DURATION DEPENDENCE AND LABOR MARKET CONDITIONS: EVIDENCE FROM A FIELD EXPERIMENT*

Kory Kroft
Fabian Lange
Matthew J. Notowidigdo

This article studies the role of employer behavior in generating "negative duration dependence"—the adverse effect of a longer unemployment spell—by sending fictitious résumés to real job postings in 100 U.S. cities. Our results indicate that the likelihood of receiving a callback for an interview significantly decreases with the length of a worker's unemployment spell, with the majority of this decline occurring during the first eight months. We explore how this effect varies with local labor market conditions and find that duration dependence is stronger when the local labor market is tighter. This result is consistent with the prediction of a broad class of screening models in which employers use the unemployment spell length as a signal of unobserved productivity and recognize that this signal is less informative in weak labor markets. *JEL* Code: J64.

I. Introduction

Does the length of time out of work diminish a worker's job market opportunities? This question attracts substantial attention from policy makers and researchers alike, reflecting the widespread belief that the adverse effect of a longer unemployment spell—what economists call "negative duration dependence"—undermines the functioning of the labor market and entails large social costs. Recently, the sharp rise in long-term unemployment has renewed interest in duration dependence; according to a recent report by the Congressional Budget Office (CBO),

* We thank Marianne Bertrand, Eric Budish, Jon Guryan, Yosh Halberstam, Larry Katz (the editor), Rob McMillan, Phil Oreopoulos, Paul Oyer, Yuanyan Wan, four anonymous referees, and many seminar participants for helpful comments. We thank Thomas Bramlage, Rolando Capote, David Hampton, Mark He, Paul Ho, Angela Li, Eric Mackay, Aaron Meyer, Nabeel Thomas, Stephanie Wu, Steven Wu, Vicki Yang, and Dan Zangri for excellent research assistance. We thank Ben Smith for excellent research assistance and exceptional project management throughout the experiment, and we thank Bradley Crocker at HostedNumbers.com for assistance with setting up the local phone numbers used in the experiment. We gratefully acknowledge the Initiative on Global Markets and the Stigler Center at the University of Chicago Booth School of Business, the Neubauer Family Assistant Professorship, and the Connaught Fund for financial support.

The Quarterly Journal of Economics (2013), 1123–1167. doi:10.1093/qje/qjt015.
Advance Access publication on April 16, 2013.

long-term unemployment may "produce a self-perpetuating cycle wherein protracted spells of unemployment heighten employers' reluctance to hire those individuals, which in turn leads to even longer spells of joblessness" (CBO 2012).

Despite this widespread interest, it has proven very difficult to credibly establish that an individual's chance of finding a job worsens with the length of his or her unemployment spell. The difficulty arises in part because workers with different unemployment spell lengths who appear (otherwise) similar to researchers may actually look very different to employers. As a result, in observational data, the job-finding probability might decline with unemployment duration either because of "true" duration dependence or because unemployment spell lengths correlate with other fixed characteristics that are observed by employers but not researchers. The state of the empirical literature is succinctly summarized by Ljungqvist and Sargent (1998, 547), who write: "It is fair to say that the general evidence for duration dependence is mixed and controversial."

In this article, we confront this challenge by estimating duration dependence using a large-scale résumé audit study. We submit fictitious résumés to real, online job postings in each of the 100 largest metropolitan areas (MSAs) in the United States, and we track "callbacks" from employers for each submission. In total, we applied to roughly 3,000 job postings in sales, customer service, administrative support, and clerical job categories, and we submitted roughly 12,000 résumés. In designing each résumé, we explicitly randomize both the employment status and the length of the current unemployment spell from 1 to 36 months (if the worker is unemployed).[1] The advantages of this experimental design are twofold. First, we observe the same information as employers at the time that employers make callback decisions. Second, the unemployment spell length is orthogonal to labor market conditions and all of the other characteristics on résumés that are observable by potential employers.

1. To be precise, from the employer's perspective, a "gap" in work experience on a résumé technically corresponds to a period of nonemployment. Nevertheless, we refer to this gap as an unemployment spell, although we recognize that this is not the conventional definition. Our view is that the current gap represents the best available information to an employer about a job seeker's current job market status; in particular, whether he or she is currently unemployed, and what that signals about his or her productivity.

Our experiment identifies duration dependence in callback rates that operates through employers' beliefs about the unobservable quality of unemployed workers. In interpreting our results, we emphasize that "true" duration dependence (i.e., the genuine causal effect of a longer unemployment duration on an individual's job-finding rate) might in fact arise precisely because there is heterogeneity in the applicant pool that is unobservable to employers. Intuitively, "true" duration dependence in callback rates may arise as a result of optimizing behavior of firms that are dealing with such unobserved heterogeneity. This calls into question the standard practice of trying to separately identify state dependence and unobserved heterogeneity, because these two sources of duration dependence in job-finding rates interact in equilibrium.[2]

Turning to the experimental results, a simple plot of the raw data displays clear visual evidence of negative duration dependence: the average callback rate sharply declines during the first eight months of unemployment and then it stabilizes. Ordinary least squares (OLS) regression results confirm the pattern from the nonparametric plot. At eight months of unemployment, callbacks are about 45% lower than at one month of unemployment, as the callback rate falls from roughly 7% to 4% over this range. After eight months of unemployment, we find that the marginal effect of additional months of unemployment is negligible. To benchmark the magnitude of this result, in their study of racial discrimination, Bertrand and Mullainathan (2004) find that black-sounding names received about 33% fewer callbacks than white-sounding names.

We next estimate how duration dependence varies with labor market conditions by exploiting cross-MSA variation. Our results indicate that duration dependence is significantly stronger when the local labor market is tight. This finding is robust across several different measures of market tightness: first, metropolitan area unemployment rates; second, metropolitan area vacancy-unemployment ratios; finally, the callback rate for a newly

2. Heckman and Singer (1984) consider the econometric problem of distinguishing state dependence from unobserved heterogeneity. Without functional form assumptions on job-finding rates, they show it is not possible to distinguish between duration dependence and unobserved heterogeneity using observational data with a single unemployment spell for each individual. Multiple-spell data can resolve this identification problem, but at the cost of strong assumptions on how job-finding rates vary across unemployment spells for a given individual.

unemployed individual in each MSA, estimated within the experiment. This result is consistent with the prediction of a broad class of screening models in which employers use the length of the unemployment spell as a signal of unobserved productivity and recognize that this signal is less informative in weak labor markets. Models of duration dependence based on skill depreciation cannot readily account for this.

Most closely related to our work is Oberholzer-Gee (2008) and Eriksson and Rooth (2011), who also investigate how employers respond to unemployment spells using a résumé audit study. Oberholzer-Gee (2008) analyzes Swiss employer responses to 628 résumé submissions. Eriksson and Rooth (2011) submit 8,466 job applications to 3,786 employers in Sweden and compare the effects of contemporary and past unemployment spells (e.g., unemployment spells at graduation). Both of these studies report results consistent with the long-term unemployed being less likely to receive callbacks. Oberholzer-Gee finds that a person out of work for two and a half years is almost 50 percentage points less likely to be invited for an interview than an employed job seeker. Eriksson and Rooth find that for low- and medium-skill jobs, being out of work for nine months or longer significantly reduces callbacks.

Our study builds on these papers in two ways. First, unlike these studies, which randomize across a small number of unemployment spell lengths, our study estimates a callback rate for each month in the interval [1,36]. This allows us to flexibly estimate the relationship between callback rate and unemployment duration, which we find is highly nonlinear. Second, neither of these papers considers how duration dependence varies with market tightness, which is a key feature of our experiment.

To allay concerns about external validity, we strove to make our fictitious résumés appear similar to real résumés we collected from various online job boards. However, it is possible that employers do not attend to information about the unemployment spell length when making callback decisions. To address this concern (both in our work and in the related papers), we administered a web-based survey to MBA students. Our survey results indicate that the length of the current unemployment spell is salient to the survey participants; in particular, subjects were able to recall a worker's employment status and unemployment spell length with roughly the same degree of accuracy and precision as other résumé characteristics, such as education and job experience. This supports our assumption that unemployment duration on résumés is salient to employers, both in our experiment and in the labor market more broadly.

The remainder of the article proceeds as follows. Section II provides some of the basic facts on the extent of negative duration dependence in the escape rate from unemployment in the United States, summarizes the basic theories for such a relationship from unobserved heterogeneity to different models of true duration dependence, and discusses what can be learned from our experimental estimates. Section III describes the experimental design, the results from the web-based survey, and the empirical models. Section IV describes our experimental results. Section V discusses alternative theoretical interpretations for our results. Section VI concludes.

II. Motivation

It is well known that the short-term unemployed find jobs at a faster rate than the long-term unemployed. For instance, Shimer (2008) shows that the job-finding rate declines across the first 12 months of an unemployment spell, using pooled data from the Current Population Survey (CPS) for 1976–2007. We extend Shimer's analysis to the period January 2008 through December 2011.[3] Because the share of long-term unemployed workers was unusually high during this time period, we are able to calculate the monthly job-finding rate for the first 24 months of an unemployment spell and can thus investigate how exit rates vary with duration for unemployment durations that Shimer could not study with any precision. Our findings, reported in Figure I, show that the job-finding rate falls sharply with the length of the unemployment spell, particularly during the first few months. Beyond one year of unemployment, however, we find a much weaker relationship between the job-finding rate and unemployment duration.

Although there is broad agreement that the escape rate from unemployment declines with duration, there is less agreement

3. The job-finding rates are lower in our sample compared to Shimer's sample. This is due to business cycle conditions and also to a slightly different methodology that we adopt. Each of these sources accounts for roughly half of the total difference. See the Online Appendix for details on the construction of job-finding rates.

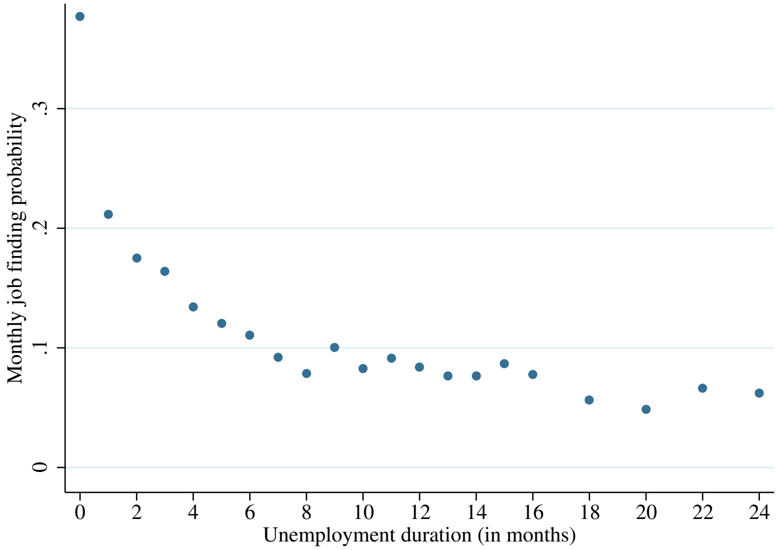

FIGURE I

Job Finding Probability and Unemployment Duration in the United States, 2008–2011

This figure reports results using pooled monthly CPS data between January 2008 and December 2011. We match observations across months, and we measure job finding as an individual exiting from unemployment and reporting being employed in each of the subsequent two months (i.e., a U-E-E spell following Rothstein 2011). See the Online Appendix for more details of the matching procedure. The CPS asks respondents for unemployment duration in weeks, and we convert weeks into months and report average job-finding probabilities by month. To preserve sample size, we merge months into two-month groups after 16 months.

about whether this represents a causal relationship or merely a correlation, driven by the composition of the unemployed pool. Machin and Manning (1999) review the large literature of Europe-based empirical studies and find little evidence of duration dependence, once one controls for observable fixed characteristics. This contrasts with the findings of Imbens and Lynch (2006), who use a sample of 5,000 young men and women from the National Longitudinal Survey Youth Cohort 1979 (NLSY79) for the years 1978–89. Controlling for a rich set of individual characteristics, they find evidence of negative duration dependence in job-finding rates. The literature also differs in its conclusions on how duration dependence varies with labor market conditions.[4] Imbens and Lynch, for instance, find that duration dependence is stronger when (local) labor markets are tight, and their findings are consistent with those reported by Sider (1985) and van den Berg and van Ours (1996). By contrast, Dynarksi and Sheffrin (1990) find that duration dependence is weaker when markets are tight. Still others find that the interaction effect between market tightness and unemployment duration varies over the length of the spell. For instance, it may be positive for some unemployment durations and negative for others (Butler and McDonald 1986; Abbring, van den Berg, and van Ours 2001).[5]

Observational studies of the type described here face the challenge of separating unobserved heterogeneity from "true" duration dependence. Individual differences in job-finding rates that are not observed by researchers will lead to declining job-finding rates in the population, even if individual job-finding rates themselves do not decline with duration. Intuitively, as durations lengthen, the pool of unemployed individuals increasingly shifts to those with permanently low job-finding rates. This dynamic selection effect can potentially explain the pattern of results we document in Figure I.

