Duration Dependence in Finding a Job: Applications, Interviews, and Job Offers

Rafael Lalive
University of Lausanne

Aderonke Osikominu
University of Hohenheim

Lorenzo Pesaresi
Bank of Italy

Jeremy Zuchuat
University of Lausanne

Josef Zweimüller
University of Zürich

2026.01.12

Abstract

The job finding rate declines with the duration of unemployment, but the relative importance of workers' search behavior and employers' recruitment behavior remains unclear. We use monthly search diaries from Swiss public employment offices to shed new light on this issue. Search diaries record each single application sent by a job seeker and indicate whether the employer followed up with an interview and a job offer. Based on more than 600,000 applications sent by 15,000 job seekers, we find that applications and interviews decrease, but job offers per interview weakly increase with duration. A theoretical framework with endogenous search effort by workers and statistical discrimination by firms replicates the duration patterns of applications, interviews and job offers closely. The estimated model predicts that over half of the decline in the job finding rate is due to structural duration dependence, with the remainder explained by dynamic selection of the unemployment pool. Falling applications by job seekers – who internalize statistical discrimination by firms – are the main driver of duration dependence.

Keywords: Job search, job finding, duration dependence, dynamic selection, search effort, job application, callback, job interview, job offer.

JEL: J24, J64

We are grateful to seminar audiences in Austin, Basel, Bergen, Berkeley, Berlin, Bonn, Cornell, Georgetown, Madrid, Mannheim, Rome, Salzburg, Stockholm, Tel Aviv, Tilburg, Uppsala, and Vienna, and especially to David Card, Pierre Cahuc, Stefano DellaVigna, Gregor Jarosch, Philipp Kircher, Pat Kline, Moritz Kuhn, Attila Lindner, Enrico Moretti, Andreas Mueller, Laura Pilossoph, Benjamin Schoefer and Simon Trenkle. We are grateful to David Liechti and Michael Morlok who collected the search diary data. Rafael Lalive acknowledges financial support through Swiss National Science Foundation Grant No. 100018_204575. Corresponding author: Rafael Lalive, University of Lausanne, rafael.lalive@unil.ch. E-mails: jeremy.zuchuat@unil.ch, rafael.lalive@unil.ch, a.osikominu@uni-hohenheim.de, lorenzo.pesaresi@econ.uzh.ch, josef.zweimueller@econ.uzh.ch.

1. Introduction

The rate at which unemployed workers find new regular jobs decreases with the duration of unemployment. While this is a well established empirical fact, the reasons are still disputed. As often, the debate is about causation versus correlation. Does the falling job finding rate reflect a causal effect of unemployment duration on the chances to find a new job? Or does it reveal negative dynamic selection, so that the long-term unemployed had weak employment prospects to begin with? Answers to these questions are crucial, as they determine the policy responses to combat long-term unemployment.

This paper sheds new light on the forces driving the falling job finding rate by using *monthly search diaries* from the Swiss public employment offices. In Switzerland, job seekers drawing unemployment insurance (UI) benefits have to document their search activities in search diaries. Search diaries not only list each single application, but also indicate whether the employer followed up with a callback for a job interview and, if so, whether the interview eventually resulted in a job offer. Search diaries are an important monitoring tool of unemployed workers' search effort, as well as the outcome of their search activities.

Our analysis makes two main contributions. Our first contribution is empirical. We digitize 58,000 search diaries containing 600,000 job applications sent by 15,000 job seekers drawing UI benefits. These data allow us to dig deeper into the various steps of the job finding process and to study jointly how applications, interviews, and job offers change with the duration of unemployment. This adds to the previous literature which has looked at search effort and employer callbacks separately. To the best of our knowledge, no previous study has yet explored how job offers change with duration.

Our second contribution is theoretical. We develop and estimate a structural model in which both job seekers' search effort and firms' recruitment behavior are endogenous. The key innovation of our framework is that it captures the evolving interaction between these two margins over the unemployment spell. The model allows us to rationalize our empirical findings and to decompose the decline in the job finding rate into duration dependence and dynamic selection.

Our empirical analysis first documents that *job applications* fall with the duration of unemployment. The average job seeker sends 11 applications in month 1 of the unemployment spell, which decreases to slightly less than 10 in months 12-15. Because applications are repeatedly observed for each job seeker, a fixed effects model can tease out duration dependence.[1] We find a strong within-individual decline from 11 applications in month 1 to 8 applications in months 12-15. Since, in the cross-section, the number of applications decreases only slightly, the strong within-individual decline of applications implies positive dynamic selection, in the sense that job seekers who eventually become long-term unemployed send more applications at all durations.

The probability that an application receives a *job interview* shows a marked decline from 5 percent for applications sent in month 1 to 2.5 percent for those sent in months 12-15 – very similar to Kroft, Lange, and Notowidigdo (2013) for fictitious job applicants in the US. Interestingly, and perhaps surprisingly, the probability to get a *job offer* after an interview weakly increases with duration, from 20 percent for applications sent in month 1 to 25 percent or more in months 12-15. The overall response to an application, the probability that an application leads to a job offer (unconditional on an interview), is falling from about 1 percent for applications sent in month 1 to less than 0.7 percent for applications sent in months 12-15.

Assessing whether falling interview rates and rising job offer rates reflect duration dependence or dynamic selection is challenging. Unlike job applications, which are repeatedly observed and vary unsystematically throughout the unemployment spell, job interviews and job offers are rare events that tend to be concentrated at the end of the spell. As a result, sample attrition is correlated with the time-varying unobservables that affect these outcomes. In such a setting of nonrandom attrition, the fixed-effects estimator of the duration effect is generally biased and inconsistent (Verbeek and Nijman, 1992; Wooldridge, 1995). Therefore, we use our rich administrative data and a double/debiased machine learning approach (Chernozhukov et al. (2018); Ahrens et al. (2024)) to estimate the duration profile in recruiter outcomes, after controlling for characteristics that we observe in the data.[2] Using this statistical approach, which holds the composition of *observable* characteristics constant, we find that the interview probability

[1]When we refer to duration dependence in what follows, we always mean the "within-individual" duration profile of the respective variable. This profile is, by construction, not driven by a change in the composition of the unemployment pool.

[2]Mueller and Spinnewijn (2023) also use a machine learning approach involving stacking to analyze the determinants of job finding/long-term unemployment, and decompose job finding into a component due to persistent and observable types vs potentially transitory unobserved heterogeneity, without explicitly modeling duration dependence. Our objective is to decompose job finding into duration dependence vs selection on observed and unobserved characteristics, and we model duration dependence explicitly.

reduces from 5 percent in month 1 to 3.5 percent in months 12-15 – a smaller decline than that observed in the cross-section, indicating negative dynamic selection on observables. The same procedure also reveals that accounting for observable characteristics does not change the estimated effect of duration on the probability to obtain a job offer after an interview.[3] Finally, with the same procedure we also look at the probability of a job offer per application (which summarizes both interviews and job offers per interview). We find that the negative duration profile in the cross-section becomes flatter when we account for observables.

Our empirical approach allows us to observe how search effort (as measured by the number of applications) changes with duration, and how recruiters respond to applications by job seekers on UI benefits. The joint observation of job seekers' applications and recruiters' responses enables us to address – at least in part – the pervasive concern of dynamic selection on *unobservables* when estimating duration dependence in interviews and job offers.[4] We do so by using the individual fixed effect in applications – a proxy for one's unobserved *search type* – as a regressor in the interview and job offer models. We find that a larger individual fixed effect in applications is associated with a significant (and quantitatively large) reduction in both the interview probability per application and the job offer probability per interview. In other words, we find that unobserved, duration-invariant determinants of the number applications sent are negatively correlated with employers' willingness to interview and hire the applicant. This empirical regularity will be important in our theoretical framework and will be a targeted moment in our structural estimation.

Overall, our empirical analysis suggests that duration dependence may play a significant role in the observed decline in the job finding rate. However, the fact that we can only imperfectly account for unobserved heterogeneity leaves our empirical analysis inconclusive. For this reason, we set up a structural model to obtain a consistent decomposition

[3] This is actually what one would expect if the information we observe in our data captures what an application reveals to the recruiter. Observables will then affect the interview decision, while unobservables will determine the job offer decision. Since our data do not only capture the relevant socio-economic characteristics but also an applicant's previous work history, our data covers many of the characteristics that applicants document in their resume.

[4] Jarosch and Pilossoph (2019) show that, in a model where employers statistically discriminate against the long-term unemployed, the negative effect of unemployment duration on employer callback rates estimated in correspondence testing studies, such as Kroft et al. (2013), can be almost entirely explained by dynamic selection on unobservables.

of the falling job finding rate into duration dependence and dynamic selection.

In our model, heterogeneous job seekers, differing in ability and search type, and heterogeneous firms, differing in ability requirements, interact in a frictional labor market.[5] The duration profile of the job finding rate is the outcome of application decisions by job seekers, and of interview and job-offer decisions by firms. Applications by job seekers of higher search type are more likely to be considered by a firm, but are subject to a higher unit cost. Workers' application behavior is driven by the net benefit of search, which varies both across job seekers, due to differences in search type and ability, and over the course of an unemployment spell for a given individual, due to firms' recruiting behavior. Firms' recruiting behavior, in turn, is driven by incomplete information about an applicant's ability, which is revealed after a costly interview (Jarosch and Pilossoph, 2019). This interaction gives rise to a model of endogenous search effort, in which firms statistically discriminate against the long-term unemployed.

This theoretical framework is able to replicate our empirical findings in equilibrium. First, individual applications reduce over an unemployment spell, due to a *discouragement effect* induced by a declining probability to obtain a job offer for a given application. Second, long-term unemployed apply more at any duration because their unit application costs are lower – *e.g.* they spend less time locating job openings, and most of their applications are unsolicited. Since recruiters are less willing to interview long-term unemployed applicants, the individual fixed effect in applications is negatively correlated with the interview probability. Finally, the probability to obtain a job offer per application declines over the unemployment spell, whereas the probability to obtain a job offer per interview increases. The job offer probability per application falls due to declining average job seeker ability, as negative dynamic selection and duration dependence reinforce each another. At the same time, the pool of job seekers becomes increasingly more homogeneous with duration. Hence, only firms willing to hire lower-ability workers continue to interview the long-term unemployed, which leads to an increasing job offer probability per interview.

---

[5]In our model, "search types" capture the heterogeneity across job seekers both in search costs and in search efficiency. We assume that search cost and search efficiency are positively correlated: Job seekers either face a high cost per application but also a high probability that the application is considered by the firm, and vice versa. This assumption allows the model to replicate the empirically observed negative correlation between the individual fixed effect in applications and the interview and job-offer probabilities.

We estimate the structural model by targeting the observed cross-sectional properties and duration profiles of applications, interviews, and job offers. To capture the relevant job seeker heterogeneity, we also target moments from the distribution of individual fixed effects in applications. We rely on a formal assignment algorithm to map each model parameter to the targeted moment to which it is most informative (Andrews et al., 2017). The estimated model allows us to quantify the relative importance of duration dependence and dynamic selection. Our decomposition analysis reveals that the decrease in the observed job finding rate (from 7 percent in month 1 to 4.5 percent in months 12-15 of the unemployment spell) is driven by both duration dependence and dynamic selection – albeit in different proportions. According to our estimates, duration dependence accounts for 61% of the observed job finding rate decline and comes about both from reduced search effort by job seekers (35%) and reduced job offers by employers (26%). Observed and unobserved heterogeneity account for the remaining 39%. Neglecting endogenous search effort would significantly underestimate the role of duration dependence.

We conduct extensive robustness checks on the decomposition of the job finding rate decline across alternative structural specifications of the job search process. Incorporating additional mechanisms behind declining application effort – such as learning from search, reference-point adaptation in consumption, and duration-dependent application costs – slightly lowers the estimated share of duration dependence relative to our baseline. Nevertheless, our conclusion that duration dependence is the dominant driver of the decline in the job finding rate remains robust to these structural modeling choices.

The remainder of the paper is organized as follows. In Section 2 we discuss the related literature. Section 3 describes the institutional context and the data we use for our empirical analysis. Section 4 studies duration dependence in applications, interviews and job offers based on the search diary data. In Section 5 we develop our theoretical framework. In Section 6 we estimate the structural model and carry out the decomposition of the duration profile of the job finding rate. Section 7 discusses alternative mechanisms that might explain our findings. Section 8 concludes.

2. Related literature

Our paper is related to a foundational literature, dating back to Lancaster (1979), Heckman and Singer (1984), and Van den Berg and Van Ours (1996), that developed suitable econometric models to disentangle duration dependence from dynamic selection. A number of recent papers have extended these approaches. Kroft, Lange, Notowidigdo, and Katz (2016) and Alvarez, Borovičková, and Shimer (2023) highlight the potential importance of duration dependence in job finding by estimating flexible models, where duration dependence is introduced in reduced form. Ahn and Hamilton (2020), Mueller and Spinnewijn (2023), and Ahn, Hobijn, and Şahin (2023) find that heterogeneity is the most important driver behind the falling job finding rate, and the dynamics of labor markets more generally.

Another related strand of literature focuses on how search effort varies with the duration of unemployment, with several studies based on repeated surveys (Krueger, Mueller, Davis, and Şahin, 2011; Mueller, Spinnewijn, and Topa, 2021; DellaVigna, Heining, Schmieder, and Trenkle, 2022) or data from online job boards (Faberman and Kudlyak, 2019; Fluchtmann, Glenny, Harmon, and Maibom, 2021). Many (though not all) of these studies find a limited role of unemployment duration on search effort of workers. Several recent papers have documented that search effort varies systematically around exhaustion of UI benefits (Marinescu and Skandalis, 2021; DellaVigna, Lindner, Reizer, and Schmieder, 2017; DellaVigna, Heining, Schmieder, and Trenkle, 2022). Other papers have explored the effects of changes in search strategies on the job finding rate (Belot, Kircher, and Muller, 2018) and of training programs on interview outcomes (Falk, Lalive, and Zweimüller, 2005). Using Swedish data on search effort and interview callbacks (but not job offers), Cederlöf and Roman (2025) report negative duration dependence in applications prior to UI exhaustion dates, and mild duration dependence in interview callbacks.

Correspondence testing studies have investigated whether callback rates are lower for long-term unemployed workers. Kroft, Lange, and Notowidigdo (2013), Oberholzer-Gee (2008), Eriksson and Rooth (2014) and Nüß (2018) find evidence in favor of that hypothesis for the US, Switzerland, Sweden and Germany, respectively. However, Farber, Silverman, and Von Wachter (2016) do not find an impact of duration on callback rates.

On the theoretical side, our paper relates to the structural literature on duration dependence in labor market outcomes. Duration dependence in the job offer rate has been explained by models of skill depreciation during unemployment (Ljungqvist and Sargent, 1998, 2008; Kospentaris, 2021), ranking by unemployment duration among multiple applicants (Blanchard and Diamond, 1994; Fernández-Blanco and Preugschat, 2018), and statistical discrimination against long-term unemployed (Vishwanath, 1989; Lockwood, 1991; Jarosch and Pilossoph, 2019; Baydur and Xu, 2024). Duration dependence in re-employment wages has been analyzed by search models with incomplete information and learning about individual job prospects (Burdett and Vishwanath, 1988; Gonzalez and Shi, 2010; Doppelt, 2016). Duration dependence in search effort has been studied by models of reference-dependent preferences, duration-dependent search costs, learning and biased beliefs about job-finding prospects (DellaVigna, Lindner, Reizer, and Schmieder, 2017; DellaVigna, Heining, Schmieder, and Trenkle, 2022; Potter, 2021; He and Kircher, 2023). We contribute to this literature by proposing a theory of endogenous search effort by workers and statistical discrimination by firms. To the best of our knowledge, our analysis is the first to quantify the role of job seekers' applications and firms' job offers for the falling job finding rate in a unified framework.

3. Institutional context and data

Our analysis focuses on job seekers in Switzerland who receive UI benefits. The Swiss UI system is rather generous. UI benefits equal 70% of previous earnings – or 80% for low income earners or job seekers with dependents – with maximum duration of 18 months. As in most UI systems, job seekers in Switzerland who receive UI benefits are obliged to actively search for new jobs. To comply with Swiss UI regulations, job seekers must document their search effort (*i.e.* applications sent) in monthly search diaries (see Figure A1 for the diary format). Search diaries are used to monitor whether job seekers have fulfilled their monthly search requirement, which caseworkers set typically at the beginning of the unemployment spell. Search requirements remain constant throughout the spell for most job seekers, and, on average, are not binding – that is, the number of applications typically exceeds the search requirement (Arni and Schiprowski, 2019). In meetings with the caseworker, search diaries are discussed and updated (to keep track of application outcomes in the current and previous unemployment-months). To verify the accuracy of the information, caseworkers review copies of the resumes and check on a random basis with employers whether the application has indeed been sent, or whether an applicant has shown up for a job interview. Non-compliance with these obligations may lead to a benefit sanction – a temporary benefit reduction or even a removal of

UI benefit payments.[6] This means that unemployed workers have a strong incentive to provide correct information in search diaries. Yet, search requirements and benefit sanctions provide a lower bound on the decline in job applications over an unemployment spell.

The search diary data used for this study were collected between April 2012 and March 2013 in five Swiss cantons (Zurich, Bern, Vaud, Zug and St. Gallen). All workers who were unemployed in April 2012 or entered unemployment between April 2012 and March 2013 are included in the analysis (combined stock-flow sample). Search diary forms contain detailed information on the number of applications made by the job seeker in each month of the unemployment spell (one diary per month). Importantly for our analysis, search diaries report information on each application's outcomes (job interview, job offer, negative or still open). Search diaries also include information on application dates, application channels (written, personal or by phone), and the work-time percentage of targeted positions (full-time or part-time).

Our dataset includes 58,755 search diaries documenting more than 600,000 job applications and their outcomes (job interview, job offer) for around 15,000 job seekers on UI benefits. Our search diary data have two key advantages. First, they can be linked to the Swiss unemployment insurance register, which reports job seekers' socio-economic and demographic characteristics, and to the Swiss social security register, which provides information on workers' previous and subsequent earnings and employment history. Second, they capture the behavior of both job seekers and recruiters simultaneously, thus allowing us to quantify the relative importance of applications, interviews, and job offers in driving the duration profile of the job finding rate.[7] We restrict our analysis samples

[6] About half of all benefit sanctions are due to failures to comply with benefit eligibility rules before the spell starts, *e.g.* job seekers who quit or do not look for jobs before they start claiming. The remaining part of benefit sanctions are due to failures to comply with search requirements. Benefit sanctions tend to raise the exit rate from unemployment (Lalive et al., 2005). Around one quarter of the job seekers in our sample face a sanction at some phase of their spell, and these job seekers have a slightly higher number of applications than those never confronted with a benefit sanction, even if differences are small and quantitatively unimportant.

[7] When interpreting applications as workers' search behavior, and interviews and job offers as firms' decisions, some caveats should be noted. The number of applications sent may be partly driven by UI compliance rules. For instance, some applications may be merely sent to fulfill search requirements or because of an assignment by the caseworkers. An employer's response to an application may be influenced by the quality of the application and a worker's behavior during the job interview. We argue that, to the extent these confounders do not vary in a systematic way with duration, they should not bias the estimated contribution of supply and demand factors to the falling job finding rate (Appendix B.2.2).

to those job applications made in months during which a job seeker receives UI benefits, *i.e.* up to 17 months of unemployment. This is motivated by data reliability: only job seekers drawing UI benefits have the legal obligation to fill in search diaries, and the recorded information is checked by caseworkers. Additionally, we focus on individuals for whom information on socio-demographic characteristics and the employment history is non-missing – these pieces of information playing an important role in our identification strategy. We remove job seekers who return to the previous employer, as job search after a temporary layoff substantially differs from job search after a permanent layoff (Nekoei and Weber, 2020).

A possible limitation of the search diary data is that some applications are censored, meaning that the outcome of the job application remains unknown. However, since censoring in applications does not vary with unemployment duration (see Figure B6 in the Appendix), it is unlikely that the estimated duration profiles are systematically biased. Moreover, we show below that the number of job offers we actually observe is very much in line with the number of people leaving unemployment (Figure 1), suggesting that censored applications would typically not have resulted in a job offer. For these reasons, we integrated censored job applications into the baseline analysis sample and code the response to these censored applications in the same way as a rejection. Our results are not sensitive to treating censored applications as rejections or to removing them from the pool of applications (see Figures B7, B8A and B8B in the Appendix).[8]

The outcome of main interest is the *job finding rate* – as measured by the probability that an individual receives at least one job offer from applications sent during a given month (note that all suitable job offers must be accepted under the Swiss UI system). Hence, our measure of the job finding rate is purely based on search diaries and links the job finding event to the month when the application was made. Table 1 reports descriptive statistics on the job finding rate, as well as on applications, job interviews, and job offers. The average monthly job finding rate is 6.1 percent. Job finding is the result of job seekers applying to jobs and firms responding to these job applications. Job seekers report an average of about 10.5 applications per search diary. Firms invite applicants to

[8]The search diary data do not provide information on the characteristics of the vacancies to which job seekers apply. In an auxiliary but smaller data set we can observe certain vacancy characteristics. Appendix Table A1 provides descriptive statistics of this auxiliary sample and compares it to the main sample. We use the vacancy information in the auxiliary data to discuss the relevance of changes in targeting of search in Section 7.

an interview with a 4.0 percent probability, and interviewees receive a job offer with a 22.5 percent probability. The probability that an application yields an interview and a job offer is 0.9 percent. This means that, on average, job seekers need to send more than 100 applications to receive one job offer.

Figure 1 shows that the job finding rate decreases with the duration of unemployment (bold line). The job finding rate is around 7 percent in the first three months in unemployment and falls below 5 percent later in the unemployment spell. We validate the information content of search diaries in two ways. First, we compare the duration profile of the job finding rate as measured in the search diary to the transition rate from unemployment to employment as observed in the social security data (Figure 1). Because search diary data can be linked to the social security data at the individual level, both graphs are conceptually similar and based on the same population at risk. Figure 1 shows that the two graphs have similar slopes, though the transition rate is located to the right of the job finding rate. The reason is that the job finding rate (as defined here) refers to the month when the application was sent, while the transition rate from unemployment to employment refers to the month when a job was actually started.[9]

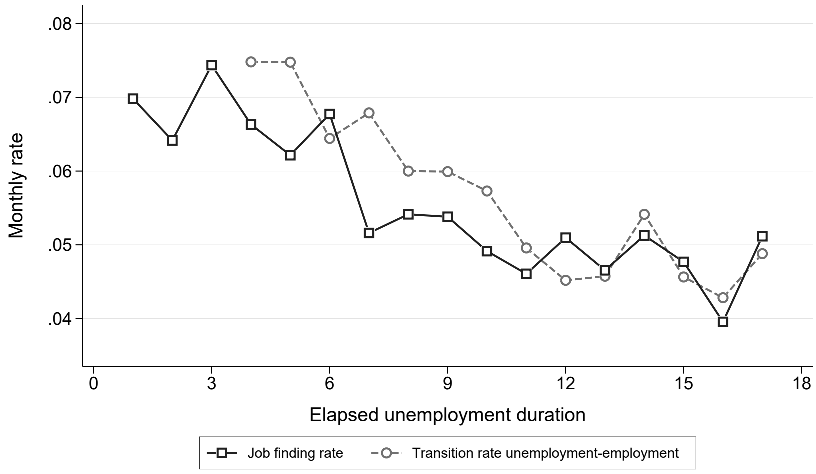
Table 1: Main outcome variables, mean (std. dev.)

A. Person-month level (search-diary level)		
Job finding rate (per month)	0.061	(0.239)
Number of applications (per month)	10.553	(4.698)
B. Application level		
Interview probability (per application)	0.040	(0.196)
Job offer probability (per interview)	0.225	(0.418)
Job offer probability (per application)	0.009	(0.095)

Note: This table reports descriptive statistics on the job finding rate, applications, interviews, and job offers. The interview probability is the probability of at last one interview for all applications in a search diary.

[9] In our definition of the job finding rate, the population at risk includes unemployed sending applications during duration month t; for the transition rate from unemployment to employment, the population at risk comprises all individuals with an elapsed duration of unemployment of t months. Search diaries contain information on the month when the application was made, but not on the month when the interview took place nor the month when the job was offered or started. Hence we assign interviews and job offers to the month when the eventually successful application was made. The job finding rate (as measured from search diary data) would be identical to the transition rate from unemployment to employment (from the social security data) under two conditions. First, a job seeker who obtains at least one job offer during month t always accepts an offered job. This condition is mostly met, since job search requirements oblige job seekers to accept suitable job offers. Second, if the successful application was made in month t of the unemployment spell, the start of the new job needs to be in the same month. This is usually not the case. Because recruitment decisions take time, the month when the application was made usually precedes the month when the job is started. As we do not know the identity of the recruiter in the search diary, we cannot directly check whether the new firm as observed in the social

Figure 1: Monthly job finding and unemployment-to-employment transition rates



Note: This figure depicts the empirical duration dependence in the job finding rate (computed from search diaries data) and the monthly unemployment-to-employment transition rate (computed from social security data).

Our second validation exercise is based on income trajectories in the social security data after the last job offer observed in the search diary data. Appendix Figure A2 confirms that labor earnings are close to zero during the months prior to the last job offer and increase sharply in the subsequent 2-3 months.[10] This makes us confident that the information on job finding in the search diaries is indeed predictive of taking up a regular job.

Figure 2 shows the empirical (cross-sectional) duration profiles of the number of job applications, the probability of a job interview (per application), and the probability of a job offer (per interview). Panel A shows that applications decrease from close to 11 in the first months of the unemployment spell to slightly less than 10 after 12 months or more. Panel B shows that the probability that an application receives a callback for a job interview declines from about 5 percent to only 2.5 percent after 15 months or more. Interestingly, the probability that an application leading to an interview results in a job offer – the job offer probability per interview – weakly *increases* with duration (Panel C). Early in the unemployment spell this probability is around 20 percent, increasing up to 30 percent at long durations.[11]

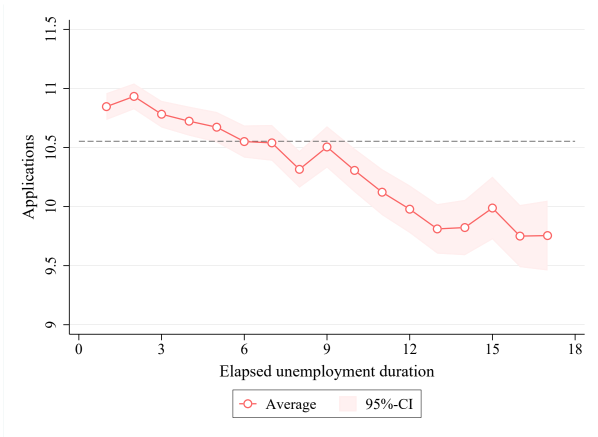
security data is identical to the employer who made a job offer to the job seeker.

[10]The vast majority of job seekers stop searching for jobs within three months after submitting the application that led to the *first* job offer, so the last job offer they receive is highly likely their first.

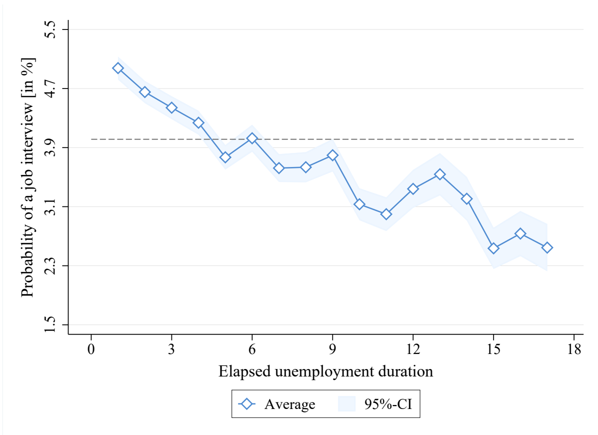
[11]If we exclude the first two months of unemployment, the estimated duration coefficient remains positive but is no longer statistically different from zero, partly reflecting the loss of precision from removing the months with the largest number of observations. For this reason, we interpret the duration profile

Figure 2: Empirical duration profiles

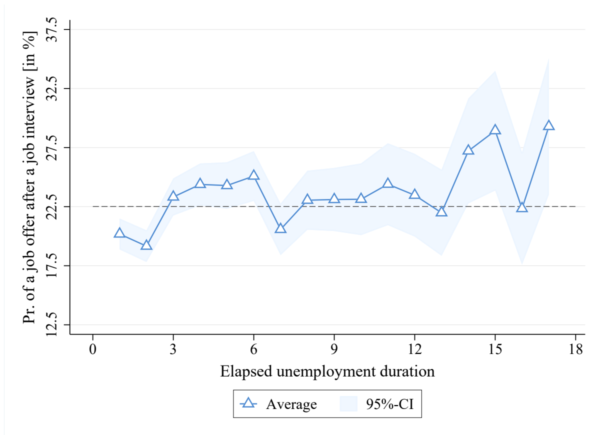
(A) Applications



(B) Interview probability



(C) Job offer prob. (after an interview)



Note: This figure depicts the empirical duration profiles of the average number of job applications submitted per person-month (Panel A), the share of applications that result in interviews, *i.e.* the application-level probability of a job interview (Panel B), and the share of all interviews that result in a job offer, *i.e.* the probability of receiving a job offer conditional on an interview (Panel C). Dashed horizontal lines indicate sample averages.

It is worth noting at this stage that, with respect to job applications and job interviews, the descriptive evidence is in the ballpark of what other studies have documented in different contexts. For instance, Faberman and Kudlyak (2019) find a decreasing profile of job applications in online job board data. The correspondence testing study of Kroft, Lange, and Notowidigdo (2013) finds callback rates of a very similar order of magnitude and a strong downward sloping duration profile. To our knowledge there is no other paper that has documented how the probability of a job offer after an interview changes with duration.

of job offers per interview as *weakly* increasing.

4. Empirical evidence on applications, interviews, and job offers

We now exploit our search diary data to study how (i) the number of job applications, (ii) the probability of an interview (per application), and (iii) the probability of a job offer (per interview) change with the duration of unemployment. These three outcomes jointly determine the job finding rate. For each stage of the job search process, we are interested in exploring whether duration dependence and dynamic selection move in the same or in opposite directions, *i.e.* whether dynamic selection reinforces, or masks, true duration dependence.

By definition, duration dependence arises within individuals. Hence, a fixed effects model that accounts for duration-invariant individual heterogeneity is the natural empirical approach. For job applications, such an approach is feasible: applications are recorded repeatedly during the unemployment spell, with varying frequency over time. Crucially, unobserved shocks to the number of applications are not directly related to attrition, *i.e.* exits from the unemployment pool (recall that less than 1% of applications lead to a job offer). This condition does not hold for interviews and job offers. Interviews are rare events that are strongly predictive of unemployment exit (22.5% of interviews lead to a job offer). Job offers mostly occur only once at the end of the spell, as job seekers must accept the first suitable job offer they get. Since unobserved shocks to interviews and job offers are strongly correlated with unemployment exit, within-spell duration dependence estimates based on the fixed-effect estimator are mechanically upwards biased (Verbeek and Nijman, 1992; Wooldridge, 1995).[12] Hence, a spell fixed effects strategy is not feasible for these outcomes.

In studying duration dependence in interviews and job offers under dynamic selection, we rely on a selection-on-observables strategy that leverages the information on job seekers and applications contained in our rich administrative data. We adopt a flexible double-debiased machine learning approach (DDML) to control for observed job seeker and application characteristics.[13] DDML consists of a two-step procedure. In the first

[12] Zuchuat (2025) provides extensive simulation evidence of severe within-estimation duration bias for interviews and job offers in a setting close to the one we study.

[13] Mueller and Spinnewijn (2023) also use machine learning methods building on the approach proposed by Einav et al. (2018) to study the determinants of long-term unemployment. While they are interested in predicting the risk of long-term unemployment, our goal is to back out duration dependence in the probability of a job interview and a job offer.

step, we regress an individual's outcome of interest (interview or job offer) and the individual's unemployment duration on observed characteristics. The residual of the outcome regression and that of the unemployment-duration regression contain only the variation unexplained by the observed characteristics. In the second step, we regress the outcome residual on the unemployment-duration residual. This yields the effect of duration on the outcome after accounting for the effect of observed characteristics on both the outcome of interest and the individual's unemployment duration.[14]

Notice that, despite our rich register data, it is unlikely that the DDML approach captures all relevant heterogeneity determining outcomes and unemployment duration, so unobserved heterogeneity may still confound our estimates. Just as we, as researchers, cannot observe all characteristics affecting interview and job-offer decisions, recruiters themselves make hiring decisions under imperfect information. After all, recruiters call back applicants for interviews precisely to figure out whether the applicant is suitable for the open position. Overall, we argue that the set of observable characteristics we can control for overlaps to a large extent with the information set of recruiters prior to job interviews.

To examine the importance of unobserved heterogeneity that may remain after controlling for observed heterogeneity, we take advantage of the individual fixed effect we obtain from the application regression. To the extent that unobserved factors influencing application behavior are correlated with those affecting search efficiency and/or employers' willingness to hire a worker, the individual fixed effect in applications should help predict interviews and job offers by capturing variation in these underlying characteristics.

Hence, we implement our analysis of duration dependence in interviews and job offers in two steps. First, we estimate residual duration profiles for both outcomes controlling for the observable characteristics of the job seeker and of the application that also a firm can likely infer from the application materials. Second, we include the estimated individual fixed effect in applications as an additional explanatory variable in the equations for

[14] This two-step approach is essentially a generalization of the Frisch-Waugh-Lovell Theorem of partitioned regression (Frisch and Waugh, 1933; Lovell, 1963). We implement it using a stacked version of DDML (Ahrens et al., 2024; Chernozhukov et al., 2018), which is robust to uncertainty about functional form and the relevance of controls, and optimally addresses potential overfitting. The stacked DDML combines an ensemble of base learners with complementary strengths – including standard parametric models, lasso and ridge regression, random forests, and gradient boosted trees – allowing the first-step estimations to be conducted in a flexible, data-driven manner that enhances robustness to misspecification and model selection errors (Ahrens et al., 2024; Breiman, 1996; van der Laan et al., 2007; Wolpert, 1992). Appendix B.1 provides further details.

interviews and job offers to explore the role of unobservable characteristics which are unknown to firms prior to a job interview.[15]

The individual fixed effect in applications may not fully capture all relevant unobserved heterogeneity influencing interview and job offer decisions – it is likely only an imperfect proxy for such heterogeneity. Therefore, our estimates of residual duration dependence in interviews and job offers may still be biased if remaining unobserved heterogeneity correlates with unemployment duration even after controlling for individual fixed effects in applications. Hence, we refrain from interpreting the estimated residual duration dependence as true duration dependence. To address this issue, Section 5 develops a structural model where the individual fixed effect explicitly serves as an imperfect proxy for the unobserved worker search type. Importantly, we estimate this structural model to match empirical moments of the distribution of individual fixed effects from the data, replicating the fixed effects estimation procedure on model-generated data – thereby incorporating potential selection bias.

Job applications. To disentangle duration dependence from dynamic selection in the number of applications, our preferred empirical strategy relies on a fixed-effects approach. Specifically, we model the number of applications, A_{it}, of job seeker i in month t of unemployment as

$$A_{it} = \alpha_i + f^A(t)\phi^A + X_{it}\beta^A + \varepsilon_{it}^A, \tag{1}$$

where α_i is the individual fixed effect, X_{it} contains a full set of calendar-quarter-by-local-labor-market dummies as well as an indicator for unemployment exit in the following month, and ε_{it}^A an idiosyncratic error term.[16] We include a dummy for unemployment exit in the following month to account for non-random attrition, *e.g.* potential anticipation of unemployment exit, that could bias our estimates through non-random attrition (see

[15] More precisely, we feed a cross-fitted and smoothed estimate of the individual fixed effect into the stacked DDML procedure. This cross-fitted and smoothed estimator is consistent for the underlying individual fixed effect. The DDML framework then allows us to partial out the influence of observed characteristics and to obtain valid inference despite the presence of the estimated individual fixed effect as a generated regressor. See Appendix B.1 for details.

[16] All the other observed variables are time-constant and therefore drop out when estimating eq. (1) with the fixed effects estimator. This approach has also been applied in other recent work studying application effort (Faberman and Kudlyak, 2019; Marinescu and Skandalis, 2021; Fluchtmann, Glenny, Harmon, and Maibom, 2021).

Appendix Table B5 for details). The function $f^A(t)\phi^A$ captures duration dependence in the number of applications after partialling out the influence of individual observed and unobserved characteristics, X_{it} and α_i. We therefore refer to it as *residual* duration dependence.

We consider two different specifications of residual duration dependence: one where $f^A(t)\phi^A = \phi^A t$ represents a linear relationship, and another where $f^A(t)$ is modeled as a step function, corresponding to a complete set of dummy variables – one for each value of elapsed unemployment duration – with ϕ^A as the associated coefficient vector. The linear specification of duration dependence serves as a summary measure of the period-by-period effects recovered with the step function, as the duration coefficient of the linear specification corresponds to a positively weighted average of the period-by-period changes (Ahrens et al., 2024).

As a comparison, we also estimate models without individual fixed effects. This alternative approach allows us to assess how job seekers' observed characteristics influence the estimated residual duration dependence (Appendix Table A2 provides an overview of these characteristics). We consider two specifications that differ in the flexibility with which they capture observable heterogeneity: (i) a basic OLS model including a hand-selected set of covariates in linear form, and (ii) a flexible machine learning model based on a comprehensive dictionary of covariates that includes the original variables along with their interactions, polynomials, and other transformations. The latter model is estimated using the stacked DDML approach that we also use for interviews and job offers (see Appendix B.1 and B.2.1).

In Table 2, we present our results for duration dependence in the number of monthly job applications. Columns 1-3 account for observed characteristics, column 4 accounts for individual fixed effects (our main specification). We report coefficients when the effect of duration is specified linearly, *i.e.* $f^A(t)\phi^A = \phi^A t$. Standard errors are clustered at the individual level and reported in parentheses. Coefficients in relative terms are reported in square brackets. According to the empirical duration profile, shown in column (1), the number of applications per month decreases by 0.078, or around 0.72%, every month. The effect is attenuated when observed characteristics are added to the model using a hand-selected set of individual control in an OLS regression (column 2). The stacked DDML estimates, in column (3), show that applications decline by about 0.048 per month, or around 0.44%, much in line with the OLS estimates in column (2). However, when we estimate a fixed effects model, the partial effect of elapsed unemployment duration sharply increases in absolute value, suggesting a decline in the number of applications of 0.21 per month, or 1.9% (see column (4)). This suggests that the individual fixed effects capture a dimension of individual heterogeneity that is quite distinct from what can be explained by the observed individual characteristics in our data.

