics, and objective values evaluated under both $\widehat{W}_{\text{diag}}$ and $\widehat{W}_{\text{full}}$.

Table D.1: Top 10 Mismatched Moments

Rank	Moment	Difference	$z-$score	Contribution z^2
Panel A: Wave (diag)				
1	Level $h = 4, d = 3$	-0.0171	-2.33	5.42
2	Level $h = 2, d = 1$	0.0171	2.18	4.77
3	Level $h = 3, d = 1$	0.0101	1.78	3.18
4	Level $h = 2, d = 0$	-0.0244	-1.74	3.03
5	Level $h = 1, d = 1$	-0.0116	-1.49	2.21
6	Level $h = 2, d = 2$	0.0085	1.22	1.49
7	Level $h = 4, d = 1$	0.0115	1.16	1.34
8	Level $h = 3, d = 3$	-0.0052	-1.08	1.17
9	Level $h = 3, d = 6$	-0.0009	-0.91	0.84
10	Level $h = 1, d = 2$	0.0058	0.87	0.75
Panel B: Macro (diag)				
1	Level $h = 4, d = 3$	-0.0155	-2.11	4.46
2	Level $h = 3, d = 1$	0.0103	1.82	3.31
3	Level $h = 2, d = 1$	0.0134	1.71	2.92
4	Level $h = 4, d = 1$	0.0139	1.40	1.95
5	Level $h = 1, d = 1$	-0.0109	-1.39	1.93
6	Level $h = 2, d = 0$	-0.0192	-1.38	1.89
7	Level $h = 1, d = 2$	0.0075	1.11	1.22
8	Level $h = 3, d = 3$	-0.0051	-1.07	1.15
9	Level $h = 2, d = 3$	-0.0060	-0.95	0.91
10	Level $h = 2, d = 2$	0.0065	0.94	0.89

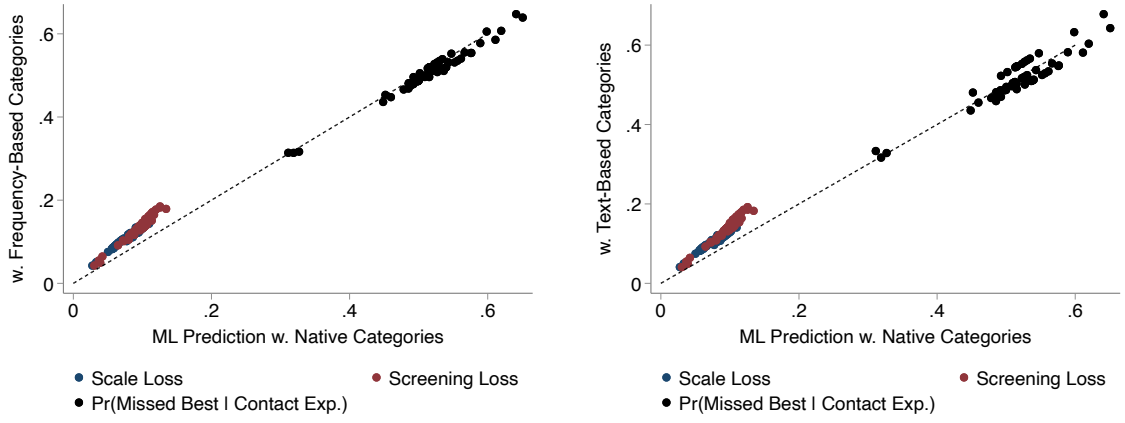
Note: This table lists the ten moments with the largest contributions to the diagonal MSM objective for wave specification and macro-series specification with diagonal weighting scheme. Each row reports the moment label, the difference between simulated values and empirical targets, the bootstrap z-score, and the squared contribution z^2. The moment set is held fixed across specifications, changes in this list reflect genuine shifts in fit with respect to the same set of empirical targets.

Figure D.1: Estimation Robustness: Main versus Benchmark Specification



Notes: This figure plots the simulated moments based on the estimates of the two specification. X-axis represents simulated moments with the macro-disciplined specification (main), y-axis presents simulated moments with the wave-specific specification.

Figure D.2: Cost-Benefit Robustness: Machine Learning Algorithms



(a) Frequency Based (b) Text Based

Notes: This figure plots the cost measures simulated based on three different machine learning algorithms: CatBoost with native categories, frequency based categories, and text based categories. Cost measures are defined as in Section 7.6.4, the machine learning algorithms are described in Section C.

E Details of Complementary Data

This appendix describes in detail how the complementary data introduced in Section 2.2 were sourced and constructed for the analysis.

E.1 Industry-Level Covariates

Industries are defined using two-digit North American Industry Classification System (NAICS) codes. In the original data released by Kline, Rose and Walters (2022), industries are reported using two-digit Standard Industrial Classification (SIC) codes. I recover NAICS-2 codes in two steps. First, for firms with Compustat coverage, I use the six-digit NAICS code reported in Compustat and convert it to the corresponding NAICS-2 industry. Second, for firms without Compustat coverage, I use the 2017 NAICS–SIC crosswalk to map SIC-2 codes into NAICS-2 codes.

Industry task measures Following Kline, Rose and Walters (2022), I use the Occupation Information Network (O*NET) database to construct task-based characteristics of U.S. industries. The ONET database provides a full list of occupations for each NAICS-2 industry. For each occupation, the database includes a projected number of job openings and the proportion of employment of that occupation within the industry. I calculate industry-level task measures as the weighted sum of occupation-level task measures, using the projected number of employment for each occupation in the industry as weights.

I construct five task-based indices intended to capture the importance of the *abilities*, *skills*, *knowledge* required to perform the work, the typical *activites* and the *context* of a job:

- **Analytical**: This measure is the average importance of 3 components: mathematical reasoning ability, mathematics knowledge, and mathematics skills. These variables measure how much mathematical skills and knowledge are required to perform the work.
- **Customer interaction**: This measure is the average importance of 4 components: 2 activities variables, establishing and maintaining interpersonal relationships and performing for or working directly with the public, 1 knowledge variable, customer and personal service, and 1 context variable contact

with others. These variables measure how much an industry requires workers to be in contact with others to perform their job.

- **Social**: This measure is the average importance of 4 skill components: social perceptiveness, coordination, persuasion, and negotiation. These variables measure to what extent developed capacities are used to work with people.
- **Routine**: This measure is the average importance of 2 context components: degree of automation and importance of repeating same tasks. These variables measure the relative amounts of routine work performed as part of a job.
- **Service**: The service measure is the average importance of 2 components: 1 activity variable, assisting and caring for others; 1 skill variable service orientation. These variables measure how important assisting people is for performing a job.

All variables are on a scale of 1 to 5. To better match the jobs sampled in the experiments, I focus on entry-level jobs, classified in the O*NET database as those in Job Zone 1 (little to no preparation needed) and 2 (some preparation needed). These jobs mostly do not require an associate degree or job-specific experiences. The five constructed measures are treated as time-invariant.

Industry employment by race and gender To measure baseline race and gender composition by industry, I use Labor Force Statistics on employment by industry, gender, and race from the Current Population Survey. For each sampled vacancy, I assign the share of Black employment and the share of female employment in the corresponding industry, measured one year prior to the year of the first experimental submission.

Monthly industry employment measures To capture time-varying conditions by industry, I use monthly (or quarterly) series on industry employment and compensation:

- **Employment size**: seasonally adjusted monthly employment by industry from the U.S. Bureau of Labor Statistics.
- **Compensation indices**: quarterly indices of *total compensation* and *total wages and salaries* for private-industry workers by industry from the Employment Cost Index. I convert the quarterly series to monthly by assigning each quar-

ter's value to the months in that quarter.

For each sampled vacancy, I compute rolling averages of these series over the 3 or 6 months prior to the month of the first experimental submission. In the analysis, I use the natural log of the averaged employment size.

E.2 State-Level Covariates

I construct three groups of state-level covariates: (i) pre-experiment income and mobility, (ii) minimum wage regulations, and (iii) monthly labor market conditions.

Income and mobility I use annual income and mobility measures from the Income Distributions and Dynamics in America (IDDA) for each state.[44] For each state, I calculate 2015-2019 averages of the following measures as pre-experiment proxies:

- **Median income and wages**: I use the median total income of the population and calculate the median wages of workers as the difference between the median total income and the median total non-wage income in each state. All variables are seasonally adjusted.
- **Income mobility**: IDDA provides a detailed yearly transition matrix across income bins $b \in \{< 25\text{pct.}, 25\text{-}50\text{pct.}, 50\text{-}75\text{pct.}, > 75\text{pct.}\}$. To construct a measure of mobility, I assign an integer $n(b_t, b_{t-1})$ equal to the number of bins moved between $t-1$ and t (e.g., moving from < 25pct. to 25–50pct. is $+1$, while moving from > 75pct. to < 25pct. is -3). I then calculate the expected value of mobility as

$$\sum_{b_t} \Pr(b_t \mid b_{t-1}) \cdot n(b_t, b_{t-1}),$$

  where $\Pr\left((b_t, b_{t-1}) \mid b_{t-1}\right)$ is the probability of moving from income bin b_{t-1} to b_t during year t, conditional on being in income bin b_{t-1} in year $t-1$. $n(b_t, b_{t-1})$ is the assigned value of moving. This procedure generates a mobility measure for each income bin b for each year.

[44] I use variables constructed from household income and earnings data aggregated from Form 1040. For details, see Kondo et al. (2023).