Several theories predict that duration has a causal effect on job-finding rates. First, employer screening models (Vishwanath 1989; Lockwood 1991) emphasize unobserved worker heterogeneity and sorting. When firms match with applicants, they receive a private signal about unobserved productivity and base their hiring decision on this signal. In equilibrium, firms expect that unemployment duration is negatively correlated with unobserved productivity, since a longer spell reveals that prior firms learned the worker was unproductive. On average, the long-term

4. The relationship between duration dependence and market tightness may be driven by the cyclical variation in the skill composition of unemployed workers (Darby, Haltiwanger, and Plant 1985). The literature has typically found little evidence for this "heterogeneity hypothesis" (Abbring, van den Berg, and van Ours 2001).

5. The contradictory findings in the literature result in part from different functional form restrictions imposed on job-finding rates. Some studies restrict the sign of the cross-derivative of tightness and duration to be constant over the unemployment spell, and others allow this sign to vary. As we argue in the Online Appendix, the cross-derivative is of limited use in distinguishing among alternative models, because it typically varies over the spell. However, these models can be tested against each other by examining how relative hiring rates vary with tightness over the spell of unemployment (comparing job-finding rates at all positive durations with the job-finding rate of the newly unemployed).

unemployed therefore have lower exit rates than the short-term unemployed. An implication of screening models is that the gap in exit rates shrinks in slack labor markets. Intuitively, workers match less often with firms in slack markets; thus, spell length is less indicative of the unobservable characteristics of workers than it is of aggregate labor market conditions.

Second, human capital models (Acemoglu 1995; Ljungqvist and Sargent 1998) focus on how a single worker's skills depreciate over their unemployment spell. In the simplest model, skill depreciation does not depend on aggregate labor market conditions. In this case, the model also generates negative duration dependence, but it is constant across the business cycle.

Third, ranking models (Blanchard and Diamond 1994; Moscarini 1997) emphasize the consequences of crowding in the labor market; in these models, vacancies potentially receive multiple applications. These models assume that if a firm meets multiple workers, it hires the worker with the shortest spell. This immediately implies that there is negative duration dependence. In addition, in tight markets, applicants for a given position are less likely to face competition from applicants with shorter durations. Therefore, under employer ranking, duration dependence is weaker in tight labor markets.

Finally, there are models of duration dependence that emphasize changes in search behavior. Workers may become discouraged over time and reduce their search intensity or they may have fewer vacancies to apply to, as in stock-flow search models (Coles and Smith 1998).

Our experiment sheds light on theories of duration dependence that emphasize employer behavior. In particular, it identifies the causal effect of duration on callbacks that arises either through employer ranking or employers' beliefs about worker quality among the population of unemployed job seekers. In turn, employers may form negative beliefs about the long-term unemployed for two reasons. First, on average, these applicants may be of lower (unobservable) quality, as would arise in a standard screening model. In this case, our experiment measures duration dependence coming from firms' beliefs that unemployment duration negatively correlates with the fixed worker characteristics that are not observable on résumés.[6] A second possibility is

6. According to this interpretation, unobserved heterogeneity indirectly causes "true" duration dependence. As such, it may not be meaningful to

that firms believe that worker skills depreciate so that the long-term unemployed are less productive. In both cases, our estimates will capture firms' attempt to screen workers on the basis of their unemployment duration. Our experiment also explores how duration dependence varies with local labor market conditions. As we discuss in more detail in Section V, this interaction is potentially useful in distinguishing among alternative theories.

Why does our experiment allow us to identify duration dependence in callbacks arising from employers' beliefs about worker unobservables? This is ensured by the random assignment of unemployment duration, which implies that duration is (by construction) orthogonal to all of the observable characteristics on the résumés. Therefore, the correlation between duration and callbacks captures firms' beliefs about the unobserved quality of the unemployment pool. If instead our analysis proceeded using real résumés combined with observational data on callbacks, then our estimation would be significantly more challenging. This is because of the inherent difficulty in controlling for all of the relevant observable information on résumés that employers use to make decisions on callbacks. If any observable characteristic that employers use to make callback decisions is omitted from the analysis, then the correlation between callbacks and duration may in part reflect a correlation between observables and unobservables. We believe that this is a very real possibility given that résumés are highly multidimensional. For example, consider the case where longer term unemployed workers have systematically "lower quality" work experience. If this is difficult to measure and quantify, then an econometrician may find that longer spells are associated with fewer callbacks, but this negative correlation would be due partly to the fact that résumés with longer spells have lower quality work experience, rather than due to employers' beliefs about worker unobservables.

We conclude with two limitations of our study. First, we cannot measure worker behavior. Worker behavior may contribute to negative duration dependence for two reasons. First, workers may become discouraged and expend less effort in job search over time, independent of firm behavior. Second, the return to search

---

distinguish between them. In the Online Appendix Section A, we develop a "mechanical model" of duration dependence and provide intuition for this interaction between unobserved heterogeneity and true duration dependence.

may fall over time if firms discriminate against the long-term unemployed, leading to declining search effort. Due to this limitation, we believe our study is complementary to recent work on job search behavior (Krueger and Mueller 2010). Second, we cannot observe employer hiring decisions, so our experiment can only shed light on negative duration dependence in callback rates. On the other hand, we note that it is much more straightforward to estimate duration dependence in callback rates. To estimate duration dependence in job-finding rates, an econometrician would need to be able to condition on the information that potential employers see at the hiring stage, in addition to the interview stage, and such information may be particularly hard to quantify.

III. Experimental Design

The design of the field experiment follows Bertrand and Mullainathan (2004), Lahey (2008), and Oreopoulos (2011) in how we generate fictitious résumés, find job postings, and measure callback rates. All of the experimental protocols (as well as the web-based survey for MBA students) were reviewed and approved by the Institutional Review Board (IRB) at the University of Chicago. The IRB placed several constraints on the field experiment.[7] First, none of the researchers involved in the study could contact the firms at any time, either during or after the experiment. Second, to ensure that the individual representatives of the prospective employers could never be identified, we were required to delete any emails or voice messages that we received from employers after ascertaining the information from the message needed for the experiment. Finally, we were not able to preserve any identifying information about the prospective employers other than the industry. By contrast, we were approved to preserve richer information on the characteristics of the job posting, such as the posted wage and required experience.

The setting for our experiment is a single major online job board in the United States. This online job board contains jobs advertised across most cities in the nation, allowing us to implement our experiment in a large set of local labor markets. Following earlier audit studies, we focus on three job categories:

7. The web-based survey instrument described herein was approved with no additional constraints.

administrative/clerical, customer service, and sales. Within these job categories, we sent roughly 12,000 fictitious résumés to roughly 3,000 job openings located in the largest 100 MSAs in the United States according to population (as measured in the 2010 census). We submitted the résumés between August 2011 and July 2012. The distribution of the jobs across the MSAs was fixed prior to the experiment and primarily reflected the population distribution across MSAs.[8] For example, we planned on submitting résumés to roughly 200 jobs to the MSA New York–Northern New Jersey–Long Island, NY–NJ–PA and roughly 15 jobs to the MSA Raleigh–Cary, NC. However, we also chose to oversample the bottom 10 and top 10 MSAs (within the set of 100) based on the unemployment rate in July 2011.[9] Within each MSA, 30% of jobs were allocated to administrative/clerical, 30% to customer service, and 40% to sales.

In choosing a job to apply to, we began by randomly sampling without replacement from the distribution of MSA and job category combinations. On being assigned an MSA and job category, we had a research assistant (RA) visit the online job board and search for jobs within the predetermined MSA for the predetermined job type. The online job board used in the experiment lists job postings by city rather than MSA, so we searched for appropriate jobs within 25 miles of the major city within the MSA. When picking jobs to apply to, we imposed several restrictions. First, we avoided independent outside sales positions (e.g., door-to-door sales). Second, we did not pick jobs that required advanced skill sets, licenses, or advanced degrees (beyond a standard four-year college degree). Typically, a job opening within a given category and MSA that satisfies these criteria was immediately available, or (in rare cases) became available within one or two weeks.

8. Our initial motivation for sampling based on population size was to achieve a nationally representative sample of job postings. As the experiment proceeded, however, we discovered a practical benefit of this decision: we found it easier to find suitable jobs for the experiment in larger cities.

9. We designed the experiment this way to help identify the interaction between market tightness and duration dependence. The 20 oversampled MSAs were the following: (high-unemployment MSAs) Miami, FL; Detroit, MI; Riverside, CA; Sacramento, CA; Las Vegas, NV; Fresno, CA; Bakersfield, CA; McAllen, TX; Stockton, CA; Modesto, CA; (low-unemployment MSAs) Washington, DC; Boston, MA; Minneapolis, MN; Oklahoma City, OK; Honolulu, HI; Tulsa, OK; Omaha, NE; Des Moines, IA; Madison, WI; Lancaster, PA.

Once a job was identified, the next step was to construct four fictitious résumés that we would customize and email to this job opening from within the online job board website; we never emailed any of the employers directly. The design of these résumés was based on roughly 1,200 real résumés that we manually collected from various online job boards. These résumés were selected based on the job categories we focused on—individuals applying to administrative/clerical, customer service, and sales positions. These résumés informed the design of our fictitious résumés in several ways. First, we found that workers do not "shroud" their unemployment spells: approximately 75% of résumés from workers who were currently unemployed listed both the year *and month* when they last worked. Second, among the currently unemployed, roughly 95% of résumés do not provide any discernible explanation for the gap (e.g., obtained a license or certificate, engaged in community service, worked as a volunteer, training); moreover, this percentage does not vary by gender or by the length of the unemployment spell.[10] Given this, we designed all of our résumés to contain both the year and the month of last employment, and we did not purposefully try to provide any information that could be seen as accounting for the gap in employment.

In total, we created 10 résumé templates that were based on the most frequent résumé formats observed in this database.[11] From this set of templates, we selected four templates (one per fictitious résumé) according the following rule: if an RA applied to a given MSA and job category combination before, she reused the templates from that application, including both the names and email addresses. Otherwise, she randomly drew 4 new templates from the 10 possible templates, drawing without replacement to ensure that no two résumés being sent to a given job share the same template. There are six more steps in designing a fictitious résumé:

(i) We decided whether each résumé would be male or female. For customer service and sales jobs, we sent

10. This finding is surprising in light of Lazear (1984), who argues that workers should attempt to shroud their job-seeking efforts, because then there would be little that could be inferred from the fact that a job was not found quickly.

11. By "template" we mean the specific formatting and layout of the items on the résumé (e.g., style of bullet points, ordering of items, margins on page, spacing).

two female and two male résumés. For administrative/clerical jobs, we sent four female résumés.[12]

(ii) We randomly generated a name for the résumé. The bank of names was generated based on common frequency census data, and the names were chosen to be minimally informative about the race of the applicant.