In Appendix B.2.1, we explore to what extent the observed individual characteristics can predict the estimated individual fixed effects. First, observed characteristics explain only 31% of the variation in the estimated fixed effects (see Table B2). Hence, unobserved individual heterogeneity accounts for the bulk of the variance in individual fixed effects in applications. Second, we find that observed individual characteristics and regional differences each contribute about 50% to the explained variation in the individual fixed effects in applications. At the same time, we find that the different methods (from OLS to linear and nonlinear machine learners) achieve similar predictive performance while attributing importance to different groups of variables. Combined with the fact that predictive performance is good but far from perfect, this makes it difficult to tie the individual fixed effect to any specific observable trait of job seekers.

Overall, our analysis reveals substantial cross-sectional variation in the individual fixed effects in applications, which is largely unexplained by observable characteristics. This supports interpreting these fixed effects as reflecting worker-level heterogeneity in unobserved *search types*, rather than compositional differences across market segments.

Table 2: Duration dependence in applications, linear specification

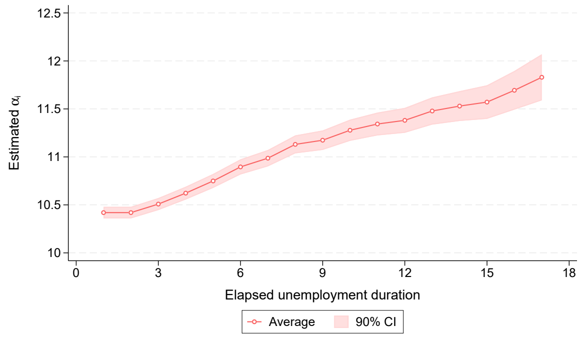
	(1)	(2)	(3)	(4)
Dependent variable: Applications per month				
Elapsed unemployment duration	-0.078***	-0.043***	-0.048***	-0.207***
	(0.008)	(0.007)	(0.008)	(0.021)
	[-0.718%]	[-0.395%]	[-0.440%]	[-1.907%]
Individual controls	No	Yes	Yes	No
Policy controls	No	Yes	Yes	No
Local labor market conditions	No	Yes	Yes	Yes
DDML	No	No	Yes	No
Individual FE	No	No	No	Yes
Mean outcome 1$^{\text{st}}$ month	10.846	10.846	10.846	10.846
Adjusted-R^2	0.005	0.158	0.002	0.041
Observations	58755	58755	58755	58755
Persons	14798	14798	14798	14798

Note: This table reports estimates of duration dependence using OLS (columns 1-2) and double debiased machine learning (3) as well as fixed effects regression as in eq. equation (1) (column 4), where duration dependence is specified linearly, *i.e.*, $f^A(t)\phi^A$. Standard errors are clustered at the individual level and reported in parentheses. Coefficients in relative terms (relative to the average in the first month of unemployment) are reported in square brackets. Stars indicate the following significance levels: * 0.1, ** 0.05 and *** 0.01.

Figure 3: Duration profile of applications

(A) Empirical and residual duration profile

(B) Average individual fixed effect by elapsed unemployment duration



Note: Panel A depicts the empirical duration profile of the number of job applications (solid line) and the estimated duration dependence obtained after controlling for time-varying observable heterogeneity and individual fixed effects (dashed line), with function $f^A(t)\phi^A$ in equation (1) modeled as a step function with one dummy for each unemployment month. The shaded area around the estimated duration dependence corresponds to the pointwise 90% confidence interval. Panel B depicts the average of the estimated individual fixed effects, α_i in equation (1), by month of elapsed unemployment. That is, the average is computed based on those individuals who are still unemployed at the respective unemployment month. Confidence intervals (shaded areas) are based on standard errors clustered at the individual level.

In Figure 3A we compare the empirical duration profile in the number of job applications with that obtained when we net out both observed and unobserved individual heterogeneity (true duration dependence), that is, after controlling for individual fixed effects and the time-varying covariates (see eq. (1)). The function $f^A(t)\phi^A$ is now specified as a step function with one dummy for every unemployment month. The duration-dependence graph is drawn such that the duration profile coincides with the empirical duration profile in month 1. In other words, the graph draws the application profile that would have emerged had the unemployment pool in any month t consisted of the same types of job seekers as the pool in month 1.[17]

Figure 3A reveals that the profile for residual duration dependence, adopting the fixed effects approach, decreases much more strongly than the empirical, cross-sectional, duration profile.[18] Since the number of applications falls much more strongly when the composition is kept constant, this means there is positive dynamic selection in application effort: those who eventually remain unemployed for longer submit more applications at all durations. Figure 3B shows the average of the estimated individual fixed effects of

[17] Figure B1 in the Appendix repeats this exercise by additionally displaying the duration profile estimated using OLS.

[18] Appendix Figure A4 provides direct evidence of a within-person decline in applications among job seekers who remain unemployed for at most 3 (or 6, 9, 12, or 15) months. Across all these sub-samples, applications decline much more sharply than in the full population of job seekers.

those job seekers who are still unemployed in the respective unemployment month. Those still in unemployment at long elapsed durations have a higher fixed effect α_i than those who leave unemployment quickly.[19]

Our interpretation that the decline in applications over the unemployment spell reflects a job seeker's intrinsic behavior would not be valid if application numbers primarily reflect search requirements or vacancy referrals imposed by caseworkers. Similarly, if the anticipation of unemployment exit leads to unusually low or zero applications in the final month of unemployment, the estimated residual duration dependence would be biased downward. To address these concerns, we conduct a series of robustness checks, which we report in detail in Appendix B.2.2. Across specifications, restricting to applications not based on referrals and to applications exceeding the minimum search requirement, we consistently find a negative residual duration dependence that becomes even steeper once individual fixed effects are controlled for. Our results are also robust to attrition and alternative treatments of the final month before exit from unemployment. We therefore conclude that there is strong negative duration dependence and strong positive dynamic selection in the number job applications.

We also estimated models that allow for heterogeneity in the duration dependence of applications with respect to observables, *e.g.* age, education or nationality (see Figure B4 in the Appendix). With few exceptions, we find that the decline in job applications, as spells lengthen, is homogeneous across age groups, education levels, and for people with different nationality, whether we look at it in the raw data or control for job seeker fixed effects.

Job interviews and job offers. We now turn to the effect of unemployment duration on the firm's response to an application – a callback for an interview or a job offer. We model the probability that an application j made by individual i in unemployment month $t \geq 1$ receives a positive response – $Y_{ijt} = 1$, where $Y = C$ in case of a callback for an interview, $Y = O$ in case of a job offer per interview, and $Y = U$ in case of a job

[19] Faberman and Kudlyak (2019) also report a strong within-person decline in applications in the US. In contrast, Marinescu and Skandalis (2021), using data on an online platform covering around 20% of all job seekers in France, find pronounced within-person increases in applications immediately prior to UI benefit exhaustion, with the duration profile otherwise relatively flat or mildly declining.

offer per application – as follows:

$$\mathbb{P}(Y_{ijt} = 1 \mid t, X_{ijt}) = H\big(f^Y(t)\phi^Y + g^Y(X_{ijt})\big), \tag{2}$$

where the function $H(\cdot)$ is a link function. The term $f^Y(t)\phi^Y$ denotes a linear (in ϕ^Y) specification of duration dependence for interviews and job offers ($Y = C, O, U$). We consider again two different specifications of residual duration dependence: a linear one with $f^Y(t)\phi^Y = \phi^Y t$ and a step function where $f^Y(t)$ represents a full set of dummy variables for each value of elapsed unemployment duration. Further, $g^Y(X_{ijt})$ denotes a potentially nonparametric function of a rich set of observed covariates X_{ijt}, including job seeker and application characteristics, as well as calendar quarter times local labor market fixed effects and regional labor market policy fixed effects. To estimate eq. (2) when $Y = O$ we restrict the sample to those applications that led to an interview, while we use the full sample (*i.e.*, the same as in the case of $Y = C$) when $Y = U$.

We exploit our data set to create a set of variables capturing the information a recruiter can typically extract from an application, *e.g.* gender, age, education, employment history (see Appendix Table A2). We consider standard logit models (where $H(\cdot)$ corresponds to the logistic function), but also linear probability models (where $H(\cdot)$ is the identity function) combined with non-parametric choices for the effects of covariates on outcomes $g^Y(\cdot)$. We use again either parametric regression or the stacked DDML approach for estimation (see Appendix B.1 and B.3 for further details).

Moreover, we estimate a variant of eq. (2) that includes the estimated individual fixed effect in applications as an additional control variable to proxy for the unobserved (to us, but potentially observed by the firm during an interview) job seeker type. We implement this augmented model using again the stacked DDML approach. Specifically, we first obtain a cross-fitted and smoothed estimate of the individual fixed effect of which we then partial out the influence of observed characteristics. Finally, we run a regression of the residualized outcome (interview or job offer) on the residualized unemployment duration and the residualized individual fixed effect in applications (see Appendix B.1 for more information on the procedure).

Table 3 shows estimates adopting a linear profile for residual duration dependence. The first two columns correspond to logit regressions, whereas the estimates shown in columns (3) and (4) correspond to a linear probability model using DDML to select covariates

Table 3: Duration dependence in job interviews and job offers

	(1)	(2)	(3)	(4)
A. Dependent variable: Application-level interview dummy				
Elapsed unemp. duration	-0.155***	-0.133***	-0.120***	-0.097***
	(0.015)	(0.015)	(0.014)	(0.015)
	[-3.117%]	[-2.668%]	[-2.414%]	[-1.950%]
Ind. fixed effect in applications (α_i)				-0.584***
				(0.099)
				[-11.724%]
Mean outcome 1st month	4.977	4.977	4.977	4.977
Goodness of fit	0.535	0.649	–	–
Observations	600323	600323	600323	600323
B. Dependent variable: Application-level job offer dummy (sample of applications that led to an interview)				
Elapsed unemp. duration	0.350***	0.357***	0.317***	0.372***
	(0.099)	(0.096)	(0.103)	(0.103)
	[1.736%]	[1.769%]	[1.572%]	[1.841%]
Ind. fixed effect in applications (α_i)				-1.687***
				(0.474)
				[8.357%]
Mean outcome 1st month	20.187	20.187	20.187	20.187
Goodness of fit	0.520	0.611	–	–
Observations	22422	22422	22422	22422
C. Dependent variable: Application-level job offer dummy				
Elapsed unemp. duration	-0.024***	-0.017***	-0.019***	-0.005
	(0.006)	(0.006)	(0.006)	(0.006)
	[-2.339%]	[-1.642%]	[-1.852%]	[-0.519%]
Ind. fixed effect in applications (α_i)				-0.346***
				(0.049)
				[-33.667%]
Mean outcome 1st month	1.027	1.027	1.027	1.027
Goodness of fit	0.518	0.640	–	–
Observations	600323	600323	600323	600323
D. Control variables and estimation strategy used in panels A-C				
Individual controls	No	Yes	Yes	Yes
Policy controls	No	Yes	Yes	Yes
Local labor market conditions	No	Yes	Yes	Yes
DDML	No	No	Yes	Yes

Note: This table reports estimates of duration effects on the probability of a job interview (A), the probability of a job offer per interview (B), and the probability of a job interview per application (C) according to equation (2) for a linear specification of residual duration dependence, *i.e.*, $f^Y(t)\phi^Y = t\phi^Y$. Columns (1) and (2) correspond to standard logit regressions with $g^Y(X_{ijt}) = X_{ijt}\beta^Y$, whereas columns (3) and (4) model $g^Y(X_{ijt})$ nonparametrically and $H(\cdot)$ is the identity link. Application-level observations are weighted by the inverse of the monthly number of applications made by individual i in month t, so as to put equal weight on all person-month observations. Point estimates correspond to average partial effects (in percentage points). In columns (1) and (2), goodness of fit is measured by the area under the receiving operating characteristic (ROC) curve. Partial effects in relative terms (relative to the average in the first month of unemployment) are reported in square brackets. In columns (1) to (4), standard errors (in parentheses) are clustered at the individual level. Stars indicate the following significance levels: * 0.1, ** 0.05 and *** 0.01.

and model the effects of observed covariates non-parametrically (Appendix B.3 reports intermediate estimation results from the stacked DDML approach). The estimated model underlying column (4) controls in addition for the standardized estimated individual fixed effect in applications, $\hat{\alpha}_i$ from eq. (1).

The probability of a job interviews decreases by 0.155 percentage points per month of unemployment (column (1) in panel A), which is about 3.1% of the mean interview probability in the first month of unemployment. Controlling for observed characteristics reduces duration dependence of interviews to -0.133 percentage points per month (column (2) in panel A).[20] DDML estimates suggest a duration dependence of -0.12 percentage points per month (column (3) in panel A). Duration dependence in interviews is -0.10 percentage points per month, holding constant job seeker fixed effects in applications (column (4) Panel A). Job seekers with higher individual fixed effects in applications tend to remain unemployed longer (column (4) in panel A), estimates suggest that a one-standard-deviation increase in the estimated individual fixed effect in applications (*i.e.* about four more applications per month) reduces the probability to be invited to an interview by 0.584 percentage points, which is around 11.7% of the baseline interview probability.

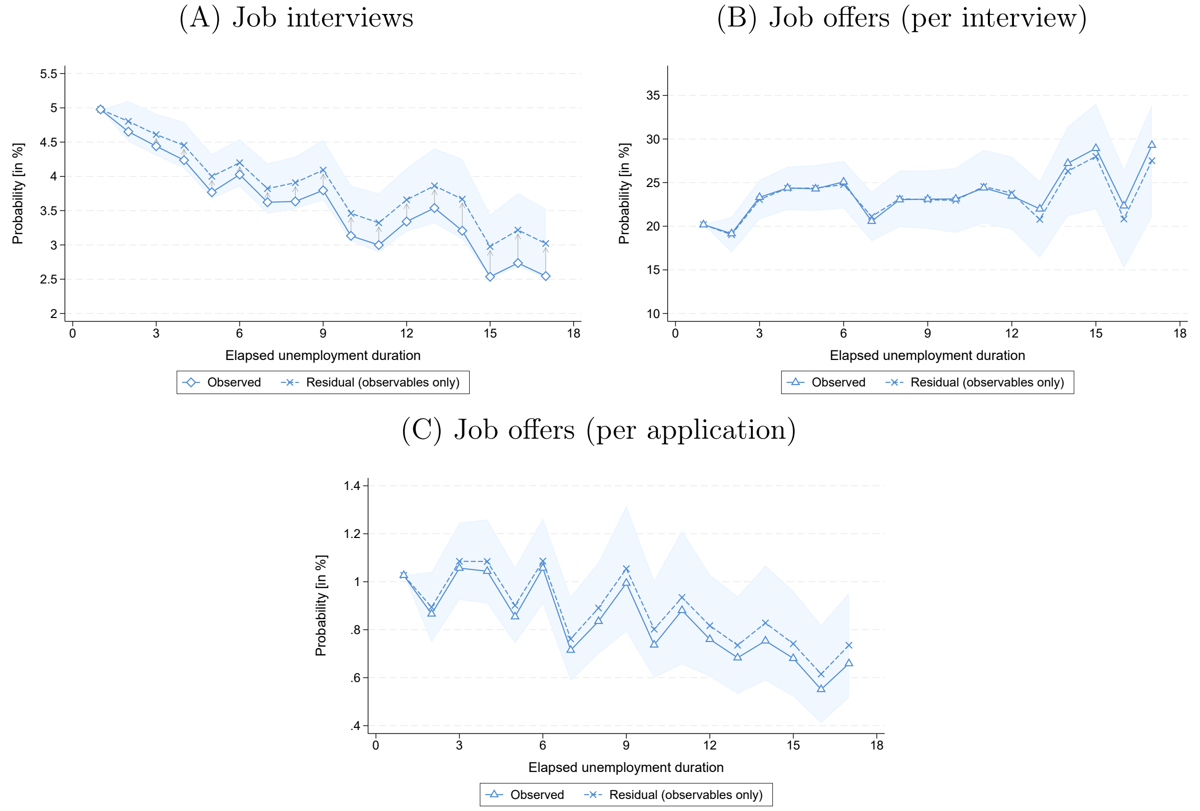
Comparing columns (1) and (4) in panel A, we see that, unlike for applications, dynamic selection is negative for job interviews, *i.e.* job seekers with lower chances to be interviewed remain unemployed longer, which introduces a negative bias in estimates of duration dependence.

Unlike the interview probability, the probability of a job offer per interview weakly *increases* with unemployment duration, by 0.32 percentage points per month, which is about 1.57% of the baseline job offer probability (panel B of Table 3, column (3)). Taken at face value, these estimates suggest that firms are more likely to offer a job to job seekers interviewed later in the spell – job interviews appear more targeted.A comparison of the estimates in columns (1) through (3) of panel B reveals that the estimated residual duration dependence declines only slightly, from 0.35 to 0.32 percentage points per month, as the conditioning on control variables becomes more comprehensive and as the impact of observed characteristics is modeled nonparametrically using stacked DDML. Like for interviews, a one-standard deviation higher individual fixed effect in applications reduces job offer chances by around 1.69 percentage points per month, or about 8.35% of the baseline job offer probability (column (4) in panel B). Hence, job seekers with a high individual fixed effect in applications have both a reduced chance of being called back to

[20] Goodness of fit, measured as the area under the receiver-operator curve (AUROC), is 0.649 in column (2). We do not report the AUROC in column (3), but Table B6 in the appendix, shows that the AUROC for the first stage model, that predicts job interviews without including duration, is 0.657.

an interview, and a reduced chance of receiving a job offer after an interview. Moreover, comparing columns (1) and (3) of panel B, we see that controlling for observed characteristics reduces duration dependence somewhat, but adding individual fixed effects in applications increases duration dependence again. Dynamic selection does not play an important role for job offers per interview.

Figure 4: Empirical and residual duration profiles of job interviews and job offers



Note: In each panel, the solid line depicts the empirical duration profile of the share of applications that result in interviews, the interview probability (panel A), the share of all interviews that result in a job offer, the job offer probability per interview (panel B), and the share of all applications that result in a job offer, the job offer probability per application (panel C). The dashed line the estimated duration dependence obtained after controlling for observable heterogeneity using double/debiased machine learning with stacking, with function $f^Y(t)\phi^Y$ in equation (2) modeled as a step function with one dummy for each month of elapsed unemployment duration. The shaded area around the estimated duration dependence corresponds to the pointwise 90% confidence interval.

Figure 4A summarizes the estimation results for the probability of a job interview, contrasting the profile of the empirical probability with the estimated one using stacked DDML. The figure shows that the empirical interview probability falls from 5 percent in the first month to 2.5 percent by month 15 of unemployment. In contrast, after adjusting for observed heterogeneity, the decline is substantially smaller, dropping from 5 to only 3 percent – equivalent to 0.13 percentage points per month (see Panel A of Table 3). These numbers suggest that 20 percent of the reduction in the interview probability can be attributed to dynamic selection on observables.

Our estimates of the decrease in the interview probability after controlling for observables are similar to those of Kroft, Lange, and Notowidigdo (2013), who document a 3.7 percentage points decline in the callback probability over a period of 36 months of unemployment, roughly 0.1 percentage point every month. Note, however, that the falling interview probability need not reflect true duration dependence. Instead, it may reflect additional heterogeneity in job seeker quality that is still unobserved at the point when the firm decides to call the job seeker back (Jarosch and Pilossoph, 2019). Controlling for individual fixed effects in applications – a proxy for job seekers' unobserved search type – allows us to alleviate this concern (at least partially). Upon doing so, the duration profile of job interviews flattens out further (see column (4) in Panel A of Table 3), hinting at negative dynamic selection on unobservable characteristics.

Figure 4B shows the corresponding results for the probability of a job offer per interview, again obtained using the stacked DDML approach. The probability of a job offer per interview increases with unemployment duration (Figure 4B) by around 0.35 percentage points per month (see Panel B of Table 3). Adjusting for dynamic selection based on characteristics observable to the firm at the time of application has minimal impact on the duration profile: the empirical, cross-sectional duration profile of job offers closely aligns with the duration profile that nets out observable factors. This suggests that job seeker and application characteristics explain less of the variation in the probability of receiving a job offer than they do in the probability of securing an interview (see also Tables B6 and B7 in Appendix B.3). This finding reflects rational behavior by firms: once a firm has decided to interview a job seeker, the hiring decision is largely based on information newly revealed during the interview, which is independent of the information already present in the job seeker's application. Similarly, allowing for individual fixed effects in applications does not affect the empirical duration profile significantly (see column (4) in Panel B of Table 3).

We have seen that the probability of a job interview declines with duration, whereas the probability of a job offer per interview weakly increases with duration. As a result, it is unclear which effect dominates when considering the chances that a job application leads to a job offer. Here, we present estimates of the probability that an application results in a job offer, accounting for all applications – including those that did not lead to an interview – using the same empirical strategy applied to interviews and job offers per interview.

Panel C of Table 3 shows estimates of residual duration dependence modeled linearly for the probability that an application results in a job offer. The job offer probability per application decreases for applications made later in the spell by 0.024 percentage points, or 2.4%, per month (column (1)). The empirical duration profile is negative, so – at a descriptive level – the negative duration dependence in interviews dominates the positive duration dependence in job offers per interview. Controlling for observables somewhat reduces the effect of one more month of unemployment on the job offer probability per application: stacked DDML estimates in column (3) show a reduction of 0.019 percentage points, or 1.9%, per month. Accounting for dynamic selection on observables in the job offer probability per application reduces duration dependence by 0.005 percentage points, or 0.5% – the difference between column (3) and (1) in panel C. Similarly, accounting for dynamic selection on observables reduces duration dependence in interviews by around 0.7% (columns (1) and (3) in panel A of Table 3), probably because the duration dependence in job offers after interviews remains unaffected by accounting for observables. The estimates in panel C, column (4) of Table 3 indicate that adding the individual fixed effect in applications to the model reduces the impact of unemployment duration by over 75%, bringing it down to 0.5 percentage points (0.5%). A one standard deviation increase in the individual fixed effect in applications significantly decreases the probability of receiving a job offer across all applications by 34.6 percentage points, or 33.7%. Again, there is negative dynamic selection on unobservables, *i.e.*, unobserved job seeker characteristics (to us, but potentially observed by the firm during an interview) play a key role in driving the negative duration profile of the job offer probability per application.

Figure 4C contrasts the empirical duration profile of the probability that an application results in a job offer with the residual duration profile that controls for observables. The empirical duration profile declines more strongly than the residual duration profile – the difference being small, though. However, column (4) in Panel C of Table 3 suggests that the residual duration profile that controls for both the individual fixed effect in applications and observables differs more strongly from the empirical profile. This happens for two reasons. First, dynamic selection introduces a particularly severe bias at the job interview stage, where it accounts for around 20% of the decline in job interviews. Second, the strong and positive duration dependence of job offers per interview, *e.g.* because these interviews are better targeted, also neutralize the negative duration dependence in job interviews.

Summary of empirical findings. We conclude this section by summarizing our main findings in the following three facts, which will guide our theoretical analysis below.

Fact 1 (Drop in individual applications)**.** *The number of applications displays negative duration dependence.*

Fact 2 (Heterogeneity in search types)**.** *The number of applications displays positive dynamic selection on unobservables. Job seekers who send more applications have a lower interview probability per application and a lower job offer probability per interview.*

Fact 3 (Differential duration profiles of job offers per application and per interview)**.** *For given observable characteristics, the job offer probability per application decreases with unemployment duration, whereas the job offer probability per interview weakly increases with duration.*

5. A theory of endogenous search effort and statistical discrimination

In this section, we develop a model of endogenous search effort by workers and statistical discrimination by firms. The model has two objectives. First, it rationalizes the three empirical facts outlined above within a consistent theoretical framework. Second, it provides a precise decomposition of the observed duration profile of the job finding rate into duration dependence and dynamic selection. Moreover, it allows us to quantify the relative importance of job seekers and recruiters in shaping overall duration dependence.

On the worker side, we build a job-search model with heterogeneous job seekers, differing in ability and search type, who optimally choose the number of applications to send over an unemployment spell. On the firm side, we consider an extended version of the Jarosch and Pilossoph (2019)'s model of statistical discrimination with a stochastic qualification probability and endogenous job creation.

The key novelty of our framework is that it allows for an interaction between workers' search behavior and statistical discrimination by firms. In our model, job seekers reduce their applications over an unemployment spell in response to a declining job offer probability per application, driven by firms' statistical discrimination. In other words, statistical discrimination by firms induces a *discouragement effect*.

Environment. We consider a discrete-time economy populated by a unit mass of workers, who differ in their on-the-job ability $x \in (\underline{x}, \overline{x})$ and search type $\gamma \in (\underline{\gamma}, \overline{\gamma})$, and by a continuum of firms differing in their productivity $y \sim G(y)$, $y \in (\underline{y}, \overline{y})$. Let $\mathcal{L}(x, \gamma)$ denote the joint distribution of ability and search type across workers. A worker's search type determines both the search efficiency of her applications, $\epsilon(\gamma)$, and the cost of drafting an application, $\psi(\gamma)$. As suggested by Fact 2, job seekers with higher search efficiency face higher unit application costs. Therefore, we assume that $\epsilon'(\gamma) > 0$ and $\psi'(\gamma) > 0$.

Both workers and firms are risk-neutral and discount the future at common rate $\beta \in (0, 1)$. Workers and firms interact in a frictional labor market under a sequential random search protocol. Search-and-matching frictions are represented by an exogenous separation probability δ_H and the endogenously determined job finding probability $f(x, \gamma, \tau)$. Following Jarosch and Pilossoph (2019), the exogenous separation probability δ_H comprises both quits to unemployment with probability δ_L and job-to-job transitions towards other identical firms with complementary probability $\delta_H - \delta_L$.

Job seekers can increase their chances to find a job by exerting application effort a. Search effort s is made up by the product between search efficiency ϵ and an increasing and iso-elastic function of application effort a, *i.e.* $s(a; \gamma) \equiv \epsilon(\gamma) a^{\chi}$, with $\chi > 0$.[21] A job seeker's job finding chances are higher either if she makes more applications (higher a) or if applications are more efficient at overcoming meeting frictions (higher ϵ).

Job finding comes as the result of a three-stage hiring process. First, job seekers decide how much application effort a to exert, subject to a convex application cost function $\sigma(a; \gamma) \equiv \psi(\gamma) \frac{a^{1+\eta}}{1+\eta}$, $\eta > 0$, whose scale depends on the search type. Search-type-specific application costs capture invariant heterogeneity in the time invested in locating job openings per application sent.[22]

[21] Application effort relates to the number of applications sent out by a job seeker in a given month, but the two concepts do not fully coincide. This is due to the sequential search protocol adopted in the model that allows for at most one worker-vacancy meeting in any period and does not restrict applications to integer numbers.

[22] Alternatively, they can be interpreted as a reduced-form modeling of other characteristics correlated with search efficiency, *e.g.* permanent income or value of leisure, which inflate the (marginal-utility-weighted) cost of an application (see Appendix C.3 for further details).

Second, job seekers come together with vacancies through a constant-return-to-scale meeting function $\mathcal{M}(S,V)$, where S denotes aggregate search effort and V the mass of outstanding vacancies. As a result, a job seeker exerting search effort s meets a vacancy with probability $s\lambda(\theta)$, where $\lambda(\theta) \equiv \frac{\mathcal{M}(S,V)}{S} = \mathcal{M}(1,\theta)$ and $\theta \equiv \frac{V}{S}$ represents labor market tightness. At the meeting stage, firms cannot observe the job seeker's type. Hence, the only relevant information released to firms upon meeting is the length of the job seeker's unemployment spell. Based on this information only, firms decide whether to call the job seeker back for a job interview at cost $\kappa > 0$.

Finally, conditional on interviewing the job seeker, the firm gets to know her ability x and decides whether to offer her a job. Let $\mathcal{Q}(x,y)$ be a qualification indicator for a worker of ability x with a firm of productivity y. To generate positive assortative matching, *i.e.* the most productive firms are the most selective with respect to workers' ability, we assume that the probability that a worker is qualified for a firm depends on the relative ranking of ability and productivity. Formally, $q(x,y) \equiv \Pr\left(\mathcal{Q}(x,y) = 1\right) = q\mathbb{1}\{x \geq y\}$. This functional form introduces a stochastic notion of qualification probability that relates directly to the scale of the job offer probability per interview.

Conditional on the worker being qualified for the job, match output is given by $p(x,y) = x + y$. A worker of higher ability is thus more likely to be qualified for a job, meaning that higher-ability job seekers enjoy a higher job offer probability per unit of search effort. Workers enjoy a flow value of leisure b while unemployed. Following Hall (2005), wages are rigid and fixed at $\omega \in (b, p(\underline{x}, \underline{y}))$ for the entire duration of the match.[23]

Workers. Workers have linear preferences over the consumption good, and they are either matched to a firm (employed) or job seekers (unemployed). Job seekers choose how much application effort a to exert at each unemployment duration τ, so as to maximize the value of unemployment. The values of unemployment and employment can be expressed recursively as:

$$U(x,\gamma,\tau) = \max_{\tilde{a}\geq 0}\ b - \sigma\left(\tilde{a};\gamma\right) + \beta\Big[U(x,\gamma,\tau+1) + s(\tilde{a};\gamma)o(x,\tau)\ \big(W(x,\gamma) - U(x,\gamma,\tau+1)\big)\Big],$$

$$W(x,\gamma) = \omega + \beta\Big[W(x,\gamma) + \delta_L\big(U(x,\gamma,0) - W(x,\gamma)\big)\Big],$$

[23]The assumption of rigid wages allows us to focus on sources of duration dependence unrelated to changes in the individual reservation wage. In Section 7 we show that our results are expected to largely go through even allowing for endogenous wages.

where $o(x, \tau)$ denotes the job offer probability per unit of search effort for a job seeker of ability x at duration τ.

Optimal application effort balances the marginal cost of exerting higher application effort to the expected marginal benefit of meeting a vacancy, *i.e.* the marginal increase in the job finding probability weighted by the discounted capital gain upon employment:

$$a(x, \gamma, \tau) : \frac{\partial \sigma(a; \gamma)}{\partial a} = \beta \, \frac{\partial s(a; \gamma)}{\partial a} o(x, \tau) \, \Big[W(x, \gamma) - U(x, \gamma, \tau + 1) \Big]. \tag{3}$$

Notice that job seekers of higher search type have both higher marginal benefit and higher marginal cost of exerting application effort (because both search effort and application costs are supermodular in application effort and search type). Fact 2 suggests that the latter force dominates in equilibrium, *i.e.* job seekers of lower search type exert more application effort for given duration. Since both search efficiency and the unit application cost are duration-invariant, the individual fixed effect in applications can thus be interpreted as an imperfect proxy for the job seeker's unobservable search type – including its correlation with ability – as previewed in the empirical analysis.

Firms. Firms can either be matched with one worker or not. Unmatched firms pay a vacancy posting cost κ_v to draw a productivity y, which allows them to meet a job seeker in the next period with probability $\lambda(\theta)/\theta$. Free entry into the labor market dictates that, in equilibrium, the labor market tightness adjusts to arbitrage out any pure profit from vacancy creation:

$$-\kappa_v + \beta \, \frac{\lambda(\theta)}{\theta} \int \Pi(y) \, dG(y) = 0, \tag{4}$$

where $\Pi(y)$ denotes expected profits of a firm with productivity y upon meeting a job seeker (derived in Appendix C.1). The value of a filled job is given by the present discounted value of flow profits, *i.e.* $J(x, y) = \frac{p(x,y) - \omega}{1 - \beta(1 - \delta_H)} \mathcal{Q}(x, y)$.

The hiring process. Upon meeting a job seeker, the firm decides whether to call her back for a job interview at cost κ, based on her elapsed unemployment duration τ only. For any (y, τ) pair, define a callback indicator as $\mathcal{C}(y, \tau) = \mathbb{1} \left\{ \int J(x, y) \; q(x, y) \; \mu(x|\tau) \; dx \geq \kappa \right\}$, where $\mu(x|\tau)$ is the search-effort-weighted density of job seekers' ability at unemployment duration τ – the key equilibrium object driving statistical discrimination. In words, a firm of productivity y calls back a job seeker with elapsed unemployment duration τ if the expected value of matching to a job seeker of that unemployment duration exceeds the interview cost κ. On the job seeker's side, this implies that the interview probability per unit of search effort only depends on unemployment duration τ: $c(\tau) = \lambda(\theta) \int \mathcal{C}(y, \tau)\, dG(y)$. Hence, cross-sectional differences in the interview probability per unit of application effort reflect heterogeneity in search efficiency and/or application behavior: $c^{app}(x, \gamma, \tau) = \epsilon(\gamma) a(x, \gamma, \tau)^{\chi-1} c(\tau)$.

After the interview takes place, the firm gets to know job seeker's ability x and makes her a job offer as long as she is qualified for its production technology, regardless of unemployment duration. On the job seeker's side, this implies that the job offer probability per interview equals $o^c(x, \tau) = \frac{\int q(x,y)\ \mathcal{C}(y,\tau)\ dG(y)}{\int \mathcal{C}(y,\tau)\ dG(y)}$. Hence, cross-sectional differences in the job offer probability per interview reflects heterogeneity in ability. Overall, the job offer probability per unit of search effort for a job seeker of ability x at duration τ is given by:

$$o(x, \tau) \equiv c(\tau)\ o^c(x, \tau) = \lambda(\theta) \int q(x, y)\ \mathcal{C}(y, \tau)\ dG(y). \tag{5}$$

Finally, the job finding probability – or, slightly abusing notation, the *job finding rate* – at duration τ reads:

$$f(x, \gamma, \tau) = s(x, \gamma, \tau)\ o(x, \tau). \tag{6}$$

Stationary equilibrium. To pin down the measure of unemployed of each type and duration, we solve the model in stationary equilibrium by imposing balance of flows:

$$u(x, \gamma, \tau) = \begin{cases} \delta_L \big(1 - \sum_{t=0}^{\infty} u(x, \gamma, t)\big) & \text{if } \tau = 0, \\ u(x, \gamma, \tau - 1)\big[1 - f(x, \gamma, \tau - 1)\big] & \text{if } \tau > 0, \end{cases} \tag{7}$$

where $1 - \sum_{t=0}^{\infty} u(x, \gamma, t)$ denotes the type-specific employment rate.

The key equilibrium object of the model is the search-effort-weighted density of job seekers' ability at each duration, $\mu(x|\tau)$, which drives firms' callback decisions:

$$\mu(x|\tau) = \frac{\int s(x, \gamma, \tau)\ u(x, \gamma, \tau)\ d\mathcal{L}(\gamma|x)}{\int \int s(\tilde{x}, \gamma, \tau)\ u(\tilde{x}, \gamma, \tau)\ d\mathcal{L}(\tilde{x}, \gamma)}, \tag{8}$$

where $\mathcal{L}(\gamma|x)$ denotes the search type distribution conditional on ability x.

Definition 1. *A stationary equilibrium is a collection* $\{a(x, \gamma, \tau), o(x, \tau), u(x, \gamma, \tau), \theta\}$*, where application effort satisfies equation (3), the job offer probability per unit of search effort satisfies equation (5), the unemployment rate satisfies equation (7), and the labor market tightness is pinned down by equation (4).*

Equilibrium characterization. We are now in the position to rationalize the three facts highlighted in the previous section through the lens of our structural model. We refer the reader to Appendix C.2 for the formal requirements underlying the equilibrium characterization.

Suppose that job seekers' ability and search types are positively correlated, *i.e.* workers who are more productive on the job submit more efficient, but also more costly, applications. Upon meeting a job seeker with unemployment duration τ, firms form an expectation about her ability based on $\mu(x|\tau)$. Since job seekers with higher ability x are more likely to find a job, the density $\mu(x|\tau)$ displays negative dynamic selection, *i.e.* expected ability declines with duration. Negative dynamic selection in ability produces a compositional and a behavioral effect on the duration profile of the job offer probability per unit of search effort. First, it implies that the *average* job offer probability per unit of search effort declines with duration, simply because low-ability job seekers are over-represented at longer unemployment durations (compositional effect). Second, negative dynamic selection in ability implies that the *individual* job offer probability per unit of search effort declines with duration, because of negative duration dependence in the interview probability (behavioral effect): since firms use elapsed unemployment duration as a signal for ability when choosing whether to call back a job seeker for an interview, some job seekers are denied interviews by firms they may have been qualified for. Hence, both the compositional and the behavioral effects contribute to the negative duration profile of the job offer probability per unit of search effort.

Negative dynamic selection in job seeker's ability further entails that the pool of job seekers becomes increasingly more homogeneous as unemployment duration lengthens, with low-ability ones accounting for a progressively larger share. As a result, the signal embedded in unemployment duration becomes more and more informative about job seekers's ability, so fewer firms will reject job seekers after inviting them to an interview.[24] This induces positive duration dependence in the job offer probability per interview. If the latter is strong enough to outweigh negative dynamic selection in the pool of job seekers, the job offer probability per application and per interview exhibit differential duration dependence, in line with Fact 3.

As long as the total reduction in the job offer probability per unit of search effort happens *smoothly* over the unemployment spell, job seekers optimally respond to negative duration dependence in their job offer probability per unit of search effort by scaling down their application effort over the unemployment spell according to equation (3).[25] Hence, our model rationalizes Fact 1 through the discouragement effect induced by firms' statistical discrimination.

Finally, if job seekers with higher search efficiency face sufficiently higher unit application costs – *e.g.* because they invest more time in locating job openings –, then the long-term unemployed submit more applications but face lower interview- and job-offer probabilities at any duration (Fact 2).

6. Quantitative analysis

To make the model amenable for quantification, we enrich the framework outlined in the previous section by allowing for coordination frictions in the form of multiple job seekers per vacancy (Blanchard and Diamond, 1994; Shimer, 2005a). Coordination frictions are a standard assumption in the existing literature as, in their presence, firms need to rank job seekers. Since in our model firms find it optimal to rank job seekers by unemployment duration (interviewing those with shorter duration first), coordination frictions induce negative duration dependence in the interview probability. We introduce coordination frictions to smooth out the duration profile of the individual job offer probability per unit of search effort, which makes sure that application effort is monotonically

[24] In Figure B5 we provide empirical validation to this mechanism by documenting that the positive duration dependence in the job offer probability per interview is driven by job seekers of lower search type, that is, with higher individual fixed effect in applications.

[25] If the job offer probability per unit of search effort followed a step-like process, job seekers would find it optimal to scale *up* their application effort during the periods when the job offer probability per unit of search effort is approximately constant, as the value of unemployment progressively depletes in anticipation of the following drop in the job offer probability per unit of search effort.

decreasing in unemployment duration.[26] Muehlemann and Strupler Leiser (2018) reports direct evidence of multiple job seekers per vacancy in administrative data from Switzerland. Namely, firms interview on average 4 applicants per vacancy before hiring one. Appendix C.5 develops the extended model.