Table E.1: Summary of Firm-Level Covariates

	Mean	SD	Min	Max
Income and Wage				
Median Income (log)	11.3232	0.1630	11.0154	11.6375
Median Wage (log)	11.2806	0.1557	10.9964	11.5965
Mobility by Income Bins				
Mobility: < 25pct.	0.3707	0.0397	0.3052	0.4971
Mobility: $25 - 50$pct.	0.1181	0.0174	0.0948	0.1655
Mobility: $50 - 75$pct.	-0.0967	0.0095	-0.1194	-0.0748
Mobility: > 75pct.	-0.3343	0.0402	-0.4357	-0.2806

Note: This table summarizes state-level covariates for the 51 states in the sample. Data are sourced from IDDA 1040 (household) data.

State minimum wage and employment To capture state-level labor regulations, I use the history of changes in state minimum wages for non-farm employment published by the U.S. Department of Labor. For states without a minimum wage, I use the federal minimum wage instead. I use the Expanded State Employment Status tables from the Local Area Unemployment Statistics. For each sampled vacancy, I assign the share of Black employment and the share of female employment in the state where the vacancy is located, measured one year prior to the year of the first experimental submission.

State labor market conditions The last set of state-level covariates are sourced from the Job Openings and Labor Turnover Survey (JOLTS). JOLTS provides a rich set of variables and state-month level estimations[45]. I use the following seasonally adjusted measures for non-farm employment:

- **Job opening rate**: The number of positions that are open on the last business day of the month divided by the sum of employment and job openings.
- **Hire rate**: The number of additions to the payroll during the month divided by employment.
- **Total separation rate**: The number of employees separated from the payroll during the calendar month divded by employment. I also include two subcategories of separations: **quit rate**, defined as the number of employees

[45]For the state-month level estimation methodology and more details of JOLTS, see offical page for details of estimation methodology.

who left voluntarily except for retirements and transfers to other locations, divided by employment; and **layoff/discharge rate**, defined as the number of involuntary separations initiated by the employer, divided by employment.

- **Tightness**: The number of unemployed persons per job opening.

For each sampled vacancy, I assign the state-month value corresponding to the state of the vacancy and the month of the first experimental submission. I also compute 3-month pre-submission averages as proxies for pre-experiment conditions. All variables are seasonally adjusted.

E.3 Firm-Level Covariates

For the 108 sampled firms, I assemble covariates capturing human resource management practices, employee satisfaction, financial performance, and within-state labor market power.[46] The main proprietary source is Revelio Labs (Revelio Labs, 2026), which aggregates publicly available workforce information from online professional profiles, job postings, and employee reviews.[47]

Matching with Revelio Labs data In the replication package of Kline, Rose and Walters (2024), firm names are revealed. For publicly traded firms, I manually collect tickers. I then match the experiment firms to Revelio Labs in two steps. First, I use tickers to match publicly traded firms. Second, I apply fuzzy matching on firm names to match all firms. I cross-validate matches from the two approaches, manually audit match quality, and finalize a crosswalk between the experiment firms and Revelio Labs identifiers. This crosswalk then bridges the experiment data and the two Revelio Labs datasets I use.

Local labor market power I use the job posting data from Revelio Labs to construct measures of firm labor market power at the state-by-month level.[48]

Specifically, I calculate:

- **Share of postings within industry-state**: the number of postings by the

[46] Covariates marked with *(std.)* are normalized to have a standard deviation of 1 in the analysis.

[47] Revelio Labs, Data Dictionary documentation (accessed February 24, 2026): https://www.data-dictionary.reveliolabs.com/data.html.

[48] Revelio Labs, *Job Postings* datasets; see https://www.data-dictionary.reveliolabs.com/data.html (accessed February 24, 2026).

firm in a state and month divided by total postings by all firms in the same NAICS-2 industry in that state and month.

- **Share of postings within state**: the number of postings by the firm in a state and month divided by total postings by all firms in that state and month.

For each measure, I also construct an entry-level version by restricting to postings in O*NET Job Zones 1 and 2.

Similar to the construction of labor market conditions with JOLTS data, for each sampled vacancy, I assign the relevant state-month measure using the vacancy's state and the month of the first experimental submission. I also compute 3-month pre-submission averages as proxies for pre-experiment conditions.

Employee satisfaction Following Kline, Rose and Walters (2022), I use worker review ratings as proxies for employee satisfaction.[49] The measures include:

- **Overall** (scale): Overall rating of the company.
- **Business outlook** (categorical): Rating of the company's business outlook.
- **Career opportunities** (half-scale): Rating of the company's career opportunities.
- **CEO** (categorical): Approval rating of the company's CEO.
- **Compensation/benefits** (half-scale): Rating of the company's compensation and benefits.
- **Cultural and values** (scale): Rating of the company's culture and values.
- **Diversity and inclusion** (scale): Rating of the company's diversity and inclusion.
- **Recommend to Friend** (categorical): Whether the reviewer would recommend the company to a friend.
- **Senior leadership** (half-scale): Rating of the company's senior management.
- **Work-life balance** (half-scale): Rating of the company's work-life balance.

Scale measures range from 1 to 5, with 5 indicating the most favorable rating. Half-scale measures also range from 1 to 5 but allow half-point increments. Categorical measures are coded as 1 for positive, 0 for neutral, and -1 for negative. For each sampled vacancy, I compute the rolling mean of each rating over the six months preceding the month of the first experimental submission and assign these values

[49] Revelio Labs, *Sentiment* datasets; see https://www.data-dictionary.reveliolabs.com/data.html (accessed February 24, 2026).

to the vacancy. All ratings are standardized in the analysis.

Human resource management I use the following measures to capture the human resource management practices of the sample firms:

- 1 (**federal contractor**): the indicator for a firm being a federal contractor.
- **Labor law violations per employee** *(std.)*: I collect the total count of employment related law violations since 2000 from the Good Job First violation tracker, and divide it by total employment from Compustat.
- **Contact centralization**: This measure was constructed by Kline, Rose and Walters (2022), defined as

$$\frac{\text{total number of unique phone numbers from a firm to contact applicants}}{\text{total number of jobs where at least one contact was made} - 1}$$

  This ratio serves as a proxy for the concentration of hiring responsibility within the firm.
- **HR per employee** *(std.)*: For each sampled firm, I collect the total number of LinkedIn profiles associated with the firm that hold a human resource management-related position and divide it by the total number of LinkedIn profiles listing the firm as the employer. The search was conducted in July 2024. The human resource service provider ADP was excluded.

Financial performance Following Kline, Rose and Walters (2022), I sourced the annual financial performance data from Compustat:

- **Log employment** *(std.)*: the nature log of the 2010-2018 average employment.
- **Total assets per employee** *(std.)*: total sales divided by employment, averaged over 2010-2018.
- **Gross profit per employee** *(std.)*: gross profit divided by employment, averaged over 2010-2018.
- **Total revenue per employee** *(std.)*: total revenue divided by employment, averaged over 2010-2018.

Firms without Compustat coverage are omitted for these measures.

Table E.2: Summary of Firm-Level Covariates

	Mean	SD	Median	No.(firms)
Human resource management				
$\mathbf{1}$(federal contractor)	0.6667	0.4736	1.0000	108
Violations per emp.	0.0004	0.0007	0.0002	94
Contact centralization	-1.1204	0.3685	-1.0842	100
HR per emp.	0.0254	0.0092	0.0251	107
Financial performance				
log(employment)	11.2432	1.0384	11.0745	94
Total assets per emp. ($M)	0.8484	2.1732	0.1979	94
Gross profit per emp. ($M)	0.1042	0.0806	0.0819	94
Total revenue per emp. ($M)	0.3563	0.3828	0.2613	94

Note: This table summarizes firm-level covariates related to **human resource management**, **employee satisfaction** and **financial performance**. The data here are not normalized. No.(firms) refers to the number of firms with non-missing values for the covariates.

F Robustness Checks

F.1 Robustness of Timing Measures

I first check the robustness of early application indicator $early^1$. I compare it with

- $early^2$: The first alternative early submission indicator, mirroring $early^1$, is defined such that the last four submissions to a job are coded as late. For jobs with no more than four submissions, all submissions are coded as late. In full sample, 40210 (48.07%) are coded as early submissions.
- $early^3$: The second alternative early submission indicator, based on the designed pair orders, is defined such that the last two pairs of submissions to a job are coded as late. For jobs with no more than two pairs of submissions (836 jobs, 7.52%; 2321 submissions, 2.77%), all submissions are coded as late. In full sample, 40354 submissions (in 20177 pairs, 48.25%) are coded as early submissions.

Table F.1 summarizes the three constructed early submission indicators, showing that they are largely consistent with each other. Table F.2 shows that replacing $early^1$ in Model (3.1) with the two alternatives does not change results.

As described in Section 2.1, applications were submitted in pairs by design. Therefore, I also consider the designed order of pairs as an alternative measure of submission order. I estimate

$$Y_i = \alpha_0 + \sum_{j=2}^{4} \beta_j \cdot \mathbf{1}\left(Order_i^{pair} = j\right) + \boldsymbol{\gamma}\mathbf{X}_i + \lambda_{j(i)} + \epsilon_i \tag{F.1}$$

where $\mathbf{1}\left(Order_i^{pair} = j\right)$ is the pair order indicator, taking the value 1 if ficticious applicant i is designed to be in the $j-$th pair of submission to a job. Results are reported in Table F.3. Early application advantages are significant, consistent with results in Figure 1.

As discussed in Section 3.1, the measure $D2L^t$ is constructed such that all submissions more than 6 days earlier than the last submission are grouped together and coded as 7. One might worry that this regrouping could bias the results. To address this concern, I reestimate Model (3.3) directly using the raw measure $D2L$ winsorized at the top 1% level to adjust for the right-skewness. The results, shown

in Table F.4, are virtually the same as those presented in Table A.3.