(iii) We chose the home address, local phone number, and email address. In general, we constructed local addresses based on addresses that were listed on the real résumés in the database of actual résumés described above, and we modified these addresses by choosing a nonexistent street number. We purchased 400 unique local phone numbers (4 per MSA) that could each receive voicemail messages, and we created roughly 1,600 unique email addresses to use in the experiment. Both the phone numbers and email addresses allowed us to track callbacks on an ongoing basis.

(iv) The next step was updating the fictitious résumé's job history, educational history, and the objective summary to match the job we applied to. Work histories were constructed from the sample of real résumés that we self-collected. For instance, if the job was for an administrative assistant position, we identified a résumé with experience as an administrative/executive assistant and used this to construct the work history. For résumés that were sent to jobs in the same MSA, we never shared work histories. In terms of education, we searched for large, local degree-granting institutions. Finally, we verified that there was not a real individual with the same name and with a similar background on any of the major social network and job network websites (e.g., Facebook and LinkedIn).

(v) We defined a measure of "quality" for each résumé. A "low-quality" résumé is one that is assigned the minimum qualifications required for the job (in terms of experience and education). A "high-quality" résumé had qualifications that exceeded these minimum

12. This design decision follows the protocol of Bertrand and Mullainathan (2004), although in hindsight we believe it would have been more appropriate to keep the same gender balance across each job category, as this would have increased our ability to detect gender differences.

requirements. Specifically, these résumés had a couple of extra years of experience and an extra level of education. For instance, if the job requires high school completion, we would list an associate's degree, or if the job requires an associate's degree, we would list a bachelor's degree. For jobs that required a bachelor's degree, we did not increase the education level for the high-quality résumés. For each job that we applied to, two résumés were low quality and two were high quality. This means we either had a set of one high-quality male, one high-quality female, one low-quality male, and one low-quality female résumé, or we had a set of two high-quality female résumés and two low-quality female résumés, depending on the gender ratio the job category calls for.

(vi) The final and most important step was to randomize employment status and the length of the current unemployment spell. We describe the randomization procedure in more detail when we introduce the empirical model. The randomly drawn length of the unemployment spell for a given résumé pins down the end date of the worker's last job, and hence the worker's prior job tenure. In most cases, we designed résumés so that the most recent job started in 2008 or earlier, so that we did not end up dropping a prior job when assigning long unemployment spells.

III.A. Measuring Salience of Résumé Characteristics

Our field experiment assumes that the information on the résumé regarding a job applicant's employment status and unemployment spell length is salient to employers. To test this assumption, we designed and conducted a web-based survey, the details of which are provided in the Online Appendix. Respondents were asked to read a hypothetical job posting and evaluate two résumés for the job opening. Respondents were then asked to recall specific information on the résumé such as total work experience, tenure at last job, level of education, current employment status, and the length of unemployment spell. We used these responses to evaluate the extent to which the various characteristics on the résumé are salient to subjects.

Our findings indicate that respondents are able to recall information about applicant's employment status and length of unemployment spell about as well as they are able to recall information about other résumé characteristics (such as education, total work experience, and tenure at last job). When we restrict the sample to those who respond that they have "high experience" reviewing résumés, respondents are more likely to correctly identify employment status and length of unemployment. Overall, these results suggest that employment status and length of unemployment are salient to those evaluating résumés, especially if they are experienced at evaluating résumés.

III.B. Measuring Callbacks

We track callbacks from employers by matching voice or email messages to résumés. We follow Bertrand and Mullainathan (2004) by defining a callback as a message from an employer explicitly asking to set up an interview. The voicemail messages were coded independently by two RAs who were not otherwise involved in the project, and they agreed virtually all of the time. We always allowed at least six weeks for a callback, although in practice the vast majority of callbacks were received in the first two weeks. Additionally, the vast majority of callbacks were voice messages; email messages from employers asking to set up an interview were extremely rare. Later, in Table IV, we report results that use an alternative definition of a callback based on whether the employer left any voice message at all, even if the message simply asked for more information.

III.C. Empirical Models

In terms of the experimental design, we created two treatment groups:

- *Treatment 1:* Individuals are randomly assigned to employment status "Employed" with probability 0.25. In this case, the résumé indicates that the applicant is still working at her current job. Let $E_{i,c}$ denote an indicator variable that equals 1 if individual i in MSA c is employed and 0 otherwise.
- *Treatment 2:* Individuals that are not assigned to the Employed treatment are unemployed and are randomly assigned an (integer) unemployment duration or "gap" (in months) according to a discrete uniform distribution

on the interval [1,36]. Let $\log(d_{i,c})$ denote the log of the unemployment duration for individual i in MSA c. Employed individuals are assigned $\log(d_{i,c}) = 0$.

To analyze the experimental data, we estimate the following linear probability model that includes, for efficiency gains, individual and MSA characteristics, $X_{i,c}$:

$$(1) \qquad y_{i,c} = \beta_0 + \beta_1 E_{i,c} + \beta_2 \log(d_{i,c}) + X_{i,c}\Gamma + \varepsilon_{i,c},$$

where $y_{i,c}$ is a callback indicator that equals 1 if individual i in MSA c receives a callback for an interview. Given our randomized design, the coefficients β_1 and β_2 provide unbiased estimates of the mean impact of being employed versus being newly unemployed and the mean impact of changes in the log of unemployment duration, conditional on being unemployed. Because the effect may differ in magnitude across different unemployment durations, we also report results using alternative functional forms for how callbacks depend on duration. In particular, we examine the data nonparametrically using local mean smoothers and plot callback rates as a function of unemployment duration.

To examine how duration dependence varies with local labor market conditions, we restrict the sample to the unemployed and pursue two complementary approaches. First, we use proxies for market tightness (x_c) to estimate the following linear probability model:

$$(2) \quad y_{i,c} = \beta'_0 + \beta'_1 \log(d_{i,c}) + \beta'_2 \log(d_{i,c}) \times x_c + \beta'_3 x_c + X_{i,c}\Gamma' + \varepsilon_{i,c}.$$

This specification includes interactions between log duration and the market tightness proxies. We explore several alternative proxies in the specifications that follow, including the metropolitan area unemployment rate and MSA-level estimates of the vacancy-unemployment ratio. We also estimate specifications that use the full sample and interact the employed indicator ($E_{i,c}$) with the market tightness proxies (x_c).

Our second approach to estimating how duration dependence varies with market tightness implements the following fixed effects model:

$$(3) \qquad y_{i,c} = \delta^c + \gamma^c \log(d_{i,c}) + X_{i,c}\Gamma + \varepsilon_{i,c}.$$

The parameter δ^c is an MSA fixed effect, and γ^c is a MSA-specific estimate of the effect of unemployment duration on callbacks.

This specification is directly motivated by the intuition that there is a one-to-one relationship between the intercept δ^c (i.e., the call-back rate for a newly unemployed individual) and the level of market tightness, as formally laid out in the mechanical model in the Online Appendix. It is worth mentioning that this relationship relies on there being no "compositional effects" over the business cycle in terms of changes in the average quality of newly unemployed workers. Empirical support for this relationship is provided in Online Appendix Figures OA.IX and OA.X, which show a strong correlation between the estimated MSA fixed effects and observed proxies for labor market tightness (such as the unemployment rate). Therefore, the covariance between δ^c and γ^c (i.e., $E[(\delta^c - \bar{\delta}^c)\gamma^c]$) indicates the extent to which duration dependence varies with market tightness.[13] We prove in the Appendix that an unbiased estimate of this covariance is given by the following expression:

$$E[(\delta^c - \bar{\delta}^c)\gamma^c] = \frac{1}{C}\sum_{c=1}^{C}\hat{\delta}^c\hat{\gamma}^c + \frac{1}{C}\sum_{c=1}^{C}\frac{\hat{\sigma}_c^2}{N^c}\frac{E_c[\log(d)]}{Var_c(\log(d))}, \tag{4}$$

where C is the total number of cities in the sample, $\hat{\delta}^c$ and $\hat{\gamma}^c$ are the estimated MSA fixed effects and MSA-specific estimates of the effect of unemployment duration, $\hat{\sigma}_c^2$ is the estimated MSA-specific residual variance, and N^c is the number of observations in the MSA. The second term in equation (4) represents a bias correction to account for the negative mechanical correlation between the MSA-specific estimates $\hat{\delta}^c$ and $\hat{\gamma}^c$. Intuitively, the slope and intercept estimates in an OLS regression are correlated, so to obtain an unbiased estimator of the covariance of the estimated intercept and slope parameters across cities, we need to adjust for this "mechanical" bias using equation (4). We then convert the covariance estimate to a correlation by dividing by the standard deviation of the estimated MSA-specific interaction terms and the standard deviation of the estimated MSA fixed effects.[14]

13. In the Online Appendix, we report results from a (correlated) random effects model. In this model, δ^c is a MSA random effect and γ^c is a MSA-specific random coefficient on unemployment duration. The covariance is estimated by specifying that δ^c and γ^c are jointly normally distributed and estimating the variance-covariance parameters of the joint normal distribution.

14. See the Appendix for details on constructing standard errors for inference.

Finally, we examine how duration dependence varies with characteristics of the résumés, employers, and job postings. We do this by reporting results for various subsamples based on these characteristics and testing whether the estimated effects across these subsamples are equal.

IV. Experimental Results

Our final sample includes 12,054 résumés submitted to 3,040 jobs.[15] Of these 12,054 résumés, 9,236 had (ongoing) unemployment spells of at least one month, with the remaining 2,818 conveying that the worker was currently employed.[16] Table I reports descriptive statistics for the sample. Roughly 4.7% of résumés received a callback from the employer for an interview. In terms of demographics, our sample is relatively young and inexperienced. Roughly two-thirds of our résumés are female, the average age is approximately 27 years, and the average years of experience is 5. Compared to the types of jobs that individuals are applying to, the résumé sample is fairly educated: around 38% of the respondents have bachelor's degrees. This is primarily due to our strategy of sending out both résumés that just match the minimum requirements and résumés that are of higher quality. In terms of MSA characteristics, we see that the average unemployment rate in 2011 across our sample is 9.4%, and this ranges from a low of 5.1% to a high of 17%. Finally, due the randomized design of the field experiment, there is balance across the covariates (across employed/unemployed and across the distribution of unemployment durations), as shown in Table II.

15. Our power calculations called for 12,000 résumé submissions. We needed to submit to more than 3,000 jobs to reach 12,000 résumés because there were several instances where the job posting was taken down before we were able to submit all four of the résumés prepared for the job. This happened on occasion because we waited one day between each résumé submission for a given job posting. In total, we were not able to send all four résumés to 46 jobs; these jobs received 78 résumés.

16. The share of résumés currently employed is 23.4%, which is less than 25% (which was the experimental protocol). The discrepancy comes from roughly 600 résumés where the employment status was randomized slightly differently (in particular, employment was chosen with $p = 1/37$ rather than $p = 1/4$). All results are robust to dropping these observations.

TABLE I
DESCRIPTIVE STATISTICS

	N	Mean	Std. dev.	Min	Max
Dependent variable					
Received callback for interview	12,054	0.047	0.212	0	1
Experimental variables					
Employed	12,054	0.234	0.423	0	1
Months unemployed | Unemployed	9,236	18.018	10.303	1	36
Résumé attributes					
College degree (bachelor's degree)	12,054	0.386	0.487	0	1
Some college (associate's degree)	12,054	0.416	0.493	0	1
High school degree only	12,054	0.198	0.398	0	1
High-quality résumé	12,054	0.502	0.500	0	1
Female	12,054	0.637	0.481	0	1
Years of experience	12,054	5.381	2.012	1	15
Age	12,054	26.908	3.048	19	40
Metropolitan area characteristics					
Unemployment rate (in 2011)	12,054	9.364	2.481	5.07	17.03
Vacanices/Unemployed (V/U) ratio (in 2011)	12,054	3.796	1.581	0.80	7.47
Unemployment rate growth (2008–2011)	12,054	3.644	1.270	1.07	7.40
Job characteristics					
Administrative/clerical job	12,054	0.293	0.455	0	1
Customer service job	12,054	0.306	0.461	0	1
Sales job	12,054	0.401	0.490	0	1

Notes. The first row reports the primary dependent variable which is whether the résumé received a callback from the employer explicitly asking to set up an interview. The experimental sample is split into résumés where the worker reports currently being employed and résumés where the worker does not report currently being employed (with the gap between when the worker last reported working and when the résumé was submitted being uniformly distributed between 1 and 36 months, inclusive).