Functional forms. We assume that worker ability x, worker search type γ and firm productivity y lie in the unit interval, *i.e.* $\text{supp}(x) = \text{supp}(\gamma) = \text{supp}(y) = [0,1]$. Worker ability and firm productivity follow flexible Beta distributions. Formally, $x \sim \mathcal{L}_x(x) = \text{Beta}(B_1, B_2)$ and $y \sim G(y) = \text{Beta}(G_1, G_2)$. We further assume that the worker search type is perfectly correlated to her ability by an increasing function $\gamma(x)$ – in line with Fact 2. This allows us to express search efficiency and the unit application cost as a function of ability. We let search efficiency be linear in ability with gradient ϕ, *i.e.* $\epsilon(x) = 1+\phi x$. We further posit that the unit application cost is an iso-elastic function of search efficiency, *i.e.* $\psi(x) = \psi_0 \epsilon(x)^\zeta$. Since our model is cast in discrete time, we adopt the meeting function of Ramey, den Haan, and Watson (2000), $\mathcal{M}(V,S) = \left(V^{-\xi} + S^{-\xi}\right)^{-\frac{1}{\xi}}$, which makes sure that contact probabilities lie in the unit interval.

Structural estimation. We estimate the structural model at monthly frequency for unemployment duration $\tau = 0, \ldots, \tilde{\tau}$. We set the grid size for ability and productivity to $N = 25$ and $\tilde{\tau} = 16$. The estimation is carried out in two steps. First, we pin down a set of parameters that have direct empirical counterparts in our data or in external sources. Then, we estimate the remaining moments internally targeting our empirical findings.

Table 4 Panel A reports the externally chosen parameters. Following Davis and von Wachter (2011), we set the discount factor to 0.996 to replicate a 5% annual interest rate. We then directly pin down the two separation probabilities from the EU and EE transition probabilities measured in our Swiss social security data. We set the wage rate to 0.985 to induce an average value of a job equal to 65% of average monthly output, as per Jarosch and Pilossoph (2019)'s proposed average across standard calibrations. We follow the same strategy for setting the flow value of leisure to 0.678.

[26]The model outlined in Section 5 is unable to replicate our empirical results quantitatively. Intuitively, in order to match the observed duration dependence in applications, optimal application effort would need to be extremely elastic to changing labor market conditions. However, because application behavior drives the pace of dynamic selection, such high responsiveness leads to an abrupt decline in job offer probabilities in equilibrium, which is inconsistent with a linear decline in individual application effort (see Proposition 2 in Appendix).

We then estimate the remaining set of parameters via indirect inference through the Simulated Method of Moments (SMM). Each such parameters conceptually relates to some moment in the data through the equilibrium conditions of the model. We make use of the cross-sectional properties and duration profiles of the number of applications, *individual* interview probability (probability that a job seeker receives at least an interview in a given unemployment month), and job finding rate from our search diary data as targeted moments to estimate the model parameters. We complement the set of targeted moments with two statistics that are informative of the heterogeneity in search types, namely the standard deviation of individual fixed effects in applications and their partial effect on the interview probability per application. To estimate interview costs, we use the breakdown of firms' hiring costs into specific components provided by Muehlemann and Strupler Leiser (2018) and take the average interview cost per hire as targeted moment.[27] Formally, let Θ be the vector of parameters still to be determined: $\Theta = \left\{B_1, B_2, G_1, G_2, \eta, \psi_0, \phi, q, \chi, \zeta, \kappa, V, \lambda\right\}$. Rather than directly estimating the vacancy posting cost, κ_v, and the elasticity of the meeting function, ξ, we include in the parameter vector the meeting probability per unit of search effort, λ, and the measure of vacancies, V, as auxiliary parameters. Even if λ and V are equilibrium objects, we can treat them as auxiliary parameters since the vacancy posting cost and elasticity of the meeting function can be chosen to rationalize any value of these variables.

We choose parameter values that minimize the sum of weighted squared percentage deviations between a set of empirical moments (μ) and model-generated moments ($\hat{\mu}$): $\Theta^* = \arg\min_{\Theta \in \mathcal{P}} \frac{1}{|\mathcal{M}|} \sum_{m \in \mathcal{M}} \left(\frac{\hat{\mu}_m(\Theta) - \mu_m}{\mu_m}\right)^2$, where $\mathcal{P}$ denotes the parameter space, $\mathcal{M}$ the set of (equally weighted) targeted moments, and $|\mathcal{M}|$ its cardinality.

Table 4 Panels B-C report the internally estimated parameters, along with their respective standard error and assigned targeted moment. Following Andrews et al. (2017) and Honoré et al. (2020), we report in Appendix D.2 the sensitivity matrix of the estimated parameters to each targeted moment. Inspection of the sensitivity matrix provides two key insights. First, the most informative moments for the estimated parameters are

[27] Muehlemann and Strupler Leiser (2018) reports administrative and representative Swiss survey data at the establishment-level. We compute average interview costs per hire based on its Table 1-2. Specifically, interview costs are defined as the difference between total search costs and vacancy posting costs, thus encompassing the time spent on job interviews and the cost of external advisors/headhunters. The number of hires is read off Table 1. The average monthly output is computed by applying a factor $1/\omega$ to the average monthly wage.

Table 4: Parameter Estimates

Parameter	Description	Estimate (Std Error)	Target/Source	Data	Model
Panel A. Externally set parameters					
β	Discount factor	0.996	5% annual interest rate in Davis and von Wachter (2011)		
δ_L	Separation prob. (workers)	0.009	Monthly EU prob. (Swiss social security)		
δ_H	Separation prob. (firms)	0.019	Monthly EE+EU prob. (Swiss social security)		
ω	Wage rate	0.985	Avg job value in Shimer (2005b), Hagedorn and Manovskii (2008), and Gertler and Trigari (2009)		
b	Value of leisure	0.678	Avg value of leisure in Shimer (2005b), Hagedorn and Manovskii (2008), and Gertler and Trigari (2009)		
Panel B. Internally estimated structural parameters					
$B1$	1^{st} shape param. Beta distr. search eff.	0.071 (0.000)	$\hat{\beta}_{c^{app}(x,\tau),\alpha(x)\|\tau}/\mathbb{E}_\tau[c^{app}(x,0)]$: partial effect ind. FE in app's on interview prob. per app.	-0.029	-0.030
$B2$	2^{nd} shape param. Beta distr. search eff.	0.376 (0.003)	$\mathbb{E}_\tau[a(x,0\!:\!11)]$: short-term duration profile applications, residual (obs.)	See Figure 5B	
$G1$	1^{st} shape param. Beta distr. prod.	0.141 (0.002)	$\mathbb{E}_\tau[a(x,12\!:\!16)]$: long-term duration profile applications, residual (obs.)	See Figure 5B	
$G2$	2^{nd} shape param. Beta distr. prod.	0.994 (0.014)	$\mathbb{E}_\tau[f(x,12\!:\!16)]$: long-term duration profile job finding rate, residual (obs.)	See Figure 6B	
ψ_0	Scalar search effort cost	0.014 (0.000)	$\sigma_{\alpha(x)}$: std. dev. ind. fixed effects in applications	4.095	3.894
ϕ	Search efficiency dispersion param.	16.34 (0.201)	$\mathbb{E}[f(x,\tau)]$: avg job finding rate	0.062	0.061
η	Convexity search effort cost	0.215 (0.007)	$\mathbb{E}_\tau[f(x,0\!:\!11)]$: short-term duration profile job finding rate, residual (obs.)	See Figure 6B	
q	Qualification prob. $x \geq y$	0.346 (0.001)	$\mathbb{E}[c^{ind}(x,\tau)]$: avg individual interview prob.	0.231	0.223
χ	App. effort elasticity search effort	0.884 (0.000)	$\mathbb{E}_\tau[c^{ind}(x,0\!:\!11)]$: short-term duration profile individual interview prob., residual (obs.)	See Figure 6A	
ζ	Search eff. elasticity app. costs	1.261 (0.005)	$\mathbb{E}_\tau[a(x,\tau)]$: duration profile applications, residual (FE)	See Figure 5A	
$\tilde{\kappa}$	Interview cost factor ($=\kappa/q$)	0.107 (0.000)	Avg interview cost per hire in Muehlemann and Strupler Leiser (2018)	0.129	0.128
Panel C. Internally estimated auxiliary parameters					
V	Measure of vacancies	0.160 (0.003)	$\mathbb{E}[a(x,\tau)]$: avg applications	10.69	10.39
λ	Meeting prob. per unit of search effort	0.014 (0.000)	$\mathbb{E}_\tau[c^{ind}(x,12\!:\!16)]$: long-term duration profile individual interview prob., residual (obs.)	See Figure 6A	
$L(\Theta^\star)$	SMM loss function	0.154%			

Note: Expectations are taken with respect to the ability distribution at the duration of the respective subscript. When the subscript is not specified, the expectation is taken with respect to the ability distribution at all durations from 0 to $\bar{\tau}$. Individual fixed effects in applications are not standardized. Standard errors of estimated parameters are reported in parenthesis (see Appendix D.2 for computation details). Numeraire: cross-sectional avg monthly output.

the duration profiles of the individual interview probability and the job finding rate at long durations, and the standard deviation of individual fixed effects in applications. Hence, we claim that the availability of data at each stage of the hiring process, along with a proxy for individual unobserved heterogeneity, are key to estimating the model parameters. Second, to construct a one-to-one mapping between targeted moments and parameters based on the sensitivity of the respective estimates, we collect the short-term (until 1 year) and long-term (more than 1 year) duration profiles into single variables and solve a linear assignment problem. In this way, we formally address the question of assigning moments to estimated parameters based on their informativeness – an exercise that is typically carried out heuristically (see Appendix D for further details on our estimation exercise).

Our data and modeling choices allow us to make progress in the statistical discrimination literature. Statistical discrimination models in the spirit of Jarosch and Pilossoph (2019) assume that application effort is exogenous and constant over the unemployment spell, attributing the entire decline in individual job finding probabilities to firms' changing interview behavior. Given functional form assumptions, the joint information contained in the duration profiles of the interview probability and the job finding rate informs the cross-sectional distribution of job seeker ability at each duration, $\mu(x|\tau)$. However, if search effort varies over the unemployment spell, models with exogenous search effort are misspecified: the observed duration profile of individual job finding probabilities reflects changes in both firms' interview behavior *and* job seekers' application behavior. Recovering a model-consistent ability distribution thus requires purging the duration profile of the job finding rate of variation driven by applications. Our empirical evidence on the cross-sectional distribution and dynamic evolution of individual applications allows us to disentangle changing application behavior from underlying job seeker heterogeneity.

Still, the resulting ability distribution would remain sensitive to the model structure and difficult to validate empirically. This is where the positive correlation between individual ability and search efficiency becomes critical. Our data suggest that the distribution of individual fixed effects in applications is informative of the unobservable ability distribution. Replicating the standard deviation of individual fixed effects in applications provides empirical discipline to the extent of job seeker heterogeneity in ability. To sum up, our approach accommodates duration dependence arising from job seekers' application behavior and provides an assessment of job seeker heterogeneity in a statistical discrimination framework.

Model fit. The estimated model is successful at replicating all the targeted moments. In Appendix D.2 we show that the model parameters are well informed by the joint information contained in the targeted moments

Figure 5A compares the duration profile of average applications controlling for observables in the data with that of average application effort in the model, $\mathbb{E}_\tau[a(x,\tau)]$. Figure 5B compares the duration profile of average applications controlling for individual fixed effects in the data with that predicted by the model when the composition of the unemployment pool is kept constant, $\mathbb{E}_0[a(x,\tau)]$. Comparison across panels reveals that, quantitatively, the divergence between the two duration profiles is slightly lower in the model than in the data.[28] Importantly, the model matches quantitatively the positive dynamic selection in applications, since low-ability job seekers apply more at any duration. Appendix Figure D6C shows that the model-implied duration profile of the average individual fixed effect in applications is virtually identical to its empirical counterpart.

Figure 6A displays the average individual interview probability controlling for observables, $\mathbb{E}_\tau[c^{ind}(x,\tau)]$, while Figure 6B shows the average job finding rate controlling for observables, $\mathbb{E}_\tau[f(x,\tau)]$. The model replicates their empirical duration profiles accurately, also capturing the slight convexity at long durations. Notice that the duration profile of the individual interview probability is steeper than that of the job finding rate at long durations. This is because, as in the data, the average individual job offer probability per interview increases with duration (see Appendix Figure D6A).

Notably, our estimated model is able to replicate all the duration profiles not only in relative terms but also in levels. It follows that the pace of dynamic selection – the driver of statistical discrimination and compositional changes – is virtually the same in the model and in the data, as governed by the observed job finding and separation probabilities.

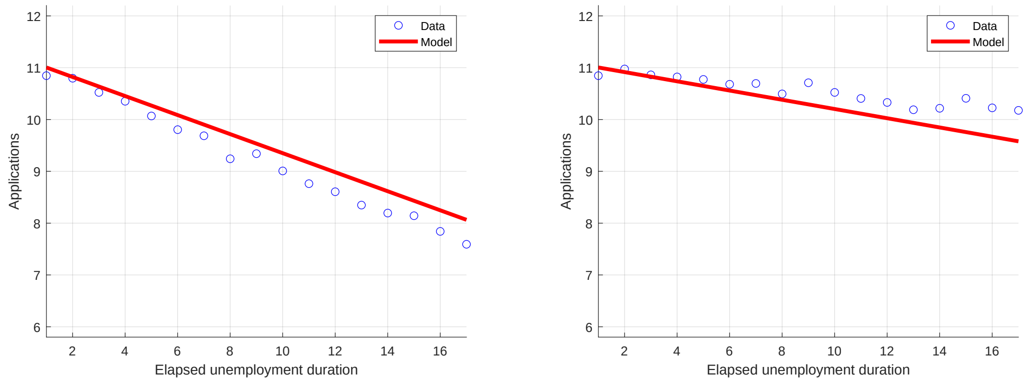
Duration dependence versus dynamic selection. We now use our model as an accounting framework to decompose the decline in the observed job finding rate into duration dependence and dynamic selection. In turn, we are interested in assessing the respective roles of workers and firms in shaping overall duration dependence.

On the workers' side, the estimated model allows us to map observed application effort into the relevant notion of search effort for the sake of job finding, *i.e.* $s(x,\tau) = \epsilon(x)a(x,\tau)^\chi$. This has two important implications. First, the negative dynamic selection in search efficiency partially offsets the positive dynamic selection in application effort. As a result, the duration profile of individual and average search effort are closer to each other than those of application effort (see Appendix Figure D6B). Second, we estimate mild decreasing returns in application effort. This implies that negative duration dependence

[28] The small discrepancy between the empirical and model-implied duration profiles of application effort is likely a product of the time lag between application records and the corresponding job finding dates.

Figure 5: Duration profile of application effort, model vs data

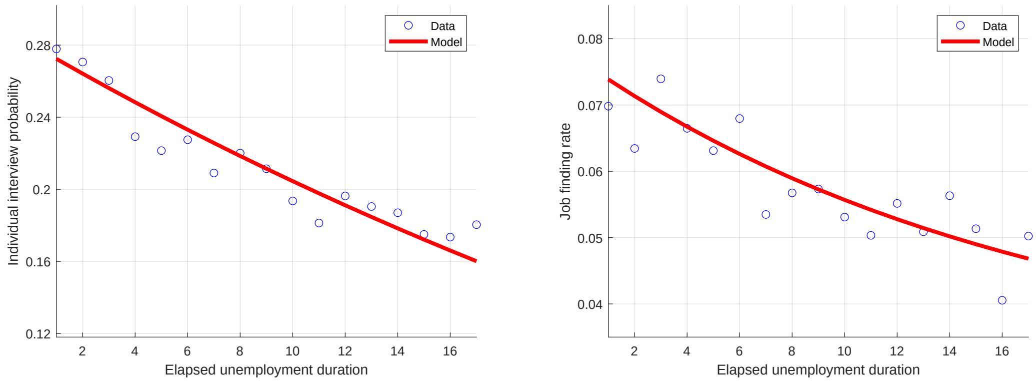
(A) Application effort, residual (FE) (B) Application effort, residual (obs.)



Note: This figure contrasts the duration profiles controlling for individual fixed effects (Panel A) and for observables (Panel B) of application effort in the data (circles) with those implied by the estimated model. The duration profiles in the model are derived by estimating individual fixed effects and duration effects from a saturated regression and computing the expected values of application effort at any unemployment duration. For the duration profile controlling for individual fixed effects, expected values are computed with respect to the ability distribution in the first month of unemployment, *i.e.* $\mathbb{E}_0[a(x,\tau)]$; for the duration profile controlling for observables, expected values are computed with respect to the ability distribution in the contemporaneous period of unemployment, *i.e.* $\mathbb{E}_\tau[a(x,\tau)]$. The distribution of observables across unemployment durations is kept the same as in the first month of unemployment in both specifications. The model-based duration profiles are fitted using a linear function estimated via weighted least squares.

Figure 6: Duration profile of interview prob. and job finding rate, model vs data

(A) Ind. interview prob., residual (obs.) (B) Job finding rate, residual (obs.)



Note: This figure contrasts the duration profiles controlling for observables of the individual interview probability (Panel A) and job finding rate (Panel B) detected in the data (circles) with those implied by the estimated models. The duration profiles in the model are derived by estimating individual fixed effects and duration effects from a saturated regression and computing the expected values of the individual interview probability and job finding rate at any unemployment duration. Expected values are computed with respect to the ability distribution in the contemporaneous period of unemployment, *i.e.* $\mathbb{E}_\tau[c^{ind}(x,\tau)]$ and $\mathbb{E}_\tau[f(x,\tau)]$.The distribution of observables across unemployment durations is kept the same as in the first month of unemployment. The model-based duration profiles are fitted using a negative exponential function estimated via weighted least squares.

in application effort gets transmitted to search effort with an almost unitary elasticity.

On the firms' side, the estimated models provide a structural decomposition of the duration profile of the job offer probability per unit of search effort (controlling for observables) into duration dependence and dynamic selection on unobservables, which goes one step further our empirical assessment (Table 3).

Recall from equation (6) that the job finding rate at duration τ is shaped both by workers' behavior (search effort) and firms' behavior (job offers), *i.e.* $f(x,\tau) = s(x,\tau)\ o(x,\tau)$.

We decompose the decline in the job finding rate (controlling for observables) into duration dependence – separating the components due to workers and firms – and dynamic selection on unobservables, as follows:

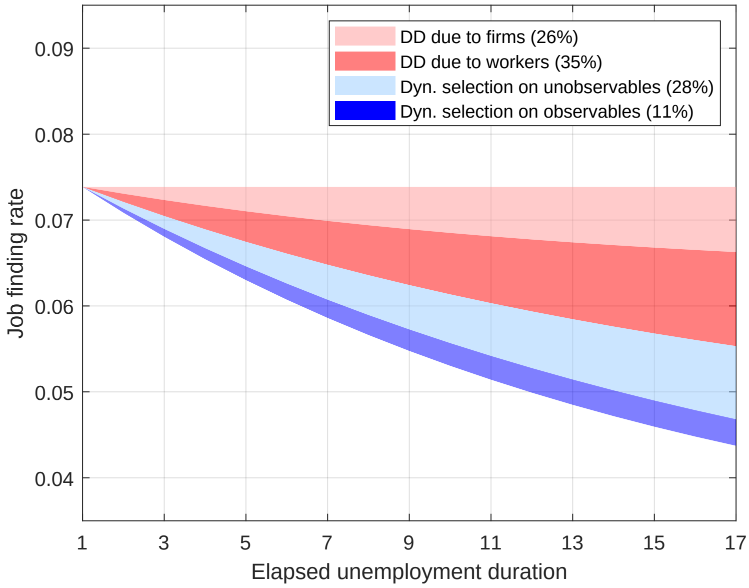
$$\underbrace{\mathbb{E}_\tau\left[f(x,\tau)\right]-\mathbb{E}_0[f(x,0)]}_{\text{Duration profile controlling for obs.}} = \underbrace{\mathbb{E}_\tau\Big[s(x,0)\ \big(o(x,\tau)-o(x,0)\big)\Big]}_{\text{DD due to firms}} + \underbrace{\mathbb{E}_\tau\Big[\big(s(x,\tau)-s(x,0)\big)\ o(x,\tau)\Big]}_{\text{DD due to workers}}$$
$$+\underbrace{\mathbb{E}_\tau\Big[s(x,0)\ o(x,0)\Big]-\mathbb{E}_0\Big[s(x,0)\ o(x,0)\Big]}_{\text{Dynamic selection on unobservables}}, \tag{9}$$

where $\mathbb{E}_t[.]$ denotes the expectation with respect to the joint distribution of workers' unobservable characteristics (search type and ability) $\mathcal{L}(x,\gamma(x))$ at duration t. "Duration dependence due to firms" captures the extent to which the reduction in the job offer probability per unit of search effort affects the job finding rate directly, while "duration dependence due to workers" captures by how much the change in application effort contributes to the reduction in the job finding rate. The dynamic selection component reflects to what extent job seekers still unemployed in month τ differ from those in the first month of the unemployment spell in terms of unobservable characteristics.

We notice that our model assumes that workers are homogeneous in terms of observable characteristics in a given labor market. Accordingly, when estimating the model, we target the duration profile of the job finding rate controlling for observables. This amounts to positing that the distribution of observables at any unemployment duration is the same as in the first month of unemployment. Let X be a vector of observable characteristics. Hence, $\mathbb{E}_t[f(x,\tau)] \equiv \tilde{\mathbb{E}}_0\left[\mathbb{E}_t[f(x,\tau)|X]\right]$, where $\tilde{\mathbb{E}}_t[\cdot]$ denotes the expectation with respect to the distribution of workers' observable characteristics at duration t. To complete the decomposition of the observed duration profile of the job finding rate, we therefore combine the model-based assessment of duration dependence versus dynamic selection on unobservables with our empirical estimate of the importance of dynamic selection on observables, reported in Appendix Figure B9C. The observed duration profile of the job finding rate can be further decomposed as follows:

$$\underbrace{\hat{\mathbb{E}}_\tau\left[f(x,\tau,X)\right]-\hat{\mathbb{E}}_0\left[f(x,0,X)\right]}_{\text{Observed duration profile}} = \underbrace{\mathbb{E}_\tau\left[f(x,\tau)\right]-\mathbb{E}_0\left[f(x,0)\right]}_{\text{Duration profile controlling for obs.}} \tag{10}$$
$$+\underbrace{\hat{\mathbb{E}}_\tau\left[f(x,\tau,X)\right]-\mathbb{E}_\tau\left[f(x,\tau)\right]}_{\text{Dynamic selection on observables}},$$

Figure 7: Duration profile of the job finding rate, decomposition



Note: This figure reports the decomposition of the duration profile of the job finding rate into the different sources of duration dependence and dynamic selection derived in equation (9) and equation (10). The duration profiles of the components of the job finding rate reported in equation (9) are derived by estimating individual fixed effects and duration effects from a saturated regression and computing the expected values of each component at any unemployment duration. Expected values are computed with respect to the ability distribution in the contemporaneous period of unemployment. The distribution of observables across unemployment durations is kept the same as in the first month of unemployment. According to equation (10), the duration profiles of the component due to dynamic selection on observables is computed as the difference between the observed duration profile of the job finding rate and the duration profile controlling for observables (see Appendix Figure B9C). Then, all duration profiles are fitted by a negative exponential function estimated via weighted nonlinear least squares. The shares of each component are computed as the frequency-weighted average shares of the respective raw components over the entire unemployment spell.

where $\hat{\mathbb{E}}_t[\cdot]$ denotes the expectation with respect to the joint distribution of job seekers' unobservable characteristics $(x, \gamma(x))$ *and* observable characteristics X at duration t, which we read off the data.

We report the decomposition results in Figure 7. Our estimated model attributes 61% of the observed decline in the job finding rate to duration dependence and 39% to dynamic selection. Duration dependence is significantly affected by both workers' search behavior, which accounts for 35% of the observed decline of the job finding rate, and firms' hiring behavior (26%). Dynamic selection happens primarily on unobservables, which accounts for 28% of the observed decline, while the role of observables is more muted (11%).

Our results indicate that duration dependence explains over half of the observed decline in the job finding rate. Accordingly, the contribution of duration dependence is substantially larger than in Mueller, Spinnewijn, and Topa (2021), Mueller and Spinnewijn (2023), and Jarosch and Pilossoph (2019), which attribute most of the decline to dynamic selection, and closer to the findings of Kroft et al. (2016) and Alvarez, Borovičková, and

Shimer (2023), who document a significant role for duration dependence.[29] It is important to note, however, that Mueller, Spinnewijn, and Topa (2021) and Mueller and Spinnewijn (2023) employ a different notion of dynamic selection than ours, namely the share of the empirical decline in job finding probabilities that can be predicted *ex ante* using time-invariant characteristics. This concept encompasses the *average* effect of duration dependence over the relevant time horizon. Their conclusion that the long-term unemployed are a highly selected subset of job seekers therefore also holds true in our framework, as well.[30] Consequently, provided that administrative data are sufficiently rich to proxy for unobserved ability, our findings are consistent with theirs.

On the other hand, our results stand in contrast with statistical discrimination models with exogenous search effort. Using the same model and estimation strategy as in Jarosch and Pilossoph (2019), the decomposition results change significantly, with dynamic selection accounting for 70% to 90% of the decline in the job finding rate (see Appendix E.1 for further details).

Hence, we argue that modeling endogenous search effort by workers is crucial for studying duration dependence in the equilibrium job finding process. The reason is twofold. First, endogenous search effort shapes the equilibrium pace of dynamic selection in ability, which in turn determines the extent of statistical discrimination in firms' interview behavior. Second, endogenous search effort amplifies the effect of statistical discrimination on the job finding rate by discouraging worker search. In Appendix F, we estimate that the general-equilibrium elasticity of the job finding rate with respect to the job offer probability per unit of search effort exceeds 3, due to induced worker discouragement (under exogenous search effort, it would be only 1). Overall, statistical discrimination determines duration dependence in the job finding rate through both firms' interview decisions (direct effect) and the equilibrium response of workers' application effort (indirect effect), with the latter playing the quantitatively larger role. These results underscore the importance of jointly analyzing workers' search behavior and firms' hiring choices to

[29] Mueller and Spinnewijn (2023), using Swedish administrative data, estimate that dynamic selection on observables accounts for at least 49% of the decline in the six-month job finding probability during the first six months of unemployment and 36% during the subsequent six months. Alvarez, Borovičková, and Shimer (2023), using Austrian social security data, find that duration dependence is the main driver of the duration profile of job finding probabilities after the first 20 weeks of unemployment.

[30] In our estimates, the lowest-ability workers represent 79% of the long-term unemployed, in the face of a 69% population share.

explain the duration profile of the job finding rate.

Robustness to model extensions. In our model, job seekers adjust application effort in response to changes in firms' recruitment policies. However, optimal application effort may also depend on other factors unrelated to firms' decisions. If it is the case, our decomposition results are likely to over-estimate the contribution of duration dependence due to firms in the decline of the job finding rate. In Appendix E, we study three extensions of our baseline model where job seekers have (i) incomplete information about one's own ability and learning it from search, (ii) reference-dependent preferences, and (iii) duration-dependent application costs – which we interpret as a reduced-form treatment of the depletion of personal networks. Each model extension features an additional parameter that relates to the duration profile of individual application effort at long durations, when the pool of remaining job seekers is largely unaffected by statistical discrimination from firms with high ability requirements. The goal of these model extensions is to provide a sensitivity analysis of our decomposition results to the structural modeling of the job search process.[31]

Performing the same decomposition exercise as in equation (10) on the estimated model extensions, the contribution of duration dependence due to firms decreases, as expected. Yet, the overall contribution of duration dependence to the decline in the job finding rate remains substantial, ranging between 52% and 57%. Therefore, our conclusion that duration dependence is the dominant driver of the decline in the job finding rate is robust to the structural modeling of the job search process.

7. Other mechanisms generating a falling job finding rate

Our framework abstracts from other potentially important drivers of duration dependence in finding a job. Here we briefly discuss some of these mechanisms.

[31] Baydur and Xu (2024) proposes to generalize the qualification function to allow for a positive qualification probability of low-ability job seekers with high-productivity firms. In principle, this generalized qualification function would allow statistical discrimination to affect the job offer probability per application of low-ability job seekers even in the absence of coordination frictions. However, enriching the baseline model of Section 5 with this generalized qualification function fails to replicate our empirical targets. Also, extending the quantitative model with the generalized qualification function does not improve its fit to the targeted moments.

Declining reservation/target wages. For simplicity, our framework assumes that wages are exogenous. In a more realistic setting, however, job seekers may reduce the reservation wage (random search) or the target wage (directed search) to improve their employment prospects. With random search and wage bargaining, a lower reservation wage would reduce the gap between the value of employment and unemployment, thus reducing optimal application effort, consistent with Fact 1. However, the downward wage adjustment needs to be small enough so that firms strictly prefer to hire shorter-duration applicants for job offers per application to decline with duration (Fact 3). Hence, falling reservation wages, while potentially relevant in practice, cannot themselves account for the empirical evidence. In directed search models (Galenianos and Kircher, 2009; Wright, Kircher, Julien, and Guerrieri, 2021; Lehmann, 2023), a declining value of unemployment, due to *e.g.* duration-dependent application costs, induces long-term unemployed job seekers to target lower-paying jobs with a higher job offer probability per application, which is at odds with Fact 3. Since the return from search rises with duration, job seekers would also scale *up* application effort, contradicting Fact 1.

Moreover, the existing empirical literature points to a rather flat reservation wage (Krueger and Mueller, 2016) and a mildly decreasing target wage (Marinescu and Skandalis, 2021; Fluchtmann et al., 2021; DellaVigna et al., 2022) over the unemployment spell. Le Barbanchon et al. (2024) provide a comprehensive overview of reservation wage and target wage changes.

We conclude that, both for theoretical and empirical reasons, abstracting from reservation/target wage adjustment over the unemployment spell does not substantially confound our decomposition analysis of the falling job finding rate.

Occupational switching and increasing commuting distances. Our analysis does not take into account occupation or location choices. However, a job seeker might consider increasing her search radius over the unemployment spell, by looking beyond her previous occupation or by applying to jobs at more distant locations. From a theoretical perspective, increasing the spatial search radius is conceptually similar to lowering the reservation wage. Indeed, in a general model where jobs provide wage and non-wage amenities (commuting distance), job seekers may decide to lower their standards either in terms of wage or amenities (or both) to improve their employment prospects. Similarly, changing the target occupation is akin to adjusting the target wage in a directed search framework, potentially subject to a reallocation cost. Manning and Petrongolo (2017), Guglielminetti et al. (2024), and Carrillo-Tudela and Visschers (2023) develop structural models to study job seeker's strategies across spatial locations and occupations.

Empirically, the relevance of changing occupations and locations with unemployment duration is mixed. Carrillo-Tudela and Visschers (2023) show that US job seekers increasingly accept jobs at different occupations – even if occupational attachment remains high. Fluchtmann et al. (2024) document that individuals progressively target less attractive jobs in terms of occupations, industries, and geography. Belot et al. (2018, 2025) show that targeted occupational advice in an online platform induces job seekers in declining occupations to broaden the set of occupations they consider. Guglielminetti et al. (2024) find that job seekers accept jobs with lower wages, but not longer commuting times, as unemployment duration lengthens.

Using a smaller search diary data set from one employment office in the city of Zurich, we can shed some light on occupational targeting. We find that the share of applications to the same occupation as before unemployment remains stable with duration (see Appendix Figure G2A). Also, the skill intensity of new occupations does not change once the analysis is conducted within individual (Appendix Figure G2B). Unfortunately, our data do not provide information on commuting times. Overall, changes in search strategies with respect to location and occupation are conceptually similar to a reduction in reservation/target wages. For the reasons mentioned above, we think that abstracting from them unlikely affects our main conclusions.

Human capital depreciation. In our theoretical framework, a job seeker's ability is exogenous and does not change with the duration of unemployment. However, a prominent argument holds that job seekers' human capital depreciates during unemployment (Ljungqvist and Sargent, 1998). As a consequence, the probability that an application receives an interviews, and eventually a job offer, will fall with duration. In our model, it is unlikely that accounting for human capital depreciation will have a strong impact on our decomposition analysis. As long as wages are constant, it does not matter much whether ability is modeled as falling, rather than staying constant, with duration. Given an ergodic ability distribution, ability depreciation would not alter firms' hiring policies and would reduce optimal application effort only for the subset of workers whose ability depreciates quickly enough to render them unqualified for the jobs they are interviewed for. Flexible wages, however, may give a larger role to ability depreciation in driving negative duration dependence in application effort.

The existing empirical literature remains inconclusive on whether – and, if so, to what extent – human capital actually depreciates over the course of unemployment (Edin and Gustavsson, 2008; Dinerstein, Megalokonomou, and Yannelis, 2022; Arellano-Bover, 2022; Cohen, Johnston, and Lindner, 2023). Unfortunately, the information available in the Swiss search diaries does not allow to address this further. Taken together, the role of ability depreciation for the falling job finding rate remains inconclusive and we consider it a fruitful avenue for future research.

Stock-flow matching. According to our analysis, duration dependence is largely driven by falling applications as job seekers internalize statistical discrimination by firms. Alternatively, negative duration dependence in application effort may stem from stock-flow sampling of vacancies (Coles and Petrongolo, 2008; Ebrahimy and Shimer, 2010). The basic idea is that job seekers initially apply to the stock of vacancies and, over time, to the (smaller) inflow of new vacancies. Consequently, stock–flow matching predicts a non-gradual decline in applications with elapsed unemployment duration. Contrary to this claim, we find that applications decrease gradually and linearly over time (Fact 1). Notice, however, that institutional rules may disguise the stock-flow matching, as minimum application requirements may make it optimal to smooth applications over the spell. Absent vacancy data, we are not able to empirically test stock-flow matching against random matching.

From a theoretical standpoint, stock-flow matching implies that a reduction in individual applications may not solely reflect a lower net benefit of search per application, but also the gradual exhaustion of available job openings. As a result, dynamic selection on unobservables may play a more important role for the declining job finding rate, at the expense of duration dependence due to workers' search behavior.

Changes in search channel. Job seekers may adjust not only their application effort but also the channel through which they apply for jobs over the course of the unemployment spell (Beaman and Magruder, 2012; Burks, Cowgill, Hoffman, and Housman, 2015; Hensvik and Nordström Skans, 2016). In our smaller data set from one Zurich employment office, we observe the search channel used for each application (written, by phone, or personal). Appendix Figure G1 shows that the relative share of each channel slightly shifts from personal to written applications, consistent with the gradual depletion of one's personal network.

From a theoretical perspective, this shift may reflect a form of *personal* stock-flow matching: as the unemployment spell lengthens, job seekers exhaust their personal connections and must apply to jobs for which they have no personal ties – typically through the written channel. Since personal applications are, on average, more successful than written ones, search efficiency declines over the unemployment spell. This mechanism is akin to an increase in unit application costs with unemployment duration (see the model extension in Appendix E), consistent with the reduction in application effort observed in the data.

8. Conclusions

This paper uses monthly search diaries from the Swiss public employment offices to better understand why the job finding rate falls with the duration of unemployment. We find that applications and interviews per application decrease, but job offers per interview weakly increase with duration. A theoretical framework with endogenous search effort by workers and statistical discrimination against the long-term unemployed by firms matches these findings closely. This theoretical framework allows for an exact decomposition of the falling job finding rate into structural duration dependence and dynamic selection (on observable and unobservable factors). We find that over half of the decline in the job finding rate is due to structural duration dependence, with the remainder explained by dynamic selection of the unemployment pool. Structural duration dependence is largely driven by a worker discouragement effect: as the unemployment spell lengthens, job seekers reduce their application effort, internalizing that firms statistically discriminate against longer-term unemployed candidates.

While our data set provides comprehensive information on job applications by job seekers and interview- and job-offer decisions by firms, important aspects that may account for the decline in the job finding rate remain unobserved. On the one hand, we do not observe changes in job seeker characteristics over the unemployment spell (other than applications and their outcomes). As a consequence, the specific reasons that induce workers to reduce application effort – such as reference-dependent preferences, learning from unsuccessful search outcomes, duration-dependent application costs, or fewer job opportunities – cannot be inferred from the data. On the other hand, our data provide no information on the pool and characteristics of suitable vacancies, nor on the subset of vacancies to which job seekers actually apply at different durations. As a result, it is difficult to disentangle the extent to which application outcomes are driven by changes in job seeker characteristics (*e.g.* ability depreciation) versus changes in the composition of vacancies targeted by applicants. Addressing these open questions is a fruitful direction for future research.

References

Ahn, H. J., Hobijn, B., and Şahin, A. The dual u.s. labor market uncovered. *NBER Working Paper no 31241*, 2023.

Ahn, H. J. and Hamilton, J. D. Heterogeneity and unemployment dynamics. *Journal of Business & Economic Statistics*, 38(3):554–569, 2020.

Ahrens, A., Hansen, C. B., and Schaffer, M. E. pystacked: Stacking generalization and machine learning in stata. *The Stata Journal*, 23(4):909–931, 2023. doi: 10.1177/1536867X231212426. URL `https://doi.org/10.1177/1536867X231212426`.

Ahrens, A., Hansen, C. B., Schaffer, M. E., and Wiemann, T. Model averaging and double machine learning, 2024.

Alvarez, F., Borovičková, K., and Shimer, R. Decomposing Duration Dependence in a Stopping Time Model. *The Review of Economic Studies*, page rdad109, 12 2023. ISSN 0034-6527. doi: 10.1093/restud/rdad109. URL `https://doi.org/10.1093/restud/rdad109`.

Andrews, I., Gentzkow, M., and Shapiro, J. M. Measuring the sensitivity of parameter estimates to estimation moments. *The Quarterly Journal of Economics*, 132(4):1553–1592, 06 2017. ISSN 0033-5533. doi: 10.1093/qje/qjx023. URL `https://doi.org/10.1093/qje/qjx023`.

Arellano-Bover, J. The Effect of Labor Market Conditions at Entry on Workers' Long-Term Skills. *The Review of Economics and Statistics*, 104(5):1028–1045, December 2022.

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

Baydur, I. and Xu, J. Statistical discrimination and duration dependence in a semistructural model. *International Economic Review*, 2024.

Beaman, L. and Magruder, J. Who gets the job referral? Evidence from a social networks experiment. *American Economic Review*, 102(7):3574–93, 2012.

Belot, M., de Koning, B. K., Fouarge, D., Kircher, P., Muller, P., and Phlippen, S. Advising job seekers in occupations with poor prospects: A field experiment. Working Paper 33819, National Bureau of Economic Research, 2025. URL `https://www.nber.org/papers/w33819`.

Belot, M., Kircher, P., and Muller, P. Providing Advice to Jobseekers at Low Cost: An Experimental Study on Online Advice. *The Review of Economic Studies*, 86(4): 1411–1447, 10 2018.