F.2 Robustness to Alternative Specifications

Wave 1 versus Wave 2-5. As discussed in the main text, the experiment differs in submission time frame between Wave 1 and Wave 2-5. In Wave 1, each pair of applications was planned to be submitted every other day, while in Waves 2-5, pairs were planned to be submitted daily. Specifically, in Wave 1, around 95.6% of all applications were submitted no more than 8 days before the last submission, whereas in Waves 2-5, around 96.6% were submitted no more than 4 days before the last submission. Figure F.1 reestimate Model (3.4) separately for Wave 1 and Wave 2-5 subsamples, Wave 1 subsample results differ, but the linear trend is mostly robust.z The patterns are similar.

Incomplete subsample In Figure F.2, I employ the same within-job specification as in Figure 1 for *Complete* jobs and *Incomplete* jobs separately. Since 94% of submissions come from *Complete* jobs, the results on *Complete* job sub-sample, shown in Panel (a) of F.2, are nearly identical to the full-sample results. For the *Incomplete* job subsample, shown in Panel (b) of Figure F.2, the within-job early application advantage remains significant and even larger in scale, but the results are noisy due to the much smaller sample size.

Alternative FEs. Kline, Rose and Walters (2022) employ the specification with wave and region fixed effects and job-level clustering. Although the main focus of this paper is within-vacancy comparision, I show in Table F.5 and F.6 that the results are robust with this alternative specification.

Pair stratification of race. In the experiment, applicants' race are randomized with pair-level stratification. That is, within each pair of submission, one of the application is randomly assigned to be black while the other white. I reestimate Model (F.1) and and add interaction terms of $\mathbf{1}\left(Order_i^{pair} = j\right)$, $j = 2, 3, 4$, the designed pair order indicators, with race and gender indicators. Table F.7 shows that the results are consistent with those in Table A.4.

In Table F.8, I comparing the full set of resume characteristics across submissions of the four designed pairs. While there are changes of effects for individual characteristics, joint tests cannot reject the null hypothesis of no interaction between the characteristics and the designed order of submission pairs, consistent with the results reported in Table 2.

F.3 Robustness of Contact History Measures

In Table F.9 and F.10, I repeat the exercise in Table A.9 and Table F.9 with indicators $\mathbf{1}\left(NoC_i^{up} > 0\right)$ and $\mathbf{1}\left(NoC_i^{low} > 0\right)$ instead of the numeric measures of contact history. The results are very consistent. Quota mechanism dominates for both black and white applicants, but black applicants experience a tighter quota. Learning mechanism dominates for both female and male applicants, and the impact is symmetric.

As discussed in Section 4.2, the two contact history measures rely on the time span between the submission and the contact for an application, conditional on it being contacted. For those not contacted, no information is available on how long it takes for a recruiter to evaluate and decide not to contact. Here, I construct an alternative measure of contact history NoC_i^{alt}, solely relying on the order of submission (see Table A.5 for an example of construction). All applications for a job submitted prior to submission i are included to calculate the contact history. Using this measure and the corresponding indicator $\mathbf{1}\left(NoC_i^{alt} > 0\right)$, I reestimate (4.2) and report results in Table F.11. These results are very consistent with those estimated using the two contact-date-based measures in Table A.9 and A.10.

Table F.1: Consistency of Early Submission Indicators

		$early^1$		$early^2$		$early^3$	
		$=0$	$=1$	$=0$	$=1$	$=0$	$=1$
Panel A: full sample							
$early^1$	$=0$			42531	0	42005	526
	$=1$			902	40210	1284	39828
$early^2$	$=0$					42755	678
	$=1$					534	39676
total		42531	41112	43433	40210	43289	40354
				$N = 83643$			
Panel B: jobs with 8 submissions							
$early^1$	$=0$			39316	0	38802	514
	$=1$			0	39316	514	38802
$early^2$	$=0$					38802	514
	$=1$					514	38802
total		39316	39316	39316	39316	39316	39316
				$N = 78632$			

Note: Full sample in Panel A, Jobs with 8 submissions in Panel B. The four early submission indictors are constructed as outlined in Section 3.1.

Table F.2: Estimation of Model (3.1), Main Specification

	(1) Full	(2) < 8	(3) Full	(4) < 8	(5) Full	(6) < 8
$early^1$	0.0116	0.0240				
	(0.0024)	(0.0127)				
$early^2$			0.0113	0.0135		
			(0.0023)	(0.0118)		
$early^3$					0.0105	0.0121
					(0.0023)	(0.0119)
Black	-0.0205	-0.0125	-0.0205	-0.0124	-0.0205	-0.0125
	(0.0018)	(0.0066)	(0.0018)	(0.0066)	(0.0018)	(0.0066)
Female	-0.0021	-0.0006	-0.0021	-0.0009	-0.0021	-0.0008
	(0.0022)	(0.0080)	(0.0022)	(0.0081)	(0.0022)	(0.0081)
Over 40	-0.0043	0.0041	-0.0043	0.0038	-0.0043	0.0040
	(0.0020)	(0.0083)	(0.0020)	(0.0083)	(0.0020)	(0.0083)
Pol. club	0.0059	0.0037	0.0059	0.0021	0.0059	0.0020
	(0.0051)	(0.0251)	(0.0051)	(0.0252)	(0.0051)	(0.0252)
Acad. club	0.0049	0.0136	0.0049	0.0135	0.0049	0.0140
	(0.0051)	(0.0175)	(0.0051)	(0.0174)	(0.0051)	(0.0175)
LGBTQ club	-0.0003	0.0152	-0.0003	0.0152	-0.0003	0.0154
	(0.0036)	(0.0171)	(0.0036)	(0.0171)	(0.0036)	(0.0171)
Typical pron.	-0.0077	0.0056	-0.0077	0.0042	-0.0077	0.0043
	(0.0051)	(0.0203)	(0.0051)	(0.0201)	(0.0051)	(0.0201)
Neutral pron.	-0.0086	0.0065	-0.0086	0.0068	-0.0085	0.0069
	(0.0051)	(0.0217)	(0.0051)	(0.0216)	(0.0051)	(0.0216)
Assoc. degree	0.0021	0.0082	0.0021	0.0080	0.0021	0.0080
	(0.0020)	(0.0085)	(0.0020)	(0.0085)	(0.0020)	(0.0085)
Constant	0.2475	0.1680	0.2477	0.1746	0.2481	0.1744
	(0.0025)	(0.0097)	(0.0024)	(0.0085)	(0.0024)	(0.0088)
Job FEs	✓	✓	✓	✓	✓	✓
N	83643	5011	83643	5011	83643	5011

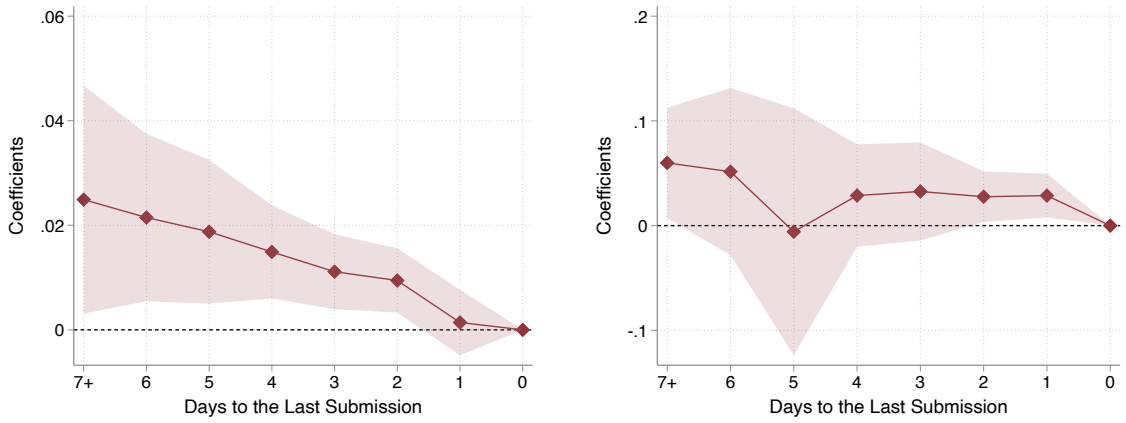
Note: This table presents the effects of early submission and other characteristics on the probability of employer contact within 30 days. Columns titled "Full" include all jobs, while columns titled "<8" include jobs with less than 8 submissions. All regressions include job fixed effects. Standard errors in parentheses are clustered at the job level.

Table F.3: Estimation of Model (F.1), w. Job Fixed Effects

	(1) Full	(2) $= 8$	(3) < 8
$\mathbf{1}(Order_i^{pair} = 2)$	-0.0036	-0.0021	-0.0211
	(0.0027)	(0.0028)	(0.0091)
$\mathbf{1}(Order_i^{pair} = 3)$	-0.0126	-0.0113	-0.0327
	(0.0031)	(0.0032)	(0.0141)
$\mathbf{1}(Order_i^{pair} = 4)$	-0.0127	-0.0116	-0.0572
	(0.0032)	(0.0033)	(0.0346)
Constant	0.2601	0.2641	0.1901
	(0.0028)	(0.0030)	(0.0090)
Controls $\mathbf{X}_i$	✓	✓	✓
Job FEs	✓	✓	✓
N	83643	78632	5011

Note: This table presents the effects of the order of submission pairs on the probability of employer contact within 30 days. The column titled "Full" includes all jobs, the column titled "=8" includes all *Complete* jobs with 8 submissions, the column titled "<8" includes only *Incomplete* jobs with less than 8 submissions. All regressions include the full set of resume characteristics and job fixed effects. Standard errors clustered at the job level are reported in parentheses, robust standard errors are reported in brackets.