IV.A. Estimating Duration Dependence

1. Nonparametric Evidence. Before turning to regression results, we begin with simple nonparametric plots of the average callback rate. Figure II reports the relationship between the callback rate and unemployment duration. The first dot corresponding to zero months of unemployment represents the callback rate for the employed. In the top figure, the remaining dots represent average callback rates for each month of unemployment, and the dashed line is a (smoothed) local mean, which is generated using an Epanechnikov kernel and a bandwidth of two months. In the bottom figure, the data are grouped into three- to four-month bins before computing average callback rates. Both the dots and the

TABLE II
RANDOMIZATION TESTS

	Sample means			Sample means		
	Employed	Unemployed	p-value of test of equality	Unemployment duration ≥18 months	Unemployment duration <18 months	p-value of test of equality
Some college	0.408	0.419	.267	0.422	0.415	.624
College degree	0.402	0.381	.217	0.389	0.373	.276
High-quality résumé indicator	0.506	0.501	.630	0.504	0.497	.533
Female	0.629	0.639	.335	0.630	0.648	.062[a]
Unemployment rate (in 2011)	9.382	9.359	.971	9.348	9.370	.749
V/U ratio (in 2011)	3.740	3.813	.373	3.797	3.829	.740
Unemp. rate growth (2008–2011)	3.628	3.648	.359	3.637	3.660	.648
Administrative/clerical job	0.298	0.291	.553	0.293	0.289	.735
Customer service job	0.304	0.307	.708	0.309	0.305	.673
Sales job	0.397	0.402	.838	0.398	0.406	.469
N	2,818	9,236		4,650	4,586	

Notes. This table reports means across subsamples of the experimental sample and presents simple randomization tests based on comparing the means across the subsamples.
[a] Significant at 10%.

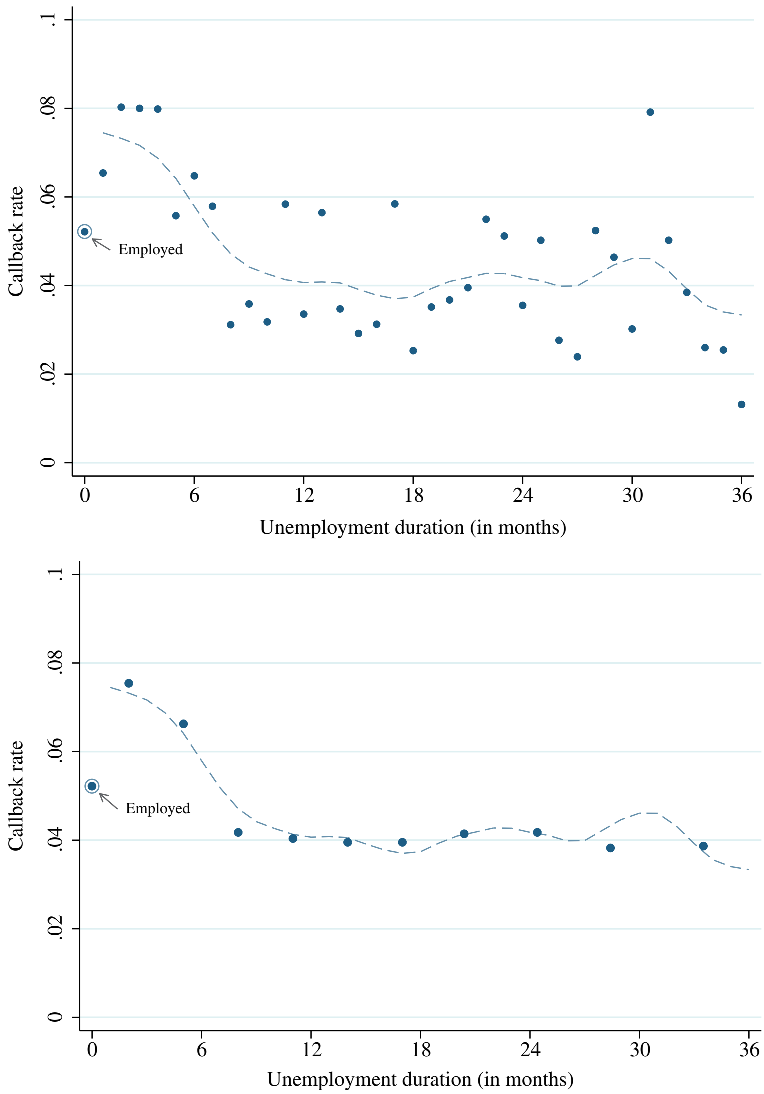

FIGURE II

Callback Rate versus Unemployment Duration

The top figure reports average callback rate by unemployment duration (in months); résumés for which the individual was currently employed are assigned unemployment duration of 0. In the bottom figure, the data are grouped into three- to four-month bins before computing the average callback rate. In both panels, the dashed line is a (smoothed) local mean, which is generated using an Epanechnikov kernel and a bandwidth of two months.

dashed line show clear visual evidence that callbacks decline sharply with unemployment duration for the first six to eight months, and then the callback rate is flat for unemployment durations beyond that. We also see that the callback rate for an employed job seeker is lower than the callback rate for a newly unemployed job seeker.

In Figure III, using the sample of unemployed individuals ($N = 9{,}236$), we report nonparametric local linear regression results that are constrained to be weakly monotonic following the rearrangement procedure of Chernozhukov, Fernandez-Val, and Galichon (2009). The bootstrapped standard errors in Figure III are uniform confidence intervals. We can visually reject the null hypothesis that there is no relationship between unemployment duration and callback rates, based on the inability to draw any horizontal line through the plotted confidence intervals. Overall, Figures II and III show a clear negative relationship between callback rates and unemployment duration, with the steepest decline coming in the first eight months of the unemployment spell. This provides some of the first experimental evidence of negative duration dependence in callback rates, and it also helps partially resolve the set of mixed and inconclusive empirical results from studies that are based on nonexperimental approaches. Interestingly, the pattern of duration dependence from the experiment reported in these figures largely mirrors the pattern based on observational data reported in Figure I.

2. Regression Results. The regression results confirm the results from the graphical analysis. Table III reports OLS regression results estimating equation (1). Longer unemployment durations are associated with lower callback rates. A 1 log point change in unemployment duration is associated with a strongly statistically significant 1.1 percentage point decline in the callback probability, from a mean of 4.7 percentage points. This corresponds to a 23 percent decline in the callback rate. The results in the second row confirm the surprising result from Figures II and III that employed applicants are actually *less* likely to receive callbacks relative to newly unemployed individuals. We discuss possible explanations for this result in Section V. In the remaining columns, we investigate alternative functional forms. Column (2) reports results from a specification with unemployment duration in levels, and column (3) reports results from a spline

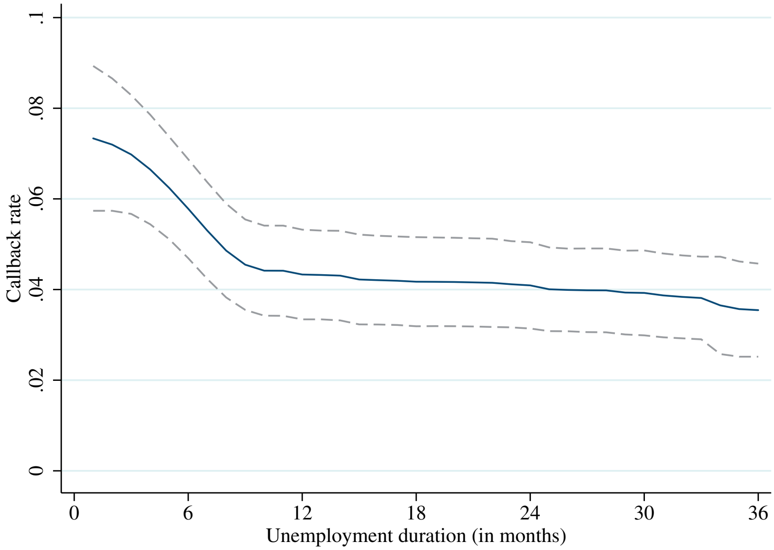

FIGURE III

Nonparametric Results

This figure reports local linear regression estimates using the experimental sample, limited to the unemployed only. The nonparametric estimates are constrained to be monotonic following the rearrangement procedure of Chernozhukov, Fernandez-Val, and Galichon (2009). The confidence intervals are bootstrapped 95% uniform confidence intervals based on 1,000 replications.

regression that allows for a structural break in the effect of unemployment duration at eight months (where the location of the structural break is determined through auxiliary regressions that choose the location of the break to maximize the R^2 of the regression). The results in this column suggest that callbacks are decreasing in unemployment duration for the first eight months and nearly flat after that. Finally, column (4) reports results using piecewise indicator variables for various groups of months (with months 0–6 as the omitted category). The pattern of coefficients in these columns suggests that callbacks are sharply decreasing initially and no longer decreasing after six months.[17]

17. Online Appendix Table OA.II reports the estimated coefficients on the control variables used in all of the main tables, such as gender and "high-quality" résumé indicator. Additionally, Table OA.II shows that when we drop city fixed

TABLE III

THE EFFECT OF UNEMPLOYMENT DURATION ON PROBABILITY OF CALLBACK

	(1)	(2)	(3)	(4)
log(Months unemployed)	−0.011			
	(0.003)			
	[.000]			
$\mathbf{1}$\{Employed\}	−0.020	−0.004	−0.035	−0.016
	(0.010)	(0.006)	(0.013)	(0.008)
	[.040]	[.556]	[.005]	[.043]
$\frac{\text{Months unemployed}}{12}$		−0.008	−0.074	
		(0.003)	(0.021)	
		[.002]	[.000]	
$\frac{\text{Months unemployed}}{12} \times \mathbf{1}\{\text{Months unemployed} > 8\}$			0.074	
			(0.022)	
			[.001]	
$6 <$ Months unemployed ≤ 12 | Unemployed				−0.032
				(0.009)
				[.000]
$12 <$ Months unemployed ≤ 24 | Unemployed				−0.030
				(0.008)
				[.000]
$24 <$ Months unemployed | Unemployed				−0.029
				(0.008)
				[.000]
Joint significance of piecewise coefficients [p-value]				[.001]
F-test of equality across piecewise coefficients [p-value]				[.933]
Average callback rate	0.047	0.047	0.047	0.047
N	12,054	12,054	12,054	12,054
R^2	0.038	0.037	0.039	0.039
Metropolitan area fixed effects	X	X	X	X
Baseline controls	X	X	X	X

Notes. Dependent variable: received callback for interview; full sample. All columns report OLS linear probability model estimates. The data are résumé submissions matched to callbacks from employers to request an interview. The baseline controls are the following: indicator variables for associate's degree, bachelor's degree, high-quality résumé, female gender, and the three job categories (administrative/clerical, customer service, and sales). Standard errors, adjusted to allow for an arbitrary variance-covariance matrix for each job posting, are in parentheses, and p-values are in brackets.

effects and include the city unemployment rate as an additional control instead, we find that the unemployment rate strongly predicts callbacks. This is consistent with a large literature in labor economics that finds that aggregate labor market variables strongly predict individual unemployment durations (Petrongolo and Pissarides 2001).

Table IV shows that results are robust to alternative specifications. First, we explore specifications where all controls are excluded or additional controls are added.[18] Second, we estimate a Probit model to address concerns about boundary effects that arise because of the low average callback rate. Last, we explore an alternative, more inclusive definition of employer callbacks. About 13% of our résumés elicit some response by employer; however, not all of these are callbacks for interviews.[19] In all cases, we find results that are extremely similar to our main results.