Blanchard, O. J. and Diamond, P. Ranking, unemployment duration, and wages. *The Review of Economic Studies*, 61(3):417–434, 1994.

Breiman, L. Stacked regressions. *Machine Learning*, 24:49–64, 1996. doi: 10.1007/BF00117832. URL `https://doi.org/10.1007/BF00117832`.

Burdett, K. and Vishwanath, T. Declining reservation wages and learning. *The Review of Economic Studies*, 55(4):655–665, 10 1988.

Burks, S. V., Cowgill, B., Hoffman, M., and Housman, M. The value of hiring through employee referrals. *The Quarterly Journal of Economics*, 130(2):805–839, 2015.

Carrillo-Tudela, C. and Visschers, L. Unemployment and endogenous reallocation over the business cycle. *Econometrica*, 91(3):1119–1153, 2023. doi: https://doi.org/10.3982/ECTA12498. URL `https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA12498`.

Cederlöf, J. and Roman, S. Duration dependence and job search over the spell: Evidence from job seeker activity reports. Manuscript, October 2025.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., and Robins, J. Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1):C1–C68, 01 2018. ISSN 1368-4221. doi: 10.1111/ectj.12097. URL `https://doi.org/10.1111/ectj.12097`.

Christensen, B. J., Lentz, R., Mortensen, D., Neumann, G. R., and Werwatz, A. On-the-job search and the wage distribution. *Journal of Labor Economics*, 23(1):31–58, 2005.

Cohen, J., Johnston, A. C., and Lindner, A. Skill depreciation during unemployment: Evidence form panel data. *Working paper*, 2023.

Coles, M. and Petrongolo, B. A test between stock-flow matching and the random matching function approach. *International Economic Review*, 49(4):1113–1141, November 2008. doi: None. URL `https://ideas.repec.org/a/ier/iecrev/v49y2008i4p1113-1141.html`.

Davis, S. J. and von Wachter, T. Recessions and the Costs of Job Loss. *Brookings Papers on Economic Activity*, 42(2 (Fall)):1–72, 2011.

DellaVigna, S., Lindner, A., Reizer, B., and Schmieder, J. Reference-dependent job search: Evidence from hungary. *The Quarterly Journal of Economics*, 132, 05 2017.

DellaVigna, S., Heining, J., Schmieder, J. F., and Trenkle, S. Evidence on job search models from a survey of unemployed workers in germany. *The Quarterly Journal of Economics*, 137(2):1181–1232, 2022.

Dinerstein, M., Megalokonomou, R., and Yannelis, C. Human capital depreciation and returns to experience. *American Economic Review*, 112(11):3725–62, November 2022.

Doppelt, R. The hazards of unemployment: A macroeconomic model of job search and résumé dynamics. *Working Paper*, 2016.

Ebrahimy, E. and Shimer, R. Stock–flow matching. *Journal of Economic Theory*, 145 (4):1325–1353, 2010. ISSN 0022-0531. Search Theory and Applications.

Edin, P.-A. and Gustavsson, M. Time out of work and skill depreciation. *ILR Review*, 61 (2):163–180, 2008. URL `https://EconPapers.repec.org/RePEc:sae:ilrrev:v:61:y:2008:i:2:p:163-180`.

Einav, L., Finkelstein, A., Mullainathan, S., and Obermeyer, Z. Predictive modeling of U.S. health care spending in late life. *Science*, 360(6396):1462–1465, June 2018.

Eriksson, S. and Rooth, D.-O. Do employers use unemployment as a sorting criterion when hiring? evidence from a field experiment. *American Economic Review*, 104(3): 1014–39, 2014.

Faberman, R. J. and Kudlyak, M. The intensity of job search and search duration. *American Economic Journal: Macroeconomics*, 11(3):327–57, 2019.

Falk, A., Lalive, R., and Zweimüller, J. The success of job applications: a new approach to program evaluation. *Labour Economics*, 12(6):739–748, 2005. ISSN 0927-5371. doi: https://doi.org/10.1016/j.labeco.2004.05.002. URL `https://www.sciencedirect.com/science/article/pii/S0927537104000727`.

Farber, H. S., Silverman, D., and Von Wachter, T. Determinants of callbacks to job applications: An audit study. *American Economic Review*, 106(5):314–18, 2016.

Fernández-Blanco, J. and Preugschat, E. On the effects of ranking by unemployment duration. *European Economic Review*, 104:92–110, 2018. ISSN 0014-2921.

Fluchtmann, J., Glenny, A. M., Harmon, N., and Maibom, J. The dynamics of job search in unemployment: Beyond search effort and reservation wages. *Working Paper*, 2021.

Fluchtmann, J., Glenny, A. M., Harmon, N., and Maibom, J. Unemployed job search across people and over time: Evidence from applied-for jobs. *Journal of Labor Economics*, 42(4):1175–1217, 2024. doi: 10.1086/725165.

Frisch, R. and Waugh, F. V. Partial time regressions as compared with individual trends. *Econometrica*, 1(4):387–401, 1933. URL `https://doi.org/10.2307/1907330`.

Galenianos, M. and Kircher, P. Directed search with multiple job applications. *Journal of economic theory*, 144(2):445–471, 2009.

Gertler, M. and Trigari, A. Unemployment fluctuations with staggered nash wage bargaining. *Journal of Political Economy*, 117(1):38–86, 2009.

Gonzalez, F. M. and Shi, S. An equilibrium theory of learning, search, and wages. *Econometrica*, 78(2):509–537, 2010.

Guglielminetti, E., Lalive, R., Ruh, P., and Wasmer, E. Job search with commuting and unemployment insurance: A look at workers' strategies in time. *Labour Economics*, 88, 2024. doi: 10.1016/j.labeco.2024.102537.

Hagedorn, M. and Manovskii, I. The cyclical behavior of equilibrium unemployment and vacancies revisited. *American Economic Review*, 98(4):1692–1706, September 2008.

Hall, R. E. Employment fluctuations with equilibrium wage stickiness. *American Economic Review*, 95(1):50–65, March 2005.

He, Q. and Kircher, P. Updating about yourself by learning about the market: The dynamics of beliefs and expectations in job search. *Working Paper*, 2023.

Heckman, J. and Singer, B. A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica*, 95(2):50–65, March 1984.

Hensvik, L. and Nordström Skans, O. Social networks, employee selection, and labor market outcomes. *Journal of Labor Economics*, 34(4):825–867, 2016.

Honoré, B., Jørgensen, T., and de Paula, A. The informativeness of estimation moments. *Journal of Applied Econometrics*, 35(7):797–813, 2020. URL `https://EconPapers.repec.org/RePEc:wly:japmet:v:35:y:2020:i:7:p:797-813`.

Jarosch, G. and Pilossoph, L. Statistical discrimination and duration dependence in the job finding rate. *The Review of Economic Studies*, 86(4):1631–1665, 2019.

Kospentaris, I. Unobserved heterogeneity and skill loss in a structural model of duration dependence. *Review of Economic Dynamics*, 39:280–303, 2021. ISSN 1094-2025. doi: https://doi.org/10.1016/j.red.2020.07.008. URL `https://www.sciencedirect.com/science/article/pii/S1094202520300648`.

Kroft, K., Lange, F., and Notowidigdo, M. J. Duration dependence and labor market conditions: Evidence from a field experiment. *The Quarterly Journal of Economics*, 128(3):1123–1167, 2013.

Kroft, K., Lange, F., Notowidigdo, M. J., and Katz, L. F. Long-term unemployment and the great recession: The role of composition, duration dependence, and non-participation. *Journal of Labor Economics*, 34(S1, Part 2):S7–S54, 2016. URL `http://www.journals.uchicago.edu/doi/abs/10.1086/682390`.

Krueger, A. B. and Mueller, A. I. A contribution to the empirics of reservation wages. *American Economic Journal: Economic Policy*, 8(1):142–79, February 2016. doi: 10.1257/pol.20140211. URL `https://www.aeaweb.org/articles?id=10.1257/pol.20140211`.

Krueger, A. B., Mueller, A., Davis, S. J., and Şahin, A. Job search, emotional well-being, and job finding in a period of mass unemployment: Evidence from high frequency longitudinal data. *Brookings Papers on Economic Activity*, pages 1–81, 2011.

Lalive, R., van Ours, J., and Zweimüller, J. The effect of benefit sanctions on the duration of unemployment. *Journal of the European Economic Association*, 3(6):1386–1417, 2005. URL `https://EconPapers.repec.org/RePEc:tpr:jeurec:v:3:y:2005:i:6:p:1386-1417`.

Lancaster, T. Econometric methods for the duration of unemployment. *Econometrica: Journal of the Econometric Society*, pages 939–956, 1979.

Le Barbanchon, T., Schmieder, J., and Weber, A. Chapter 6 - job search, unemployment insurance, and active labor market policies. volume 5 of *Handbook of Labor Economics*, pages 435–580. Elsevier, 2024. doi: https://doi.org/10.1016/bs.heslab.2024.11.006. URL `https://www.sciencedirect.com/science/article/pii/S1573446324000087`.

Lehmann, T. How directed is search in the labor market? Evidence from an online job board. *Working paper*, 2023.

Lise, J. On-the-Job Search and Precautionary Savings. *Review of Economic Studies*, 80 (3):1086–1113, 2013.

Ljungqvist, L. and Sargent, T. J. The European Unemployment Dilemma. *Journal of Political Economy*, 106(3):514–550, June 1998.

Ljungqvist, L. and Sargent, T. J. Two Questions about European Unemployment. *Econometrica*, 76(1):1–29, January 2008.

Lockwood, B. Information externalities in the labour market and the duration of unemployment. *The Review of Economic Studies*, 58(4):733–753, 1991.

Lovell, M. C. Seasonal adjustment of economic time series and multiple regression analysis. *Journal of the American Statistical Association*, 58(304):993–1010, 1963. doi: 10.1080/01621459.1963.10480682. URL `https://doi.org/10.1080/01621459.1963.10480682`.

Manning, A. and Petrongolo, B. How local are labor markets? evidence from a spatial job search model. *American Economic Review*, 107(10):2877–2907, October 2017. doi: 10.1257/aer.20131026. URL `https://www.aeaweb.org/articles?id=10.1257/aer.20131026`.

Marinescu, I. and Skandalis, D. Unemployment insurance and job search behavior. *The Quarterly Journal of Economics*, 136(2):887–931, 2021.

Muehlemann, S. and Strupler Leiser, M. Hiring costs and labor market tightness. *Labour Economics*, 52(C):122–131, 2018. URL `https://EconPapers.repec.org/RePEc:eee:labeco:v:52:y:2018:i:c:p:122-131`.

Mueller, A. I. and Spinnewijn, J. Predicting long-term unemployment risk. *Working paper*, 2023.

Mueller, A. I., Spinnewijn, J., and Topa, G. Job seekers' perceptions and employment prospects: Heterogeneity, duration dependence, and bias. *American Economic Review*, 111(1):324–63, 2021.

Nekoei, A. and Weber, A. Seven Facts about Temporary Layoffs. CEPR Discussion Papers 14845, C.E.P.R. Discussion Papers, June 2020.

Nüß, P. Duration dependence as an unemployment stigma: Evidence from a field experiment in germany. Technical report, Economics Working Paper, 2018.

Oberholzer-Gee, F. Nonemployment stigma as rational herding: A field experiment. *Journal of Economic Behavior & Organization*, 65(1):30–40, 2008.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

Pissarides, C. A. *Equilibrium Unemployment Theory, 2nd Edition*, volume 1 of *MIT Press Books*. The MIT Press, February 2000.

Potter, T. Learning and job search dynamics during the great recession. *Journal of Monetary Economics*, 117:706–722, 2021. ISSN 0304-3932. doi: https://doi.org/10.1016/j.jmoneco.2020.04.006. URL `https://www.sciencedirect.com/science/article/pii/S0304393220300532`.

Ramey, G., den Haan, W. J., and Watson, J. Job destruction and propagation of shocks. *American Economic Review*, 90(3):482–498, June 2000.

Robinson, P. M. Root-n-consistent semiparametric regression. *Econometrica*, 56(4):931–954, 1988. doi: https://doi.org/10.2307/1912705. URL `http://www.jstor.org/stable/1912705`.

Shimer, R. The Assignment of Workers to Jobs in an Economy with Coordination Frictions. *Journal of Political Economy*, 113(5):996–1025, 2005a.

Shimer, R. The cyclical behavior of equilibrium unemployment and vacancies. *American Economic Review*, 95(1):25–49, March 2005b.

Spinnewijn, J. Unemployed but optimistic: Optimal insurance design with biased beliefs. *Journal of the European Economic Association*, 13(1):130–167, 2015.

Van den Berg, G. J. and Van Ours, J. C. Unemployment dynamics and duration dependence. *Journal of Labor Economics*, 14(1):100–125, 1996.

van der Laan, M. J., Polley, E. C., and Hubbard, A. E. Super learner. *Statistical Applications in Genetics and Molecular Biology*, 6(1), 2007. doi: doi:10.2202/1544-6115.1309. URL `https://doi.org/10.2202/1544-6115.1309`.

Verbeek, M. and Nijman, T. Testing for selectivity bias in panel data models. *International Economic Review*, 33(3):681–703, 1992.

Vishwanath, T. Job search, stigma effect, and escape rate from unemployment. *Journal of Labor Economics*, 7(4):487–502, 1989.

Wolpert, D. H. Stacked generalization. *Neural Networks*, 5(2):241–259, 1992. ISSN 0893-6080. doi: https://doi.org/10.1016/S0893-6080(05)80023-1. URL `https://www.sciencedirect.com/science/article/pii/S0893608005800231`.

Wooldridge, J. M. Selection corrections for panel data models under conditional mean independence assumptions. *Journal of Econometrics*, 68(1):115–132, 1995.

Wright, R., Kircher, P., Julien, B., and Guerrieri, V. Directed search and competitive search equilibrium: A guided tour. *Journal of Economic Literature*, 59(1):90–148, 2021.

Yashiv, E. The determinants of equilibrium unemployment. *American Economic Review*, 90(5):1297–1322, December 2000.

Zuchuat, J. Estimating duration dependence in job search: the within-estimation duration bias, 2025. URL `https://arxiv.org/abs/2512.06928`.

Appendix

Table of Contents

A. Data and empirical measurements

In this Section, we provide further details on the contents of our main search diary data, that includes information from the cantons Bern (BE), St. Gallen (SG), Vaud (VD), Zug (ZG), and Zurich (ZH), as well as of the auxiliary data, that includes information from one employment office in Zurich.

Table A1: Job seekers' outcomes and selected observed characteristics, *main* and *auxiliary samples*

	Main sample			***Auxiliary sample***		
	Mean	St. Dev.	N	Mean	St. Dev.	N
A. Outcomes						
Person-month level (search-diary level)						
Job finding rate	0.061	(0.239)	58755	0.078	(0.269)	2783
Number of applications	10.553	(4.698)	58755	8.900	(4.597)	2783
Job interview rate	0.226	(0.418)	58755	0.289	(0.453)	2783
Application-level						
Interview Probability	0.040	(0.196)	600323	0.074	(0.262)	24770
Conditional Job Offer Probability	0.225	(0.418)	22422	0.206	(0.404)	1559
Unconditional Job Offer Probability	0.009	(0.095)	600323	0.015	(0.122)	24770
B. Individual characteristics						
Age	39.372	(11.898)	14798	39.307	(10.651)	655
1 = Female	0.458	(0.498)	14798	0.487	(0.500)	655
1 = Swiss	0.545	(0.498)	14798	0.539	(0.499)	655
1 = Primary education	0.269	(0.444)	14798	0.351	(0.478)	655
1 = Secondary education	0.588	(0.492)	14798	0.377	(0.485)	655
1 = Tertiary education	0.143	(0.350)	14798	0.189	(0.392)	655
1 = Manager	0.054	(0.225)	14798	0.092	(0.289)	655
1 = Specialist	0.598	(0.490)	14798	0.475	(0.500)	655
1 = Auxiliary	0.331	(0.471)	14798	0.423	(0.494)	655
C. Sample structure						
Time-period		04.2012 - 03.2013			07.2007 - 03.2008	
Region		BE, SG, VD, ZG, ZH			ZH	
Number of applications		600323			24770	
Person-month observations		58755			2699	
Number of individuals		14798			655	

Note: This table reports means and standard deviations on job seekers' outcomes, socio-demographic characteristics and sample information, for the *main sample* and *auxiliary sample*. The *main sample* is described in Section 3 of the main text. The *auxiliary sample* consists of approximately 30,000 job applications digitized from job search diaries at one PES in the Zurich region, during July 2007–March 2008. Data entry was carried out by a team of nine students over more than one year.

Table A2: Variables on application and job seeker characteristics

Characteristics of application and targeted job	
Application channel	variable with 3 categories: written, phone, personal
Work hours	variable with 2 categories: full-time, part-time
Caseworker referral	variable with 2 categories indicating whether application is to a job suggested by the caseworker
Rank	rank of application in a given month
Demographic characteristics	
Age	age in years and age category (9 categories)
Sex	variable with 2 categories: male, female
Education	two variables indicating highest educational degree (one with 3 categories, one with 6)
Nationality/residence permit	variable with 4 categories indicating Swiss nationality or type of residence permit if foreigner
Marital status	variable with 3 categories indicating marital status
Employment prospects	
Desired occupation	three variables indicating the first desired occupation at three different levels of aggregation (level 1 distinguishes 85 categories, level 2 38 categories and level 3 9 categories); two dummies indicating whether job-seeker has a second or third desired occupation
Health status	dummy indicating whether job seeker experienced sickness days during the unemployment spell
Employability	variable with 4 categories indicating caseworker's assessment of employability
Mobility	variable with 5 categories indicating degree of regional mobility
Employment history	
Previous position	variable with 10 categories indicating position in previous job
Previous occupation	variable with 9 categories indicating occupation in previous job
Previous wage	metric variable indicating average monthly wage in the year before the beginning of the unemployment spell and its logarithm
Unemployment history	dummy variable indicating whether someone has been unemployed in the up to five years before the start of the unemployment spell; variable indicating the number of unemployment months; variable indicating the number of unemployment episodes
Nonemployment history	dummy variable indicating whether someone has been nonemployed in the up to five years before the start of the unemployment spell; variable indicating the number of nonemployment months; variable indicating the number of nonemployment episodes
Length of observable history	variable indicating the length of the observable employment history (max. 60 months)

Note: This table documents the dictionary of variables considered in the estimation. In addition, we control for calendar time effects, local policy effects and time-varying local labor market conditions.

Figure A1: Job search diaries

Assurance-chômage

A remettre à l'ORP
au plus tard le 5 du mois suivant

Date de réception / du timbre postal

Preuves des recherches personnelles effectuées en vue de trouver un emploi

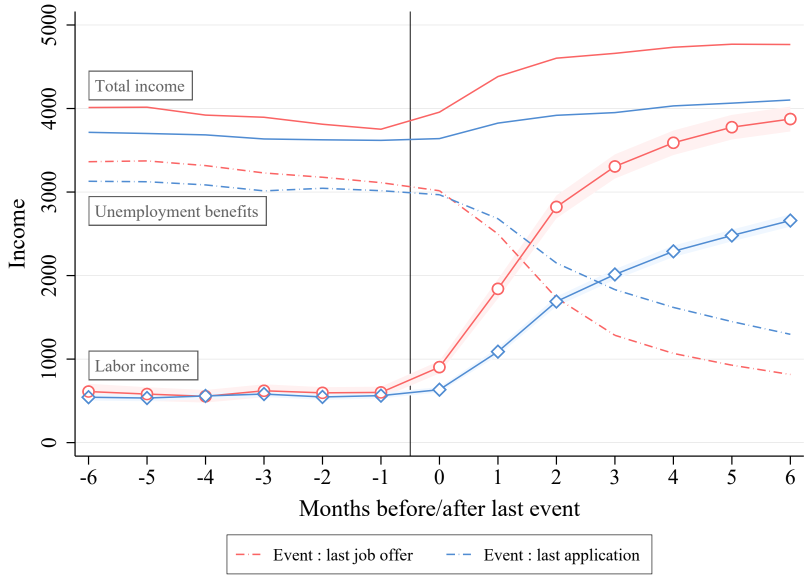
Nom et prénoms	No AVS	Mois et année

Date de l'offre de services (jour / mois)	Entreprise, adresse Personne contactée, numéro de tél.	Description du poste	Assignation ORP	Activité: à plein temps	Activité: à temps partiel (%)	Offre de service: par lettre / électronique	Offre de service: visite personnelle	Offre de service: par téléphone	Résultat de l'offre de service: en suspens	Résultat de l'offre de service: entretien	Résultat de l'offre de service: engagement	Résultat de l'offre de service: négatif	Motif

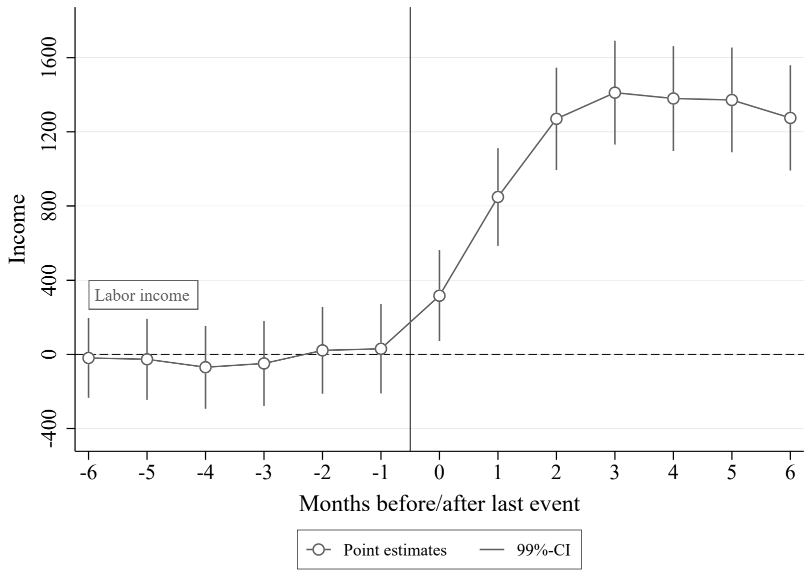
Note: This figure presents the job search diary form that job seekers have to use to document their search activities.

Figure A2: Job offers and income trajectories

(A) Observed average income trajectories

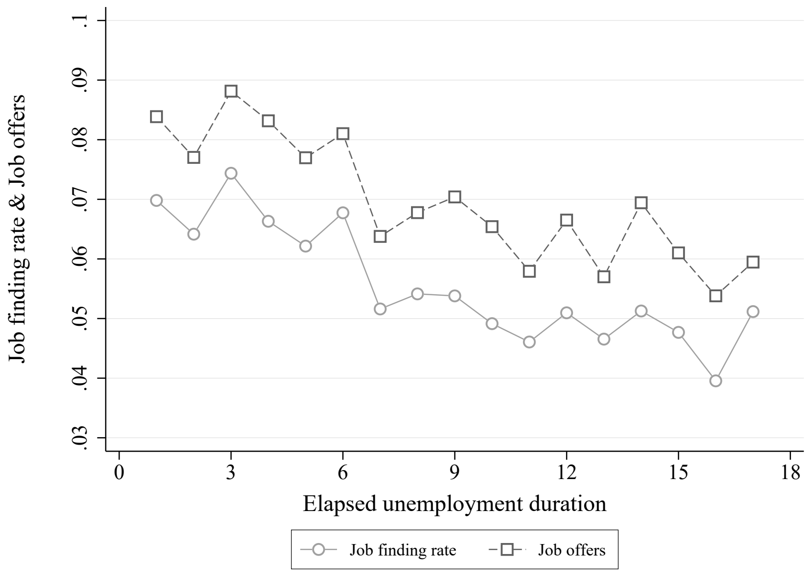


(B) Δ in labor income trajectories (accounting for heterogeneity)



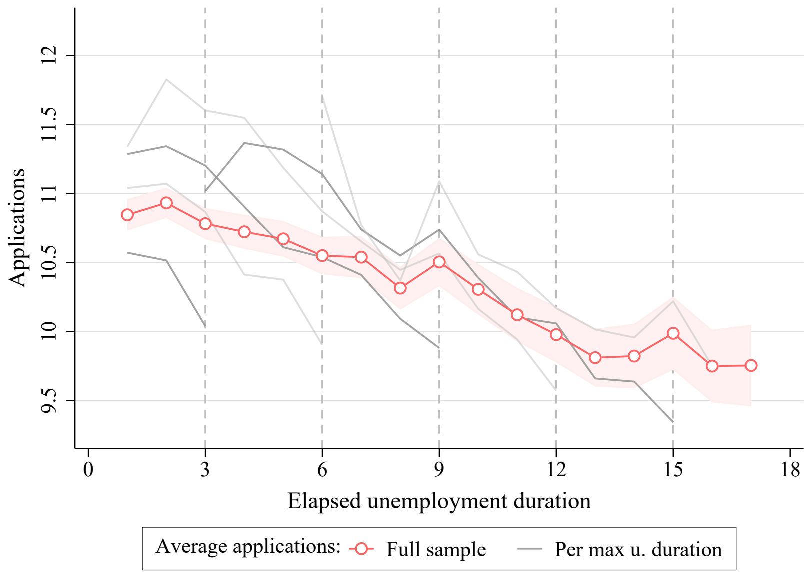
Note: This figure presents an event-study analysis, contrasting information from the search diary data and the social security data. It highlights the informational content of the search diaries. Panel A shows the average evolution of total income, labor income and unemployment benefits in months before and after individual-specific events. For each individual, the event is either the last month when a job offer is recorded (in red, if at least one job offer is recorded in the observed data) or the last month when search diaries are reported (in blue, if no job offer is recorded). Panel B presents the results of a two-way fixed effects specification, to measure the differences in the labor income trajectories of the two above mentioned groups.

Figure A3: Monthly job finding rate and number of job offers



Note: This figure plots the average monthly probability of a job offer together with the average monthly number of job offers.

Figure A4: Monthly overall applications, and conditional on duration



Note: This figure plots the average number of applications for the whole sample (in red), and conditional on the completed unemployment duration (in grey). The first grey line, which ends at 3 months, is for job seekers who were unemployed for at most three months, the second line for job seekers unemployed for at most 6 months, etc..

B. Details of the empirical analysis and further empirical results

B.1 Implementation of the double/debiased machine learning (DDML) approach

To estimate equation (2) and an analogous model for applications when duration dependence is modeled linearly and the function $g^Y(\cdot)$ is allowed to be high-dimensional and nonparametric we proceed as follows. Let the conditional expectation function of outcome $Y = A, C, O, U$ (*i.e.*, applications, callbacks/interviews, job offers per interview, and job offers per application) denote as

$$\mathbb{E}(Y_{\iota t} \mid t, X_{\iota t}) = \phi^Y t + g^Y(X_{\iota t}) \tag{B.1}$$

where the index $\iota \in \{i : 1, \ldots, N\}$ in the case of applications ($Y = A$) and $\iota \in \{i : 1, \ldots, N\} \times \{j : 1, \ldots, J\}$ in the case of the other outcomes, $Y = C, O, U$; the index t indicates the unemployment month. The vector $X_{\iota t}$ captures the observed covariates listed in Table A2 and their transformations (*e.g.* squares and interactions). To recover the residual effect of unemployment duration on outcome Y, ϕ^Y, we proceed in two steps building on the double/debiased machine learning approach proposed by Chernozhukov et al. (2018) and the short stacking approach described in Ahrens et al. (2024).

In step one, we estimate the conditional expectation functions $m^Y(X_{\iota t}) \equiv \mathbb{E}(Y_{\iota t} \,|\, X_{\iota t})$ and $l^Y(X_{\iota t}) \equiv \mathbb{E}(t \,|\, X_{\iota t})$, $Y = A, C, O, U$, on the relevant subsamples, *i.e.*, all person-month observations if $Y = A$, all person-application-month observations if $Y = C, U$, and only those person-application-month observations that led to an interview if $Y = O$.[32] We obtain predictions of these conditional expectation functions using the short stacking algorithm of Ahrens et al. (2024). Stacking is an ensemble method that averages over multiple base learners to obtain the final prediction model (Wolpert, 1992; Breiman, 1996; van der Laan et al., 2007). It improves in predictive performance over pre-selected single machine learners and, for the purpose of causal machine learning, offers additional robustness against biases due to misspecification of the conditional expectation function (Ahrens et al., 2024).[33] Stacking does not require the researcher to pre-select one particu-

[32] These conditional expectation functions are auxiliary estimands in the two-stage estimation procedure, see Robinson (1988) and Chernozhukov et al. (2018) for more information on the partially linear regression model and its estimation.

[33] For instance, whereas penalized regression techniques such as ridge or lasso perform well if the con-

lar machine learner. Instead, they can specify as base learners a range of different learning methods (including standard parametric models) as well as different tuning parameters and/or predictor dictionaries for a given learner. In the case of applications, $Y = A$, we consider as base learners for $m^A(\cdot)$ OLS regressions with two different specifications of the covariates, ridge and lasso regressions, as well as random forests and gradient-boosted trees. For the binary outcomes interviews and job offers, $Y = C, O, U$, we use standard, ridge and lasso logit regressions as well as random forests and gradient boosted trees as base learners for $m^Y(\cdot)$. For $l^Y(\cdot)$, we proceed analogously. To avoid overfitting we rely on cross-validated out-of-sample predictions of the base learners to obtain the final stacked learner (Ahrens et al., 2024).[34]

Specifically, we specify the following six base learners:

1. Baseline logit/OLS: uses a hand-selected set of control variables from the dictionary of variables in Table A2, fixed effects for calendar quarter times local labor markets and fixed effects for regional labor market policies (*i.e.*, 87 slope coefficients in total).
2. Flexible logit/OLS: uses the same variables as before plus their polynomials and interactions between them (*e.g.* 340 variables in total for interviews and 183 variables in total for job offers per interview).
3. Lasso: uses as dictionary the same variables as the flexible logit. The penalty term is chosen via cross-validation during the estimation (grid sizes vary between 4 and 100, depending on the outcome variable).
4. Ridge: uses as dictionary the same variables as the flexible logit. The penalty term is chosen via cross-validation during the estimation (grid sizes vary between 4 and 100, depending on the outcome variable).
5. Random forest: uses as dictionary all available predictor variables in their original form (see Table A2). Number of bootstrap replications and number of selected predictors at each replication are pre-selected through optimizing the predictive performance out of sample over a grid consisting of three to ten alternatives for

ditional expectation function can be well approximated by a linear combination of the predictors, tree-based methods show a superior performance if the conditional expectation function is highly non-linear in the predictors and/or involves complex interactions between them.

[34] We implement the estimations using the 'pystacked'-package in Stata written by Ahrens et al. (2023). This package calls the 'scikit-learn' suite in Python, see Pedregosa et al. (2011).

each tuning parameter.

6. Gradient boosted trees: uses as dictionary all available predictor variables in their original form. Number of bootstrap replications, number of splits, and learning rate are pre-selected through optimizing predictive performance out of sample over a grid consisting of two to five alternatives for each tuning parameter

We split the data into five folds such that the observations on a given job seeker appear in only one fold. Four folds are used as the training sample and one fold as the validation sample. We fit the six base learners on the training sample and use the held-out fold to obtain cross-validated predictions of $m^Y(\cdot)$ for each base learner k, $k = 1, \ldots, 6$. After iterating five times, so that every fold takes on the role of the validation sample once, we have six crossvalidated predictions for every observation in the estimation sample, $\hat{m}^Y_{\iota tk}$, $k = 1, \ldots, 6$. At this estimation stage, we also crossvalidate the penalty terms of the ridge and lasso estimators. The final stacked learner is a convex combination of the six base learners: $\hat{m}^Y_{\iota t} = \sum_{k=1}^{6} w_k \hat{m}^Y_{\iota tk}$, where w_k, with $0 \leq w_k \leq 1$, is the stacking weight, *i.e.*, the contribution of the k-th base learner to the stacked learner. To determine the stacking weights we fit the following constrained least squares regression on the full sample

$$\min_{w_1,\ldots,w_6} \sum_{\iota} \sum_{t} \left(Y_{\iota t} - \sum_{k=1}^{6} w_k \, \hat{m}^Y_{\iota tk} \right) \tag{B.2}$$

subject to the constraints $w_k \geq 0$ and $\sum_k w_k = 1$. We implement the same procedure to predict $l^Y(\cdot)$.

In the second step, we compute the residuals from the first step estimation, *i.e.*, $\check{Y}_{\iota t} \equiv Y_{\iota t} - \hat{m}^Y_{\iota t}$ and $\check{t}_{\iota t} \equiv t - \hat{l}^Y_{\iota t}$, which are then used in the regression

$$\check{Y}_{\iota t} = \phi^Y \check{t}_{\iota t} + \varepsilon^Y_{\iota t}\,. \tag{B.3}$$

This way we recover the residual effect of elapsed unemployment duration, $\hat{\phi}^Y$ that remains after partialling out the influence of observed covariates.

When residual duration dependence is modeled as a step function we implement the second step by regressing $\check{Y}_{\iota t}$ on a set of dummies for each value of elapsed unemployment duration and $\hat{l}^Y_{\iota t}$.

Moreover, we estimate versions of equation (2) that include in addition the estimated individual fixed effect, $\hat{\alpha}_i$, from the application equation, eq. (1) as right hand side variable, see column (4) of Table 3.[35] To account for the fact that $\hat{\alpha}_i$ is a generated regressor that is unbiased but inconsistent under large N, fixed T asymptotics, we implement the following approach that feeds a smoothed, cross-fitted estimate of α_i into the DDML approach.

In the first step, we use cross-fitting with five folds to estimate α_i. Let training set $k = 1, \ldots, 5$ be $\mathcal{T}_k = \{1, \ldots, N\} \backslash \mathcal{I}_k$, where $\mathcal{I}_k$ is the held-out fold. Iterating over k, we estimate eq. (1) on training set $\mathcal{T}_k$ and compute the individual fixed effect in applications as $\hat{\alpha}_i = \bar{A}_i - \overline{f^A(t)}\hat{\phi}^A_{\mathcal{T}_k} - \bar{X}_i \hat{\beta}^A_{\mathcal{T}_k}$ for $i \in \mathcal{I}_k$ ($\bar{Z}_i \equiv \sum_{t=1}^T Z_{it}$). Note that the individual fixed effects in applications from this first step are cross-fitted, i.e. the individual fixed effects in applications in the held out set, $\mathcal{I}_k$, are estimated using parameter estimates from the training set $\mathcal{T}_k$.

After we have obtained estimates of α_i for all i, we rank the estimates in ascending order and assign each i its percentile rank τ_i. Using the fold assignments from before, we estimate the τ-quantile $Q(\tau)$ of α on training set $\mathcal{T}_k$ as

$$\hat{Q}^{(-k)}(\tau) \;=\; \frac{\sum_{i \in \mathcal{T}_k} K_h(\tau - \tau_i)\, \hat{\alpha}_i}{\sum_{i \in \mathcal{T}_k} K_h(\tau - \tau_i)} \,,$$

where $K_h(\cdot)$ is the epanechnikov kernel, and evaluate $\hat{Q}^{(-k)}(\tau_j)$ at all τ_j such that $j \in \mathcal{I}_k$. The estimated quantiles of α_i are consistent for $Q(\tau)$ and can therefore be used as inputs in the DDML approach. This step smooths the cross-fitted estimates of the individual fixed effects in applications in the held out fold, $\mathcal{I}_k$, using a kernel estimator that averages over an increasing number of observations as the number of individuals i tends to infinity.

On the same training set $\mathcal{T}_k$, we estimate the conditional expectation functions $m^Y(X_{ijt})$ and $l^Y(X_{ijt})$, $Y = C, O, U$, as well as $q(X_{ijt}) \equiv \mathbb{E}(Q(\tau_i) \,|\, X_{ijt})$ and form residuals $\check{Y}_{ijt}$, $\check{t}_{ijt}$, and $\check{q}_{ijt}$. Pooling the residuals across folds, we run the regression of $\check{Y}_{ijt}$ on $\check{t}_{ijt}$ and $\check{q}_{ijt}$.

[35] Note that the DDML approach also works for a low-dimensional vector of target parameters (here: the partial effects of elapsed unemployment duration and the individual fixed effect in applications), see footnote 1 in Chernozhukov et al. (2018).

B.2 Job applications

B.2.1 Additional estimation results

Table B1 summarizes the performance of the stacking approach used to estimate the conditional expectation functions $m^A(X_{it})$ and $l^A(X_{it})$ needed to estimate duration dependence in eq. (B.1). Specifically, they show the predictive performance of the individual base learners as well the final stacked learner (columns (2) and (4)) along with the the stacking weight of each base learner (columns (1) and (3)). Specifically, we can see that learners relying on a linear index of the predictors contribute 37.6% (column (1)) and 25.5% (column (3)) to the final learner, while the two nonlinear learners, random forest and especially gradient boosted trees, contribute the rest. The best performing base learner is gradient boosted trees both for the conditional expectation $m(\cdot)$ (column (1)) as well as for $l(\cdot)$ (column (3)). Gradient boosted trees also have the by far highest stacking weight in both cases. Overall, the predictive performance of gradient boosted trees is somewhat worse than that of the stacked learner.

Table B1: Predictive performance and stacking weights, eq. (B.1) for applications

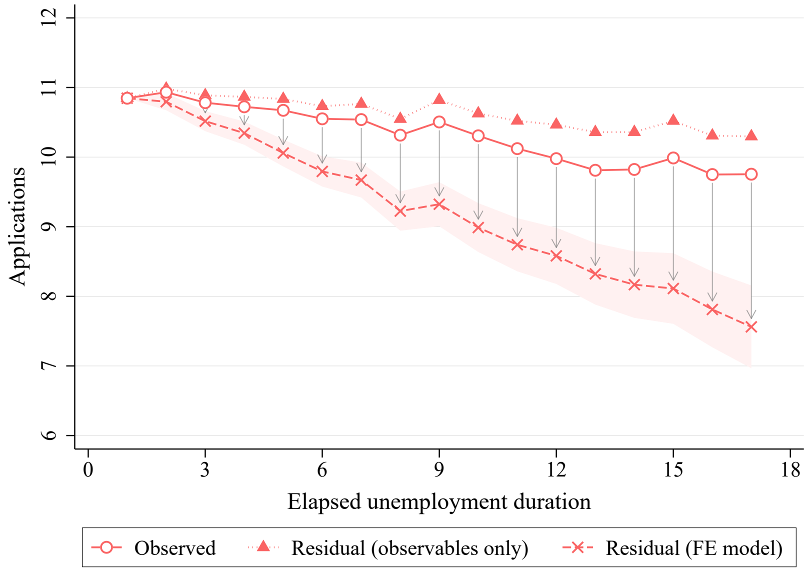
Dependent variables:	**Applications**		**Elapsed duration**	
	(1)	(2)	(3)	(4)
	Stacking weights	R squared	Stacking weights	R squared
Hand-curated	0.108	0.171	0.050	0.052
Very flexible	0.064	0.173	0.000	0.080
Lasso	0.204	0.175	0.067	0.083
Ridge	0.000	0.157	0.137	0.079
Random forest	0.130	0.154	0.140	0.073
Boosting	0.494	0.185	0.605	0.103
Stacked	1.000	0.191	1.000	0.107

Note: This Table summarizes intermediate results in the estimation of eq. (B.1) for applications. It reports the stacking weights, *i.e.*, the contributions of the base learners to the final stacked learner, along with the predictive performance of each of the base learners as well as the final learner (in the last row). Predictive performance is measured by the R^2. Total sample size is 58755 observations.