Figure F.1: Contact Gap and $D2L^t$, w. Job Fixed Effects



(a) Wave 1 ($N = 14541$) (b) Wave 2-5 ($N = 69102$)

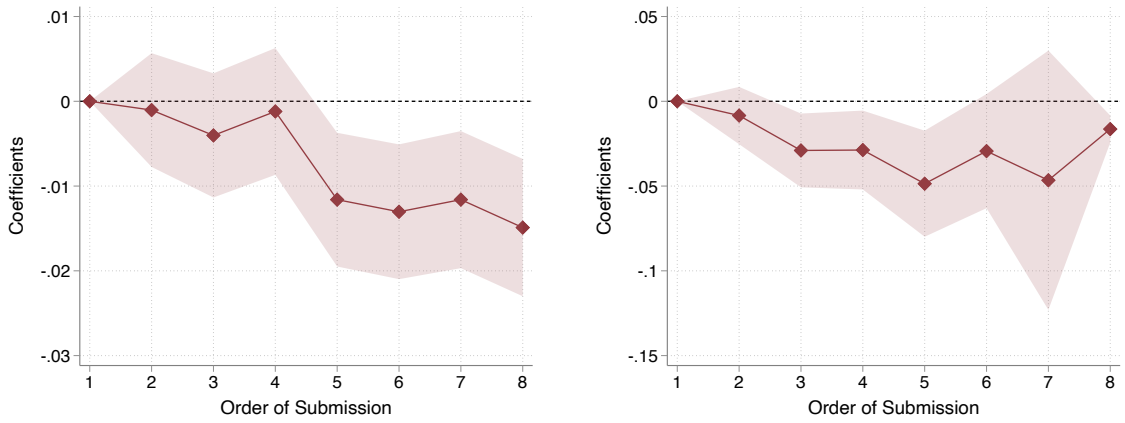
Notes: The figures plot the estimated coefficients $\hat{\delta}_j, j \in \{1, \cdots, 7\}$ of Model (3.4) with days to the last submission $D2L^t$ on the x-axis, controlling for resume characteristics. The left panel shows Wave 1 results. The right panel, shows Wave 2-5 results. The regression includes job fixed effects, the standard errors are clustered at the job level. The area depicts the 95% confidence intervals.

Table F.4: Estimation of Model (3.3), Winsorized $D2L$

	(1)	(2)	(3)	(4)	(5)	(6)
	All		Wave 1		Wave 2-5	
	Full	< 8	Full	< 8	Full	< 8
$D2L$	0.0039	0.0084	0.0034	0.0079	0.0042	0.0092
	(0.0008)	(0.0032)	(0.0013)	(0.0038)	(0.0010)	(0.0064)
Black	-0.0205	-0.0125	-0.205	-0.0112	-0.0205	-0.0132
	(0.0018)	(0.0066)	(0.0041)	(0.0113)	(0.0020)	(0.0080)
Female	-0.0021	-0.0007	0.0013	-0.0044	-0.0028	0.0014
	(0.0022)	(0.0080)	(0.0051)	(0.0131)	(0.0024)	(0.0102)
Over 40	-0.0043	0.0034	-0.0041	0.0036	-0.0043	0.0034
	(0.0020)	(0.0083)	(0.0048)	(0.0147)	(0.0022)	(0.0100)
Pol. club	0.0060	0.0026			0.0060	0.0027
	(0.0051)	(0.0253)			(0.0051)	(0.0252)
Acad. club	0.0050	0.0136			0.0050	0.0136
	(0.0051)	(0.0174)			(0.0051)	(0.0173)
LGBTQ club	-0.0002	0.0154			-0.0002	0.0153
	(0.0036)	(0.0171)			(0.0036)	(0.0171)
Typical pron.	-0.0078	0.0040			-0.0078	0.0040
	(0.0051)	(0.0202)			(0.0051)	(0.0202)
Neutral pron.	-0.0085	0.0062			-0.0085	0.0059
	(0.0051)	(0.0216)			(0.0050)	(0.0216)
Assoc. degree	0.0021	0.0084	-0.0013	0.0040	0.0028	0.0108
	(0.0020)	(0.0085)	(0.0050)	(0.0150)	(0.0022)	(0.0103)
Constant	0.2454	0.1669	0.1932	0.1743	0.2561	0.1624
	(0.0026)	(0.0092)	(0.0061)	(0.0148)	(0.0030)	(0.0121)
Job FEs	✓	✓	✓	✓	✓	✓
N	83643	5011	14541	1821	69102	3190

Note: This table presents the effects of early submission measured by days before the last submission and other characteristics on the probability of employer contact within 30 days. Columns titled "Full" include all jobs, while columns titled "<8" include jobs with less than 8 submissions. Column (1) and (2) are on full sample, Column (3) and (4) are on Wave 1 where the variables "Pol. club", "Acad. club", "LGBTQ club", "Typical pron." and "Neutral pron." were not designed to vary, Column (5) and (6) are on Wave 2-5. All regressions include job fixed effects. Standard errors in parentheses are clustered at the job level.

Figure F.2: Contact Gap and Order of Submission, *Complete* v.s. *Incomplete* Jobs



(a) *Complete* Jobs $(N = 78632)$ (b) *Incomplete* Jobs $(N = 5011)$

Notes: The figures plot the estimated coefficients $\hat{\beta}_j, j \in \{2, \cdots, 8\}$ of Model (3.2) with the order of submission on the x-axis, controlling for resume characteristics. In the left panel, the sample consists of jobs with 8 submissions. In the right panel, the sample consists of jobs with less than 8 submissions. The regression includes job fixed effects, the standard errors are clustered at the job level. The area depicts the 95% confidence intervals.

Table F.5: Estimation of Model (3.1), w. Wave and Region Fixed Effects

	(1) Full	(2) < 8	(3) Full	(4) < 8	(5) Full	(6) < 8
$early^1$	0.0123	0.0048				
	(0.0023)	(0.0160)				
$early^2$			0.0141	0.0028		
			(0.0023)	(0.0145)		
$early^3$					0.0131	0.0061
					(0.0023)	(0.0150)
Black	-0.0206	-0.0118	-0.0206	-0.0118	-0.0205	-0.0118
	(0.0017)	(0.0059)	(0.0017)	(0.0059)	(0.0017)	(0.0059)
Female	0.0001	0.0134	0.0001	0.0134	0.0001	0.0134
	(0.0030)	(0.0107)	(0.0030)	(0.0107)	(0.0030)	(0.0107)
Over 40	-0.0058	0.0071	-0.0058	0.0071	-0.0058	0.0071
	(0.0030)	(0.0113)	(0.0030)	(0.0113)	(0.0030)	(0.0113)
Pol. club	-0.0018	-0.0440	-0.0018	-0.0440	-0.0018	-0.0440
	(0.0074)	(0.0296)	(0.0074)	(0.0295)	(0.0074)	(0.0295)
Acad. club	0.0096	0.0185	0.0095	0.0185	0.0095	0.0186
	(0.0076)	(0.0339)	(0.0076)	(0.0340)	(0.0076)	(0.0340)
LGBTQ club	-0.0052	0.0317	-0.0052	0.0317	-0.0052	0.0317
	(0.0055)	(0.0228)	(0.0055)	(0.0228)	(0.0055)	(0.0228)
Typical pron.	-0.0138	-0.0174	-0.0139	-0.0175	-0.0138	-0.0174
	(0.0073)	(0.0294)	(0.0073)	(0.0294)	(0.0073)	(0.0294)
Neutral pron.	-0.0105	-0.0322	-0.0105	-0.0323	-0.0105	-0.0323
	(0.0076)	(0.0301)	(0.0076)	(0.0301)	(0.0076)	(0.0301)
Assoc. degree	0.0012	0.0176	0.0012	0.0175	0.0012	0.0176
	(0.0030)	(0.0110)	(0.0030)	(0.0110)	(0.0030)	(0.0110)
Constant	0.2009	0.1854	0.2003	0.1866	0.2008	0.1859
	(0.0113)	(0.0324)	(0.0113)	(0.0322)	(0.0113)	(0.0322)
Wave FEs	✓	✓	✓	✓	✓	✓
Region FEs	✓	✓	✓	✓	✓	✓
N	83643	5011	83643	5011	83643	5011

Note: This table presents the effects of early submission and other characteristics on the probability of employer contact within 30 days. Columns titled "Full" include all jobs, while columns titled "<8" include jobs with less than 8 jobs. All regressions include wave and region fixed effects. Standard errors in parentheses are clustered at the job level.

Table F.6: Estimation of Model (3.3), w. Wave and Region Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)
	All		Wave 1		Wave 2-5	
	Full	< 8	Full	< 8	Full	< 8
$D2L^t$	0.0054	0.0161	0.0056	0.0254	0.0053	-0.0034
	(0.0010)	(0.0051)	(0.0014)	(0.0066)	(0.0014)	(0.0073)
Black	-0.0206	-0.0118	-0.0216	-0.0136	-0.0204	-0.0108
	(0.0017)	(0.0059)	(0.0038)	(0.0099)	(0.0019)	(0.0074)
Female	0.0002	0.0131	0.0045	0.0279	-0.0007	0.0047
	(0.0030)	(0.0106)	(0.0065)	(0.0172)	(0.0034)	(0.0134)
Over 40	-0.0059	0.0063	-0.0024	0.0167	-0.0067	-0.0000
	(0.0030)	(0.0113)	(0.0066)	(0.0195)	(0.0033)	(0.0138)
Pol. club	-0.0018	-0.0436			-0.0018	-0.0431
	(0.0074)	(0.0298)			(0.0074)	(0.0296)
Acad. club	0.0095	0.0184			0.0096	0.0195
	(0.0076)	(0.0339)			(0.0076)	(0.0341)
LGBTQ club	-0.0051	0.0316			-0.0051	0.0310
	(0.0055)	(0.0228)			(0.0055)	(0.0228)
Typical pron.	-0.0140	-0.0181			-0.0139	-0.0168
	(0.0073)	(0.0294)			(0.0073)	(0.0295)
Neutral pron.	-0.0104	-0.0335			-0.0103	-0.0332
	(0.0076)	(0.0301)			(0.0076)	(0.0301)
Assoc. degree	0.0012	0.0180	0.0026	0.0028	0.0009	0.0278
	(0.0030)	(0.0110)	(0.0068)	(0.0196)	(0.0034)	(0.0131)
Constant	0.1902	0.1604	0.1844	0.1274	0.2518	0.1806
	(0.0116)	(0.0326)	(0.0195)	(0.0488)	(0.0130)	(0.0469)
Wave FEs	✓	✓	✓	✓	✓	✓
Region FEs	✓	✓	✓	✓	✓	✓
N	83643	5011	14541	1821	69102	3190

Note: This table presents the effects of early submission measured by days before the last submission and other characteristics on the probability of employer contact within 30 days. Columns titled "Full" include all jobs, while columns titled "<8" include jobs with less than 8 submissions. Column (1) and (2) are on full sample, Column (3) and (4) are on Wave 1 where the variables "Pol. club", "Acad. club", "LGBTQ club", "Typical pron." and "Neutral pron." were not designed to vary, Column (5) and (6) are on Wave 2-5. All regressions include region and wave fixed effects. Standard errors in parentheses are clustered at the job level.