IV.B. Duration Dependence and Labor Market Conditions

1. Nonparametric Evidence. We next turn to the question of how the relationship between callback rates and unemployment duration varies with market tightness. We begin by providing graphical evidence. Figure IV shows a plot analogous to Figure II, but it divides the sample depending on whether the local unemployment rate is above or below 8.8% (the median unemployment rate across cities in the experiment). This figure shows that callback rates always decline more rapidly in markets with lower unemployment.

These patterns are robust to other proxies for labor market tightness. For example, Figure V shows similar results when the sample is divided based on median ratio of vacancies to unemployment (V/U ratio), and Figure VI shows similar results when the sample is split based on whether the unemployment rate increased by more than 3.6 percentage points between 2008 and 2011 (the median percentage point increase across the cities in the experiment).

2. Regression Results. The regression evidence confirms the patterns in these figures. We begin by estimating equation (2), using three proxies for local labor market tightness: the unemployment rate (columns (1), (4), and (7)), the vacancy-unemployment (V/U) ratio (columns (2), (5), and (8)), and the

18. The additional control variables that we add include the following: résumé template and résumé font fixed effects, year $\times$ week fixed effects, metropolitan area $\times$ job type fixed effects, and year $\times$ week $\times$ job type fixed effects.

19. Additionally, we explore a specification that drops 83 jobs (comprising 330 résumés) that were posted by employers that we later deemed "questionable." These employers were flagged because we found evidence online that the employer was engaging in dishonest, deceptive, or illegal behavior.

TABLE IV

SENSITIVITY TO ALTERNATIVE CONTROLS AND ALTERNATIVE SPECIFICATIONS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
log(Months unemployed)	−0.011	−0.012	−0.012	−0.010	−0.011	−0.011	−0.011	−0.011	−0.010	−0.014
	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.001]	[.000]	[.001]	[.001]
1{Employed}	−0.020	−0.023	−0.025	−0.017	−0.020	−0.020	−0.019	−0.019	−0.017	−0.009
	(0.010)	(0.010)	(0.010)	(0.006)	(0.007)	(0.010)	(0.010)	(0.009)	(0.009)	(0.013)
	[.040]	[.019]	[.012]	[.005]	[.002]	[.041]	[.045]	[.048]	[.063]	[.491]
Average callback rate	0.047	0.047	0.047	0.047	0.047	0.047	0.047	0.047	0.044	0.126
N	12,054	12,054	12,054	12,054	12,054	12,054	12,054	12,054	11,724	12,054
R^2	0.038	0.015	0.002			0.039	0.054	0.093	0.036	0.069
Alternative controls and specifications										
Dependent variable: callback for interview	X	X	X	X	X	X	X	X	X	
Linear probability model	X	X	X			X	X	X	X	X
Baseline controls	X	X		X		X	X	X	X	X
Metropolitan area fixed effects	X					X	X	X	X	X
Probit (reported marginal effects at mean)				X	X					
Resume template and résumé font fixed effects						X	X	X		
Year × week fixed effects							X	X		
Metropolitan area × job type fixed effects								X		
Year × week × job type fixed effects								X		
Drop job postings from questionable employers									X	
Dependent variable: receive any callback										X

Notes. Dependent variable: received callback for interview; full sample. The baseline controls are the following: indicator variables for Associate degree, Bachelor's Degree, High quality résumé, Female gender, and the three job categories (Administrative/Clerical, Customer Service, and Sales). The week fixed effects indicate the week that the résumé was submitted to the employer. The "questionable employers" restriction drops 83 jobs (comprising 330 résumés) that were for employers that we later deemed to be inappropriate because of evidence online that the employers were engaging in dishonest, deceptive, or illegal behavior. The alternative dependent variable in column (10) is an indicator for whether the employer made any contact at all (whether or not the employer asked explicitly to set up an interview). Standard errors, adjusted to allow for an arbitrary variance-covariance matrix for each employment advertisement, are in parentheses, and p-values are in brackets.

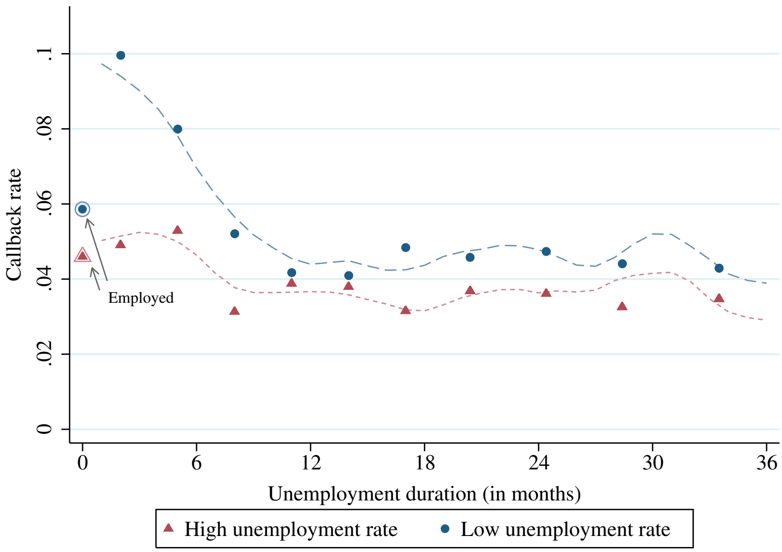

FIGURE IV

Callback Rate versus Unemployment Duration, by Unemployment Rate

This figure is generated by computing the average callback rate for each three- to four-month bin for two subsamples of the experimental data: data from cities with low unemployment rates in July 2011 ($u < 8.8\%$) and cities with high unemployment rates ($u \geq 8.8\%$). The dashed lines are (smoothed) local means, which are generated using an Epanechnikov kernel and a bandwidth of two months.

difference in unemployment rates between 2008 and 2011 (columns (3), (6), and (9)). For the unemployment rate, we use the monthly unemployment rate in the metropolitan area at the start of the experiment (July 2011) from the Bureau of Labor Statistics (BLS). The V/U ratio is constructed using 2011 data on vacancies from the Help Wanted Online Index and 2011 data on the number of unemployed from the BLS. The change in the unemployment rate is calculated using the July 2008 and 2011 BLS unemployment rates.

Table V reports the OLS estimates of equation (2) using these proxies for labor market tightness, based on the sample of unemployed individuals. In columns (1) to (3), we consistently estimate that the effect of unemployment duration is stronger when the local labor market is relatively tight (i.e., either the unemployment rate is relatively low, the V/U ratio is relatively high, or the growth rate in unemployment is relatively low). In

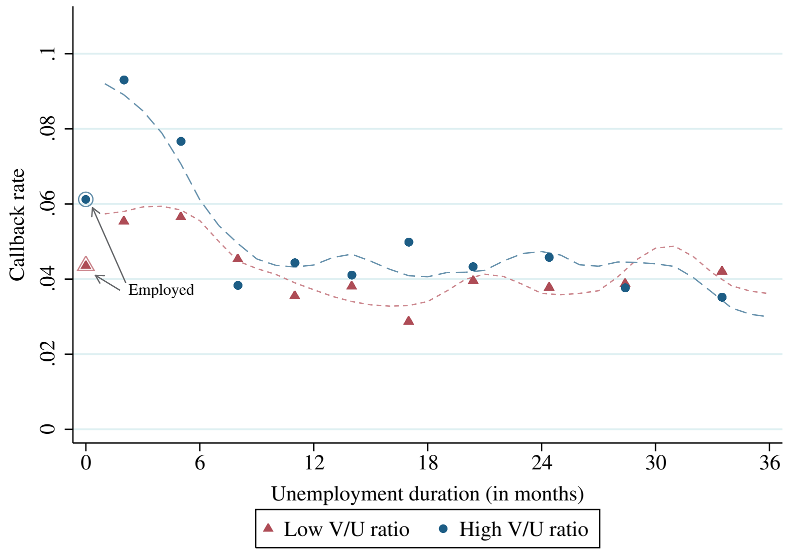

FIGURE V

Callback Rate versus Unemployment Duration, by Vacancy/Unemployment Ratio

This figure is generated by computing the average callback rate for each three- to four-month bin for two subsamples of the experimental data: data from cities with high vacancy/unemployment (V/U) ratios (V/U > 3.25) and cities with low V/U ratios (V/U ≤ 3.25). The dashed lines are (smoothed) local means, which are generated using an Epanechnikov kernel and a bandwidth of two months.

column (1) of Table V, the standardized effect of the unemployment rate implies that a 1 standard deviation increase in the unemployment rate reduces the callback rate by 3.5 percentage points (from a mean of 4.7%). This same change reduces the magnitude of coefficient on unemployment duration by 0.011 (i.e., from –0.012 to –0.001).

The remaining columns in Table V verify that the estimated interaction terms are robust to including both MSA fixed effects (columns (4)–(6)) as well as a wide range of interactions between unemployment duration and MSA characteristics (columns (7)–(9)).[20] The robustness of the results to including these

20. The MSA characteristics include population, median income, fraction of population with a bachelor's degree, and fraction of employed in information industries, professional occupations, service sectors, public administration, construction, manufacturing, wholesale/retail trade, and transportation.

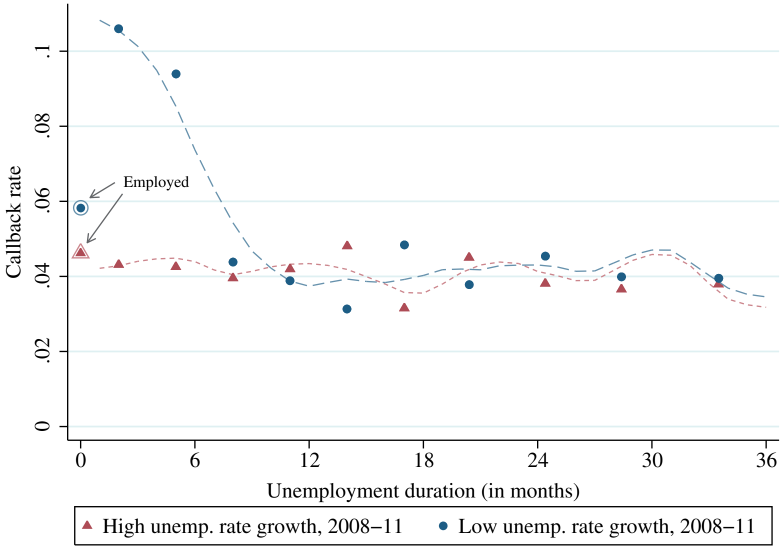

FIGURE VI

Callback Rate versus Unemployment Duration, by Unemployment Rate Growth

This figure is generated by computing the average callback rate for each three- to four-month bin for two subsamples of the experimental data: data from cities with low unemployment rate growth (<3.6 percentage points between 2008 and 2011) and cities with high unemployment rate growth (≥3.6 percentage points). The dashed lines are (smoothed) local means, which are generated using an Epanechnikov kernel and a bandwidth of two months.

additional interactions suggests that the interaction effects of interest are not confounded by other metropolitan area characteristics that are correlated with labor market tightness.

In the Online Appendix, we examine several additional robustness tests. Table OA.III replaces the unemployment rate, the V/U ratio and the difference in unemployment rates with the logarithm of these variables and finds a similar pattern of results. In Table OA.IV, we report analogous results replacing the OLS (linear probability) model with a Probit model. The Probit results show that the estimated marginal effects are very similar to the OLS results.

Next, we turn to our second approach of estimating how duration dependence varies with market tightness. We begin with a simple test of whether there is heterogeneity in duration dependence across labor markets. Table VI reports results that test

TABLE V

How Does Duration Dependence Vary with Labor Market Conditions?