In Figure B1, we report results for a version of equation (1) that models duration dependence as a step function as well as an analogous OLS regression the controls for observable job-seeker characteristics. The figure distinguishes between the effect of elapsed unemployment duration on applications as observed in the data, the effect of duration net of observable heterogeneity and the effect of duration after controlling for observable heterogeneity and individual fixed effects.

To better understand what characterizes job seekers with higher values of the individual fixed effect, we predict the cross-fitted and smoothed estimate of the individual fixed

Figure B1: Duration profile of application effort, alternative prediction models



Note: This figure depicts the empirical profile of duration dependence in the number of job applications (solid line) and the estimated duration dependence that controls for observable heterogeneity and fixed effects (dashed line), where function $f^A(t)\phi^A$ in equation (1) is modeled as a step function with one dummy for each month of elapsed unemployment duration. The shaded area around the estimated duration dependence corresponds to the 95% confidence interval.

effect, $\hat{\alpha}_i$, from equation (1), using the independent variables given in Table A2. We fit OLS, lasso and ridge regressions as well as a random forest and gradient boosted trees and also consider an ensemble learner that consists of convex combination of the base learners. Table B2 summarizes the overall predictive performance of the models, whereas Table B3 shows the importance of the different groups of control variables in predicting the estimated individual fixed effect. According to Table B2 gradient boosted trees is the best performing single learner. It achieves an out-of-sample R^2 of 30.3%, implying a correlation of 55% in absolute value between actual and predicted values of the dependent variable, which is only marginally better than that of the best linear learner, the lasso, which is R^2 of 29.2%. The R^2 of the ensemble learner is 31%.

Table B3 shows that region fixed effects, measured using indicators for local labor market offices, are the most important predictors of $\hat{\alpha}_i$.[36] In the basic OLS specification without transformations of the predictors (e.g., interactions or higher-order terms), region fixed effects account for nearly 75% of the explained variation. Their contribution falls below 50% once transformations are introduced, as in the flexible OLS specification or the post-lasso OLS model, where the set of predictors is restricted to those selected in a first-round lasso. In these richer specifications (columns (2) and (3)), age-education

[36] Using dummies for cantons instead of local labor market offices yields similar results.

interactions, followed by indicators for past and desired occupations and flexible measures of the employment history, each explain between 20% and 11% of the variation in $\hat{\alpha}_i$. By contrast, variables capturing a job seeker's employability, most notably the caseworker's assessment of reemployment chances, contribute almost nothing.

Comparing these linear models with nonlinear approaches such as random forests and gradient boosted trees, displayed in columns (4) and (5), shows that conclusions about variable importance depend strongly on the modeling strategy. In the random forest model, employment history contributes the largest share (37%) to reducing the sum of squared residuals, about twice as much as region fixed effects. In the gradient boosted trees model, by contrast, region fixed effects dominate with a 55% share, followed by occupation indicators at 15%. These patterns highlight that with many discrete and correlated predictors, different methods can achieve similar predictive performance while attributing importance to different groups of variables. Combined with the fact that predictive performance is good but far from perfect, this makes it difficult to tie the individual fixed effect to any specific observable trait of job seekers. In fact, unobserved individual differences account for the bulk of the variance in the individual fixed effect in applications. Taken together, these findings support our modeling choice to treat the individual fixed effect in applications, net of observed characteristics, as capturing worker-level search efficiency rather than differences across market segments.

Table B2: Prediction of the estimated individual fixed effect in eq. (1)

	(1)	(2)
	Stacking weights	R-squared
Basic OLS	0.000	0.270
Flexible OLS	0.292	0.288
Lasso	0.089	0.292
Ridge	0.000	0.290
Random Forest	0.007	0.263
Boosted Trees	0.612	0.303
Stacked	1.000	0.310

Note: The cross-fitted and smoothed individual fixed effect in eq. (1), $\hat{\alpha}_i$ is predicted based on OLS, lasso and ridge regressions as well as a random forest, and gradient boosted trees using the independent variables given in Table A2. This table summarizes the predictive performance of each base learner as well as the stacked learner that is a convex combination of the base learners. It reports the stacking weights, *i.e.*, the contributions of the base learners to the final stacked learner, along with the predictive performance of each of the base learners as well as the stacked learner (in the last row). Total sample size is 11602 observations.

B.2.2 Robustness Checks

We perform several robustness checks to assess the validity of our finding that unemployment duration negatively affects the number of applications per month, with a steeper gradient once individual fixed effects are accounted for.

Table B3: Variable importance statistics for predicting the estimated individual fixed effect

	(1)	(2)	(3)	(4)	(5)
	Basic OLS	Flexible OLS	Post-Lasso OLS	Random Forest	Grad. Boosted Trees
Age, education	0.090	0.195	0.200	0.145	0.088
Employment history	0.012	0.147	0.110	0.368	0.086
Occupation	0.081	0.156	0.148	0.161	0.152
Demographic characteristics	0.001	0.010	0.009	0.051	0.013
Employability	0.008	0.012	0.007	0.039	0.007
Regional heterogeneity	0.748	0.468	0.484	0.175	0.553
Calendar time at start	0.061	0.014	0.043	0.062	0.100

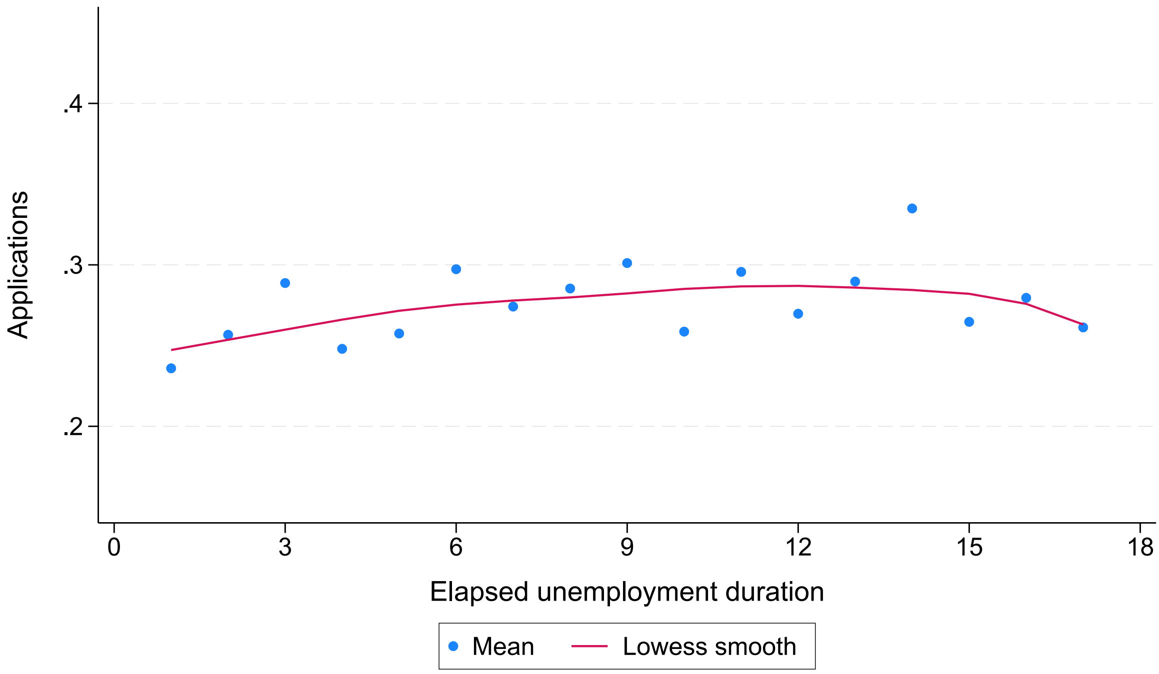
Note: The cross-fitted and smoothed individual fixed effect in eq. (1), $\hat{\alpha}_i$, is predicted based on OLS regressions, a random forest, and gradient boosted trees using the independent variables given in Table A2. This Table shows the importance of the different groups of control variables for predicting $\hat{\alpha}_i$. Group 'Employment history' includes variables capturing the previous unemployment, employment and nonemployment history as well as the previous wage. Group 'Occupation' includes indicators for the desired and the previous occupation as well as the previous position. Group 'Demographic characteristics' includes other demographic characteristics besides age and education. Group 'Employability' includes indicators for the job seekers health, regional mobility and the caseworker's assessment of reemployment chances. Group 'Regional heterogeneity' includes indicators for the local employment offices. Group 'Calendar time at start' includes indicators for the quarter and year when the unemployment spell began. Columns (1) to (3) show the contributions to the (in-sample) R^2 measured as the average change in the R^2 across all possible combinations of the groups of control variables in the full model, and expressed relative to the total contribution. The variable importance measures shown in columns (2) and (3) correspond to the average decrease in the sum of squared residuals, expressed relative to the total decrease.

First, we consider alternative measures of the number of applications per month to verify that the number of applications per month truly captures the intrinsic effort of the job seeker rather than external factors. In fact, there are two reasons why our baseline application variable may be confounded by external factors. On the one hand, not all applications directly result from the job seeker's own initiative, but some occur in response to a suggestion by the caseworker (a referral). On the other hand, caseworkers set a minimum search requirement for every job seeker which specifies the minimal number of job applications that a job seeker has to make every month.

In a first robustness check, we redefine the application variable to exclude those applications made in response to a suggestion by the caseworker. Caseworker referrals affect 12% of the observations in our sample. Figure B2 plots the mean number of applications to a vacancy referral by the caseworker by elapsed unemployment duration. The mean number of applications follows an inverse U-shaped pattern peaking at 12 months. The overall average is 0.27 applications per month. Columns (1) and (2) of Table B4 report the estimation results of this exercise that considers only the applications made on the job seeker's own initiative. The estimates of duration dependence are very similar to our baseline results. Panel C of Figure B3 confirms this visually: after controlling for individual fixed effects the duration gradient becomes steeper.

In a second robustness check, we draw on an additional administrative register to impute search requirements for the job seekers in our analysis sample. Alongside the pen-and-paper documentation of job search used in our main analysis, the federal employment

Figure B2: Applications to vacancy referrals



Note: The figure shows the mean number of applications to vacancy referrals by the caseworker by elapsed unemployment duration.

agency also maintained an electronic register in which caseworkers could summarize key information from the paper forms. However, use of this centrally managed register varied across cantons. While about three-quarters of all caseworkers across the five cantons have ever entered search requirement data into the central system, only in two of the five cantons, Berne and St. Gallen, does a large majority of unemployment spells contain at least one nonzero entry.[37]

According to the data for 2012 and 2013 in the central register, the median of search requirements is ten applications per month in the cantons of Zurich and Vaud and between six and eight in the other three cantons. Search requirements tend to be somewhat higher for low skilled job seekers and lower for older job seekers.

We conduct two separate robustness checks based on two different extracts from the central register: (i) all unemployment spells recorded in 2012–13 in Berne and St. Gallen, and (ii) all nonzero search requirements in 2012–13 in all five cantons. For each robustness check, we use the relevant extract to compute, in each canton, the median number of required applications in 96 cells defined by calendar year, elapsed unemployment duration,

[37] From the database, it is unclear whether a search requirement of zero indicates no requirement or missing data for that month.

education, and age.[38] We then match the cell-specific search requirements to job seekers in our analysis sample and calculate the excess number of applications as the difference between the actual and imputed required applications.

Columns (3)–(6) of Table B4 and panels (A) and (B) of Figure B3 present the resulting duration profiles based on these alternative application measures. The results closely mirror our baseline findings.

Table B4: Duration dependence in applications, alternative application measures

Dependent variables:	App. on own initiative		Excess applications			
			Berne, St. Gallen		All cantons	
	(1)	(2)	(3)	(4)	(5)	(6)
Elapsed unemployment duration	-0.081***	-0.188***	0.011	-0.142***	-0.009	-0.174***
	(0.008)	(0.021)	(0.011)	(0.030)	(0.008)	(0.021)
	[-0.760%]	[-1.768%]	[0.379%]	[-5.114%]	[-0.548%]	[-10.562%]
Individual controls	No	No	No	No	No	No
Policy controls	No	No	No	No	No	No
Local labor market conditions	No	Yes	No	Yes	No	Yes
Individual FE	No	Yes	No	Yes	No	Yes
Mean outcome 1st month	10.610	10.610	2.778	2.778	1.647	1.647
Adjusted R^2	0.006	0.036	0.000	0.031	0.000	0.033
Observations	58755	58755	24023	24023	58755	58755
Persons	14798	14798	7091	7091	14798	14798

Note: This table reports estimates of equation (1) for our alternative measures of applications (applications on own initiative and excess applications), where the duration function $f^A(t)\phi^A$ is specified linearly. Models are estimated using OLS. For each independent variable, we consider either a bivariate model or the full specification. Standard errors clustered at the individual level are reported in parentheses. Stars indicate the following significance levels: * 0.1, ** 0.05 and *** 0.01.

Second, we discuss sensitivity of our results to potentially nonrandom attrition and missing or unusually low numbers of applications in the final month before exit.

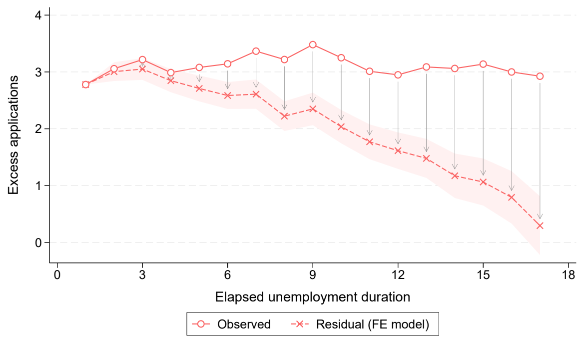
In principle, all job seekers need to document their search efforts on a standardized form that they need to return to the local employment office by the fifth of the next month. Caseworkers review the forms and check the truthfulness of applications. They also follow up on the success of applications documented on previous forms. Failure to demonstrate sufficient search effort or untruthful information on search efforts may entail a benefit sanction (*i.e.*, a reduction or suspension of benefits) of up to three months.

There are several reasons why a job seeker may not complete a search diary in a given month. First, the law grants job seekers the right to temporary exemptions from job search. For every three months of monitored search activity, a job seeker may request one week off without being required to document job search efforts. Second, job

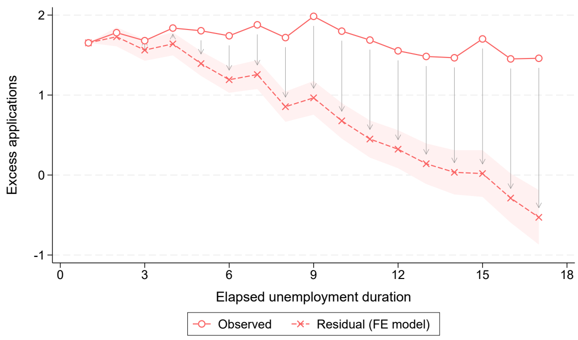
[38] In the small canton of Zug, we omit the education dimension, yielding 23 cell medians, whereas in the large canton of Zurich, we use more detailed age groups, yielding 128 cell medians.

Figure B3: Empirical and residual duration profiles of alternative application measures

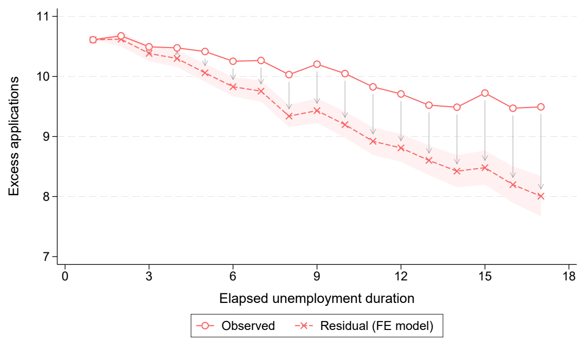
(A) Excess applications (BE, SG)



(B) Excess applications (all cantons)



(C) Applications on own initiative



Note: In each panel, the solid line depicts the empirical duration profile of excess applications (panels A and B) and applications on own initiative (panel C) and the dashed line the estimated duration dependence obtained after controlling for individual fixed effects, with function $f^A(t)\phi^A$ in equation (2) modeled as a step function with one dummy for each month of elapsed unemployment duration. The shaded area around the estimated duration dependence corresponds to the pointwise 90% confidence interval.

seekers are exempt from search requirements during periods of sickness or when granted temporary exemptions by their caseworker. In particular, job seekers are not required to continue searching once they have secured employment and know they will soon exit unemployment.

In the data, the share of search diaries reported every month is around 91% and constant throughout most of the unemployment spell, except in the first month of unemployment, when it attains 100%, and in the month immediately preceding exit from unemployment, when it drops to 35%. To address potential concerns about the effects of nonrandom attrition and missing or unusually low numbers of applications in the final month before exit, we conducted several sensitivity analyses, reported in Table B5

In columns (1) through (3) of Table B5, we report a summary of the sensitivity analyses that examine the robustness of our results to anticipation of future exit from unemployment. The results in column (1), which corresponds to our baseline specification also reported in Table 2, suggest that job seekers anticipate their exit from unemployment and hence dropout from the panel. Specifically, unemployment exit in the next month reduces search effort by two applications in the current month. This effect is not surprising as job seekers usually know some time in advance they are going to take up a new job. Omitting the indicator for future unemployment exit has a noticeable effect on the estimated duration dependence which increases by about 25% once dropout in the near future is accounted for (compare the estimated coefficient on elapsed unemployment duration in columns (1) and (2)). The estimation results reported in column (3) are based on a restricted sample that excludes terminal unemployment months. The estimated duration effect is very similar to that in column (1). The way how we deal with the anticipation of exit from unemployment hardly affects the sample distribution of the estimated individual fixed effects (summarized through its mean and standard deviation), supporting the idea that it is unbiased.

To address concerns about nonrandom attrition in months before unemployment exit, we conducted further sensitivity analyses, reported in columns (4) and (5) of Table B5. In these additional robustness checks, we restricted the estimation sample to unemployment spells with search diaries reported in every month (column (4)), and imputed zero applications in non-terminal months with missing search diaries (column (5)). None of these modifications materially affect the estimated duration dependence in applications nor the sample distribution of the individual specific effect. Note that we do not report results for a specification that imputes zero applications in all months with missing search diaries, including also the terminal months, because this would mechanically produce even more negative estimates of duration dependence.

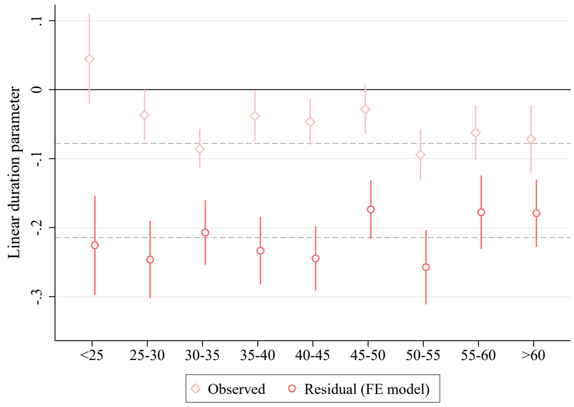
Table B5: Duration dependence in applications, robustness checks

	(1)	(2)	(3)	(4)	(5)
Dependent variable: Applications					
Elapsed unemployment duration	-0.207***	-0.251***	-0.205***	-0.211***	-0.221***
	(0.021)	(0.021)	(0.021)	(0.027)	(0.024)
	[-1.907%]	[-2.315%]	[-1.893%]	[-1.934%]	[-2.074%]
Exit next month	-2.055***			-2.164***	-1.411***
	(0.125)			(0.151)	(0.132)
Mean of $\hat{\alpha}_i$	11.506	11.789	11.503	11.942	10.471
St. dev. of $\hat{\alpha}_i$	4.141	4.185	4.198	4.240	4.363
Mean outcome 1^{st} month	10.846	10.846	10.846	10.919	10.671
Adjusted-R^2	0.041	0.030	0.026	0.044	0.027
Observations	58755	58755	56684	32834	65165
Persons	14798	14798	14366	8595	14798

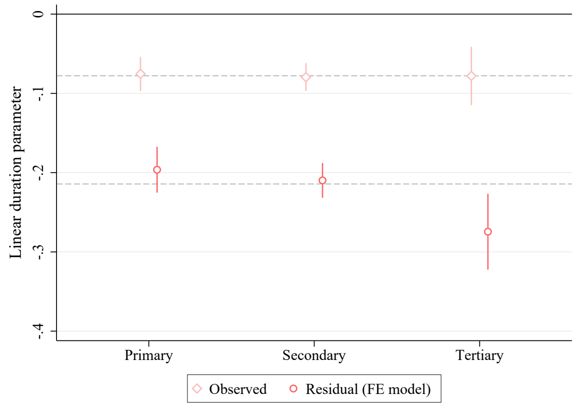
Note: This table reports estimates of duration dependence using fixed effects regression as in eq. (1), where duration dependence is specified linearly. Specification (1) corresponds to the baseline specification shown in Table 2. Specification (2) excludes the control indicating an exit from unemployment in the subsequent month. Specification (3) is estimated from a restricted sample that drops the observations preceding unemployment exit. Specification (4) is estimated on a restricted sample that drops the observations from incomplete search spells. Specification (5) is estimated from an augmented sample that imputes zero applications for search months with no reported application information. Standard errors are clustered at the individual level and reported in parentheses. Coefficients in relative terms (with respect to the average in the first month of unemployment) are indicated in square brackets. Stars indicate the following significance levels: * 0.1, ** 0.05 and *** 0.01.

Figure B4: Heterogeneity in the effect of duration on applications

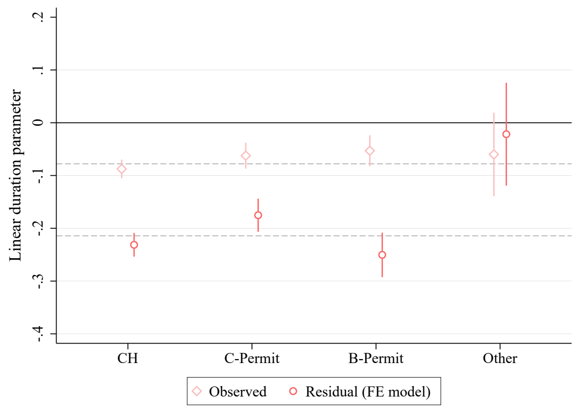
(A) Age



(B) Education



(C) Nationality/Residential status



Note: This figure depicts the estimation results of Equation (1) on sub-samples based on various observables, where the $f(t)\phi^A$ is specified linearly. The estimated coefficient is reported together with pointwise 90% confidence intervals.

B.3 Job interviews and job offers

Tables B6 through B8 provide summary information on the performance of the stacking approach used to estimate the conditional expectation functions $m^Y(X_{ijt})$ and $l^Y(X_{ijt})$ for the outcomes $Y = C, O, U$ (callbacks for interviews, job offers per interview, job offers per application), see Section B.1. Specifically, they show the predictive performance of the individual base learners as well the final stacked learner (columns (2), (4), and (6)) along with the the stacking weight of each base learner (columns (1), (3), and (5)). Turning to Table B6, we can see that learners relying on a linear index of the predictors contribute 36% (column (1)) and 34% (column (3)) to the final learner, while the two nonlinear learners, random forest and gradient boosted trees, contribute the rest. The predictive performance of the best performing linear learner (*i.e.*, the ridge model) is much worse than that of the best performing nonlinear learner (*i.e.*, gradient boosted trees) in columns (1) and (3). A different pattern emerges for job offers per interview, as can be seen in Table B7. Here, learners that rely on a linear index of the predictors contribute 56% (column (1)) and 44% (column (3)) to the final learner, and the predictive performance of the best linear learner is comparable to that of the best nonlinear learner, see columns (1) and (3) of Table B7. Both in Table B6 and in Table B7, the predictive performance of the best base learner is somewhat worse than that of the stacked learner. The results for job offers per application shown in Table B8 in columns (3) through (6) are by definition identical to those for interviews in Table B6. According to column (2) of Table B8 the best linear and nonlinear learners exhibit a similar predictive performance as in the case of job offers per interview, cf. column (2) of Table B7.

Figure B5 shows the average partial effects of elapsed unemployment duration evaluated at different deciles of the distribution of the individual fixed effect in applications (eq. (1)). Specifically, the average partial effects are based on estimates of the following modified version of equation (2):

$$\mathbb{P}(Y_{ijt} = 1 \mid t, \hat{\alpha}_i, X_{ijt}, Z_{ijt} = 1) = H\big(\phi_1^Y t + \phi_2^Y t \times \hat{\alpha}_i + \gamma^Y \hat{\alpha}_i + g^Y(X_{ijt})\big)\,, \qquad \text{(B.4)}$$

with $Y = C, O$.

Panels (A) and (B) of Figure B5 display estimates of the average partial effect of unemployment duration on the probability that an application leads to an interview,

Table B6: Predictive performance and stacking weights, eq. (2) for interviews

Dependent variables:	**Job interview**		**Elapsed duration**		**Est. indi. fixed effect**	
	(1)	(2)	(3)	(4)	(5)	(6)
	Stacking weights	Area under ROC	Stacking weights	R squared	Stacking weights	R squared
Hand-curated	0.107	0.643	0.066	0.047	0.826	0.356
Very flexible	0.000	0.472	0.022	0.077	0.163	0.245
Lasso	0.000	0.608	0.282	0.082	0.000	0.252
Ridge	0.258	0.648	0.000	0.078	0.000	0.248
Random forest	0.156	0.618	0.191	0.077	0.010	0.211
Boosting	0.479	0.655	0.440	0.092	0.000	0.254
Stacked	1.000	0.657	1.000	0.100	1.000	0.361

Note: This Table summarizes intermediate results in the estimation of eq. (2) for job interviews. It reports the stacking weights, *i.e.*, the contributions of the base learner to the final stacked learner, along with the predictive performance of each of the base learners as well as the final learner (in the last row). For the binary outcome job offer, predictive performance is measured as the area under the receiving operating characteristic (ROC) curve, for elapsed unemployment duration and the estimated individual fixed effect as the outcomes it is measured as the R^2. Total sample size is 600323 observations.

Table B7: Predictive performance and stacking weights, eq. (2) for job offers per interview

Dependent variables:	**Job offer**		**Elapsed duration**		**Est. indi. fixed effect**	
	(1)	(2)	(3)	(4)	(5)	(6)
	Stacking weights	Area under ROC	Stacking weights	R squared	Stacking weights	R squared
Hand-curated	0.271	0.597	0.076	0.023	0.254	0.192
Very flexible	0.000	0.568	0.031	0.040	0.149	0.195
Lasso	0.002	0.586	0.429	0.045	0.054	0.201
Ridge	0.290	0.610	0.000	0.041	0.074	0.199
Random forest	0.182	0.601	0.273	0.042	0.011	0.171
Boosting	0.255	0.616	0.191	0.040	0.457	0.208
Stacked	1.000	0.627	1.000	0.054	1.000	0.223

Note: This Table summarizes intermediate results in the estimation of eq. (2) for job offers per interview. It reports the stacking weights, *i.e.*, the contributions of the base learner to the final stacked learner, along with the predictive performance of each of the base learners as well as the final learner (in the last row). For the binary outcome job offer, predictive performance is measured as the area under the receiving operating characteristic (ROC) curve, for elapsed unemployment duration and the estimated individual fixed effect as the outcomes it is measured as the R^2. Total sample size is 22422 observations.

while panels (C) and (D) show the corresponding results for the probability of a job offer conditional on an interview. The findings suggest that the negative duration dependence in interviews tends to be less negative for job seekers with a high individual fixed effect in applications. Similarly, the positive duration dependence in job offers conditional on an interview also tends to be stronger for job seekers with a high individual fixed effect in applications. While these patterns emerge consistently using both logit and DDML the differences in duration dependence across the distribution of the individual fixed effects in applications are not statistically significant.

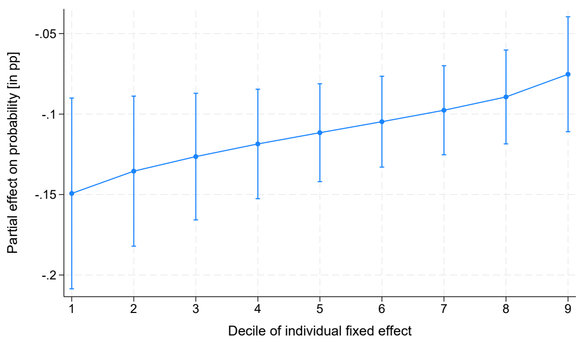
Table B8: Predictive performance and stacking weights, eq. (2) for job offers per application

Dependent variables:	Job offer (uncond.)		Elapsed duration		Est. indi. fixed effect	
	(1)	(2)	(3)	(4)	(5)	(6)
	Stacking weights	Area under ROC	Stacking weights	R squared	Stacking weights	R squared
Hand-curated	0.508	0.638	0.066	0.047	0.826	0.356
Very flexible	0.000	0.493	0.022	0.077	0.163	0.245
Lasso	0.000	0.494	0.282	0.082	0.000	0.252
Ridge	0.239	0.647	0.000	0.078	0.000	0.248
Random forest	0.176	0.615	0.191	0.077	0.010	0.211
Boosting	0.077	0.663	0.440	0.092	0.000	0.254
Stacked	1.000	0.654	1.000	0.100	1.000	0.361

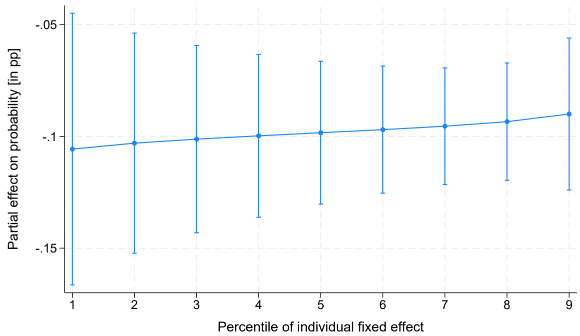
Note: This Table summarizes intermediate results in the estimation of eq. (2) for job offers per application. It reports the stacking weights, *i.e.*, the contributions of the base learner to the final stacked learner, along with the predictive performance of each of the base learners as well as the final learner (in the last row). For the binary outcome job offer, predictive performance is measured as the area under the receiving operating characteristic (ROC) curve, for elapsed unemployment duration and the estimated individual fixed effect as the outcomes it is measured as the R^2. Total sample size is 600323 observations.

Figure B5: Average partial effects of duration by decile of the individual fixed effect

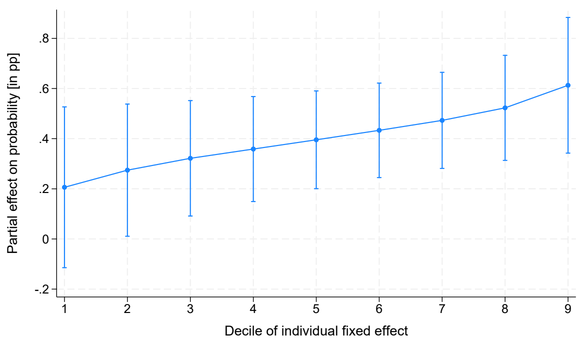
(A) Effect on interview probability (logit)



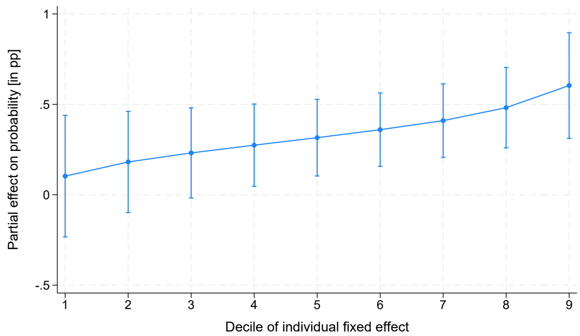
(B) Effect on interview probability (DDML)



(C) Effect on job offer probability (logit)



(D) Effect on job offer probability (DDML)



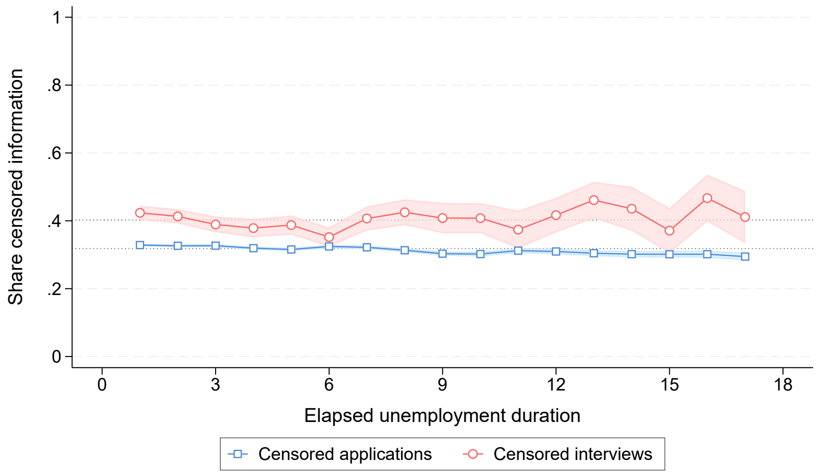
Note: This figure shows estimates of duration effects on the probability of a job interview and a job offer per interview, respectively, based on equation (B.4). Panels (A) and (C) correspond to a standard logit regression with $g^I(X_{ijt}) = X_{ijt}\beta^I$, whereas panels (B) and (D) model $g^I(X_{ijt})$ nonparametrically and $H(\cdot)$ is the identity link (linear probability model). Application-level observations are weighted by the inverse of the monthly number of applications made by individual i in month t, so as to put equal weight on all person-month observations. Point estimates correspond to average partial effects (in percentage points).

B.4 Robustness of results to censoring

In our data, we observe at each stage of the job search process whether an employer responded. Applications with an unknown employer response at any later stage of the job search process (*i.e.* interview or job offer) are classified as censored. Likewise, interviews for which it is unknown whether they resulted in a job offer are also classified as censored. Missing information on application outcomes could pose a concern if the incidence of censoring changed systematically over the course of unemployment. Figure B6 plots the share of censored applications and interviews by elapsed unemployment duration, showing that both shares remain constant throughout the unemployment spell.

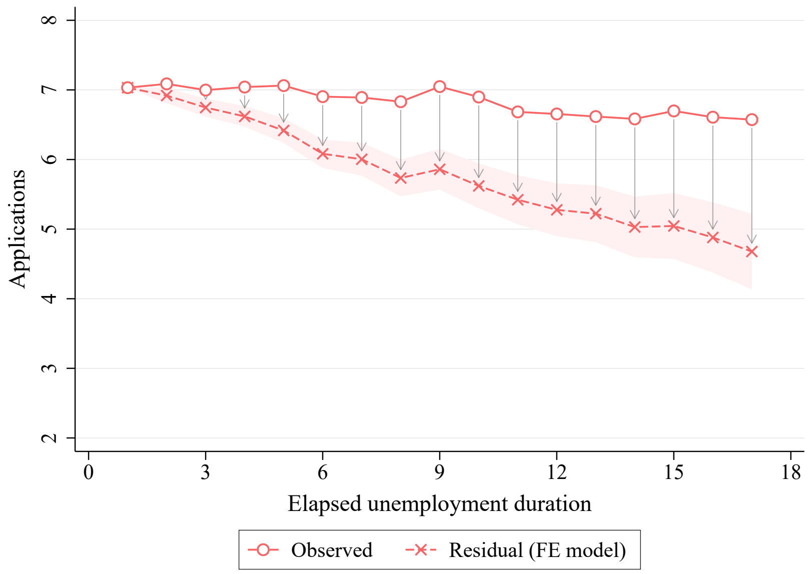
In additional analyses that restrict the sample to non-censored applications, we confirm that the duration dependence patterns are essentially identical to those reported in the main text, which are based on all applications, compare Figures B7, B8A and B8B in this subsection to Figures 3A, 4A and 4B in Section 4.

Figure B6: Incidence of censoring by elapsed unemployment duration



Note: This figure plots the share of censored applications at the interview and job-offer stages by elapsed unemployment duration, along with pointwise 95% confidence intervals. The line labeled "Censored applications" shows the share of applications for which it is unknown whether they resulted in an interview. The line labeled "Censored interviews" refers to the share of interviews for which it is unknown whether they resulted in a job offer (out of all applications that led to an interview).

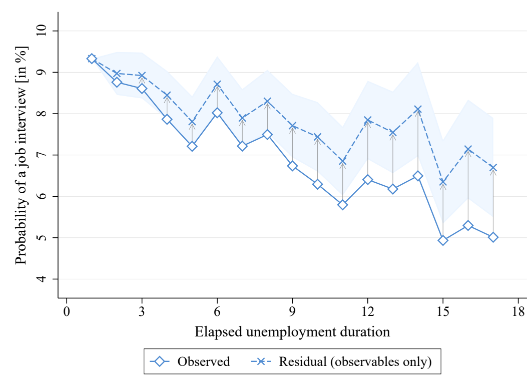
Figure B7: Duration profile of applications, non-censored applications only



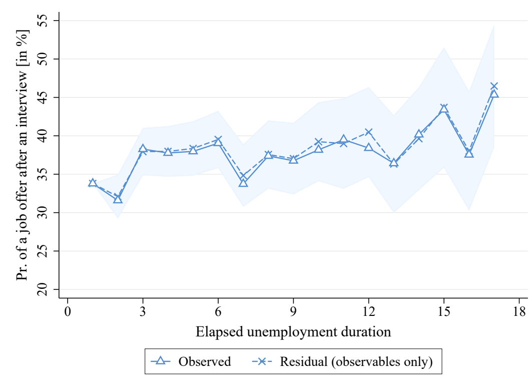
Note: This Figure depicts the empirical duration dependence in the number of job applications (solid line) and the estimated duration dependence obtained after controlling for observable heterogeneity and individual fixed effects (dashed line), with function $f^A(t)\phi^A$ in equation (1) modeled as a step function with one dummy for each month of elapsed unemployment duration. The shaded area around the estimated duration dependence corresponds to the 95% confidence interval. Only non-censored applications are considered.