Table F.7: Application Pair Order, Race and Gender

	(1)	(2)	(3)
$\mathbf{1}(Order_i^{pair} = 2)$	-0.0023	-0.0015	0.0055
	(0.0036)	(0.0038)	(0.0052)
$\mathbf{1}(Order_i^{pair} = 3)$	-0.0136	-0.0113	-0.0107
	(0.0039)	(0.0040)	(0.0054)
$\mathbf{1}(Order_i^{pair} = 4)$	-0.0121	-0.0091	-0.0040
	(0.0040)	(0.0043)	(0.0057)
Black	-0.0200	-0.0205	-0.0089
	(0.0032)	(0.0018)	(0.0049)
$\mathbf{1}(Order_i^{pair} = 2) \times$ Black	-0.0026		-0.0142
	(0.0046)		(0.0071)
$\mathbf{1}(Order_i^{pair} = 3) \times$ Black	0.0020		-0.0011
	(0.0046)		(0.0071)
$\mathbf{1}(Order_i^{pair} = 4) \times$ Black	-0.0011		-0.0102
	(0.0046)		(0.0073)
Female	-0.0021	0.0012	0.0123
	(0.0022)	(0.0041)	(0.0057)
$\mathbf{1}(Order_i^{pair} = 2) \times$ Female		-0.0041	-0.0156
		(0.0055)	(0.0078)
$\mathbf{1}(Order_i^{pair} = 3) \times$ Female		-0.0025	-0.0055
		(0.0054)	(0.0077)
$\mathbf{1}(Order_i^{pair} = 4) \times$ Female		-0.0071	-0.0161
		(0.0057)	(0.0081)
Black $\times$ Female			-0.0223
			(0.0078)
$\mathbf{1}(Order_i^{pair} = 2) \times$ Black $\times$ Fem.			0.0232
			(0.0109)
$\mathbf{1}(Order_i^{pair} = 3) \times$ Black $\times$ Fem.			0.0059
			(0.0108)
$\mathbf{1}(Order_i^{pair} = 4) \times$ Black $\times$ Fem.			0.0182
			(0.0111)
Controls $\mathbf{X}_i$	✓	✓	✓
Job FEs	✓	✓	✓
N	83643	83643	83643

Note: This table presents the effects of designed submission pair orders interacted with race and gender on the probability of employer contact within 30 days, controlling for other characteristics. $\mathbf{1}(Order_i^{pair} = j), j = 2, 3, 4$ indicates whether application i was designed to be submitted in Pair j. All regressions are on the full sample, including job fixed effects and all resume characteristics. Standard errors in parentheses are clustered at the job level.

Table F.8: Order of Submission Pairs and Resume characteristics

	(1) Pair 1 - 2	(2) Pair 1 - 3	(3) Pair 1 - 4	(4) Pair 2 - 3	(5) Pair 2 - 4	(6) Pair 3 - 4
Black	0.0132	-0.0024	0.0135	-0.0118	-0.0041	0.0059
	(0.0078)	(0.0080)	(0.0083)	(0.0080)	(0.0085)	(0.0080)
Female	0.0132	0.0014	0.0168	-0.0091	0.0010	0.0062
	(0.0087)	(0.0090)	(0.0095)	(0.0089)	(0.0094)	(0.0088)
Black × Female	-0.0216	0.0012	-0.0242	0.0150	0.0054	-0.0059
	(0.0121)	(0.0127)	(0.0131)	(0.0127)	(0.0132)	(0.0124)
Over 40	0.0037	0.0038	-0.0043	-0.0001	-0.0076	-0.0012
	(0.0061)	(0.0064)	(0.0067)	(0.0063)	(0.0067)	(0.0063)
Pol. club	-0.0086	0.0077	-0.0036	0.0180	0.0172	-0.0058
	(0.0168)	(0.0169)	(0.0181)	(0.0160)	(0.0169)	(0.0159)
Acad. club	0.0156	-0.0035	-0.0104	0.0028	-0.0145	-0.0123
	(0.0159)	(0.0168)	(0.0174)	(0.0155)	(0.0171)	(0.0160)
LGBTQ club	0.0030	-0.0019	-0.0187	0.0006	-0.0130	-0.0098
	(0.0109)	(0.0118)	(0.0120)	(0.0115)	(0.0121)	(0.0114)
Typical pron.	-0.0004	-0.0104	-0.0028	-0.0087	-0.0003	0.0164
	(0.0164)	(0.0166)	(0.0171)	(0.0165)	(0.0167)	(0.0159)
Neutral pron.	0.0000	0.0035	0.0120	0.0056	0.0103	0.0025
	(0.0158)	(0.0164)	(0.0173)	(0.0151)	(0.0178)	(0.0154)
Assoc. degree	-0.0079	-0.0044	-0.0035	0.0036	0.0008	-0.0019
	(0.0063)	(0.0066)	(0.0068)	(0.0063)	(0.0067)	(0.0063)
Constant	-0.0029	0.0126	0.0086	0.0138	0.0140	-0.0018
	(0.0076)	(0.0079)	(0.0084)	(0.0077)	(0.0085)	(0.0078)
Job FEs	✓	✓		✓	✓	
N	43433	42627	41873	41770	41016	40210
$F-$ stats, joint	0.70	0.17	0.76	0.42	0.52	0.34
$p-$value	0.7237	0.9980	0.6644	0.9376	0.8798	0.9704

Note: This table presents the effects of designed submission pairs interacted with resume characteristics on the probability of employer contact within 30 days. Column (1)-(3) display the differences of between submissions of Pair 2 to 4 and Pair 1; Column (4)-(5) display the differences of between submissions of Pair 3 to 4 and Pair 2; Column (6) displays the differences of between submissions of Pair 3 and Pair 4. All regressions include job fixed effects. Standard errors in parentheses are clustered at the job level. χ^2 statistics and joint $p-$values come from tests that all differences in reported coefficients other than the constant term are zero.

Table F.9: Race-Specific Contact History Dummy

	(1)	(2)	(3)	(4)	(5)	(6)
	History : $\mathbf{1}\left(NoC^{up} > 0\right)$			History : $\mathbf{1}\left(NoC^{low} > 0\right)$		
Black	-0.0189	-0.0175	-0.0179	-0.0193	-0.0183	-0.0185
	(0.0019)	(0.0018)	(0.0018)	(0.0018)	(0.0018)	(0.0018)
History$^{\text{Black}}$	-0.1295		-0.0689	-0.1803		-0.1054
	(0.0094)		(0.0117)	(0.0102)		(0.0144)
History$^{\text{Black}}$ $\times$ Black	-0.0386		-0.0449	-0.0230		-0.0120
	(0.0073)		(0.0140)	(0.0082)		(0.0144)
History$^{\text{White}}$		-0.1521	-0.1183		-0.1777	-0.1086
		(0.0094)	(0.0110)		(0.0098)	(0.0138)
History$^{\text{White}}$ $\times$ Black		-0.0012	0.0222		-0.0189	-0.0121
		(0.0072)	(0.0137)		(0.0080)	(0.0140)
Controls $\mathbf{X}_i$	✓	✓	✓	✓	✓	✓
Job FEs	✓	✓	✓	✓	✓	✓
Order FEs	✓	✓	✓	✓	✓	✓
N	83643	83643	83643	83643	83643	83643

Note: This table presents the effects of race-specific prior contact histories, controlling for resume characteristics on the probability of employer contact within 30 days. In Column (1) - (3), History refers to $\mathbf{1}\left(NoC_i^{up} > 0\right)$, the indicator of positive number of submissions already contacted no later than the day when a submission was made, by race (Black/White); In Column (4) - (6), History refers to $\mathbf{1}\left(NoC_i^{low} > 0\right)$, the indicator of positive number of submissions already contacted at one day before a submission was made, by race (Black/White). All regressions include submission order fixed effects and job fixed effects. Standard errors in parentheses are clustered at the job level.