	Baseline controls only			Baseline controls + MSA fixed effects			Baseline controls + MSA fixed effects + MSA characteristics × log(Months unemployed)		
Interaction term formed using proxy for local labor market conditions, $\boldsymbol{X}$ = ...	(1) u_{2011}	(2) $\frac{-V_{2011}}{U_{2011}}$	(3) $u_{2011}-u_{2008}$	(4) u_{2011}	(5) $\frac{-V_{2011}}{U_{2011}}$	(6) $u_{2011}-u_{2008}$	(7) u_{2011}	(8) $\frac{-V_{2011}}{U_{2011}}$	(9) $u_{2011}-u_{2008}$
log(d = Months unemployed)	−0.012	−0.012	−0.012	−0.011	−0.011	−0.011	−0.011	−0.011	−0.011
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]
log(d) × $\boldsymbol{X}$	0.429	0.007	0.790	0.427	0.007	0.783	0.451	0.007	0.853
	(0.142)	(0.002)	(0.252)	(0.142)	(0.002)	(0.253)	(0.201)	(0.004)	(0.332)
	[.002]	[.003]	[.002]	[.003]	[.004]	[.002]	[.025]	[.068]	[.010]
$\boldsymbol{X}$ [Local labor market conditions proxy]	−1.397	−0.024	−2.516						
	(0.415)	(0.007)	(0.744)						
	[.001]	[.002]	[.001]						
Standardized effect of log(d) interaction term	0.011	0.012	0.010	0.011	0.011	0.010	0.011	0.011	0.011
Standardized effect of $\boldsymbol{X}$	−0.035	−0.038	−0.032						
Joint significance of the MSA interactions [p-value]							[.016]	[.029]	[.011]

Notes. Dependent variable: received callback for interview; unemployed sample only. N = 9,236. All columns report OLS linear probability model estimates. All regressions include the same controls listed in Table III. In each column, the proxy for local labor market conditions is indicated in the column heading: u_{2011} corresponds to the metropolitan area (MSA) unemployment rate in July 2011; V_{2011}/U_{2011} corresponds to the MSA vacancy/unemployment ratio in 2011; and $u_{2011}-u_{2008}$ corresponds to the difference in unemployment rates between 2008 and 2011 (both measured in July). In columns (7) through (9), the MSA characteristics include population, median income, fraction of population with a bachelor's degree, and fraction of employed in information industries, professional occupations, service sectors, public administration, construction, manufacturing, wholesale/retail trade, and transportation. The standardized effects reported at the bottom are computed by multiplying the estimated coefficients by the (cross-MSA) standard deviation of the labor market conditions proxy. Standard errors, adjusted to allow for an arbitrary variance-covariance matrix for each employment advertisement, are in parentheses, and p-values are in brackets.

TABLE VI

DURATION DEPENDENCE BY LOCAL LABOR MARKET: FIXED EFFECTS ESTIMATES

	Covariate $\boldsymbol{X}$ = ...				
	(1) log(d)	(2) Female	(3) High-quality résumé	(4) Customer service job	(5) Sales job
Point estimate on $\boldsymbol{X}$	−0.011	0.002	0.011	0.029	0.057
(Reported estimates from model	(0.003)	(0.004)	(0.004)	(0.006)	(0.007)
with no interaction terms)	[.000]	[.529]	[.010]	[.000]	[.000]
F-test of equality for interaction terms (p-value)	[.001]	[.311]	[.910]	[.000]	[.000]
(MSA fixed effect × $\boldsymbol{X}$)					
Correlation between MSA fixed	−0.783	0	0	−0.329	−0.109
effect and MSA-specific	(0.159)			(0.213)	(0.200)
interaction term; $corr(\delta^c,\ \gamma^c)$	[.000]			[.123]	[.586]
R^2	0.083	0.083	0.083	0.083	0.083

Notes. Dependent variable: received callback for interview; unemployed sample only. $N = 9{,}236$. All columns report OLS linear probability model estimates. Each column reports results from two separate regressions. The first row reports the point estimate on the covariate included in the column heading, when the effect is constrained to be the same across all MSAs. The second and third rows report results from an alternative specification which estimates a full set of interaction terms formed by multiplying indicator variables for each MSA with the variable listed in the column heading. The second row reports p-values from a test of equality across all of the estimated interaction terms, and the third row reports a bias-corrected estimate of the correlation between the estimated interaction terms and the MSA fixed effects. All regressions include same controls listed in Table III. In the third row, when a cell entry has 0 with no standard error or p-value, this implies that the model does not reject the null that the effect of the variable in the column is the same in all MSAs. Standard errors, adjusted to allow for an arbitrary variance-covariance matrix for each employment advertisement, are in parentheses, and p-values are in brackets.

whether the effect of unemployment duration on callbacks is the same across all metropolitan areas based on our fixed effects estimates from equation (3). We interact a full set of metropolitan area fixed effects with the log of unemployment duration and conduct an F-test of equality across all of the estimated coefficients for these interaction terms. Based on the results in column (1), we confidently reject the null hypothesis that the effect of unemployment duration is the same across all metropolitan areas ($p = .001$). To test exactly how the effect of unemployment duration varies with market tightness, we construct an estimate of the correlation between the estimates of the MSA-specific interaction terms and the MSA fixed effects based on the covariance expression in equation (4). Consistent with the results in Figures IV through VI, we estimate a statistically significant negative correlation between δ^c and γ^c: $\widehat{corr(\delta^c, \gamma^c)} =$

-0.783; std. err. = 0.159.[21] Under the assumption that the MSA fixed effects are valid proxies for market tightness, these results imply that duration dependence is stronger (i.e., more negative) in tight labor markets. These results are also consistent with the pattern shown in Online Appendix Figure OA.VIII, which graphs the relationship between the estimated MSA-specific coefficient on unemployment duration (from the fixed effects regression estimated in column (1) of Table V) against the MSA unemployment rate at the start of the experiment. The positive relationship in the figure implies that MSAs with lower unemployment rates have stronger (i.e., more negative) duration dependence.

The remaining columns in Table VI repeat this same exercise, reporting the covariance in equation (4), but replacing $\log(d_{i,c})$ with other covariates in $X_{i,c}$.[22] Columns (2) and (3) test for similar heterogeneous effects across labor markets for the effect of gender and "skill" (as measured by whether résumé was "high-quality"), and we find no evidence that the effect of either of these covariates varies across cities. Columns (4) and (5) show that the callback rate of customer service jobs and sales jobs (relative to administrative/clerical jobs) varies strongly across cities. However, these effects are correlated with the average callback rate within the experiment to a much lesser extent; moreover, the sign of this correlation is not consistent across the two types of jobs. In particular, cities with higher average callback rates are not relatively more likely to call back applicants to customer service jobs or sales jobs, even though these jobs have higher average callback rates. We interpret this as evidence against a "mechanical" interpretation of our results in column (1): specifically, these results are inconsistent with low average callback rates in a MSA being associated with simply attenuating the effect of all covariates. In this case, one would expect that cities with higher average callback rates to also have higher callback rates for customer service jobs and sales jobs relative to administrative/clerical jobs, and we do not find evidence that this is the case.

21. Online Appendix Table OA.V reports similar results based on estimating a correlated random coefficients model. We find results that are similar: for unemployment duration we estimate a significant negative correlation ($\widehat{corr(\delta^c, \gamma^c)} = -0.802$; std. err. = 0.092), which implies that cities with higher average callback rates within the experiment have stronger duration dependence.

22. When we replace $\log(d_{i,c})$ with one of the covariates in $X_{i,c}$, we place $\log(d_{i,c})$ in the $X_{i,c}$ vector.

Our findings indicate that duration dependence is stronger when the labor market is relatively tight. Interestingly, this implies that market tightness might have little or no effect on callback rates for the long-term unemployed. Periods of high unemployment will lead to lower callback rates among those with short durations, but they might not adversely affect the callback rates for those with long durations. Whether duration dependence weakens sufficiently to insulate the callback rates of the long-term unemployed from market conditions is of course an empirical question. To investigate this, we examine Figures IV, V, and VI to determine whether there is a duration beyond which callback rates are the same in weak and in strong markets. These figures use three different proxies for labor market tightness. In Figure IV (which uses the metropolitan area unemployment rate as a proxy for labor market tightness), the callback rate is lower in weaker labor markets at all unemployment durations. By contrast, in Figures V and VI (which use the vacancy-unemployment ratio and the growth in the unemployment rate as proxies, respectively), the callback rates converge across strong and weak labor markets at around 8–10 months of unemployment. Therefore, whether duration dependence is strong enough in tight labor markets to (eventually) overturn the positive direct effect of tighter labor market conditions is somewhat sensitive to how we proxy for labor market conditions. Across all figures, however, it is clear that the expected difference in callback rates between strong and weak labor markets is declining in unemployment durations. In other words, across all of our measures of local labor market conditions, we find that duration dependence is stronger in tight labor markets.

IV.C. Heterogeneity by Résumé Characteristics and Employer/Job Characteristics

Table VII explores whether duration dependence varies with résumé characteristics and employer/job characteristics, respectively. The point estimates in Table VII suggest similar levels of duration dependence across education categories (columns (4) and (5)) and ages (columns (6) and (7)), and somewhat larger duration dependence estimates for women compared to men (columns (2) and (3)), although this difference is not statistically significant at conventional levels. Turning to employer/job characteristics, our estimates in columns (8) and (10) indicate that

TABLE VII

HETEROGENEITY IN DURATION DEPENDENCE BY RÉSUMÉ AND JOB CHARACTERISTICS

	(1) Full sample	(2) Men	(3) Women	(4) Bachelor's degree	(5) No bachelor's degree	(6) Older workers (age $\geq$ 27)	(7) Younger workers (age < 27)	(8) Admin/ clerical jobs	(9) Customer cervice jobs	(10) Sales jobs
log(d = Months unemployed)	−0.011	−0.006	−0.014	−0.010	−0.011	−0.012	−0.010	−0.010	−0.003	−0.017
	(0.003)	(0.005)	(0.003)	(0.005)	(0.004)	(0.005)	(0.004)	(0.004)	(0.005)	(0.006)
	[.000]	[.276]	[.000]	[.031]	[.005]	[.007]	[.016]	[.007]	[.464]	[.005]
1{Employed}	−0.020	0.003	−0.032	−0.021	−0.017	−0.030	−0.012	−0.016	0.001	−0.036
	(0.010)	(0.016)	(0.011)	(0.016)	(0.012)	(0.015)	(0.013)	(0.011)	(0.014)	(0.019)
	[.040]	[.850]	[.003]	[.174]	[.151]	[.044]	[.328]	[.163]	[.952]	[.060]
log(d) equal across columns [p-value]		[.178]		[.982]		[.465]			[.185]	
Average callback rate in sample	0.047	0.057	0.041	0.048	0.047	0.048	0.046	0.016	0.044	0.072
N	12,054	4,380	7,674	4,653	7,401	5,918	6,136	3,531	3,690	4,833
R^2	0.038	0.043	0.046	0.045	0.050	0.044	0.052	0.066	0.057	0.048
MSA fixed effects	X	X	X	X	X	X	X	X	X	X
Baseline controls	X	X	X	X	X	X	X	X	X	X

Notes. Dependent variable: received callback for interview; full sample. All columns report OLS linear probability model estimates. All regressions include same controls listed in Table III. Standard errors, adjusted to allow for an arbitrary variance-covariance matrix for each employment advertisement, are in parentheses, and p-values are in brackets.

duration dependence is strongest for workers applying to sales jobs (as compared to administrative/clerical jobs and customer service jobs), although these differences are not statistically significant.

In Online Appendix Table OA.VI, we explore heterogeneity across low- and high-quality résumés and do not find any significant differences. We also compare estimates of duration dependence across industries and find that our overall duration dependence estimates are concentrated in construction, manufacturing, and wholesale and retail trade, with no significant evidence of duration dependence in business/financial services or professional/personal services. The remaining columns in Table OA.VI explore differences according to whether the job posting mentions that experience is required and whether the posting explicitly mentions that the employer is an Equal Opportunity Employer. In both of these cases, there are no significant differences across the subsamples.

V. Alternative Theoretical Interpretations

This section discusses possible theoretical explanations for our findings. In Section II, we argued that employer screening, human capital, and ranking models all predict negative duration dependence in job-finding rates. Our main empirical finding that callback rates decline with unemployment duration indicates that these mechanisms are at work. We can therefore reject the hypothesis that duration dependence in job-finding rates is entirely due to compositional effects based on characteristics that are observable to employers, but not researchers, when making callback decisions. We can also reject the hypothesis that duration dependence is entirely due to changes in job seeker behavior over the spell due to discouragement effects.