Figure B8: Duration profiles of job interviews and job offers, applications with known outcome only

(A) Duration profile of job interviews, applications with known interview outcome only



(B) Duration profile of job offers, applications with known job offer outcome only



Note: This figure depicts the empirical duration dependence in the probability of a job interview/job offer per interview (solid line) and the estimated duration dependence obtained after controlling for observable heterogeneity (dashed line), with function $f^C(t)\phi^C$ in equation (2) modeled as a step function with one dummy for each month of elapsed unemployment duration. The shaded area around the estimated duration dependence corresponds to the pointwise 90% confidence interval. Only non-censored applications are considered.

B.5 Job interviews, job offers, and job finding at the person-month level

In this section, we report evidence on the individual interview probability, individual job offer probability per interview, and job finding rate at the person-month level that are used as targets in the structural estimation (see Section D). For this purpose, we aggregate the application-level information to the person×month (or search diary) level (see Panel A of Table A1 for descriptive statistics). In contrast, the evidence discussed in Section 4 and Appendix B is based on data at the application level.

Specifically, we estimate three versions – one for each outcome, "at least one job interview" ($Y = I$), "at least one job offer conditional on at least one interview" ($Y = O$), and "at least one job offer" ($Y = F$) – of the following model for individual i in month t of unemployment:

$$\mathbb{P}(Y_{it} = 1 \mid t, X_{it}) = H\big(f^Y(t)\phi^Y + g^Y(X_{it})\big) \tag{B.5}$$

where the function $H(\cdot)$ is a link function. The term $f^Y(t)\phi^Y$ denotes a linear (in ϕ^Y) specification of duration dependence and $g^Y(X_{it})$ a potentially nonparametric function of a rich set of observed covariates X_{it}, including job seeker characteristics as well as calendar quarter times local labor market fixed effects and regional labor market policy fixed effects. We estimate eq. (B.5) using the double/debiased machine learning approach, analogously to eq. (2) for interviews and job offers. Figure B9 reports the observed duration profiles of the individual interview probability, the individual job offer probability per interview, and the job finding rate, along with the profiles controlling for observable characteristics.

C. Details of the structural model

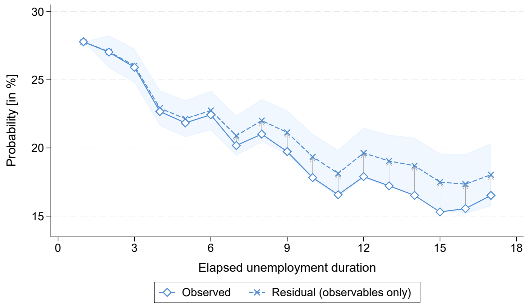
C.1 Baseline model

The *individual interview probability* – the probability that a job seeker exerting optimal search effort receives an interview at duration τ – equals $c^{ind}(x, \gamma, \tau) = s(x, \gamma, \tau)\ c(\tau)$.

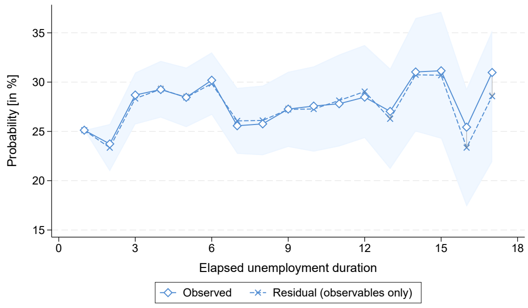
As a result of the two-stage recruitment process, expected profits of a firm with productivity y upon meeting a job seeker equal $\Pi(y) = \sum_{\tau=0}^{\infty} r(\tau) \int \left(J(x, y)q(x, y) - \kappa\right)\ \mu(x|\tau)dx\, \mathcal{C}(y, \tau)$, where $r(\tau)$ denotes the probability of meeting a job seeker with unemployment duration τ.

Figure B9: Duration profiles at the person-month level

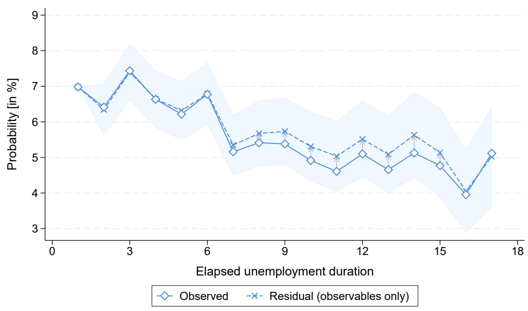
(A) Individual interview prob.



(B) Individual job offer prob. per interview



(C) Job finding rate



Note: This figure reports the empirical duration dependence in the individual interview probability (probability of at least one job interview in a month) in Panel A, individual job offer probability per interview (probability of at least one job offer conditional on at least one job interview in a month) in Panel B, and job finding rate in Panel C, along with the estimated duration dependence obtained after controlling for observable heterogeneity (dashed lines). To control for observable heterogeneity, we use double/debiased machine learning with stacking, with duration dependence modeled as a step function with one dummy for each month of elapsed unemployment duration. The shaded area around the estimated duration dependence corresponds to the pointwise 90% confidence interval.

C.2 Equilibrium characterization

For the equilibrium to feature statistical discrimination, we restrict our attention to equilibria where the job finding rate is increasing in ability at any duration (accounting for the correlation between ability and search types).

Proposition 1 provides a sufficient condition for the job offer probability per application and per interview to exhibit differential duration dependence. All the proofs are relegated to Appendix C.4.

Proposition 1. *If* $\int \max\{J(x,y),0\}\,\mu(x|0)\,dx > \kappa\ \forall y$ *and* $G(y \in \mathcal{Y} : J(\underline{x},y) < \kappa) > 0$, *then the job offer probability per unit of search effort exhibits negative duration dependence, i.e.* $do(x,\tau)/d\tau \leq 0\ \forall\tau$ *and* $\exists\hat{\tau} : do(x,\hat{\tau})/d\tau < 0$.

If, on top, the callback indicator $\mathcal{C}(y,\tau)$ *is monotonically decreasing in* y *and* τ, *then the job offer probability per interview exhibits positive duration dependence, i.e.* $do^c(x,\tau)/d\tau \geq 0\ \forall(x,\tau)$ *and* $\exists\hat{\tau} : do^c(x,\hat{\tau})/d\tau > 0$ *for some* x.

Proposition 2 reports a sufficient condition on the path of the individual job offer probability per unit of search effort for application effort to exhibit negative duration dependence under statistical discrimination.

Proposition 2. *Let* $\{o(x,\tau)\}_{\tau=0}^{\tilde{\tau}}$ *be the sequence of job offer probability per unit of search effort for job seekers of ability* x *for any unemployment duration* $\tau \leq \tilde{\tau}$*, where* $do(x,\tau)/d\tau \leq 0$*. For every worker ability* x*,* $\exists!\ \underline{D}(x) \in \left[0, \frac{1}{\tilde{\tau}-1}\right]$ *such that, if* $\frac{\Delta o(x,\tau)}{o(x,\tilde{\tau})-o(x,0)} > \underline{D}(x)\ \forall \tau \in \{1,\ldots,\tilde{\tau}-1\}$*, then application effort exhibits negative duration dependence, i.e.* $da(x,\gamma,\tau)/d\tau \leq 0\ \forall \tau$ *and* $\exists \hat{\tau} : da(x,\gamma,\hat{\tau})/d\tau < 0$*.*

Proposition 3 provides a sufficient condition for application effort to be decreasing in search type, which entails that application effort displays positive dynamic selection.

Proposition 3. *Let application costs be iso-elastic in search efficiency with elasticity* ζ*. Suppose that the average search type declines with duration, i.e.* $d\mathbb{E}_{\gamma|\tau}[\gamma]/d\tau < 0$*, and that the return on a unit of search effort,* $o(x,\tau)\left[W(x,\gamma) - U(x,\gamma,\tau+1)\right]$*, is jointly continuous in* x *and* γ*. Then, there exists a finite constant* $\zeta^\star \in (1,+\infty)$ *such that if* $\zeta > \zeta^\star$*, then application effort exhibits positive dynamic selection, i.e.* $\mathbb{E}_0[a(x,\gamma,\tau)] < \mathbb{E}_\tau[a(x,\gamma,\tau)]\ \forall \tau > 0$*, where* $\mathbb{E}_t[.]$ *denotes the expectation with respect to the joint distribution of ability* x *and search type* γ *at duration* t*.*

C.3 Microfoundation for application cost function

In this section, we propose two potential microfoundations for the application cost function increasing in search type adopted in the main text, *i.e.* $\sigma(a;\gamma)$, $\sigma_\gamma > 0$.

Permanent income increasing in search type (wealth effect). We set out by proposing an interpretation of our baseline model as a reduced-form for an extended model where workers of the same search type belong to risk-averse households. Accordingly, the (positive) dependence of application costs on search type is the result of the positive correlation between permanent income and search type, *i.e.* a wealth effect.[39]

We consider the same economy studied in Section 5 up to two tweaks. First, we assume that workers of the same search type belong to risk-averse households, which have

[39]Even if wages are rigid, our baseline model itself features a positive correlation between permanent income and search type because workers of higher search type spend less time, on average, in the unemployment state.

access to the asset market. Second, application costs are independent of search type. Hence, households make their consumption-savings decisions by solving the following utility maximization problem:

$$
\begin{aligned}
V(A, n, u_\tau; \gamma) = \max_{c, A'} \ & \frac{c^{1-\vartheta}}{1-\vartheta} - \tilde{\sigma}(a) + \beta \mathbb{E}\left[V\left(A', n', u'_{\tau'}; \gamma\right)\right] \\
\text{s.t. } & c + A' = RA + \omega n + bu, \\
& A \geq \underline{A},
\end{aligned}
$$

where ϑ is the coefficient of relative risk aversion, and n and u_τ are dummies for the respective labor market state. The solution to the utility maximization problem is represented by the Euler equation, $c^{-\vartheta} = \beta R \mathbb{E}[(c')^{-\vartheta}]$, and a transversality condition. In a stationary equilibrium, $\beta = \frac{1}{R}$ and $c' = c$. As long as initial asset holdings are zero, households find it optimal to consume their current income in each period, *i.e.* $c(\gamma) = \omega n(\gamma) + bu(\gamma)$. Since ability and search type are positively correlated, the stationary share of employed workers $n(\gamma)$ is increasing in γ, and so is lifetime consumption.

The values of employment and unemployment obtain by taking the envelope condition with respect to the respective labor market states, *i.e.* $\tilde{W} \equiv V_n$, $\tilde{U}(\tau) \equiv V_{u_\tau}$:

$$
\begin{aligned}
\tilde{U}(x, \gamma, \tau) &= \max_{\hat{a} \geq 0} \ \frac{b}{c(\gamma)^\vartheta} - \tilde{\sigma}(\hat{a}) + \beta \Big[\tilde{U}(x, \gamma, \tau+1) + s(\hat{a}, \gamma) o(x, \tau) \left(\tilde{W}(x, \gamma) - \tilde{U}(x, \gamma, \tau+1)\right)\Big], \\
\tilde{W}(x, \gamma) &= \frac{\omega}{c(\gamma)^\vartheta} + \beta \Big[\tilde{W}(x, \gamma) + \delta_L \big(\tilde{U}(x, \gamma, 0) - \tilde{W}(x, \gamma)\big)\Big].
\end{aligned}
$$

Hence, the value of unemployment and employment are the same as in the baseline model, up to the fact that the flow utility is weighted by the marginal utility of consumption. Let the ratio between capital gains upon employment in the two models at each duration be $\alpha(x, \gamma, \tau) \equiv \frac{\tilde{W}(x,\gamma) - \tilde{U}(x,\gamma,\tau+1)}{W(x,\gamma) - U(x,\gamma,\tau+1)}$. Let both application costs and search effort be iso-elastic in application effort with elasticities $1 + \eta$ and χ, respectively. If the marginal contribution of application effort to search effort is low enough, the values of unemployment and employment are linear in the marginal utility of consumption. Hence, $\lim_{\chi \to 0} \alpha(x, \gamma, \tau) = c(\gamma)^{-\vartheta} \equiv \tilde{\alpha}(\gamma)$. Since workers of higher search type enjoy higher consumption, $\tilde{\alpha}(\gamma)$ is decreasing in γ. This means that the capital gain upon employment, $\tilde{W}(x, \gamma) - \tilde{U}(x, \gamma, \tau+1)$, is lower than in the baseline model – the more so, the higher the job seeker's lifetime consumption and, therefore, the higher the search type. In general, the ratio between capital gains upon employment in the two models can be expressed as the product between the marginal utility of consumption $\tilde{a}(\gamma)$ and a correction term, *i.e.* $\alpha(x,\gamma,\tau) = \tilde{a}(\gamma)v(x,\gamma,\tau)$.

Optimal application effort solves:

$$a(x,\gamma,\tau) : \frac{\partial\tilde{\sigma}(a)}{\partial a}\frac{1}{\tilde{\alpha}(\gamma)} = \beta\,\frac{\partial s(a,\gamma)}{\partial a}o(x,\tau)\;v(x,\gamma,\tau)\Big[W(x,\gamma) - U(x,\gamma,\tau+1)\Big].$$

Comparing this optimality condition to equation (3), we conclude that the application cost function increasing in search type of the baseline model can be interpreted as capturing a wealth effect in reduced-form, which stems from a positive correlation between permanent income and search type when households are risk averse. Of course, the two models are exactly isomorphic only if the variance of the correction term $v(x,\gamma,\tau)$ approaches zero. In that case, $\sigma(a;\gamma) \equiv \frac{\tilde{\sigma}(a)}{\tilde{\alpha}(\gamma)}$.

Value of leisure increasing in search efficiency (time allocation effect) We now review a common microfoundation for the search effort cost function adopted in standard models of endogenous search effort (Pissarides, 2000) and extend it to our specific framework. Consider the problem of a job seeker who gains utility from consumption and social leisure in an additively separable fashion. The job seeker is endowed with one unit of time each period, which can be spent either exerting search effort s or in social activities. Formally,

$$\begin{aligned}&\max_{s,\ell}\; u(b) + \nu(\ell) + \beta so(W-U),\\ &\text{s.t. } h(s) + \ell = 1,\end{aligned}$$

where $h(s)$ denotes the hours it takes to exert s units of search effort (normalized by the unitary amount of total hours) and o represents the job offer probability per unit of search effort. Optimal time allocation trades off higher current utility from social leisure against higher expected discounted utility from finding a job:

$$\nu'(\ell)h'(s) = \beta o(W-U).$$

By assuming linear utility from social leisure, *i.e.* $\nu(\ell) = \gamma\ell$, and convex and isoelastic search effort hours function, *i.e.* $h(s) = \psi\frac{s^{1+\eta}}{1+\eta}$, this time allocation model is isomorphic to standard models of endogenous search effort subject to convex costs, with cost function $\sigma(s) = \nu'(\ell)h(s)$.

Our main innovation with respect to standard models of endogenous search effort is modeling search effort as the product between individual search efficiency ϵ and endogenous application effort a. In what follows, we extend the previous microfoundation to a model where job seekers are heterogeneous in their character (*search type*), which determines both how much they value social leisure and their search efficiency.

Assume that workers differ in their character, which ranges from introverted to outgoing. Outgoing workers draw higher utility from spending time in social relations and therefore have a larger personal network which allows them to overcome meeting frictions more easily when looking for a job. Formally, we identify a worker's character as the marginal utility she gains from social leisure, γ. It follows that workers of character γ have search efficiency $\epsilon = \epsilon(\gamma)$, where $\epsilon' > 0$.

We are interested in the time allocation decisions made by job seekers of different character γ. Optimal application effort solves the following utility-maximization problem:

$$\begin{aligned} &\max_{a,\ell}\ u(b) + \nu(\ell) + \beta s(a;\gamma) o(W - U), \\ &\text{s.t. } h(a) + \ell = 1. \end{aligned}$$

Notice that, differently from the previous case, $h(a)$ denotes the hours it takes to exert a units of application effort – not of total search effort. Taking the first-order condition with respect to a yields:

$$\nu'(\ell)h'(a) = \beta\frac{\partial s(a;\gamma)}{\partial a}o(W-U) \iff \gamma h'(a) = \beta\frac{\partial s(a;\gamma)}{\partial a}o(W-U).$$

Following the same argument as before, we assume convex and iso-elastic application effort hours function, *i.e.* $h(a) = \psi_0\frac{a^{1+\eta}}{1+\eta}$. Under this functional form assumptions, this time allocation model is isomorphic to the model of endogenous application effort adopted in the main text with cost function $\sigma(a;\gamma) = \gamma h(a)$.

C.4 Proofs

Proof Proposition 1. Consider a job seeker of ability x, whose job offer probability per unit of search effort is given by equation (5). As shown in Jarosch and Pilossoph (2019), the callback indicator $\mathcal{C}(y,\tau)$ is monotonically decreasing in τ. Hence, $\exists \hat{x}$: for $x \geq \hat{x}, \exists$ at least one unemployment duration $\hat{\tau}$ s.t.

$$\begin{cases} \mathcal{C}(y,\hat{\tau}-1)\mathcal{Q}(x,y) = 1 \\ \mathcal{C}(y,\hat{\tau})\mathcal{Q}(x,y) = 0 \end{cases} \iff o(x,\hat{\tau}) < o(x,\hat{\tau}-1)$$

For $x < \hat{x}$, $o(x,\tau) = o(x,0)\ \forall \tau$.
Hence, we conclude that the job offer probability per unit of search effort is nonincreasing in unemployment duration.

Since the callback indicator is monotonically decreasing in y, one can define as $y^*(\tau)$ the productivity of the firm that is just indifferent between calling back a job seeker with duration τ or not. Formally,

$$y^*(\tau) : \int J\left(x, y^*(\tau)\right)\ q\left(x, y^*(\tau)\right)\ \mu(x|\tau)dx = \kappa$$

Therefore, the interview probability reads:

$$c(\tau) \equiv \lambda(\theta)\mathbb{P}\left(y \leq y^*(\tau)\right) = \lambda(\theta)G(y^*(\tau))$$

In turn, the job offer probability per interview is given by:

$$o^c(x,\tau) \equiv \frac{\mathbb{P}\left(y \leq \min\{x, y^*(\tau)\}\right)}{\mathbb{P}\left(y \leq y^*(\tau)\right)} = \frac{G(\min\{x, y^*(\tau)\})}{G(y^*(\tau))}$$

The duration profile of the job offer probability per interview obeys:

$$\frac{do^c(x,\tau)}{d\tau} = \frac{1}{G\left(y^*(\tau)\right)}\left[\frac{dG\left(\min\{x, y^*(\tau)\}\right)}{d\tau} - \frac{G\left(\min\{x, y^*(\tau-1)\}\right)}{G\left(y^*(\tau-1)\right)}\frac{dG\left(y^*(\tau)\right)}{d\tau}\right] \tag{C.1}$$

where $dG\left(\min\{x, y^*(\tau)\}\right)/d\tau \equiv G\left(\min\{x, y^*(\tau)\}\right) - G\left(\min\{x, y^*(\tau-1)\}\right)$. In order to pin down the sign of equation (C.1), we distinguish two cases.

$$\text{CASE 1}: \; x \geq y^*(0) \iff \min\{x, y^*(\tau)\} = y^*(\tau)$$

$$\implies \frac{do^c(x,\tau)}{d\tau} = 0$$

$$\text{CASE 2}: \; x < y^*(0)$$

$$\text{Monotonicity of } \mathcal{C}(y,\tau) \text{ in } \tau \text{ entails that } x \begin{cases} < y^*(\tau) & \text{if } \tau < T \\ \geq y^*(\tau) & \text{if } \tau \geq T \end{cases} \quad \text{for some } T < \infty$$

$$\text{For } \tau < T, \; \min\{x, y^*(\tau)\} = x$$

$$\implies \frac{do^c(x,\tau)}{d\tau} \propto -\frac{dG\left(y^*(\tau)\right)}{d\tau} \geq 0$$

$$\text{For } \tau = T, \; \min\{x, y^*(\tau)\} = y^*(\tau), \; \min\{x, y^*(\tau-1)\} = x$$

$$\implies \frac{do^c(x,\tau)}{d\tau} \propto -\left[G(y^*(\tau-1)) - G(x)\right] > 0$$

$$\text{For } \tau \geq \tilde{\tau}, \; \min\{x, y^*(\tau)\} = y^*(\tau)$$

$$\implies \frac{do^c(x,\tau)}{d\tau} = 0$$

Hence, we conclude that the job offer probability per interview is nondecreasing in unemployment duration. □

Proof Proposition 2. For analytical transparency, we prove the proposition for $\tilde{\tau} = 2$.[40] Let $X \equiv [x, \gamma]$ denote a worker type in terms of ability and search type. Optimal

[40] The same proof strategy holds for longer time horizons, but the $\tilde{\tau} = 2$ case is the only one that can be worked out analytically.

application effort, $a(X,\tau)$, solves:

$$\frac{\partial\sigma(a(X,\tau);X)}{\partial a} = \beta\ \frac{\partial s(a(X,\tau),X)}{\partial a} o(X,\tau)\ \Big[W(X) - U(X,\tau+1)\Big].$$

By exploiting the fact that the job offer probability per unit of search effort is constant from unemployment duration $\tilde{\tau}$ onward, the relevant workers' value functions read:

$$W(X) = \omega + \beta[W(X) - \delta_L(W(X) - U(X,0))], \tag{C.2}$$

$$U(X,0) = b - \sigma(a(X,0);X) + \beta[U(X,1) + s(X,0)o(X,0)(W(X) - U(X,1))], \tag{C.3}$$

$$U(X,1) = b - \sigma(a(X,1);X) + \beta[U(X,2) + s(X,1)o(X,1)(W(X) - U(X,2))], \tag{C.4}$$

$$U(X,2) = b - \sigma(a(X,2);X) + \beta[U(X,2) + s(X,2)o(X,2)(W(X) - U(X,2))]. \tag{C.5}$$

Let $\Delta_\tau(X) \equiv W(X) - U(X,\tau)$ be the capital gain upon employment at duration $\tau - 1$. Application effort exhibits negative duration dependence if and only if $a(X,0) \geq a(X,1) \geq a(X,2) \iff o(X,0)\Delta_1(X) \geq o(X,1)\Delta_2(X) \geq o(X,2)\Delta_2(X)$. Since $o(X,2) \leq o(X,1)$ by assumption, then $a(X,1) \geq a(X,2)$ always holds. Hence, to establish negative duration dependence in application effort, it suffices to prove that $o(X,0)\Delta_1(X) \geq o(X,1)\Delta_2(X)$.

Let $\Delta^u_\tau(X) \equiv \omega - (b - \sigma(a(X,\tau)))$ be the excess flow value of employment over unemployment at duration τ. By subtracting equation (C.2) from equation (C.3)-equation (C.5), we are left with a 3-equation system in $(\Delta_0(X), \Delta_1(X), \Delta_2(X))$:

$$\begin{cases} \Delta_0(X) = & \frac{\Delta^u_0(X) + \beta(1-f(X,0))\Delta_1(X)}{1+\beta\delta_L}, \\ \Delta_1(X) = & \Delta^u_1(X) - \beta\delta_L\Delta_0(X) + \beta(1 - f(X,1))\Delta_2(X), \\ \Delta_2(X) = & \frac{\Delta^u_2(X) - \beta\delta_L\Delta_0(X)}{1-\beta(1-f(X,2))}. \end{cases}$$

Solving the system yields:

$$\Delta_2(X) = \frac{1}{1-\beta(1-f(X,2))}\left[\Delta^u_2(X) - \frac{\beta\delta_L}{1+\beta\delta_L}\Delta^u_0 - \frac{\beta\delta_L}{1+\beta\delta_L}\beta(1 - f(X,0))\Delta_1(X)\right], \tag{C.6}$$

$$\Delta_1(X) = \frac{1}{1+\frac{\beta\delta_L}{1+\beta\delta_L}\beta(1-f(X,0))\frac{1-\beta(f(X,1)-f(X,2))}{1-\beta(1-f(X,2))}}\left[\Delta^u_1(X) - \frac{\beta\delta_L}{1+\beta\delta_L}\frac{1-\beta(f(X,1)-f(X,2))}{1-\beta(1-f(X,2))}\Delta^u_0(X) + \frac{\beta(1-f(X,1))}{1-\beta(1-f(X,2))}\Delta^u_2(X)\right]. \tag{C.7}$$

Without loss of generality, let $o(X,1) = \alpha(X,1)\bar{o}(X) + (1-\alpha(X,1))\underline{o}(X)$, where $\bar{o}(X) = o(X,0)$ and $\underline{o}(X) = o(X,2)$.

Suppose by contradiction that $a(X,0) < a(X,1)$. From equation (C.6), it follows that:

$$\Delta_1(X) < \tfrac{\tilde{\alpha}(X,1)(1+\beta\delta_L)}{(1+\beta\delta_L)(1-\beta(1-f(X,2)))+\tilde{\alpha}(X)\beta\delta_L\beta(1-f(X,0))}\left(\Delta_2^u(X) - \tfrac{\beta\delta_L}{1+\beta\delta_L}\Delta_0^u(X)\right), \tag{C.8}$$

where $\tilde{\alpha}(X,1) \equiv \alpha(X,1) + (1-\alpha(X,1))\frac{\underline{o}(X)}{\bar{o}(X)}$.

Let $\tilde{\Delta}_{ij}^u(X) \equiv \Delta_i^u(X) - \frac{\beta\delta_L}{1+\beta\delta_L}\Delta_j^u(X)$. Rearranging equation (C.7) yields:

$$\Delta_1(X) = \frac{1}{1+\frac{\beta\delta_L}{1+\beta\delta_L}\beta(1-f(X,0))\frac{1-\beta(f(X,1)-f(X,2))}{1-\beta(1-f(X,2))}}\left[\tilde{\Delta}_{10}^u(X) + \frac{\beta(1-f(X,1))}{1-\beta(1-f(X,1))}\tilde{\Delta}_{20}^u(X)\right].$$

Plugging this expression into equation (C.8) yields:

$$\tilde{\Delta}_{10}^u < \mathcal{M}(\alpha(X,1))\tilde{\Delta}_{20}^u, \tag{C.9}$$

$$\begin{aligned}\mathcal{M}(\alpha(X,1)) \equiv \tfrac{1}{1-\beta(1-\underline{f}(X))}\Bigg[&\tilde{\alpha}(\alpha(X,1))\left(1+\beta\delta_L\beta(1-\bar{f}(X))\tfrac{1-\beta(\varphi(\alpha(X,1))\tilde{\alpha}(\alpha(X,1))\bar{f}(X)-\underline{f}(X))-\tilde{\alpha}(\alpha(X,1))}{(1+\beta\delta_L)(1-\beta(1-\underline{f}(X)))+\tilde{\alpha}(\alpha(X,1))\beta\delta_L\beta(1-\bar{f}(X))}\right)\\ &-\beta(1-\varphi(\alpha(X,1))\tilde{\alpha}(\alpha(X,1))\bar{f}(X))\Bigg],\end{aligned} \tag{C.10}$$

where $f(X,1) = \varphi(\alpha(X,1))\tilde{\alpha}(\alpha(X,1))\bar{f}(X)$ and $\varphi(\alpha(X,1)) \equiv \frac{s(\tilde{\alpha}(X,1))}{s(X,0)}$ denotes relative search effort with respect to duration 0. Moreover, $\bar{f}(X) \equiv f(X,0)$, $\underline{f}(X) \equiv f(X,2)$.

Since $\tilde{\Delta}_{10}^u > \tilde{\Delta}_{30}^u$, if $\mathcal{M}(\alpha(X)) < 1$, then equation (C.8) never holds. We now study how the multiplier (C.10) varies with $\alpha(X)$. We start by analyzing its limiting behavior as $\alpha(X)$ approaches 1 and 0.

Result 1. $\mathcal{M}(1) > 1$, $\mathcal{M}(0) < 1$.

Proof. Evaluating equation (C.10) at $\tilde{\alpha}(1) = 1$ yields:

$$\mathcal{M}(1) = \tfrac{1}{1-\beta(1-\underline{f}(X))}\left[1 - \tfrac{\beta\delta_L\beta(1-\bar{f}(X))\beta(\varphi(1)\bar{f}(X)-\underline{f}(X))}{(1+\beta\delta_L)(1-\beta(1-\underline{f}(X)))+\beta\delta_L\beta(1-\bar{f}(X))} - \beta(1-\varphi(1)\bar{f}(X))\right].$$

It follows that:

$$\begin{aligned}\mathcal{M}(1) > 1 \iff &(1+\beta\delta_L)(1-\beta(1-\underline{f}(X)))(1-\underline{f}(X)) + (1-\varphi(1)\bar{f}(X))[2\delta_l\beta(1-\bar{f}(X))\beta+\\ &(1+\beta\delta_L)(1-\beta(1-\underline{f}(X)))] > 0.\end{aligned}$$

Since the latter expression is always positive, $\mathcal{M}(1) > 1$.

Evaluating equation (C.10) at $\tilde{\alpha}(0) = \frac{\underline{o}(X)}{\bar{o}(X)}, \varphi(0) = \frac{s(X,2)}{s(X,2)}$ yields:

$$\mathcal{M}(0) = \frac{1}{1-\beta(1-\underline{f}(X))}\left[\tilde{\alpha}(0)\left(1+\beta\delta_L\beta(1-\bar{f}(X))\frac{1-\tilde{\alpha}(0)}{(1+\beta\delta_L)(1-\beta(1-\underline{f}(X)))+\tilde{\alpha}(0)\beta\delta_L\beta(1-\bar{f}(X))}\right)-\beta(1-\underline{f}(X))\right].$$

From it, we can check that:

$$\mathcal{M}(0) < 1 \iff \tilde{\alpha}(0) < 1.$$

Since, by construction, $\tilde{\alpha}(0) < 1$, it follows that $\mathcal{M}(0) < 1$. □

Hence, if the drop in the job offer probability per unit of search effort happens fully between duration 1 and 2 ($\alpha(X,1) = 1$), then application effort may be non-monotonic, *i.e.* job seekers may exert more application effort at duration $\tau = 1$ than at duration $\tau = 0$.[41] On the contrary, if the drop in the job offer probability per unit of search effort happens fully between duration 0 and 1 ($\alpha(X) = 0$), then application effort is always monotonic, *i.e.* application effort displays negative duration dependence.

Finally, we study how the multiplier C.10 behaves within the two bounds.

Result 2. $\frac{\partial\mathcal{M}(\tilde{\alpha}(X))}{\partial\tilde{\alpha}(X)} > 0$.

Proof. Taking the derivative of equation (C.10) with respect to $\tilde{\alpha}(X)$ yields:

$$\frac{\partial\mathcal{M}(\tilde{\alpha}(X,1))}{\partial\tilde{\alpha}(X,1)} \propto \left(1-\frac{\tilde{\alpha}(X)\beta\delta_L\beta(1-\bar{f}(X))}{(1+\beta\delta_L)(1-\beta(1-\underline{f}))+\tilde{\alpha}(X,1)\beta\delta_L\beta(1-\bar{f}(X))}\right)\times$$
$$\left(1+\frac{\beta\delta_L\beta(1-\bar{f}(X))(1-\tilde{\alpha}(X,1)-\beta(\varphi(\tilde{\alpha}(X,1))\tilde{\alpha}))}{(1+\beta\delta_L)(1-\beta(1-\underline{f}(X)))+\tilde{\alpha}(X,1)\beta\delta_L\beta(1-\bar{f}(X))}+\beta(\varphi(\tilde{\alpha}(X,1))+\varphi'(\tilde{\alpha}(X,1)))\bar{f}(X)\right).$$

The first term in brackets is positive if and only if:

$$(1+\beta\delta_L)(1-\beta(1-\underline{f}(X))) > 0,$$

[41] Non-monotonicity happens for sure if the application cost function is such that $dU(\tau)/d\tau \leq 0$ if $do(\tau)/d\tau \leq 0$.

which is always the case. The second term in brackets is positive if and only if:

$$(1+\beta\delta_L)(1-\beta(1-\underline{f}(X)))+\beta(\varphi(\tilde{\alpha}(X))+\varphi'(\tilde{\alpha}(X))\tilde{\alpha}(X))\bar{f}(X)[(1+\beta\delta_L)(1-\beta(1-\underline{f}(X)))+ \tilde{\alpha}(X)\beta\delta_L\beta(1-\bar{f}(X))]+\beta\delta_L\beta(1-\bar{f}(X))[1-\beta(\varphi(\tilde{\alpha}(X))\tilde{\alpha}(X)\bar{f}(X)-\underline{f}(X))]>0,$$

which similarly always holds. □

Since $\tilde{\alpha}(X)$ is a continuous function of $\alpha(X)$, there exists a threshold $\bar{\alpha}^T(X) \in (0,1]$ such that, if $\alpha(X,1) < \bar{\alpha}^T(X)$, then $a(X,0) > a(X,1)$ and $a(X,\tau)$ exhibits negative duration dependence for type X. Equivalently, we can think of the threshold as defining a lower bound $\underline{D}(X)$ to $\frac{\Delta\hat{o}(X,\tau)}{\hat{o}(X,\bar{\tau})-\hat{o}(X,0)}$. Formally, $\underline{D}(X) \equiv 1-\bar{\alpha}^T(X)$.

□

Proof Proposition 3. From equation (3), optimal application effort at duration τ is decreasing in search type if the elasticity of the marginal cost with respect to search efficiency $\epsilon(\gamma)$ exceeds that of the marginal benefit. Hence, if $\zeta > 1+\frac{\partial \ln\left(\mathbb{E}_{x|\gamma,\tau}[o(x,\tau)[W(x,\gamma)-U(x,\gamma,\tau+1)]]\right)}{\partial \ln(\epsilon(\gamma))}$ $\forall \tau,\gamma$, then job seekers with higher search type exert less application effort at any duration. Since γ lies in the compact set $[0,1]$, and both $\epsilon(\gamma)$ and $\mathbb{E}_{x|\gamma,\tau}[o(x,\tau)(W(x,\gamma)-U(x,\gamma,\tau+1))]$ are continuous functions in γ, the composed elasticity function is continuous over a compact domain.[42] Hence, by the Extreme Value Theorem, this elasticity is bounded from above by a finite number $\zeta^\star \equiv 1+\max\left\{\frac{\partial \ln\left(\mathbb{E}_{x|\gamma,\tau}[o(x,\tau)[W(x,\gamma)-U(x,\gamma,\tau+1)]]\right)}{\partial \ln(\epsilon(\gamma))}\right\}$. Since average search type decrease with duration, average application effort at duration τ, $\mathbb{E}_\tau[a(x,\gamma,\tau)]$, exceeds the counterfactual average application effort if the joint distribution of ability and search type were constant over the unemployment spell, $\mathbb{E}_0[a(x,\gamma,\tau)]$. □

C.5 Quantitative model

In this section we describe the quantitative model used for structural estimation. The quantitative model extends the baseline model outlined in Section 5 by allowing for multiple job seekers per vacancy. Job seekers and vacancies are assumed to get together through an urn-ball meeting process generating coordination frictions.[43]

[42] To derive this result, we implicitly assume that the joint distribution of ability and search type is well-behaved. Namely, we need that the conditional density $\ell(x|\gamma,\tau)$ is continuous in γ $\forall x$.

[43] The urn-ball meeting process gives rise to a distribution of the number of job seekers that each firm meets in each period, being the average number of such meetings still determined by the meeting function.

The hiring process has the following timing: (1) upon meeting at least one job seeker, the firm decides whether to call back a job seeker at cost κ; (2) conditional on calling back a job seeker, the firm gets to know her ability x and, based on that, decides whether to interview another job seeker; (3) if any of the interviewees meet the firm ability requirement, *i.e.* $x \geq y$, the firm offers a job to the highest-ability one with probability q.

In the presence of coordination frictions, firms need to sort potentially multiple job seekers. Since average job seeker's ability is decreasing with duration, when faced with multiple job seekers, firms find it optimal to rank them according to their unemployment duration starting with the shortest. Upon calling back the shortest-duration job seeker (as long as it is profitable, *i.e.* $\mathcal{C}(y,\tau) = 1$), the firm calls back the next job seeker, as well, if:[44]

$$\int \max\left\{J(x,y) - J(\hat{x},y), 0\right\} q(x,y)\ \mu(x|\tau)\ dx \geq \kappa, \tag{C.11}$$

where $\hat{x}$ represents the ability of the previous job seeker, which is revealed at the interview stage. Denoting as $z^c(x,y,\tau)$ the search-effort-weighted measure of job seekers crowding out a job seeker with ability x and unemployment duration τ in contact with a firm of productivity y at the callback stage (derived in Appendix C.6), the interview probability per unit of search effort writes:

$$c(x,\tau) = \lambda(\theta) \int \mathcal{C}(y,\tau) \exp\left\{-\frac{z^c(x,y,\tau)}{V}\right\}\ dG(y),$$

where $\exp\left\{-\frac{z^c(x,y,\tau)}{V}\right\}$ equals the probability that firm y is not in contact with any job seeker with shorter duration than τ that does not warrant an interview to a $(x,\tau)-$job seeker in the sense of equation (C.11).

Denoting as $z(x,y,\tau)$ the search-effort-weighted measure of job seekers crowding out a job seeker with ability x and unemployment duration τ in contact with a firm of productivity y in hiring (derived in Appendix C.6), the job offer probability per interview

[44] Following Jarosch and Pilossoph (2019), we assume that by interviewing another candidate the firm does not lose the option of hiring any of the previous interviewees.

writes:

$$o^c(x,\tau) = \frac{\int q(x,y)\ \mathcal{C}(y,\tau)\ \exp\left\{-\frac{z(x,y,\tau)}{V}\right\}\ dG(y)}{\int \mathcal{C}(y,\tau)\ \exp\left\{-\frac{z^c(x,y,\tau)}{V}\right\}\ dG(y)}.$$

In words, with probability q, a firm makes a job offer to the highest-ability job seeker that grants it positive flow profits, conditional on discovering her ability type during the interview. Hence, the job offer probability per unit of search effort is defined as:

$$o(x,\tau) \equiv c(x,\tau)\cdot o^c(x,\tau) = \lambda(\theta)\int q(x,y)\ \mathcal{C}(y,\tau)\exp\left\{-\frac{z(x,y,\tau)}{V}\right\}\ dG(y).$$

For given labor market tightness, expected profits upon drawing productivity y read:

$$\begin{aligned}\mathbb{E}[\Pi(y)|\theta] = \sum_{m=1}^{\infty}\mathbb{P}(N=m|\theta)\sum_{\tau_1=0}^{\infty}\mathbb{P}\left(\underline{\mathbf{t}}_{1,N}=\tau_1\right)\sum_{\tau_2=\tau_1}^{\infty}\mathbb{P}\left(\underline{\mathbf{t}}_{2,N}=\tau_2|\underline{\mathbf{t}}_{1,N}=\tau_1\right)\dots\\ \sum_{\tau_N=\tau_{N-1}}^{\infty}\mathbb{P}\left(\tau_N=\tau_N|\underline{\mathbf{t}}_{N-1,N}=\tau_{N-1}\right)\int\cdots\int\Big[\sum_{k=1}^{N}\big(J(x,y)\ q(x,y)\ \mu(\bar{\mathbf{x}}_{1,k}=x|\mathbf{t}_{1,k})\\ -k\kappa\big)\mathbb{1}\{t_{k+1}>\bar{\tau}(\bar{\mathbf{x}}_{1,k},y)\ \&\ t_s\le\bar{\tau}(\bar{\mathbf{x}}_{1,s-1},y)\ \forall s\le k\}\Big]\ dx_1\dots dx_N\ \mathcal{C}(y,\tau_1),\end{aligned}$$

where $\mathbb{P}(N=m|\theta) = \left(\frac{\lambda(\theta)S}{V}\right)^m\frac{1}{m!}\exp\left\{-\frac{\lambda(\theta)S}{V}\right\}$, $\underline{\mathbf{t}}_{n_1,n_2}\equiv\min\{t_{n_1},\dots,t_{n_2}\}$ and $\bar{\mathbf{x}}_{1,k}\equiv\max\{x_1,\dots,x_k\}$. In the presence of coordination frictions, the number of job seekers N met by a firm in each period is not restricted to $\{0,1\}$ (as in the baseline model) but follows a Poisson distribution. As discussed above, the firm finds it optimal to rank such N job seekers by unemployment duration, with τ_1 denoting the shortest and τ_N the longest. If the duration of the first job seeker warrants a job interview, *i.e.* $\mathcal{C}(y,\tau_1)=1$, the firm calls back as many job seekers n as warranted by equation (C.11) at cost κ each. Upon selecting the highest-ability job seeker among them (as long as she meets the ability requirement, *i.e.* $\bar{\mathbf{x}}_{1,n}\ge y$), the firm offers her a job with probability q – the job seekers' selection process being therefore independent of the actual qualification realization.