Table F.10: Gender-Specific Contact History Dummy

	(1)	(2)	(3)	(4)	(5)	(6)
	History : $\mathbf{1}\left(NoC^{up} > 0\right)$			History : $\mathbf{1}\left(NoC^{low} > 0\right)$		
Female	-0.0058	0.0005	-0.0031	-0.0056	-0.0003	-0.0035
	(0.0022)	(0.0022)	(0.0022)	(0.0022)	(0.0022)	(0.0022)
History$^{\text{Female}}$	-0.1515		-0.1318	-0.1941		-0.1457
	(0.0102)		(0.0121)	(0.0109)		(0.0131)
History$^{\text{Female}} \times$ Female	0.0094		0.0324	0.0127		0.0197
	(0.0092)		(0.0129)	(0.0106)		(0.0138)
History$^{\text{Male}}$		-0.1205	-0.0572		-0.1685	-0.0887
		(0.0105)	(0.0122)		(0.0111)	(0.0134)
History$^{\text{Male}} \times$ Female		-0.0062	-0.0323		-0.0023	-0.0137
		(0.0091)	(0.0127)		(0.0104)	(0.0136)
Controls $\mathbf{X}_i$	✓	✓	✓	✓	✓	✓
Job FEs	✓	✓	✓	✓	✓	✓
Order FEs	✓	✓	✓	✓	✓	✓
N	83643	83643	83643	83643	83643	83643

Note: This table presents the effects of gender-specific prior contact histories, controlling for resume characteristics on the probability of employer contact within 30 days. In Column (1) - (3), History refers to $\mathbf{1}\left(NoC_i^{up} > 0\right)$, the indicator of positive number of submissions already contacted no later than the day when a submission was made, by gender (female/male); In Column (4) - (6), History refers to $\mathbf{1}\left(NoC_i^{low} > 0\right)$, the indicator of positive number of submissions already contacted at one day before a submission was made, by gender (female/male). All regressions include submission order fixed effects and job fixed effects. Standard errors in parentheses are clustered at the job level.

Table F.11: Group-Specific Contact History, Alternative Measure

	(1)	(2)	(3)	(4)	(5)	(6)
	History : NoC^{alt}			History : $\mathbf{1}\left(NoC^{alt} > 0\right)$		
Panel A: Race-specific contact history						
Black	-0.0205	-0.0124	-0.0153	-0.0202	-0.0085	-0.0125
	(0.0018)	(0.0016)	(0.0016)	(0.0018)	(0.0016)	(0.0015)
History$^{\text{Black}}$	-0.0592		-0.0366	-0.1160		-0.0909
	(0.0027)		(0.0053)	(0.0073)		(0.0096)
History$^{\text{Black}}$ × Black	-0.0202		0.0048	-0.0256		0.0055
	(0.0026)		(0.0066)	(0.0055)		(0.0101)
History$^{\text{White}}$		-0.0721	-0.0414		-0.1304	-0.0922
		(0.0032)	(0.0061)		(0.0069)	(0.0086)
History$^{\text{White}}$ × Black		0.0018	-0.0120		-0.0215	-0.0322
		(0.0026)	(0.0065)		(0.0054)	(0.0100)
Panel B: Gender-specific contact history						
Female	-0.0119	0.0071	-0.0028	-0.0115	0.0059	-0.0036
	(0.0021)	(0.0021)	(0.0020)	(0.0021)	(0.0021)	(0.0019)
History$^{\text{Female}}$	-0.0618		-0.0517	-0.1386		-0.1343
	(0.0028)		(0.0035)	(0.0075)		(0.0087)
History$^{\text{Female}}$ × Female	0.0060		0.0172	0.0229		0.0561
	(0.0026)		(0.0034)	(0.0066)		(0.0092)
History$^{\text{Male}}$		-0.0537	-0.0307		-0.0998	-0.0562
		(0.0030)	(0.0036)		(0.0074)	(0.0083)
History$^{\text{Male}}$ × Female		-0.0051	-0.0166		-0.0207	-0.0545
		(0.0028)	(0.0036)		(0.0066)	(0.0092)
Controls $\mathbf{X}_i$	✓	✓	✓	✓	✓	✓
Job FEs	✓	✓	✓	✓	✓	✓
Order FEs	✓	✓	✓	✓	✓	✓
N	83643	83643	83643	83643	83643	83643

Note: This table presents the effects of group-specific prior contact histories with the alternative measure based on submission order, controlling for resume characteristics on the probability of employer contact within 30 days. In Column (1) - (3), History refers to NoC^{alt}, the number of applications submitted prior to a submission was made, by race (Black/White) or gender (female/male); In Column (4) - (6), History refers to $\mathbf{1}\left(NoC_i^{alt} > 0\right)$, the indicator of positive number of applications submitted prior to a submission was made, by race (Black/White) or gender (female/male). All regressions include submission order fixed effects and job fixed effects. Standard errors in parentheses are clustered at the job level.

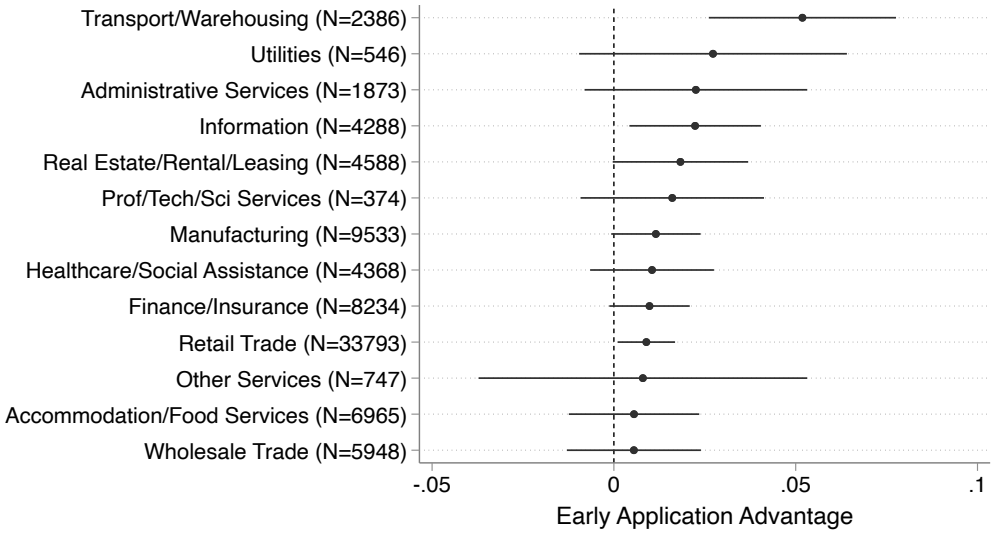
G Additional Heterogeneity Results

In this section, I show additional heterogeneity of the early application advantage. I use the early submission indicator $early_i^1$ as the main regressor for statistical power.

G.1 Industry Heterogeneity

First, I estimate the early application advantage for each of the thirteen industries and plot the estimated early application advantage $\hat{\beta}$ in Figure G.1. There is significant heterogeneity across industries with respect to the early application advantage. In the transport/warehousing industry, an earlier submission experiences a 5 percentage point higher contact rate. The early application advantage is smallest in the wholesale trade and accommodation/food service industries, though still positive, it is not statistically significant. In the industry of retail trade, where 40% of the fictitious applications were made, there is a significant early application advantage of about 1 percentage point.

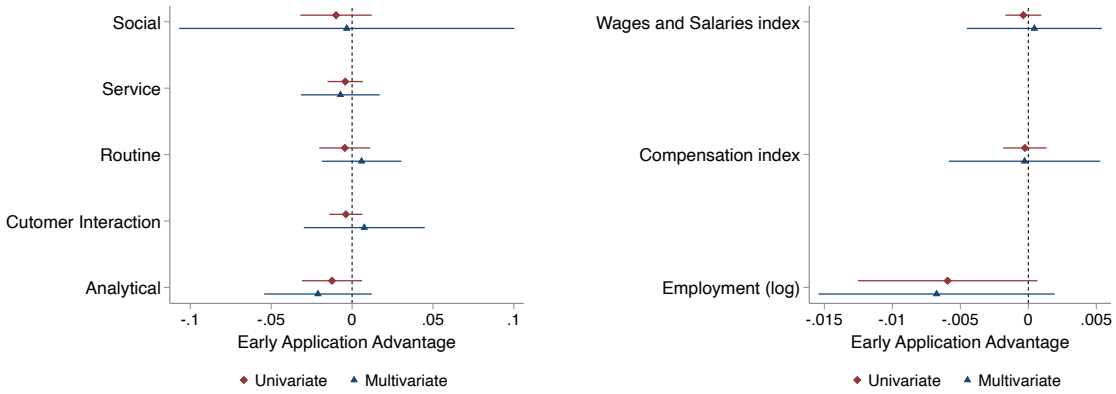
Figure G.1: Early Application Advantage by Industry



Notes: The figure plots the estimated coefficients $\hat{\beta}$ of Model (3.1) for each industry, with the early indicator $early^1$ which equals 1 for the first four applications to each job, controlling for resume characteristics. The regressions include job fixed effects, the standard errors are clustered at the job level. The line depicts the 95% confidence intervals. The nubmer of applications in each industry is reported in parentheses.

In the left panel of Figure G.2, I plot the estimated coefficients of interest $\hat{\nu}_m$ and $\hat{\psi}_m$ for the five task-based industry characteristics: social, service, routine, customer interaction, analytical. Both univariate and multivariate estimates show that industries requiring more analytical skills have a smaller early application advantage, pointing to the potential effect of screening cost k in the model.