Our second key empirical finding is that duration dependence is stronger in tight labor markets. This finding is consistent with a broad class of employer screening models. To establish this, in the Online Appendix we develop a general screening model, similar to Lockwood (1991). As an intuitive measure of the strength of duration dependence, we define the *relative* callback rate, the ratio of the population callback probability evaluated at some positive duration to the population callback probability among the newly unemployed. The relative callback rate is below 1 if there is negative duration dependence. We show that the relative callback rate—and hence duration dependence—varies negatively with market tightness.[23] Intuitively, in screening models, market tightness pins down the share of productive types in the population at a given unemployment duration. In tight markets, this share is lower at all positive durations because the market sorts workers more quickly. Therefore, the relative callback rate declines with market tightness in screening models.[24] As an alternative to the screening model, we also develop a simple model of human capital depreciation in the Online Appendix and show that it predicts no relationship between the relative callback rate and market tightness if the rate at which a worker's skills depreciate doesn't vary with aggregate labor market conditions.[25]

Our empirical results in Section IV.B show that the slope of the callback function (with respect to duration) is greater in tight markets. We also report complementary results in the Online Appendix that correspond more closely to the relationship between relative callback rates and market tightness. Figure OA.IV corresponds to Figure II but instead plots the relative callback rate against unemployment duration for the full sample (i.e., the callback rate at each unemployment duration divided by the callback rate of a newly unemployed worker). Figures OA.V, OA.VI, and OA.VII (corresponding to Figures IV, V, and VI) similarly plot the relative callback rates separately for tight and slack labor markets using three proxies for labor market tightness: the metropolitan area unemployment rate, the vacancy-unemployment ratio, and the growth in unemployment rate, respectively. The graphical evidence in all of these figures indicates that the relative callback rate is declining and falls sharply in tighter labor markets. To quantify how the relative callback

23. An alternative measure of how duration dependence varies with market tightness is the cross-derivative between duration and market tightness. However, the cross-derivative is a local measure of duration dependence, and it can be positive for some values of duration and negative for others. As it turns out, our model has no general implications for such local measures of duration dependence. Instead, we use a global measure that holds for all positive values of duration.

24. This requires that the composition of newly unemployed job seekers does not change over the business cycle.

25. We also consider the ranking model of Blanchard and Diamond (1994) in the Online Appendix and show that it predicts a positive relationship between duration dependence and market tightness.

rate varies with market tightness, Table OA.VII reports results from estimating a fixed-effects Poisson model. The parameter estimates of the Poisson model can be interpreted as "proportional effects" and therefore capture the effect of duration and market tightness on relative callbacks.[26] The results in this table verify that the magnitude of duration dependence declines proportionally with the unemployment rate, which is consistent with the graphical evidence in Figures OA.V through OA.VII. Overall, we emphasize that our empirical evidence is more consistent with employer screening than pure human capital depreciation with unemployment duration.

Finally, we return to our results comparing the newly unemployed to workers who are currently employed. In Table III, we found that a currently employed worker is less likely to be called back for an interview than a newly unemployed individual. These results are perhaps surprising given the widespread media attention toward firms that expressed hiring preference for workers who are currently employed.[27] We investigate this result further in Table VIII by estimating a richer model that interacts the employed indicator with the proxies for labor market tightness from Table V. In columns (2)–(4), we find that across each of the proxies for labor market conditions, the callback "gap" between employed workers and newly unemployed workers shrinks when labor market conditions are poor.[28]

We evaluate these results using existing models of asymmetric information in the labor market (Greenwald 1986; Gibbons and Katz 1991) that emphasize the signaling value of unemployment: a worker who is unemployed is likely to be of lower quality compared to a worker who is employed. Therefore, the

26. More specifically, the estimates of the Poisson model are informative on how the elasticity of callbacks with respect to duration varies with market tightness (i.e., how $\frac{d\log(y)}{d\log(d)}$ varies with x). Under a constant elasticity assumption, this sheds light on how the relative callback rate varies with market tightness. See the Online Appendix for more details.

27. See, for example, "The Unemployed Need Not Apply," *New York Times* (editorial), February 19, 2011.

28. To conserve space, we only report the variables relevant for comparing currently employed to newly unemployed workers, even though these regression results are based on the full sample. Online Appendix Table OA.VIII reports results for the full set of specifications and also reports estimates of the coefficients on unemployment duration and the interaction of this with market tightness.

TABLE VIII

CALLBACKS FOR THE CURRENTLY EMPLOYED AND LABOR MARKET CONDITIONS

	Baseline controls only		
Interaction term formed using proxy for local labor market conditions, $\boldsymbol{X} = \ldots$	(1) u_{2011}	(2) $\frac{-V_{2011}}{U_{2011}}$	(3) $u_{2011} - u_{2008}$
1{Employed}	−0.023	−0.022	−0.023
	(0.010)	(0.010)	(0.010)
	[.019]	[.024]	[.018]
1{Employed} × ***X***	0.990	0.015	1.747
	(0.434)	(0.008)	(0.785)
	[.023]	[.052]	[.026]
X [Local labor market conditions proxy]	−1.403	−0.024	−2.530
	(0.415)	(0.007)	(0.744)
	[.001]	[.001]	[.001]
Standardized effect of **1**{Employed} interaction term	0.025	0.024	0.022
Standardized effect of ***X***	−0.035	−0.037	−0.032

Notes. Dependent variable: received callback for interview; full sample. $N = 12{,}054$. All columns report OLS linear probability model estimates. All regressions include the same controls listed in Table III. See Table V for notes on the labor market conditions proxies. The standardized effects reported at the bottom are computed by multiplying the estimated coefficients by the (cross-MSA) standard deviation of the labor market conditions proxy. Standard errors, adjusted to allow for an arbitrary variance-covariance matrix for each employment advertisement, are in parentheses, and *p*-values are in brackets.

unemployed are likely to suffer relatively worse labor market outcomes, on average.[29] If workers who become unemployed during recessions are higher average quality than workers who become unemployed during normal economic times, then the signaling value of unemployment will be diminished. This theory therefore predicts that the callback rates of the employed and the unemployed should converge as the labor market worsens (Nakamura 2008).

Our empirical evidence provides mixed support for these theories. Taken at face value, our finding that the employed receive fewer callbacks then the newly unemployed is inconsistent with the theory. However, there are reasons individuals engaged in an

29. Consistent with this theory, Blau and Robins (1990) find that employed searchers are much more likely to get job offers than unemployed searchers, although they cannot rule out unobserved heterogeneity as an explanation. On the other hand, Holzer (1987) finds the opposite.

on-the-job search might not be attractive job candidates to firms. First, these individuals might be intrinsically less loyal and especially prone to job hopping. In informal discussions with human resources professionals, we have learned that some employers express the concern that workers who are currently employed are not serious job seekers and, as a result, some employers are less likely to invite them for an interview. Additionally, we suspect that it may also be easier for firms to bargain and negotiate with unemployed job seekers because they have a lower outside option compared to job seekers who have a job. Finally, we speculate that our findings could also be caused by the fact that some jobs require workers to start immediately. In this case, it seems plausible that the lag in recruiting a worker who is currently employed exceeds the lag in hiring a job seeker who is currently out of work, which may be particularly relevant for the set of less-skilled jobs that are posted on the online job board we used for the experiment. On the other hand, we find the callback gap is smaller when local labor market conditions are poor, which is consistent with Nakamura (2008). Although the comparative static is consistent with the theory, on average we find the opposite sign to what the theory predicts. Thus, one should exercise caution in interpreting this evidence as providing strong support for the theory.

VI. Conclusion

This article reports results from a field experiment studying duration dependence. Our results indicate that the likelihood of receiving a callback from employers sharply declines with unemployment duration. This effect is quantitatively large and especially pronounced during the first eight months after becoming unemployed. Additionally, we find that duration dependence is stronger in tight labor markets. This result is consistent with the prediction of a broad class of screening models in which employers use the unemployment spell length as a signal of unobserved productivity and recognize that this signal is less informative in weak labor markets. This result is not easily generated by a model of human capital depreciation when the rate of human capital depreciation is steady and the same across labor markets. Although we emphasize that we do not rule out a role for human capital depreciation, our results are most consistent with employer screening playing an important role in generating duration dependence.

We speculate that there are close connections between the screening models that are supported by our data and rational herding models that bear exploring. For example, our screening model in the Online Appendix assumes that employers meet workers sequentially and then use the information about prior actions of other firms (embedded in the duration of unemployment) to learn about worker productivity. However, the employers do not observe the private signals received by the other firms. This setup maps closely to the structure of a standard rational herding model.[30] We believe that it is important to investigate the optimal design of unemployment insurance in a setting with asymmetric information and social learning—features typically omitted from the standard analysis of unemployment insurance (Baily 1978; Gruber 1997; Chetty 2008).

More broadly, our article is part of a growing literature that exploits variation in labor market conditions to inform theories of the labor market. Davis and von Wachter (2011) find that the cost of job loss is higher during recessions and argue that this is inconsistent with a standard Mortensen and Pissarides (1994) model of the labor market. Kroft and Notowidigdo (2011) investigate how the moral hazard cost and consumption smoothing benefit of unemployment insurance varies with labor market conditions, and they use these results to calibrate and assess a standard job search model. Crépon et al. (2013) conduct a clustered randomized control trial of job placement assistance and find that the negative spillover effects of the experiment (i.e., crowd-out onto untreated individuals) are larger when the labor market is slack. They interpret this evidence as consistent with a model of job rationing (Landais, Michaillat, and Saez 2013). Under job rationing, workers will remain unemployed longer due to congestion effects. Our article suggests that this will lead to even more unemployment through duration dependence, which varies depending on the strength of the labor market.

30. For a review of herding models, see Bikhchandani, Hirshleifer, and Welch (1998). The main difference is an asymmetry in the learning process that is present in our model: once a worker is hired by a firm, the public learning process stops. Additionally, we note that the informational assumptions in our setup might be more realistic, since they require that firms merely observe the length of the spell, and not the ordering of actions of past firms.

Last, the results in our experiment suggest several additional areas for future research. Empirically, we think it is important to examine whether our results generalize to the economy more broadly. Our results speak most directly to younger job seekers with relatively little work experience. Future audit studies should explore whether our results transfer to a broader set of occupations, to different modes of searching for jobs, and to older workers. At the time of writing, long-term unemployment in the United States remains at unprecedented levels. How job-finding rates relate to unemployment duration will be an important factor shaping the recovery from the Great Recession. Our findings suggest that the interaction between duration dependence and labor market conditions needs to be taken into account when analyzing the recovery in the labor market.

Appendix

A. *Data Sources*

This section describes the various MSA-level data used in the empirical analysis.

MSA Unemployment Data. Source: U.S. Bureau of Labor Statistics, http://data.bls.gov/cgi-bin/dsrv?la. Monthly data on number of unemployed persons, number of persons in the labor force, the number of employed persons, and the unemployment rate in the given MSA (not seasonally adjusted). For New England states, the BLS provides NECTA (New England City and Town Area) data instead of MSA data, so for a few metropolitan areas NECTA-level data were used.

Vacancy Data. We purchased vacancy data from Wanted Analytics (WA), which is part of Wanted Technologies. WA collects hiring demand data and is the exclusive data provider for the Conference Board's Help-Wanted OnLine Data Series, which is a monthly economic indicator of hiring demand in the United States. WA gathers its data from the universe of online vacancies posted on Internet job boards or online newspapers. In total, it covers roughly 1,200 online job boards, although the vast majority of the ads appear on a small number of major job boards. When the same job ad appears on multiple job boards, WA uses a deduplication procedure to identify unique job ads on the basis of company name, job title, and description and MSA or State. Sahin et al. (2011) document potential measurement issues related to these data: first, the data set records a single vacancy per ad, although it is possible that multiple positions are listed in a single ad; second, it is possible that multiple locations within a state are listed in a single ad for a given position. The data we received contain the total number of job postings by MSA, six-digit occupation code, and year. Our sample spans 2008 through 2012.