Finally, the free entry condition pins down the labor market tightness θ such that vacancy posting costs equalize discounted *ex ante* expected profits as per equation (4):

$$\kappa_v = \beta\int\mathbb{E}[\Pi(y)|\theta]\ dG(y).$$

C.6 Model derivations

According to the urn-ball meeting process between job seekers and vacancies, each period $\lambda(\theta)S$ job seekers (balls) sort into V vacancies (urns). Following Jarosch and Pilossoph (2019), we scale the measure of aggregate search effort $S \equiv \sum_{\tau=0}^{\infty} \int \int s(x,\gamma,\tau)u(x,\gamma,\tau)$ $d\mathcal{L}(x,\gamma)$ by the extent of meeting frictions $\lambda(\theta)$ faced by job seekers to obtain effective applications, *i.e.* the measure of job seekers' search effort that does not get lost because of meeting frictions (or the output of the meeting function). Since we consider a continuum of job seekers and vacancies, the binomial distribution of effective applications at a given vacancy converges to a Poisson distribution (Blanchard and Diamond, 1994). As a result, each vacancy receives zero effective applications with probability $\exp\{-\frac{\lambda(\theta)S}{V}\}$. Throughout, we assume that firms, whenever faced with equivalent job seekers at each stage of the hiring process, randomize among them.

The search-effort-weighted measure of job seekers crowding out a job seeker with ability x and unemployment duration τ in contact with a firm of productivity y at the interview stage reads:

$$z^c(x,y,\tau) \equiv \lambda(\theta) \sum_{t=0}^{\tau} \left(1 - \frac{1}{2}\mathbb{1}\{t=\tau\}\right) \int \int \mathbb{1}\{\bar{\tau}(x',y) < \tau\} \left(1 - \frac{1}{2}\mathbb{1}\{x'=x\}\right)$$
$$s(x',\gamma,t)\ u(x',\gamma,t)\ d\mathcal{L}_t(x',\gamma),$$

where $\bar{\tau}(x',y)$ denotes the highest duration τ such that equation (C.11) holds. Intuitively, a job seeker with ability x and unemployment duration τ is not interviewed by a firm she is in contact with if there is at least another job candidate with shorter unemployment duration whose interview is successful and has ability high enough to make interviewing a $(x,\tau)-$job seeker unprofitable.

The search-effort-weighted measure of job seekers crowding out a job seeker with ability

x and unemployment duration τ in contact with a firm of productivity y in hiring reads:

$$z(x,y,\tau) \equiv \lambda(\theta)\Bigg(\sum_{t=0}^{\tau}\left(1-\frac{1}{2}\mathbb{1}\{t=\tau\}\right)\int\int \mathbb{1}\left\{(\bar{\tau}(x',y)<\tau)\cup(\bar{\tau}(x',y)\geq\tau, x'\geq x)\right\}$$
$$\left(1-\frac{1}{2}\mathbb{1}\{\bar{\tau}(x',y)\geq\tau, x'\geq x\}\right) s(x',\gamma,t)\ u(x',\gamma,t)\ d\mathcal{L}_t(x',\gamma) + \int\int\sum_{t=\tau}^{\bar{\tau}(x',y)}$$
$$\left(1-\frac{1}{2}\mathbb{1}\{t=\tau\}\right)\mathbb{1}\{x'\geq x\}\left(1-\frac{1}{2}\mathbb{1}\{x'=x\}\right) s(x',\gamma,t)\ u(x',\gamma,t)\ d\mathcal{L}_t(x',\gamma)\Bigg).$$

Intuitively, a job seeker with ability x and unemployment duration τ is not hired by a firm she is in contact with for two main reasons. First, she will not be hired if there is at least another job candidate with shorter unemployment duration whose interview is successful and either has ability high enough to make interviewing a $(x,\tau)-$job seeker unprofitable or is of higher ability than x (first summation). Second, she will not be hired if there is at least another job candidate with unemployment duration between hers and the longest unemployment duration such that another candidate is interviewed after her who has higher ability than hers (second summation).

D. Details of structural estimation

In this section we discuss our model estimation strategy and comment the estimation results.

D.1 Moments selection

Since workers in the model differ in unobservable characteristics only, we first notice that the correct counterparts of the unconditional duration profiles in the model are the duration profiles controlling for observables in the data. Moreover, the sequential search protocol of our model requires that we select individual-level targets – rather than application-level ones – from search diaries (see Table A1 for the respective descriptive statistics, and Figure B9 for the duration profiles). With these requirements in mind, we select empirical moments that are informative for the model parameters to be estimated.

Since our interest lies in the role of duration at each stage of the job search process, we target – in an indirect inference sense – the entire duration profiles of applications, individual interview probability, and job finding rate at each duration controlling for observables.[45] We also target the average of these variables across all durations as informative of the pace of dynamic selection. To correctly disentangle dynamic selection from duration dependence in applications, we further target the duration profile of applications controlling for individual fixed effects. For the estimated model to capture the relevant heterogeneity in search types, we target the standard deviation in the individual fixed effects in applications. To isolate the extent of decreasing returns in applications from heterogeneity in search efficiency, we target the partial effect of the interview probability per application with respect to individual fixed effects in application (normalized by the average interview probability per application in the first month of unemployment). We target the average interview cost per hire reported by Muehlemann and Strupler Leiser (2018) to inform the interview cost factor. For convenience, we estimate the interview cost *factor*, $\tilde{\kappa} = \kappa/q$, rather than the parameter κ. All targeted duration profiles are smoothed using a moving average filter to minimize noise.

D.2 Parameter identification and moment informativeness

We establish local identification of the model parameters under our relative minimum distance estimator by verifying that the Jacobian of the relative deviations between model-implied and empirical moments with respect to the parameter vector has full column rank at the estimated values. This condition ensures that small perturbations in the empirical moments generate unique local adjustments in the parameter estimates.

Figure D1 and Figure D2 illustrate the behavior of the loss function in the vicinity of the estimated parameters. Each plot shows a one-dimensional slice of the loss function while holding all other parameters fixed. Overall, the loss function exhibits clear curvature with respect to the parameters, indicating that the solution lies in an isolated local minimum.[46] Compared to the other parameters, the loss function is relatively flat with

[45]By targeting the duration profiles of both the individual interview probability and the job finding rate, we also implicitly target the duration profile of the individual job offer probability per interview. Given the importance of this variable, we introduce an additional penalization in the loss function, equal to the percentage deviation between the linear duration coefficient of the (smoothed) empirical and model-implied average job offer probability per interview. This penalization effectively amounts to targeting the *covariance* between the duration profiles of the job finding rate and individual interview probability as an additional moment.

[46]While we do not formally establish global identification—that is, uniqueness of the minimum over the full parameter space—the overall shape of the loss function, along with multiple robustness checks with respect to the initial guess of the parameter vector in our nonlinear programming solver, point towards a unique global minimum.

respect to the measure of vacancies. This is reasonable, as the measure of vacancies affects the interview and job offer probabilities only through the exponential crowding-out probabilities.

To gauge the informativeness of each empirical targets for the estimated parameters, we follow the same steps as in Andrews et al. (2017) to derive the sensitivity matrix from our relative simulated minimum distance estimator.

Let $\hat{\mu}(\Theta) \in \mathbb{R}^M$ be the simulated model moments, $\mu \in \mathbb{R}^M$ the targeted empirical moments, and $\Theta \in \mathbb{R}^P$ the model parameters. Define the vector of *relative moment deviations* as:

$$g(\Theta) := \frac{\hat{\mu}(\Theta) - \mu}{\mu} \in \mathbb{R}^M.$$

The minimum distance estimator solves:

$$\Theta^* = \arg\min_{\Theta} L(\Theta), \quad \text{where} \quad L(\Theta) := g(\Theta)' \mathcal{W} g(\Theta),$$

for some positive definite weighting matrix $\mathcal{W}$. In our case, $\mathcal{W}$ is the identity matrix.

The first-order condition (FOC) for the optimal parameter vector Θ^* is:

$$F(\Theta, \mu) := \nabla_\Theta L(\Theta) = 2\nabla_\Theta g(\Theta)' \mathcal{W} g(\Theta) = 0.$$

We define the Jacobian of the relative moment deviations as:

$$J(\Theta, \mu) := \nabla_\Theta g(\Theta) = \text{diag}\left(\frac{1}{\mu}\right) \nabla_\Theta \hat{\mu}(\Theta).$$

To compute the sensitivity of the estimator Θ^* to changes in the empirical moments μ, we totally differentiate the FOC:

$$\frac{dF}{d\mu} = \frac{\partial F}{\partial \Theta} \cdot \frac{d\Theta^*}{d\mu} + \frac{\partial F}{\partial \mu} = 0,$$

which implies

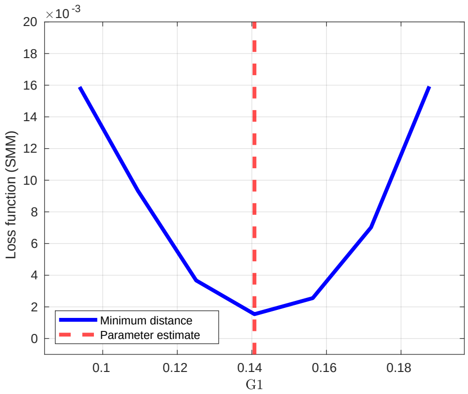
$$\frac{d\Theta^*}{d\mu} = -\left(\frac{\partial F}{\partial \Theta}\right)^{-1} \cdot \frac{\partial F}{\partial \mu}.$$

By applying a first-order approximation, the Hessian simplifies to:

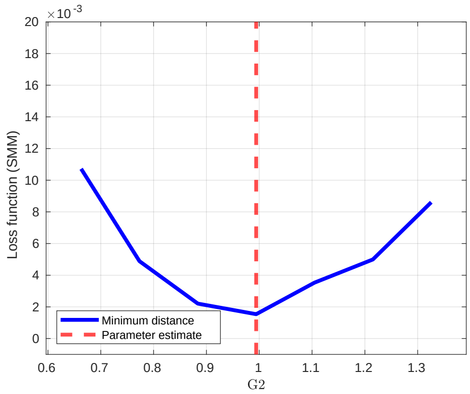
$$\frac{\partial F}{\partial \Theta} \approx 2J' \mathcal{W} J.$$

Figure D1: Minimum Distance

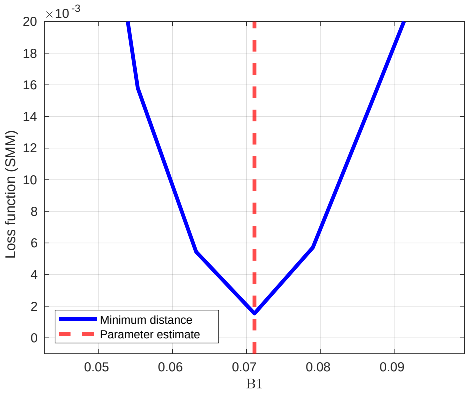
(A) $G1$



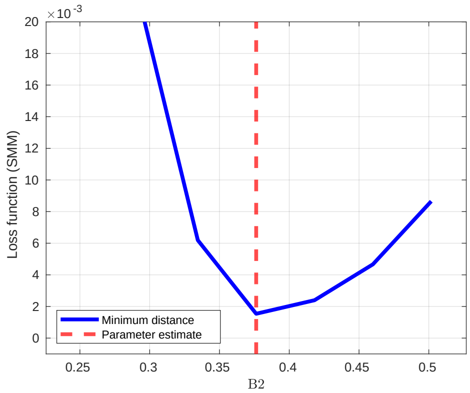
(B) $G2$



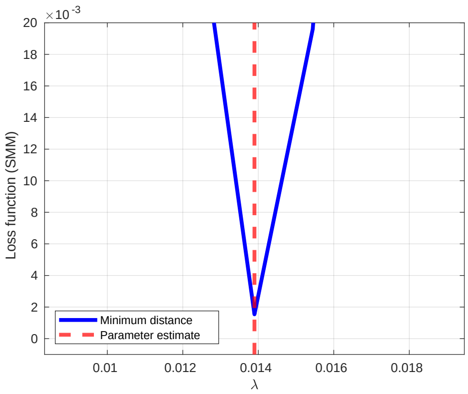
(C) $B1$



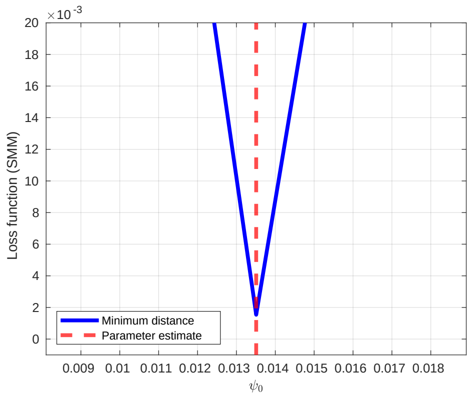
(D) $B2$



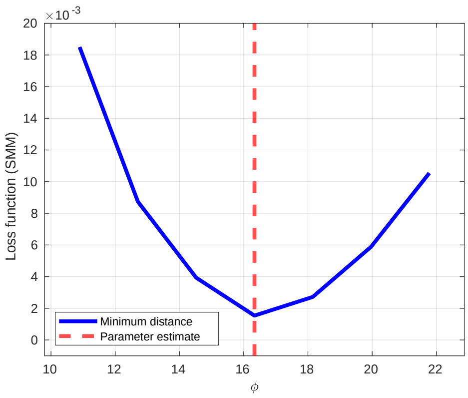
(E) λ



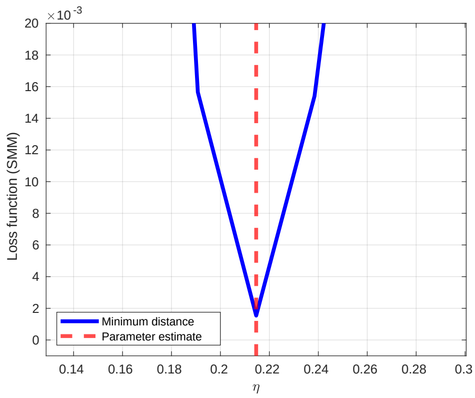
(F) ψ_0



(G) ϕ



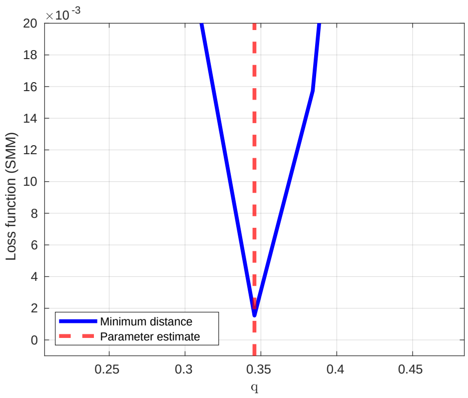
(H) η



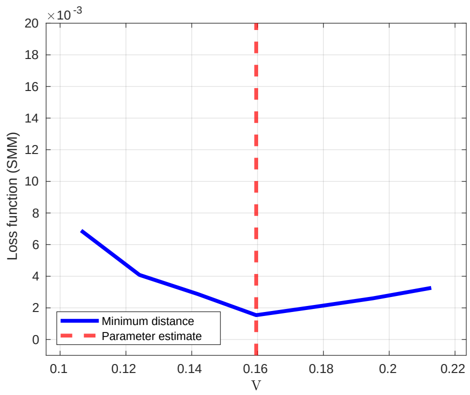
Note: The figures report the behavior of the SMM loss function $L(\Theta)$ to perturbations in each internally estimated parameter.

Figure D2: Minimum Distance, Cont'd

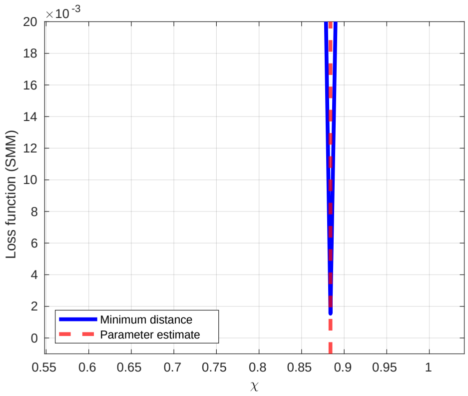
(A) q



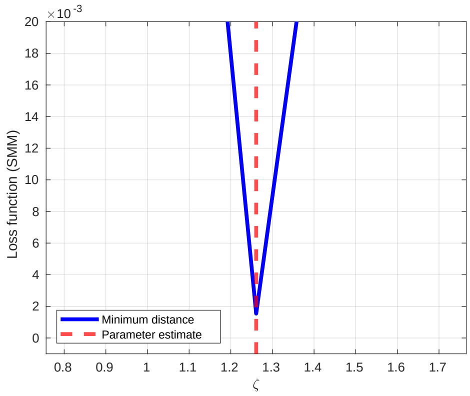
(B) V



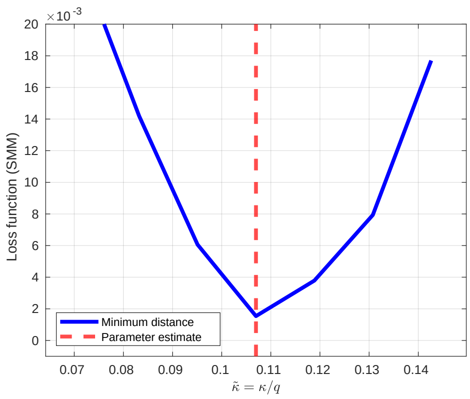
(C) χ



(D) ζ



(E) κ



Note: The figures report the behavior of the SMM loss function $L(\Theta)$ to perturbations in each internally estimated parameter.

Similarly, the derivative of F with respect to μ becomes:

$$\frac{\partial F}{\partial \mu} \approx 2J'\mathcal{W} \cdot \frac{\partial g}{\partial \mu},$$

where

$$\frac{\partial g}{\partial \mu} = -\operatorname{diag}\left(\frac{\hat{\mu}(\Theta)}{\mu^2}\right).$$

Putting all together, the sensitivity of Θ^* with respect to μ is:

$$M_1 \equiv \frac{d\Theta^*}{d\mu} \approx \left(J'\mathcal{W}J\right)^{-1} J'\mathcal{W} \cdot \operatorname{diag}\left(\frac{\hat{\mu}(\Theta)}{\mu^2}\right).$$

To facilitate interpretation and comparison across parameters, we report the elasticity of estimated parameters with respect to moments:

$$\Lambda = M_1 \frac{\mu}{\Theta^\star},$$

which expresses how a 1% change in a moment affects each parameter estimate in percentage terms. Elasticities with respect to moments pertaining to the short-term and long-term duration profile of application effort, individual interview probability, and job finding rate are averaged out, in order for the number of moments to equal the number of parameters. Table D1 reports the signed elasticities, in order to interpret the directionality of the effects. Table D2 corresponds to the *sensitivity matrix* suggested by Andrews

Table D1: Directional Sensitivity Matrix

	$\mathbf{E}_0[a(1:\tilde{\tau})]$	$\mathbf{E}_\tau[a(0:\tilde{\tau}-4)]$	$\mathbf{E}_\tau[c^{ind}(0:\tilde{\tau}-4)]$	$\mathbf{E}_\tau[f(0:\tilde{\tau}-4)]$	$\mathbf{E}[c^{ind}(0:\tilde{\tau})]$	$\mathbf{E}_\tau[c^{ind}(\tilde{\tau}-4:\tilde{\tau})]$	$\mathbf{E}[f(0:\tilde{\tau})]$	$\mathbf{E}_\tau[f(\tilde{\tau}-4:\tilde{\tau})]$	$\mathbf{E}[a(0:\tilde{\tau})]$	$\mathbf{E}_\tau[a(\tilde{\tau}-4:\tilde{\tau})]$	σ_{α_i}	$\beta_{c^{app},\alpha_i\mid\tau}$	(K/H)
G1	-0.90	0.19	-0.47	0.30	0.65	1.25	0.18	-1.44	0.33	1.52	2.87	-1.33	0.00
G2	-0.62	0.17	-1.21	0.34	0.96	3.08	0.51	-1.42	0.31	1.00	1.84	-0.80	0.00
B1	0.00	0.15	-0.00	-0.01	0.07	-0.01	0.00	0.07	0.13	-0.05	-0.14	0.68	-0.00
B2	0.25	0.01	-0.33	-0.06	0.09	0.81	0.10	0.09	0.01	-0.38	-0.73	0.55	-0.00
λ	-0.12	-0.02	0.40	-0.04	0.21	-0.80	0.16	0.07	0.02	0.13	0.19	-0.28	0.00
ψ_0	-1.11	1.61	0.79	-0.20	0.38	-1.97	0.90	1.05	1.35	0.17	-4.31	0.58	-0.00
ϕ	0.36	-0.27	-0.77	0.05	0.36	1.92	0.40	-0.14	0.23	-0.48	0.22	0.22	-0.00
η	1.22	-1.86	-0.54	0.25	0.24	1.34	0.93	-1.09	1.58	-0.14	4.21	-0.74	0.00
q	-0.14	0.01	-0.08	0.10	0.12	0.00	0.06	-0.16	0.03	0.25	0.54	-0.27	0.00
V	-0.62	0.91	0.49	-0.24	0.26	-1.04	0.65	0.81	0.78	0.30	-2.71	0.50	-0.00
χ	0.01	-0.01	0.00	0.00	0.00	-0.01	0.00	-0.00	0.01	-0.02	-0.02	0.01	-0.00
ζ	0.09	-0.25	-0.03	0.02	0.05	0.07	0.07	-0.16	0.21	0.03	1.24	-0.30	0.00
$\tilde{\kappa} = \kappa/q$	0.16	-0.05	0.01	-0.03	0.07	-0.04	0.01	0.17	0.07	-0.25	-0.41	0.24	0.99

Note: The entries of the table are the signed elasticities of estimated parameters to data targets.

et al. (2017). The sensitivity matrix provides a local approximation to the mapping from moments to estimated parameters. Therefore, the entry values can be used as a measure of moment informativeness for parameter estimates. Based on this interpretation, we address the question of linking moments and estimated parameters. Formally, we solve the

Table D2: Absolute Sensitivity Matrix

	$\mathbf{E}_0[a(1:\tilde{\tau})]$	$\mathbf{E}_\tau[a(0:\tilde{\tau}-4)]$	$\mathbf{E}_\tau[c^{ind}(0:\tilde{\tau}-4)]$	$\mathbf{E}_\tau[f(0:\tilde{\tau}-4)]$	$\mathbf{E}[c^{ind}(0:\tilde{\tau})]$	$\mathbf{E}_\tau[c^{ind}(\tilde{\tau}-4:\tilde{\tau})]$	$\mathbf{E}[f(0:\tilde{\tau})]$	$\mathbf{E}_\tau[f(\tilde{\tau}-4:\tilde{\tau})]$	$\mathbf{E}[a(0:\tilde{\tau})]$	$\mathbf{E}_\tau[a(\tilde{\tau}-4:\tilde{\tau})]$	σ_{α_i}	$\beta_{c^{app},\alpha_i\|\tau}$	(K/H)
G1	0.90	0.55	1.62	0.62	0.65	1.27	0.18	1.44	0.33	1.52	2.87	1.33	0.00
G2	0.63	0.17	2.24	1.63	0.96	3.09	0.51	1.42	0.31	1.00	1.84	0.80	0.00
B1	0.13	0.15	0.22	0.10	0.07	0.05	0.00	0.08	0.13	0.05	0.14	0.68	0.00
B2	0.25	0.30	0.45	0.52	0.09	0.85	0.10	0.22	0.01	0.38	0.73	0.55	0.00
λ	0.12	0.19	0.47	0.53	0.21	0.86	0.16	0.20	0.02	0.13	0.19	0.28	0.00
ψ_0	1.70	1.61	1.46	3.39	0.38	2.68	0.90	1.24	1.35	0.51	4.31	0.58	0.00
ϕ	0.36	0.47	0.94	1.38	0.36	2.03	0.40	0.46	0.23	0.48	0.22	0.22	0.00
η	2.01	1.86	1.62	3.44	0.24	2.23	0.93	1.22	1.58	0.52	4.21	0.74	0.00
q	0.14	0.11	0.21	0.11	0.12	0.08	0.06	0.16	0.03	0.25	0.54	0.27	0.00
V	0.86	0.91	0.97	2.20	0.26	1.52	0.65	0.84	0.78	0.33	2.71	0.50	0.00
χ	0.01	0.01	0.02	0.01	0.00	0.01	0.00	0.01	0.01	0.02	0.02	0.01	0.00
ζ	0.24	0.25	0.15	0.30	0.05	0.25	0.07	0.16	0.21	0.06	1.24	0.30	0.00
$\tilde{\kappa}=\kappa/q$	0.16	0.12	0.22	0.09	0.07	0.12	0.01	0.17	0.07	0.25	0.41	0.24	0.99

Note: The entries of the table are the elasticities of estimated parameters to data targets.

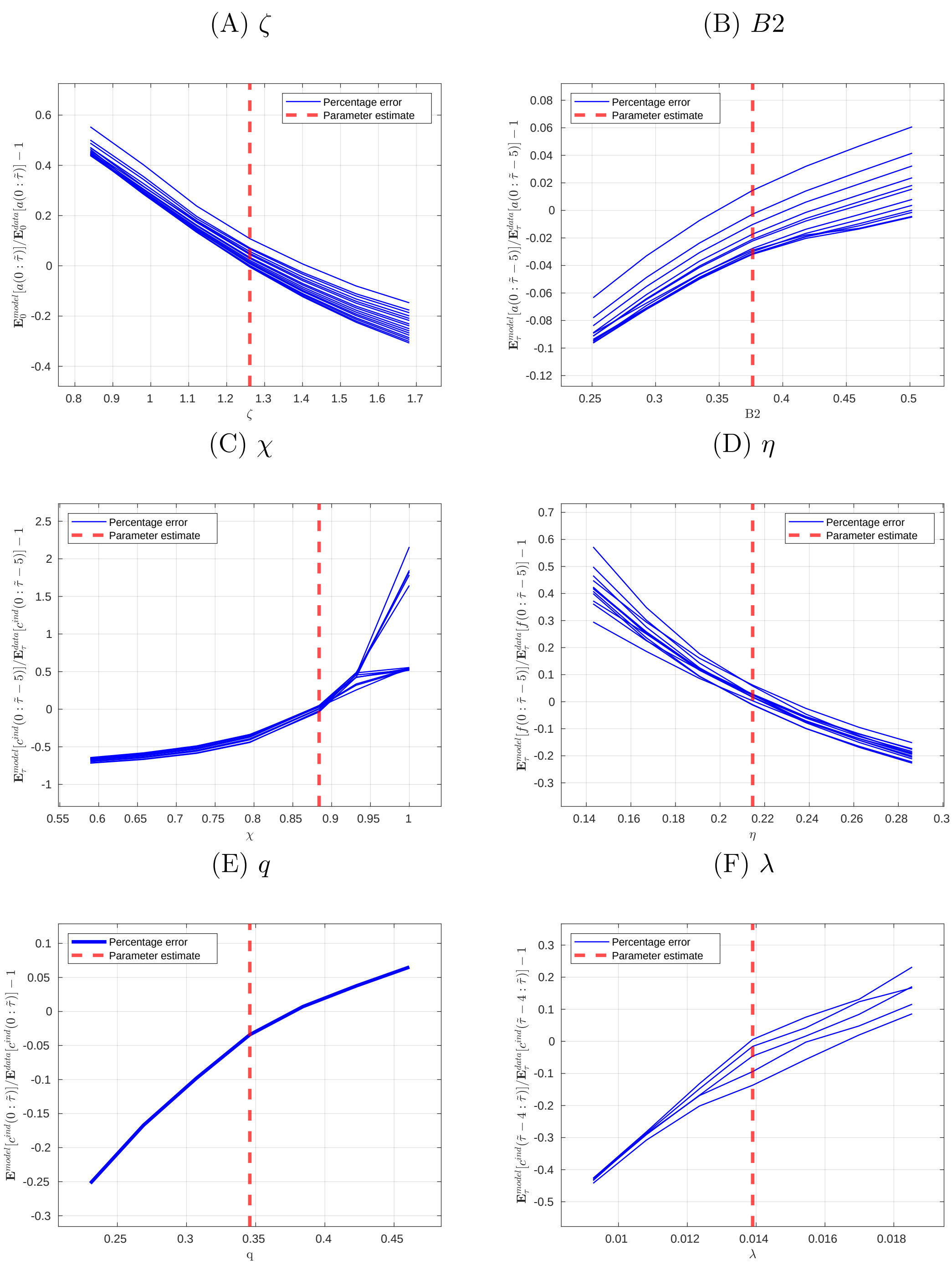
following linear sum assignment problem to optimally match moments $m = \{m_1, \dots, m_K\}$ to parameters $\Theta = \{\Theta_1, \dots, \Theta_J\}$:

$$
\begin{aligned}
\min_{X\in\{0,1\}^{K\times J}} \quad & -\sum_{k=1}^{K}\sum_{j=1}^{J}\left|\Lambda^{-1}_{kj}\right| \mathbb{1}_{\{\partial g_k(\Theta)/\partial\theta_j \text{monotonic}\}} X_{kj} \\
\text{s.t.} \quad & \sum_{j=1}^{J} X_{kj} = 1 \quad \text{for all } k = 1,\dots,K \\
& \sum_{k=1}^{K} X_{kj} = 1 \quad \text{for all } j = 1,\dots,J
\end{aligned}
$$

where $\left|\Lambda^{-1}\right| \in \mathbb{R}^{K\times J}$ is a cost matrix, whose entries $\left|\Lambda^{-1}_{kj}\right|$ measure the inverse informativeness of assigning moment m_k to parameter θ_j. For the assignment to be valid, we further require that the relative error of each moment is monotonic against perturbations in the candidate linked parameter, holding the others fixed. The binary assignment matrix X indicates the optimal one-to-one matching, where $X_{kj} = 1$ if m_k is matched to θ_j, and 0 otherwise.

Figure D3 and Figure D4 show that the assignment algorithm is successful at linking moments and informative parameters. As one may expect, the interview cost is assigned to the average interview cost per hire, the degree of decreasing returns in application effort to the short-term duration profile of the individual interview probability, and the scalar of the search effort cost function to the standard deviation in individual fixed effects in applications. Based on this analysis, Table 4 in the main text reports each parameter linked to one moment.

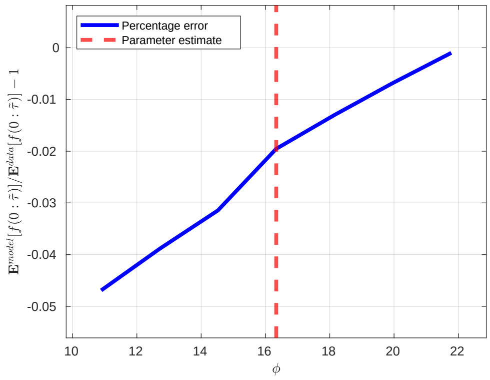
Figure D3: Target-Parameter Map



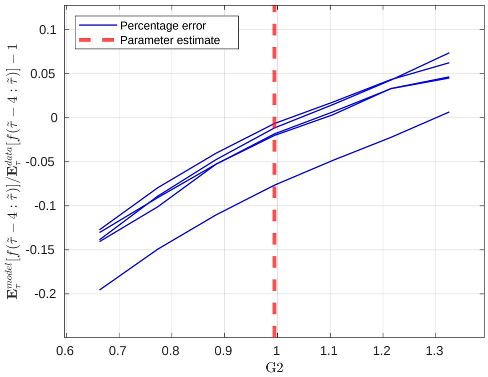
Note: The figures report the behavior of the relative error of each (set of) moment(s) to perturbations in the linked parameter.

Figure D4: Target-Parameter Map, Cont'd

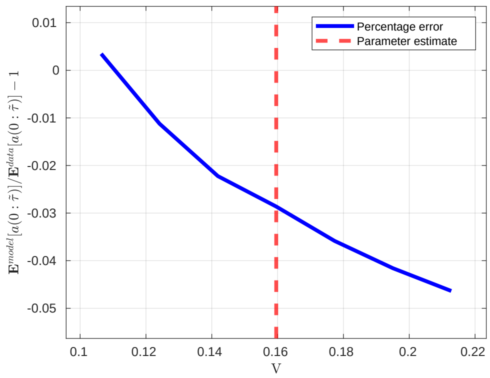
(A) ϕ



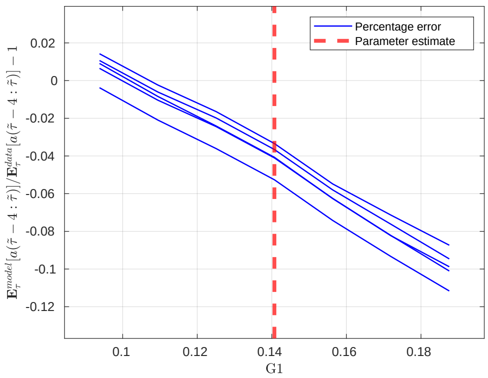
(B) $G2$



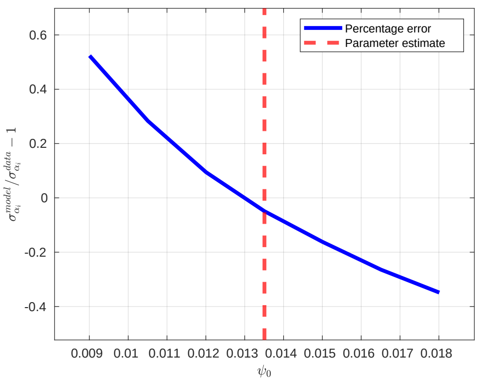
(C) V



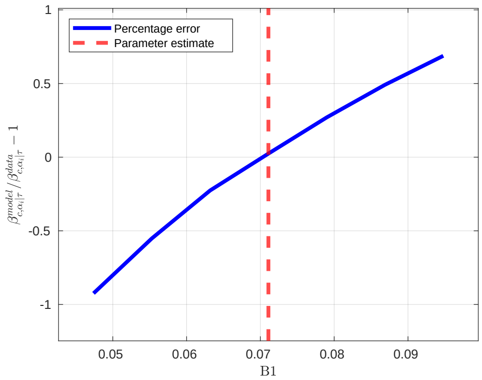
(D) $G1$



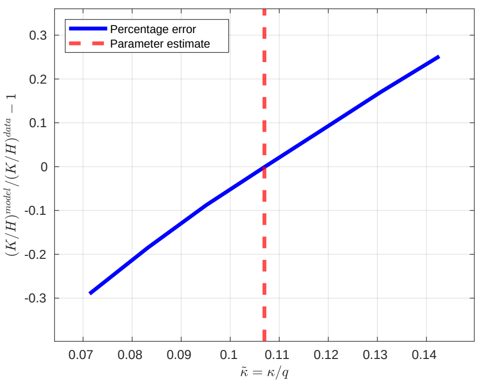
(E) ψ_0



(F) $B1$



(G) κ



Note: The figures report the behavior of the relative error of each (set of) moment(s) to perturbations in the linked parameter.

Computing standard errors. The standard errors of our relative minimum distance estimator follow from the covariance matrix of estimated parameters, computed using the standard *sandwich formula*:

$$\text{se}(\Theta^\star) = \text{diag}\left(\Sigma(\Theta^\star)^{1/2}\right), \quad \text{with} \quad \Sigma(\Theta^\star) = M_1 \cdot V(\Theta^\star, \mu) \cdot M_1',$$

where $V(\Theta^\star, \mu) := \text{Cov}[g(\Theta^\star, \mu)]$ is the covariance matrix of the relative deviations between model-implied and empirical moments.

To compute $V(\Theta^\star, \mu)$, we treat empirical moments μ as fixed and derive the covariance matrix across model-implied moments $\hat{\mu}(\Theta^\star)$. Because $g(\Theta^\star) = \text{diag}\,(1/\mu) \cdot [\hat{\mu}(\Theta^\star) - \mu]$, the implied covariance of relative deviations is:

$$V(\Theta^\star, \mu) = \text{diag}\left(\frac{1}{\mu}\right) \cdot \text{Cov}[\hat{\mu}(\Theta^\star)] \cdot \text{diag}\left(\frac{1}{\mu}\right).$$

We adopt this approach because our model explicitly characterizes the joint distribution of moments.

In practice, we estimate $\text{Cov}[\hat{\mu}(\Theta^\star)]$ by stacking individual-level observations on application effort, interview probability, and job finding rate, and estimating a joint OLS regression with individual and duration fixed effects. The resulting residual covariance matrix provides the empirical basis for computing moment variability.

We then analytically characterize the mappings from estimated fixed effects to the moments targeted in estimation—specifically, the duration profiles of application effort, interview probability, and job finding rate, as well as the cross-sectional standard deviation of individual fixed effects in applications. Similarly, we compute a first-order approximation to map individual fixed effects in applications to the semi-elasticity of interview probability with respect to these fixed effects. For the average interview cost per hire – which is uncorrelated with the other moments – we apply the Delta method to compute its variance.