Figure G.2: Early Application Advantage and Industry-Level Covariates



(a) Task-Based Characteristics (b) Employment Statistics

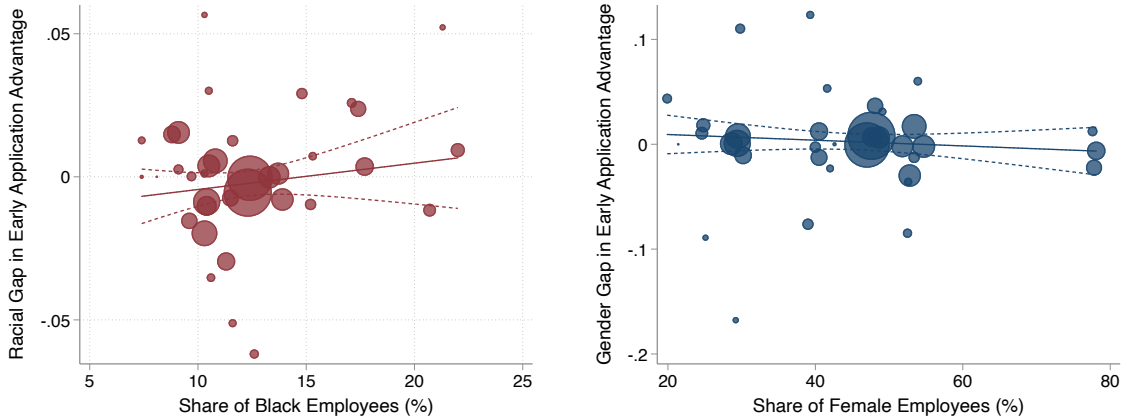
Notes: The figures plot how industry level covariates affect the early application advantage. The left panel plots the estimated effects of task-based characteristics; The right panel plots the estimated effects of industry-level employment statistics. "Univariate" refers to estimates of Model (8.1); "Multivariate" refers to estimates of Model (8.2). "3-Month Average" and "6-month Average" refers to the the two choices of time windows to calculate the average employment statistics, only covariates calculated with the same time-window choice enter the same multivariate regression. All regressions include job fixed effects and resume characteristics, the standard errors are clustered at the job level. The lines depict the 95% confidence intervals of the estimates. The construction of covariates is described in Section E.1.

In the right panel of Figure G.2, I estimate the effects of three covariates to capture the pre-experiment industry-level labor market conditions: the Wages and Salaries index, the Total Compensation index from the Employment Cost Index (ECI), and the seasonally adjusted monthly employment size prior to the month of the first submission to a job in the experiment (see Section E.1 for more details). The results show that the employment size of industries is negatively associated with the size of the early application advantage (also shown in Figure 13).

Figure G.2b suggests that application pool size is related to screening urgency. Using annual data on the demographic composition of employment for each industry sourced from the Current Population Survey (see Section E.1), I test whether similar effects can be observed for the racial and gender gap heterogeneity For each industry-year unit, I estimate the racial and gender gap in the early application ad-

vantage and plot the estimates against the share of Black and female employees in the industry the year prior to the first submission to a job in the experiment. Figure G.3a shows that in industries with a higher share of Black employees, Black applicants experience a larger early application advantage compared to their White counterparts; On the contrary, Figure G.3b shows a negligible negative correlation between the share of female employees and the gender gap in the early application advantage.

Figure G.3: Early Application Advantage and Industry Demographics



(a) Racial Gap (b) Gender Gap

Notes: The figures plot the gender and race gap in the early application advantage against the share of Black/female employees in an industry in the year prior to the year of first submission to a job in the experiment. In both panels, each dot represents an industry-year unit, weighted by the number of submissions in that unit during the experiment. The racial/gender gap is estimated as in Column (1) and (2) of Table A.4 for each industry-year unit. All regressions include job fixed effects and resume characteristics.

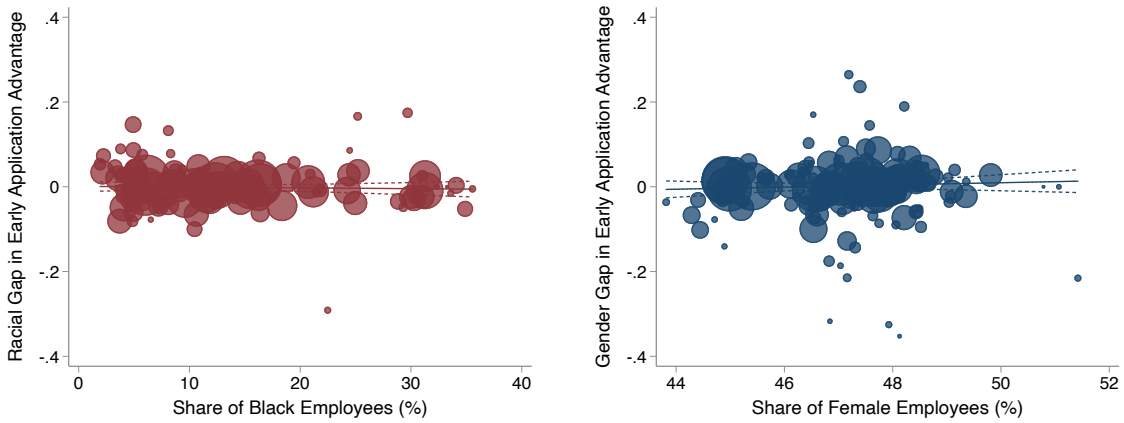
G.2 State Heterogeneity

Similarly to Figure G.1, I estimate Model (3.1) for each state and plot the estimated early application advantage $\hat{\beta}$ in Figure G.9a. A significant early application advantage is observed in most states, with only 12 states out of the 51 in the sample showing zero or negative advantage. Figure G.9b plots the estimated coefficients of the interaction terms between the early submission indicator and the indicators for Black and female applicants.

In Figure G.4, I repeat the exercise in Figure G.3 with Local Area Unemployment Statistics (LAUS) annual data of state-level demographic composition of the

employment and show a similar null effect of the shares of Black/female employees on the racial/gender gaps in the early application advantage.

Figure G.4: Early Application Advantage and State Demographics



(a) Racial Gap (b) Gender Gap

Notes: The figures plot the gender and race gap in the early application advantage against the share of Black/female employees in a state in the year prior to the year of first submission to a job in the experiment. In both panels, each dot represents a state-year unit, weighted by the number of submissions in that unit during the experiment. The racial/gender gap is estimated as in Column (1) and (2) of Table A.4 for each state-year unit. All regressions include job fixed effects and resume characteristics.

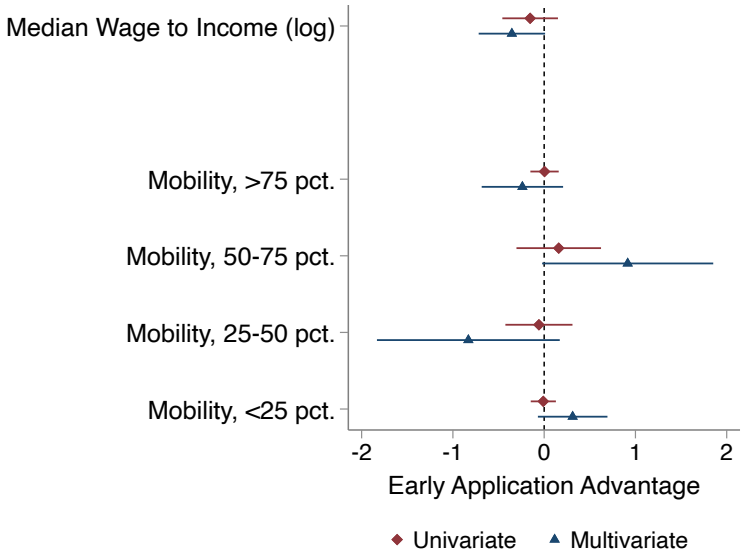
Next, I examine the explanatory power of three sets of covariates on the observed state-level heterogeneity. The construction of these covariates is detailed in Section E.2.

First, using the newly available granular Income Distributions and Dynamics in America (IDDA) data (Kondo et al., 2023), I construct the pre-experiment median income and wage, as well as the mobility of four income bins, $\{< 25\text{pct.}, 25-50\text{pct.}, 50-75\text{pct.}, > 75\text{pct.}\}$ for each state by averaging 2015-2019 annual data. I then use these constructed covariates to estimate $\hat{\nu}_m$ in Model (8.1) and $\hat{\psi}_m$ in Model (8.2). As shown in Figure 13, in states where wages constitute a larger component of income for median households, recruiters tend to wait and compile more applications rather than moving forward quickly to contact applicants.

Figure G.5 shows that the mobility of the four income bins does not have a significant impact on the early application advantage individually. In the multivariate specification, I observe a positive correlation between the early application advantage and the mobility of both the $50-75$pct. and the <25pct. bins However, none of these results are statistically significant potentially due to the small

number of observation units (51 states).[50]

Figure G.5: Early Application Advantage and State Income/Mobility Conditions



Notes: The figure plots how state pre-experiment income and mobility conditions affect the early application advantage. "Univariate" refers to estimates of Model (8.1); "Multivariate" refers to estimates of Model (8.2). The two income measures and the four mobility measures are estimated separately as two groups of covariates in two multivariate regressions. All regressions include job fixed effects and resume characteristics, with standard errors clustered at the job level. The line depicts the 95% confidence intervals.

The next covariate, the minimum wage in the state where a sampled job was located in the year when the first application for the job was submitted, reflects the impact of state-level labor protection on the early application advantage. I group jobs by states and the year of the first submission, and estimate Model (3.1) for each state-year group. The estimated coefficients are then plotted against the corresponding minimum wage in Figure G.6. The results show no correlation between the early application advantage and state minimum wages in the context of the experiment. [51] However, these results do not rule out the impact of minimum

[50] As summarized in Table E.1, the mobility of the $50-75$pct. and the >75pct. bins is negative (*downward mobility*), while the mobility of the <25pct. and the $25-50$pct. bins is positive (*upward mobility*). Thus, one way to interpret the results is: for states where richest (>75pct.) and the poorest (<25pct.) groups are more likely to converge to the middle, the early application advantage is larger; for states where the middle groups ($50-75$pct. and $25-50$pct.) are more likely to converge to the middle, the early application advantage is smaller.

[51] In unreported results, I estimate Model (8.1) with the minimum wages as the covariate and obtain a robust null result.

wages on the application timing effects. In the experiment, all firms were sampled specifically for having branches in all states. They are competitive employers in the labor market, typically offering above-minimum compensations. For firms that are actually bounded by minimum wage regulations, the results could be very different.