Median Income, Metropolitan Population, Share of Population (25 Years and Above) with Bachelor's Degree, Share of Civilian Employed Population in Various Industries. Source: 2011 American Community Survey (ACS) one-year estimates from the U.S. Census Bureau.

Covariance between City Fixed Effects and City-Specific Effect of Unemployment Duration

Recall the following estimating equation from the main text:

$$y_{i,c} = \delta^c + \gamma^c \log(d_{i,c}) + X_{i,c}\Gamma + \varepsilon_{i,c},$$

where δ^c is a metropolitan area fixed effect and γ^c is an MSA-specific estimate of the effect of unemployment duration. We test for whether duration dependence varies with labor market conditions by treating δ^c as a proxy measure of labor market tightness and then estimating the covariance between δ^c and γ^c; that is, $E[(\delta^c - \bar{\delta^c})\gamma^c]$. We compute this by first computing the covariance between the estimates; that is, $E[\hat{\delta^c}\hat{\gamma^c}]$. Defining $\hat{\eta}^c_\delta$ as estimation error for $\hat{\delta^c}$ (i.e., $\hat{\delta^c} = \delta^c + \hat{\eta}^c_\delta$) and $\hat{\eta}^c_\gamma$ as estimation error for $\hat{\gamma}^c$, we can compute $E[\hat{\delta^c}\hat{\gamma^c}]$ as follows:

$$\begin{aligned}
E[\hat{\delta}^c\hat{\gamma}^c] &= \frac{1}{C}\sum_{c=1}^{C}\hat{\delta}^c\hat{\gamma}^c \\
&= \frac{1}{C}\sum_{c=1}^{C}(\delta^c + \hat{\eta}^c_\delta)(\gamma^c + \hat{\eta}^c_\gamma) \\
&= \frac{1}{C}\sum_{c=1}^{C}\delta^c\gamma^c + \frac{1}{C}\sum_{c=1}^{C}\delta^c\hat{\eta}^c_\gamma + \frac{1}{C}\sum_{c=1}^{C}\hat{\eta}^c_\delta\gamma^c + \frac{1}{C}\sum_{c=1}^{C}\hat{\eta}^c_\delta\hat{\eta}^c_\gamma,
\end{aligned}$$

where C is the total number of cities in the sample. We can rewrite this using expectations as follows (using the fact that $E_c[\hat{\eta}^c_\gamma] = 0$ and $E_c[\hat{\eta}^c_\delta] = 0$):

$$E[\hat{\delta}^c \hat{\gamma}^c] = E[\delta^c \gamma^c] + \frac{1}{C}\sum_{c=1}^{C} E_c[\hat{\eta}^c_\delta \hat{\eta}^c_\gamma].$$

Next, we can compute $E_c[\hat{\eta}^c_\delta \hat{\eta}^c_\gamma]$ using standard statistical results:

$$E_c[\hat{\eta}^c_\delta \hat{\eta}^c_\gamma] = -\frac{\sigma^2_c}{N^c}\frac{E_c[\log(d)]}{Var_c(\log(d))},$$

where σ^2_c is the residual variance for MSA c, and N^c is the number of observations in the MSA. Combining the above gives us the following expression for the unbiased estimate of $E[\delta^c \gamma^c_2]$:

$$E[(\delta^c - \bar{\delta}^c)\gamma^c] = \frac{1}{C}\sum_{c=1}^{C} \hat{\delta}^c \hat{\gamma}^c + \frac{1}{C}\sum_{c=1}^{C} \frac{\hat{\sigma}^2_c}{N^c}\frac{E_c[\log(d)]}{Var_c(\log(d))}. \tag{5}$$

In other words, there is a negative bias in estimated covariance if one simply computes the empirical covariance based on the regression estimates $\hat{\delta}^c$ and $\hat{\gamma}^c$. Intuitively, this bias comes from the fact that the sampling errors in the estimates for these two parameters for a given MSA are negatively correlated. Although this bias goes away asymptotically, it requires both that $C \to \infty$ and $N^C \to \infty$. In Monte Carlo simulations resembling our experimental data, we find substantial bias unless we use the bias correction.

We conduct inference on the estimated covariance by computing the following standard error estimate, and we have verified that these standard errors are reliable using Monte Carlo simulations:

$$se(\widehat{E[\delta^c \gamma^c]}) = \sqrt{\frac{1}{C}\left(\frac{1}{C}\sum_{c=1}^{C} (\hat{\delta}^c)^2(\hat{\gamma}^c)^2\right)}.$$

UNIVERSITY OF TORONTO
MCGILL UNIVERSITY, IZA, AND CESIFO
UNIVERSITY OF CHICAGO BOOTH SCHOOL OF BUSINESS AND NBER

Supplementary Material

An Online Appendix for this article can be found at QJE online (qje.oxfordjournals.org).

References

Abbring, Jaap H., Gerard J. van den Berg, and Jan C. van Ours, "Business Cycles and Compositional Variation in U.S. Unemployment," *Journal of Business & Economic Statistics*, 19 (2001), 436–448.

Acemoglu, Daron, "Public Policy in a Model of Long-term Unemployment," *Economica*, 62 (1995), 161–178.

Baily, Martin N., "Some Aspects of Optimal Unemployment Insurance," *Journal of Public Economics*, 10 (1978), 379–402.

Bertrand, Marianne, and Sendhil Mullainathan, "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," *American Economic Review*, 94 (2004), 991–1013.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch, "Learning from the Behavior of Others," *Journal of Economic Perspectives*, 12 (1998), 151–170.

Blanchard, Olivier J., and Peter Diamond, "Ranking, Unemployment Duration, and Wages," *Review of Economic Studies*, 61 (1994), 417–434.

Blau, David M., and Philip K. Robins, "Job Search Outcomes for the Employed and Unemployed," *Journal of Political Economy*, 98 (1990), 637–655.

Butler, Richard J., and James B. McDonald, "Trends in Unemployment Duration Data," *Review of Economics and Statistics*, 68 (1986), 545–557.

Chernozhukov, Victor, Ivan Fernandez-Val, and Alfred Galichon, "Improving Estimators of Monotone Functions by Rearrangement," *Biometrika*, 96 (2009), 559–575.

Chetty, Raj, "Moral Hazard versus Liquidity and Optimal Unemployment Insurance," *Journal of Political Economy*, 116 (2008), 234.

Coles, Melvyn G., and Eric Smith, "Marketplaces and Matching," *International Economic Review*, 39 (1998), 239–254.

Congressional Budget Office, "Understanding and Responding to Persistently High Unemployment," 2012, available at http://www.cbo.gov/publication/42989.

Crépon, Bruno, Esther Duflo, Marc Gurgand, Roland Rathelot, and Philippe Zamora, "Do Labor Market Policies Have Displacement Effects? Evidence from a Clustered Randomized Experiment," *Quarterly Journal of Economics*, 168 (2013), 531–580.

Darby, Michael R., John Haltiwanger, and Mark Plant, "Unemployment Rate Dynamics and Persistent Unemployment iunder Rational Expectations," *American Economic Review*, 75 (1985), 614–637.

Davis, Steven J., and Till M. von Wachter, "Recessions and the Costs of Job Loss," *Brookings Papers on Economic Activity*, 43 (2011), 1–72.

Dynarski, Mark, and Steven M. Sheffrin, "The Behavior of Unemployment Durations over the Cycle," *Review of Economics and Statistics*, 72 (1990), 350–356.

Eriksson, Stefan, and Dan-Olof Rooth, "Do Employers Use Unemployment as a Sorting Criterion When Hiring? Evidence from a Field Experiment," unpublished manuscript, 2011.

Gibbons, Robert, and Lawrence F. Katz, "Layoffs and Lemons," *Journal of Labor Economics*, 9 (1991), 351–380.

Greenwald, Bruce C., "Adverse Selection in the Labour Market," *Review of Economic Studies*, 53 (1986), 325–347.

Gruber, Jonathan, "The Consumption Smoothing Benefits of Unemployment Insurance," *American Economic Review*, 87 (1997), 192–205.

Heckman, James J., and Burton Singer, "The Identifiability of the Proportional Hazard Model," *Review of Economic Studies*, 51 (1984), 231–241.

Holzer, Harry J., "Job Search by Employed and Unemployed Youth," *Industrial and Labor Relations Review*, 40 (1987), 601–611.
Imbens, Guido, and Lisa Lynch, "Re-employment Probabilities over the Business Cycle," *Portuguese Economic Journal*, 5 (2006), 111–134.
Kroft, Kory, and Matthew J. Notowidigdo, "Should Unemployment Insurance Vary with the Unemployment Rate? Theory and Evidence," NBER Working Paper No. w17173, 2011.
Krueger, Alan B., and Andreas Mueller, "Job Search and Unemployment Insurance: New Evidence from Time Use Data," *Journal of Public Economics*, 94 (2010), 298–307.
Lahey, Joanna N., "Age, Women, and Hiring: An Experimental Study," *Journal of Human Resources*, 43 (2008), 30–56.
Landais, Camille, Pascal Michaillat, and Emmanuel Saez, "Optimal Unemployment Insurance over the Business Cycle," NBER Working Paper No. w16526, 2013.
Lazear, Edward, "Raids and Offermatching," NBER Working Paper No. w1419, 1984.
Ljungqvist, Lars, and Thomas J. Sargent, "The European Unemployment Dilemma," *Journal of Political Economy*, 106 (1998), 514–550.
Lockwood, Ben, "Information Externalities in the Labour Market and the Duration of Unemployment," *Review of Economic Studies*, 58 (1991), 733–753.
Machin, Stephen, and Alan Manning, "The Causes and Consequences of Long-Term Unemployment in Europe," in *Handbook of Labor Economics*, Orley Ashenfelter, and David Card, eds. (Amsterdam: North-Holland, 1999), 3085–3139.
Mortensen, Dale T., and Christopher A. Pissarides, "Job Creation and Job Destruction in the Theory of Unemployment," *Review of Economic Studies*, 61 (1994), 397–415.
Moscarini, Giuseppe, "Unobserved Heterogeneity and Unemployment Duration: A Fallacy of Composition," Yale University Working Paper, 1997.
Nakamura, Emi, "Layoffs and Lemons over the Business Cycle," *Economics Letters*, 99 (2008), 55–58.
Oberholzer-Gee, Felix, "Nonemployment Stigma as Rational Herding: A Field Experiment," *Journal of Economic Behavior & Organization*, 65 (2008), 30–40.
Oreopoulos, Philip, "Why Do Skilled Immigrants Struggle in the Labor Market? A Field Experiment with Thirteen-Thousand Resumes," *American Economic Journal: Economic Policy*, 3, no. 4 (2011), 148–171.
Petrongolo, Barbara, and Christopher A. Pissarides, "Looking into the Black Box: A Survey of the Matching Function," *Journal of Economic Literature*, 39 (2001), 390–431.
Rothstein, Jesse, "Unemployment Insrance and Job Search in the Great Recession," *Brookings Papers on Economic Activity*, 43 (2011), 143–213.
Sahin, Aysegul, Joseph Song, Giorgio Topa, and Giovanni Violante, "Measuring Mismatch in the U.S. Labor Market," New York Federal Reserve Working Paper, 2011.
Shimer, Robert, "The Probability of Finding a Job," *American Economic Review: Papers & Proceedings*, 98 (2008), 268–273.
Sider, Hal, "Unemployment Duration and Incidence: 1968–1982," *American Economic Review*, 75 (1985), 461–472.
van den Berg, Gerard J., and Jan C. van Ours, "Unemployment Dynamics and Duration Dependence," *Journal of Labor Economics*, 14 (1996), 100–125.
Vishwanath, Tara, "Job Search, Stigma Effect, and Escape Rate from Unemployment," *Journal of Labor Economics*, 7 (1989), 487–502.