Figure G.6: Early Application Advantage and State Minimum Wages



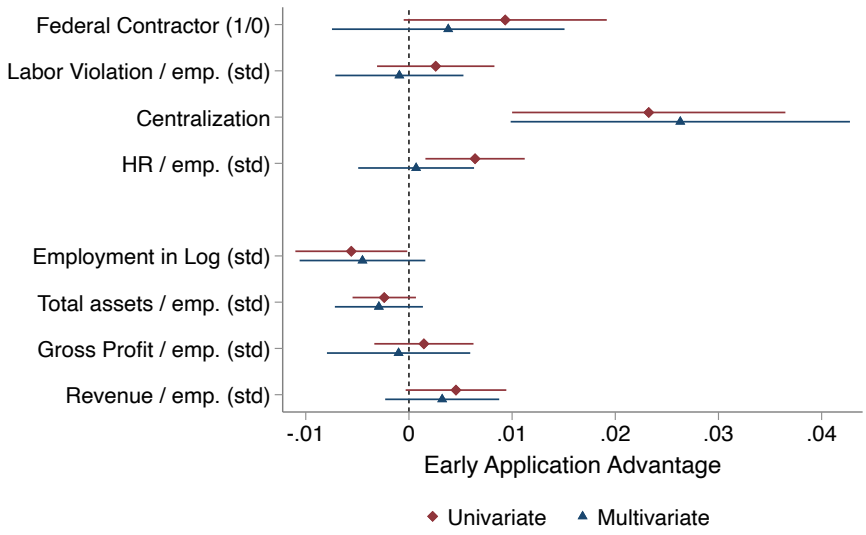
Notes: The figure plots how state minimum wages affect the early application advantage. Each dot represents the estimated early application advantage for a state-year group. The size of the dots represents the number of submissions in the corresponding state-year group. The dots in grey represent states without minimum wages. All regressions include job fixed effects and resume characteristics, with standard errors clustered at the job level.

G.3 Firm Heterogeneity

Lastly, I focus on firm-level heterogeneity in the early application advantage. For each of the sampled firms, I collect a set of pre-experiment covariates on human resource management practices, employee satisfaction, and financial performance (see Section E.3 for detailed descriptions) and plot both univariate and multivariate estimates in Figure G.7.

In addition to the results shown in Figure 13, Figure G.7 shows that the number of labor law violations per employee, the size of the human resource management

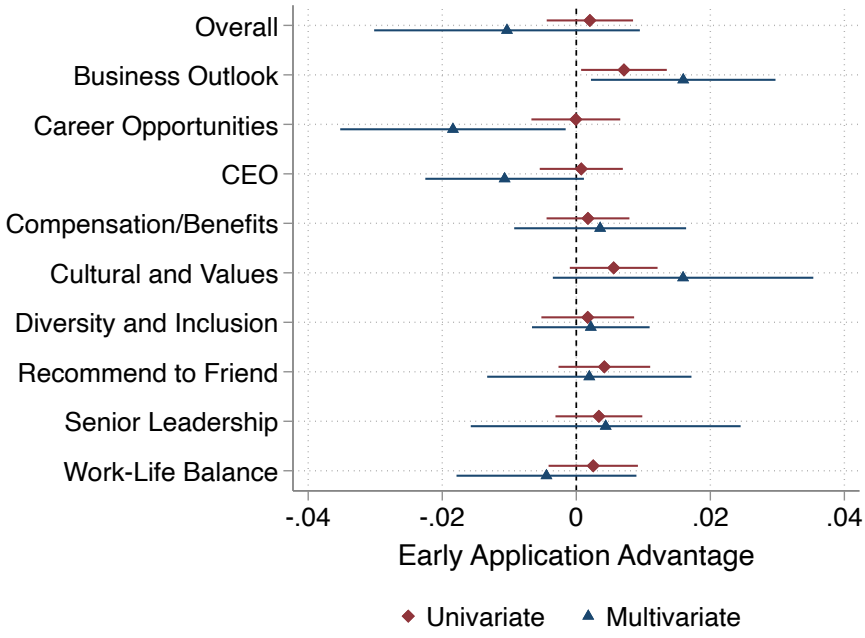
Figure G.7: Early Application Advantage and Firm characteristics



Notes: The figure plots how pre-experiment firm characteristics affect the early application advantage. "Univariate" refers to estimates of Model (8.1); "Multivariate" refers to estimates of Model (8.2). The definition of the covariates is described in Section E.3. All covariates are normalized to have a standard deviation of 1. All regressions include job fixed effects and resume characteristics, with standard errors clustered at the job level. The line depicts the 95% confidence intervals.

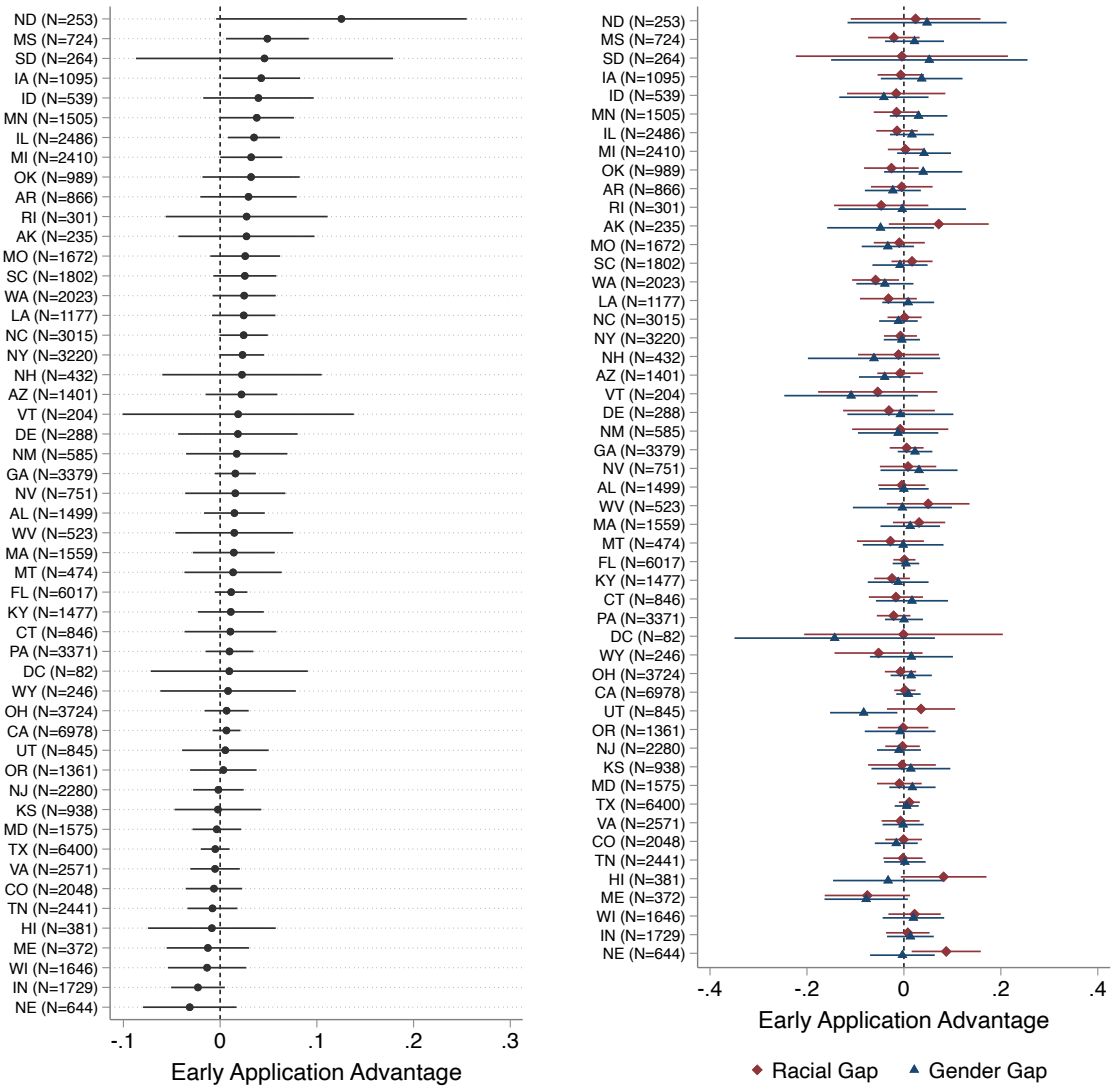
department relative to the total employment of firms, and the gross profit per employee are not significant correlated with early application advantages. Figure G.8 shows that employees' ratings on various aspects of firm management do not explain the firm-level heterogeneity of the early application advantage.

Figure G.8: Early Application Advantage and Firm characteristics



Notes: The figure plots how pre-experiment firm GlassDoor ratings affect the early application advantage. "Univariate" refers to estimates of Model (8.1); "Multivariate" refers to estimates of Model (8.2). The definition of the covariates is described in Section E.3. All covariates are normalized to have a standard deviation of 1. All regressions include job fixed effects and resume characteristics, with standard errors clustered at the job level. The line depicts the 95% confidence intervals.

Figure G.9: Early Application Advantage by State



(a) State-Level Heterogeneity (b) Gender/Racial Gap Heterogeneity

Notes: The figures plot the geographic heterogeneity of the early application advantage at the state level. The left panel plots the estimated coefficients $\hat{\beta}$ of Model (3.1) with the early indicator $early^1$ which equals 1 for the first four applications to each job; The right panel plots the estimated coefficients of the interaction term of $early^1$ and the Black (in red)/female applicant indicator (in blue) in Model (3.1). All regressions include job fixed effects and resume characteristics, the standard errors are clustered at the job level. The line depicts the 95% confidence intervals. The nubmer of applications in each state is reported in parentheses.