Standard errors for each estimated parameter are reported in parentheses in Table 4 Panels B-C in the main text.

Finally, knowledge of both the parameter covariance matrix $\Sigma(\Theta^\star)$ and the moment covariance matrix $V(\Theta^\star, \mu)$ allows us to assess the *sensitivity of estimation precision* to variation in individual moments. Following Honoré et al. (2020), we compute the measure:

$$M_3^{(k)} := M_1 \cdot O_{kk} \cdot M_1',$$

where O_{kk} is a selection matrix with a 1 in the (k,k)-entry and zeros elsewhere. Each entry of $M_3^{(k)}$ reflects how sensitive the precision of a given parameter estimate is to additional noise in moment k. For interpretation, we report the elasticity of estimated parameter precision with respect to noise in data moments (*scaled sensitivities*):

$$\tilde{\Lambda}^{(j,k)} = M_3^{(j,k)} \frac{V(\Theta^{\star}, \mu)^{(k,k)}}{\Sigma(\Theta^{\star})^{(j,j)}},$$

where $M_3^{(j,k)}$ equals the jth diagonal element of $M_3^{(k)}$. The scaled sensitivities reported in Table D3 can be interpreted as the percent changes in the variance of each estimated parameter that would result from a 1% increase in the variance (noise) of a given moment. (For readability, all entries are multiplied by 100.)

Table D3: Precision Sensitivity Matrix (Honoré et al., 2020)

	$\mathbf{E}_0[a(1:\tilde{\tau})]$	$\mathbf{E}_\tau[a(0:\tilde{\tau}-4)]$	$\mathbf{E}_\tau[c^{ind}(0:\tilde{\tau}-4)]$	$\mathbf{E}_\tau[f(0:\tilde{\tau}-4)]$	$\mathbf{E}[c^{ind}(0:\tilde{\tau})]$	$\mathbf{E}_\tau[c^{ind}(\tilde{\tau}-4:\tilde{\tau})]$	$\mathbf{E}[f(0:\tilde{\tau})]$	$\mathbf{E}_\tau[f(\tilde{\tau}-4:\tilde{\tau})]$	$\mathbf{E}[a(0:\tilde{\tau})]$	$\mathbf{E}_\tau[a(\tilde{\tau}-4:\tilde{\tau})]$	σ_{α_i}	$\beta_{c^{app},\alpha_i\mid\tau}$	(K/H)
G1	1.74	0.31	1.17	0.68	0.01	2.84	0.00	8.99	0.00	2.86	15.48	10.58	0.00
G2	0.99	0.02	1.55	2.77	0.02	13.74	0.01	8.30	0.00	1.09	5.09	3.10	0.00
B1	0.81	0.27	0.41	0.19	0.00	0.10	0.00	0.45	0.01	0.07	0.59	46.62	0.00
B2	0.28	0.25	0.38	1.27	0.00	5.22	0.00	1.06	0.00	0.57	3.35	6.01	0.00
λ	0.10	0.16	0.63	1.91	0.01	8.15	0.01	1.50	0.00	0.13	0.39	2.56	0.00
ψ_0	2.41	0.26	0.17	2.37	0.00	2.14	0.01	1.04	0.01	0.06	5.61	0.32	0.00
ϕ	0.41	0.22	0.51	2.43	0.00	8.22	0.01	1.28	0.00	0.29	0.09	0.29	0.00
η	2.80	0.29	0.15	2.12	0.00	1.23	0.01	0.89	0.01	0.06	4.69	0.46	0.00
q	0.86	0.33	0.52	0.40	0.01	0.34	0.01	3.17	0.00	2.20	16.16	12.98	0.00
V	2.02	0.27	0.22	2.86	0.00	2.01	0.01	1.56	0.01	0.10	6.67	0.71	0.00
χ	2.27	0.37	0.84	0.66	0.00	3.75	0.00	0.96	0.01	1.73	5.66	3.93	0.00
ζ	1.92	0.31	0.09	1.01	0.00	0.76	0.00	0.90	0.01	0.05	23.82	4.48	0.00
$\tilde{\kappa} = \kappa/q$	1.43	0.33	0.55	0.27	0.00	0.73	0.00	3.22	0.00	1.84	7.50	8.02	26.76

Note: The entries of the table are the elasticities of estimated parameters precision to noise in data targets multiplied by 100, i.e., $100 \cdot \tilde{\Lambda}^{(j,k)}$.

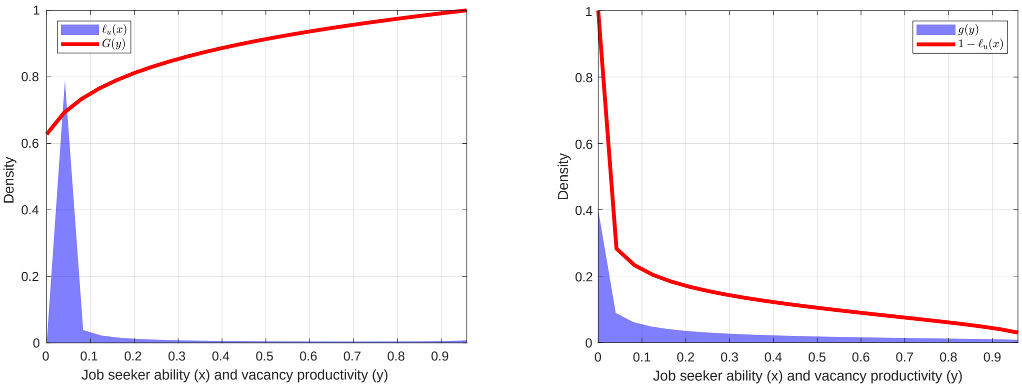
D.3 Estimation results

Our estimation results provide some insights into the structure of the Swiss labor market. First of all, we notice that the firm productivity distribution $G(y)$ displays a spike at $y = 0$, where 64% of the mass is concentrated.[47] This is in line with Jarosch and Pilossoph (2019), which finds the same spike with density ranging from 40% to 64% across different model specifications. Instead of the uniform pattern imposed by Jarosch and Pilossoph

[47] For comparability with Jarosch and Pilossoph (2019), we shift each discretized y value leftward by one discretization step in order to allow for a positive mass at $y = 0$.

Figure D5: Matching frictions, estimated model

(A) Matching probability, job seekers (B) Matching probability, vacancies



Note: This figure reports the after-meeting matching probability faced by job seekers across the ability distribution (Panel A) and by firms across the vacancy productivity distribution (Panel B). The red solid line represent the probability that a worker meets a firm she is qualified for conditional on meeting one (Panel A) and the probability that a firm meets a qualified worker conditional on meeting one (Panel B); the blue areas display the density of the job seeker ability distribution $\ell_u(x)$ (Panel A) and the density of the vacancy productivity distribution $g(y)$ (Panel B).

(2019) for the rest of the distribution, we estimate a monotonically decreasing density.[48] Similarly, the equilibrium job seeker ability distribution displays a spike at the lowest positive grid point accounting for 69.5% of the total mass. The rest of the distribution is instead relatively close to uniform with a modest final increase in correspondence to the highest ability level. The relative shape of the ability and productivity distribution is informative of the extent of matching frictions faced by searching agents. Figure D5A plots the matching probability faced by job seekers across the ability distribution, *i.e.* the probability of meeting a firm they may be qualified for (conditional on meeting one). As a result of the functional form of the qualification function, this matching probability is increasing in ability. Figure D5B reports the same graph under the firms' perspective. Unlike for workers, firms' matching probability is decreasing in productivity, with the highest-productivity firms being the most selective.

The substitution parameter of the meeting function, ξ, is estimated to be 0.25, which entails a moderate amount of complementarity between aggregate search effort and vacancies. As a result, our estimated meeting function looks closer to the standard Cobb-Douglas specification ($\xi = 0$) than to that estimated by Ramey, den Haan, and Watson (2000) ($\xi = 1.27$). According to our results, the application effort cost function displays a scalar, ψ_0, equal to 1.35% of average monthly output and a mild convexity ($\eta = 0.21$),

[48] Allowing for a flexible productivity distribution is critical for our results because the thickness of the right tail of the distribution is directly related to the extent of duration dependence in the interview probability, being high-productivity firms the most prone to statistical discrimination.

which implies an elasticity of application effort to the job offer probability per unit of search effort larger than 4. It follows that our estimated implied elasticity is markedly higher than the unitary elasticity implied by the quadratic search cost function commonly used in the literature (Yashiv, 2000; Christensen, Lentz, Mortensen, Neumann, and Werwatz, 2005), but remarkably close to that estimated by Lise (2013).[49] The search efficiency dispersion parameter ϕ is estimated to be 16.3, meaning that the most efficient workers at search are 10 times more likely to obtain an interview than the lowest efficient ones, for given application effort. Such significant cross-sectional heterogeneity in search efficiency is the reason why our estimated model is able to replicate the simultaneous patterns of positive dynamic selection and negative duration dependence in application effort detected in the data, since job seekers with higher search efficiency (and ability) find it optimal to exert less application effort in equilibrium. Importantly, we tie our hands tightly in terms of admissible dispersion in search efficiency by targeting the empirical standard deviation of individual fixed effects in applications for the sake of indirect inference. The estimated vacancy posting cost, κ_v, equals 2.4% of average monthly output, lending support to the hypothesis that most of hiring costs arises from interview costs (3.7% of average monthly output) rather than entry costs. The qualification probability q equals 35%, supporting an important role of idiosyncratic matching frictions in the hiring process. The elasticity of search effort with respect to application effort, χ, is estimated to be 0.88, meaning that decreasing returns in applications are relatively mild. The elasticity of application costs with respect to search efficiency, ζ, equals 1.26. This means that the unit application cost of job seekers with the highest search efficiency is about 19 times larger than that of job seekers with the lowest search efficiency, reflecting *e.g.* the cost of locating suitable job openings rather than sending unsolicited applications.

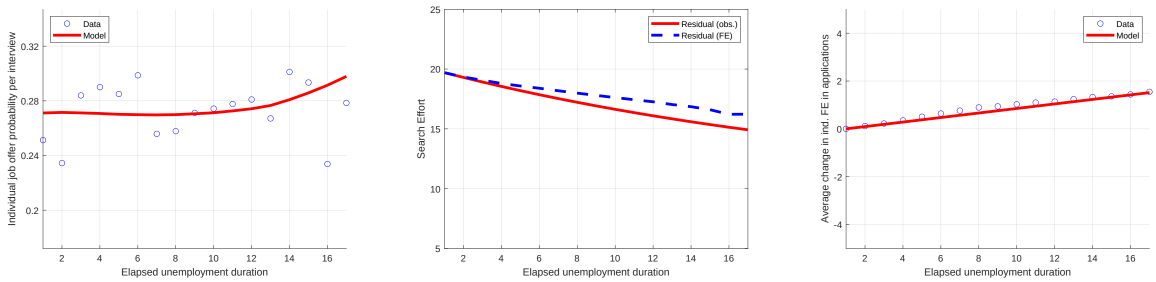
[49] This is the mirror image of our empirical finding of a significantly higher duration dependence in application effort than commonly thought.

Figure D6: Duration profile of selected variables

(A) Ind. job offer prob. per interview, model vs data

(B) Search effort, model

(C) Ind. fixed effects in applications, model vs data



Note: Panel A contrasts the duration profile controlling for observables of the individual job offer probability per interview detected in the data (circles) with that implied by the estimated model. The model-implied duration profile is estimated through a moving average filter. Panel B reports the fitted duration profiles controlling for observables (in solid red) and fixed effects (in dashed blue) of search effort in the estimated baseline model. The model-based duration profile is fitted using a negative exponential function estimated via weighted least squares. Panel C contrasts the duration profile controlling for observables of the individual fixed effect in applications detected in the data (circles) with that implied by the estimated model. The model-based duration profile is fitted using a linear function estimated via weighted least squares.

E. Alternative models

In this section we develop and estimate alternative models of job search under statistical discrimination. The goal of this section is to highlight and quantify the role of endogenous search effort as a source of duration dependence in the job finding rate.

E.1 Exogenous search effort

We estimate the model of statistical discrimination with exogenous search effort of Jarosch and Pilossoph (2019). This model is nested by our baseline model as job seekers share the same search type, $\gamma = \bar{\gamma}$, application effort is constant and costless, $a(x, \gamma, \tau) = \bar{a}$ and $\psi(\gamma) = 0$, and the qualification probability function is deterministic, $q = 1$. Moreover, the wage rate equals the flow value of leisure, *i.e.* $\omega = b$. We estimate this restricted model by using the same estimation strategy as in Jarosch and Pilossoph (2019). Specifically, we estimate negative exponential relationships between the relative (*i.e.* normalized with respect to the first month of unemployment) duration profiles of the individual interview probability and job finding rate (controlling for observables) and unemployment duration (in months), and use them as targeted moments for the sake of indirect inference. The vector of targeted moments further features the average interview cost per hire and the average job finding rate (in levels). Using the same SMM approach as in the main text, we estimate the parameters governing the worker ability distribution and firm productivity distribution, the interview cost, and the worker meeting probability.

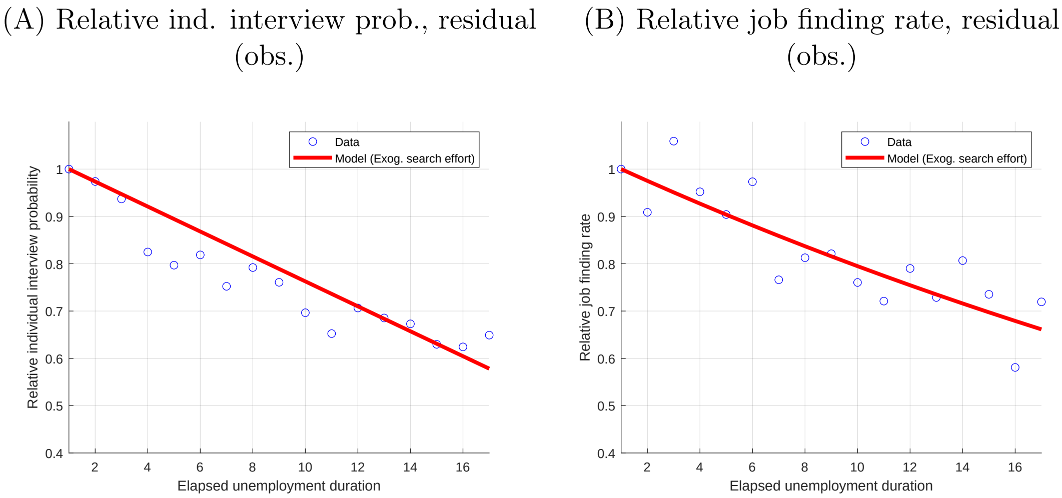
Estimation results are reported in Table E1.

Table E1: Parameter Estimates (Exogenous Search Effort Model)

Parameter	Description	Estimate	Target/Source	Data	Model
Panel A. Externally set parameters					
β	Discount factor	0.996	5% annual interest rate in Davis and von Wachter (2011)		
δ_L	Separation prob. (workers)	0.009	Monthly EU prob. (Swiss social security)		
δ_H	Separation prob. (firms)	0.019	Monthly EE+EU prob. (Swiss social security)		
b	Value of leisure	0.985	Avg job value in Shimer (2005b), Hagedorn and Manovskii (2008), and Gertler and Trigari (2009)		
Panel B. Internally estimated structural parameters					
$B1$	1^{st} shape param. Beta distr. search eff.	0.041	$\mathbb{E}_\tau[f(x,12\!:\!16)/f(x,0)]$: long-term duration profile rel. job finding prob. (residual)	See Figure E1B	
$B2$	2^{nd} shape param. Beta distr. search eff.	0.113	$\mathbb{E}_\tau[f(x,0\!:\!11)/f(x,0)]$: short-term duration profile rel. job finding prob. (residual)	See Figure E1B	
$G1$	1^{st} shape param. Beta distr. prod.	0.038	$\mathbb{E}_\tau[c^{ind}(x,0\!:\!11)/c^{ind}(x,0)]$: short-term duration profile rel. ind. interview prob. (residual)	See Figure E1A	
$G2$	2^{nd} shape param. Beta distr. prod.	0.020	$\mathbb{E}_\tau[c^{ind}(x,12\!:\!16)/c^{ind}(x,0)]$: long-term duration profile rel. ind. interview prob. (residual)	See Figure E1A	
κ	Interview cost	0.071	Avg interview cost per hire in Muehlemann and Strupler Leiser (2018)	0.129	0.129
Panel C. Internally estimated auxiliary parameters					
λ	Meeting prob.	0.153	$\mathbb{E}[f(x,\tau)]$: avg job finding prob.	0.062	0.061
$L(\Theta^\star)$	SMM loss function	0.193%			

Note: Expectations are taken with respect to the ability distribution at the duration of the respective subscript. When the subscript is not specified, the expectation is taken with respect to the ability distribution at all durations from 0 to $\bar{\tau}$. Numeraire: cross-sectional avg monthly output.

Figure E1: Duration profile of relative interview prob. and job finding rate, exog. search effort model vs data



Note: This figure contrasts the duration profiles controlling for observables of the relative individual interview probability (Panel A) and relative job finding rate (Panel B) detected in the data (circles) with those implied by the estimated model with exogenous search effort. The duration profiles in the model are derived by estimating individual fixed effects and duration effects from a saturated regression and computing the expected values of the individual interview probability and job finding rate at any unemployment duration. Expected values are computed with respect to the ability distribution in the contemporaneous period of unemployment, *i.e.* $\mathbb{E}_\tau[c^{ind}(x,\tau)]$ and $\mathbb{E}_\tau[f(x,\tau)]$. The distribution of observables across unemployment durations is kept the same as in the first month of unemployment. The model-based duration profiles are fitted using a negative exponential function estimated via weighted least squares.

E.2 Learning from search

In the first model extension we assume that job seekers know their search type but not their ability. Just like firms, job seekers form beliefs about their own ability based on their elapsed unemployment duration. Due to negative dynamic selection, beliefs are revised downward as unemployment spell lengthens (Gonzalez and Shi, 2010; Doppelt, 2016; Potter, 2021). In turn, job seekers scale down their application effort over the unemployment spell – even if the individual job offer probability per application remains constant (He and Kircher, 2023).

Formally, we assume that every time a worker of search type γ separates from a job, nature draws a new ability x from an exogenous distribution $\mathcal{H}(x|\gamma, \tau = 0)$, where $\tau \in \mathbb{N}$ stands for elapsed unemployment duration, and $\partial \mathbb{E}[x|\gamma]/\partial\gamma > 0$.[50] Workers do not observe their own ability draw. However, the underlying distribution $\mathcal{H}(x|\gamma, \tau = 0)$ is common knowledge.

The values of unemployment and employment (perceived by workers) can be expressed recursively as:

$$U(\gamma, \tau) = \max_{\tilde{a} \geq 0}\ b - \sigma(\tilde{a}; \gamma) + \beta\Big[U(\gamma, \tau+1) + s(\tilde{a}, \gamma)\hat{o}(\gamma, \tau)\ \big(W(\gamma) - U(\gamma, \tau+1)\big)\Big],$$

$$W(\gamma) = \omega + \beta\Big[W(\gamma) + \delta_L\big(U(\gamma, 0) - W(\gamma)\big)\Big],$$

where $\hat{o}(\gamma, \tau) = \int o(x, \tau)\ d\hat{\mathcal{H}}(x|\gamma, \tau)$ denotes the expected job offer probability per unit of search effort for a job seeker of search type γ at duration τ according to the belief function $\hat{\mathcal{H}}(x|\gamma, \tau)$. It follows that the expected job finding rate equals $\hat{f}(\gamma, \tau) \equiv s(\gamma, \tau)\ \hat{o}(\gamma, \tau)$.

Optimal application effort solves:

$$a(\gamma, \tau) : \ \frac{\partial\sigma(a; \gamma)}{\partial a} = \beta\frac{\partial s(a; \gamma)}{\partial a}\hat{o}(\gamma, \tau)\left[W(\gamma) - U(\gamma, \tau+1)\right].$$

The key additional equilibrium object of the learning model is the belief function about the job seeker's ability, $\hat{h}(x|\gamma, \tau) = \hat{\mathcal{H}}'(x|\gamma, \tau)$, which drives job seekers' application decisions. For given $\hat{h}(x|\gamma, 0) = h(x|\gamma, 0)$, the belief function about job seeker's ability

[50] New ability draws following job separations are meant to capture stochastic evolution in one worker's breadth of qualification for jobs in the marketplace. In the model, this implies that past labor market experience is not informative about worker's ability in her current spell.

evolves according to Bayesian updating:

$$\hat{h}(x|\gamma,\tau) = \frac{(1-f(x,\gamma,\tau))\,\hat{h}(x|\gamma,\tau-1)}{\int (1-f(x,\gamma,\tau))\ d\hat{\mathcal{H}}(x|\gamma,\tau-1)}, \quad \forall \tau > 0. \tag{E.1}$$

Intuitively, job seekers adjust their belief about their own ability as unemployment duration lengthens, by assigning increasingly higher density to ability levels with a lower-than-average job finding rate.[51] In equilibrium, job seekers' belief function about own ability equals the type-specific ability distribution at each duration, *i.e.* $\mathcal{H}(x|\gamma,\tau) = \hat{\mathcal{H}}(x|\gamma,\tau)$.[52]

In this extended model, job seekers optimally respond to negative duration dependence in their *expected* job offer probability per unit of search effort by scaling down their application effort over the unemployment spell. Learning from search entails that negative duration dependence in application effort (Fact 1) may also stem from negative dynamic selection in the job offer probability per unit of search effort – not just from genuine negative duration dependence.

To take the model to the data, we discretize worker ability, search type, and firm productivity on an equally-spaced grid with N grid points. Similarly, we discretize search efficiency on N grid points defined by $\epsilon(\gamma_j) = 1 + \phi\gamma_j$, $\forall j = 1,\ldots,N$. We then posit that the initial discretized density of job seekers' ability for given search type is given by $h(x_j|\gamma_j,\tau=0) = \rho \quad \text{if } \gamma = \gamma_j$, and $h(x_j|\gamma,\tau=0) = \frac{1-\rho}{N-1} \quad \text{else}$. The parameter ρ governs the correlation between ability and search type values which are equally ranked. This is a parsimonious way to get a positive correlation between ability and search efficiency through a single parameter.

We estimate the learning model through the SMM using the same estimation strategy as in our baseline. Table E2 reports our estimated parameters. We estimate a correlation between equally-ranked ability and search type grid points, ρ, of 99%, suggesting that

[51] Our learning process implicitly assumes that job seekers understand how their ability affects the job offer probability per application, but not how it influences the intermediate stages of the hiring process – i.e., interviews and job offers per interview –separately. This information structure implies that job seekers update their beliefs about their ability type only based on final search outcomes. Alternatively, we may assume that job seekers have full information about the hiring process, so beliefs are updated only after unsuccessful *interviews*. Under this alternative information structure, optimal application effort would depend on the full history of interviews throughout the unemployment spell. For parsimony, we adopt the former specification, where search type and unemployment duration are sufficient statistics to determine optimal application effort.

[52] Notice that job seekers in the model hold *unbiased* beliefs based on their information set. Biased beliefs, *e.g.* due to misperceptions about some structural elements of the economy, may further affect the duration profile of search effort Spinnewijn (2015); He and Kircher (2023).

job seekers hold rather precise beliefs about their ability.

Table E2: Parameter Estimates (Learning Model)

Parameter	Description	Estimate	Target/Source	Data	Model
Panel A. Internally estimated structural parameters					
$B1$	1^{st} shape param. Beta distr. search eff.	0.097	$\hat{\beta}_{c^{app}(x,\tau),\alpha(x)\mid\tau}/\mathbb{E}_\tau[c^{app}(x,0)]$: partial effect ind. FE in app's on interview prob. per app.	-0.029	-0.029
$B2$	2^{nd} shape param. Beta distr. search eff.	0.429	$\mathbb{E}_\tau[a(x,0\!:\!11)]$: short-term duration profile applications, residual (obs.)	See Figure E2B	
$G1$	1^{st} shape param. Beta distr. prod.	0.201	$\mathbb{E}_\tau[a(x,12\!:\!16)]$: long-term duration profile applications, residual (obs.)	See Figure E2B	
$G2$	2^{nd} shape param. Beta distr. prod.	0.700	$\mathbb{E}_\tau[f(x,12\!:\!16)]$: long-term duration profile job finding rate, residual (obs.)	See Figure E3B	
ψ_0	Scalar search effort cost	0.013	$\sigma_{\alpha(x)}$: std. dev. ind. fixed effects in applications	4.095	4.283
ϕ	Search efficiency dispersion param.	12.81	$\mathbb{E}[f(x,\tau)]$: avg job finding rate	0.062	0.061
η	Convexity search effort cost	0.250	$\mathbb{E}_\tau[f(x,0\!:\!11)]$: short-term duration profile job finding rate, residual (obs.)	See Figure E3B	
q	Qualification prob. $x \geq y$	0.402	$\mathbb{E}[c^{ind}(x,\tau)]$: avg individual interview prob.	0.231	0.223
χ	App. effort elasticity search effort	0.921	$\mathbb{E}_\tau[c^{ind}(x,0\!:\!11)]$: short-term duration profile individual interview prob., residual (obs.)	See Figure E3A	
ζ	Search eff. elasticity app. costs	1.376	$\mathbb{E}_\tau[a(x,\tau)]$: duration profile applications, residual (FE)	See Figure E2A	
$\tilde{\kappa}$	Interview cost factor ($=\kappa/q$)	0.090	Avg interview cost per hire in Muehlemann and Strupler Leiser (2018)	0.129	0.128
ρ	Equally-ranked ability-search type correlation	0.987	$\mathbb{E}_\tau[a(x,12:16)]$: long-term duration profile applications, residual (FE)	See Figure E2A	
Panel B. Internally estimated auxiliary parameters					
V	Measure of vacancies	0.193	$\mathbb{E}[a(x,\tau)]$: avg applications	10.69	10.32
λ	Meeting prob. per unit of search effort	0.014	$\mathbb{E}_\tau[c^{ind}(x,12\!:\!16)]$: long-term duration profile individual interview prob., residual (obs.)	See Figure E3A	
$L(\Theta^\star)$	SMM loss function	0.128%			

Note: Expectations are taken with respect to the ability distribution at the duration of the respective subscript. When the subscript is not specified, the expectation is taken with respect to the ability distribution at all durations from 0 to $\tilde{\tau}$. Individual fixed effects in applications are not standardized. Numeraire: cross-sectional avg monthly output.

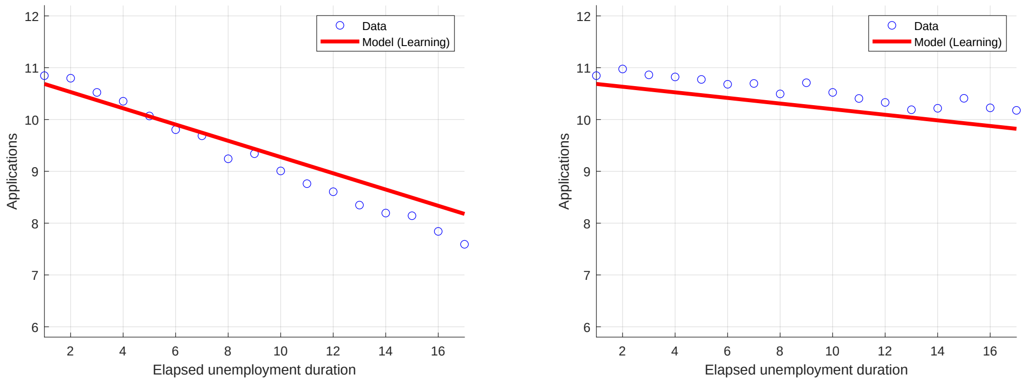
E.3 Reference-dependent preferences

In the secnod model extension we assume that workers exhibit reference-dependent preferences, in the form of loss aversion with respect to an adaptive reference point for consumption (DellaVigna et al., 2017, 2022). Specifically, workers set their reference point equal to the average consumption over the past $\tilde{\tau}$ periods. It follows that job seekers scale down their application effort over the unemployment spell as they adapt to a lower reference point for consumption. Reference-dependent preferences imply that applications should spike in correspondence to drops in consumption, *e.g.* job loss events, to then decline as agents get used to the new consumption level (DellaVigna, Lindner, Reizer, and Schmieder, 2017; DellaVigna, Heining, Schmieder, and Trenkle, 2022).[53]

[53] Reference dependence has the distinctive prediction that job seekers scale up applications as UI benefits approach exhaustion, to then decrease again. Since we lack statistical power to document search

Figure E2: Duration profile of application effort, learning model vs data

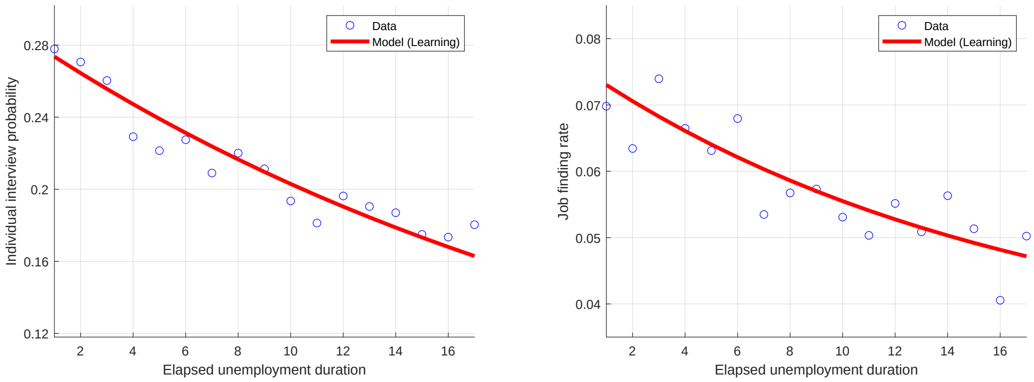
(A) Application effort, residual (FE) (B) Application effort, residual (obs.)



Note: This figure contrasts the duration profiles controlling for individual fixed effects (Panel A) and for observables (Panel B) of application effort in the data (circles) with those implied by the estimated learning model. The duration profiles in the model are derived by estimating individual fixed effects and duration effects from a saturated regression and computing the expected values of application effort at any unemployment duration. For the duration profile controlling for individual fixed effects, expected values are computed with respect to the ability distribution in the first month of unemployment, *i.e.* $\mathbb{E}_0[a(x,\tau)]$; for the duration profile controlling for observables, expected values are computed with respect to the ability distribution in the contemporaneous period of unemployment, *i.e.* $\mathbb{E}_\tau[a(x,\tau)]$. The distribution of observables across unemployment durations is kept the same as in the first month of unemployment in both specifications. The model-based duration profiles are fitted using a linear function estimated via weighted least squares.

Figure E3: Duration profile of interview prob. and job finding rate, learning model vs data

(A) Ind. interview prob., residual (obs.) (B) Job finding prob., residual (obs.)



Note: This figure contrasts the duration profiles controlling for observables of the individual interview probability (Panel A) and job finding rate (Panel B) detected in the data (circles) with those implied by the estimated learning models. The duration profiles in the model are derived by estimating individual fixed effects and duration effects from a saturated regression and computing the expected values of the individual interview probability and job finding rate at any unemployment duration. Expected values are computed with respect to the ability distribution in the contemporaneous period of unemployment, *i.e.* $\mathbb{E}_\tau[c^{ind}(x,\tau)]$ and $\mathbb{E}_\tau[f(x,\tau)]$.The distribution of observables across unemployment durations is kept the same as in the first month of unemployment. The model-based duration profiles are fitted using a negative exponential function estimated via weighted least squares.

Formally, we assume that workers have linear and reference-dependent preferences over consumption represented by the following utility function:

$$u(c_t; r_t) = \begin{cases} c_t + \Upsilon(c_t - r_t) & \text{if } c_t < r_t, \\ c_t & \text{if } c_t \geq r_t, \end{cases}$$

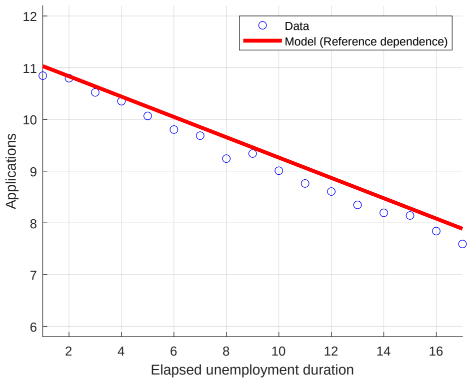
behavior around UI benefit exhaustion reliably, we are unable to provide a definitive test for reference dependence.

where $\Upsilon \geq 0$ represents the utility weight of consumption losses with respect to the reference point r_t, *i.e.* loss aversion. Following DellaVigna et al. (2022), we let the reference point be the average consumption over the $\tilde{\tau}$ previous periods, *i.e.* $r_t = \frac{1}{\tilde{\tau}} \sum_{j=1}^{\tilde{\tau}} c_{t-j}$.[54] Importantly, loss aversion induces the utility-relevant capital gain upon employment to decline with duration, as job seekers progressively adapt to a lower consumption standard (*reference-point adaptation*).

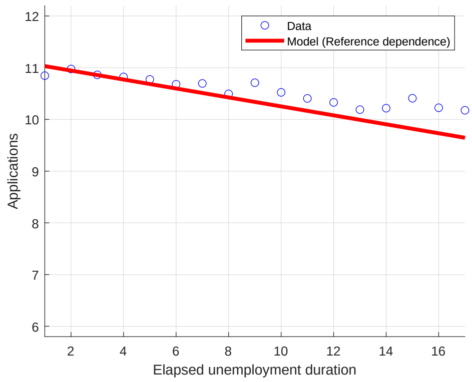
We estimate the reference dependence model through the SMM using the same estimation strategy as in our baseline. Table E3 reports our estimated parameters. We estimate a loss aversion coefficient, Υ, of 0.55, meaning that workers suffer a utility loss of 55% the gap between current consumption and reference point (provided that the former is lower than the latter).[55]

Figure E4: Duration profile of application effort, reference dependence model vs data

(A) Application effort, residual (FE)



(B) Application effort, residual (obs.)



Note: This figure contrasts the duration profiles controlling for individual fixed effects (Panel A) and for observables (Panel B) of application effort in the data (circles) with those implied by the estimated learning model. The duration profiles in the model are derived by estimating individual fixed effects and duration effects from a saturated regression and computing the expected values of application effort at any unemployment duration. For the duration profile controlling for individual fixed effects, expected values are computed with respect to the ability distribution in the first month of unemployment, *i.e.* $\mathbb{E}_0[a(x,\tau)]$; for the duration profile controlling for observables, expected values are computed with respect to the ability distribution in the contemporaneous period of unemployment, *i.e.* $\mathbb{E}_\tau[a(x,\tau)]$. The distribution of observables across unemployment durations is kept the same as in the first month of unemployment in both specifications. The model-based duration profiles are fitted using a linear function estimated via weighted least squares.

[54] Conditional on elapsed unemployment duration, history dependence in the utility function arises only if an individual transitions at least once into employment between two unemployment spells within $\tilde{\tau}$ periods. To limit the state space, we assume that the relevant consumption standards for computing the reference point of newly unemployed in all the $\tilde{\tau}$ previous periods equal the wage rate. It follows that unemployment duration is a sufficient statistic for the flow utility of unemployment.

[55] Notice that our estimate of the loss aversion coefficient is not directly comparable with those of of existing reference-dependence models where the loss aversion parameter multiplies a reference-dependence weight, such as DellaVigna et al. (2022). In our model, any $\Upsilon > 0$ means that workers value losses more than gains.

Table E3: Parameter Estimates (Reference Dependence Model)

Parameter	Description	Estimate	Target/Source	Data	Model
Panel A. Internally estimated structural parameters					
$B1$	1^{st} shape param. Beta distr. search eff.	0.075	$\hat{\beta}_{c^{app}(x,\tau),\alpha(x)\mid\tau}/\mathbb{E}_{\tau}[c^{app}(x,0)]$: partial effect ind. FE in app's on interview prob. per app.	-0.029	-0.029
$B2$	2^{nd} shape param. Beta distr. search eff.	0.389	$\mathbb{E}_{\tau}[a(x,0\!:\!11)]$: short-term duration profile applications, residual (obs.)	See Figure E4B	
$G1$	1^{st} shape param. Beta distr. prod.	0.136	$\mathbb{E}_{\tau}[a(x,12\!:\!16)]$: long-term duration profile applications, residual (obs.)	See Figure E4B	
$G2$	2^{nd} shape param. Beta distr. prod.	0.974	$\mathbb{E}_{\tau}[f(x,12\!:\!16)]$: long-term duration profile job finding rate, residual (obs.)	See Figure E5B	
ψ_0	Scalar search effort cost	0.012	$\sigma_{\alpha(x)}$: std. dev. ind. fixed effects in applications	4.095	3.907
ϕ	Search efficiency dispersion param.	16.21	$\mathbb{E}[f(x,\tau)]$: avg job finding rate	0.062	0.061
η	Convexity search effort cost	0.227	$\mathbb{E}_{\tau}[f(x,0\!:\!11)]$: short-term duration profile job finding rate, residual (obs.)	See Figure E5B	
q	Qualification prob. $x \geq y$	0.341	$\mathbb{E}[c^{ind}(x,\tau)]$: avg individual interview prob.	0.231	0.222
χ	App. effort elasticity search effort	0.880	$\mathbb{E}_{\tau}[c^{ind}(x,0\!:\!11)]$: short-term duration profile individual interview prob., residual (obs.)	See Figure E5A	
ζ	Search eff. elasticity app. costs	1.274	$\mathbb{E}_{\tau}[a(x,\tau)]$: duration profile applications, residual (FE)	See Figure E4A	
$\tilde{\kappa}$	Interview cost factor ($=\kappa/q$)	0.108	Avg interview cost per hire in Muehlemann and Strupler Leiser (2018)	0.129	0.130
Υ	Loss aversion coefficient	0.545	$\mathbb{E}_{\tau}[a(x,12:16)]$: long-term duration profile applications, residual (FE)	See Figure E4A	
Panel B. Internally estimated auxiliary parameters					
V	Measure of vacancies	0.723	$\mathbb{E}[a(x,\tau)]$: avg applications	10.69	10.44
λ	Meeting prob. per unit of search effort	0.013	$\mathbb{E}_{\tau}[c^{ind}(x,12\!:\!16)]$: long-term duration profile individual interview prob., residual (obs.)	See Figure E5A	
$L(\Theta^{\star})$	SMM loss function	0.126%			

Note: Expectations are taken with respect to the ability distribution at the duration of the respective subscript. When the subscript is not specified, the expectation is taken with respect to the ability distribution at all durations from 0 to $\bar{\tau}$. Individual fixed effects in applications are not standardized. Numeraire: cross-sectional avg monthly output.

E.4 Duration-dependent application costs

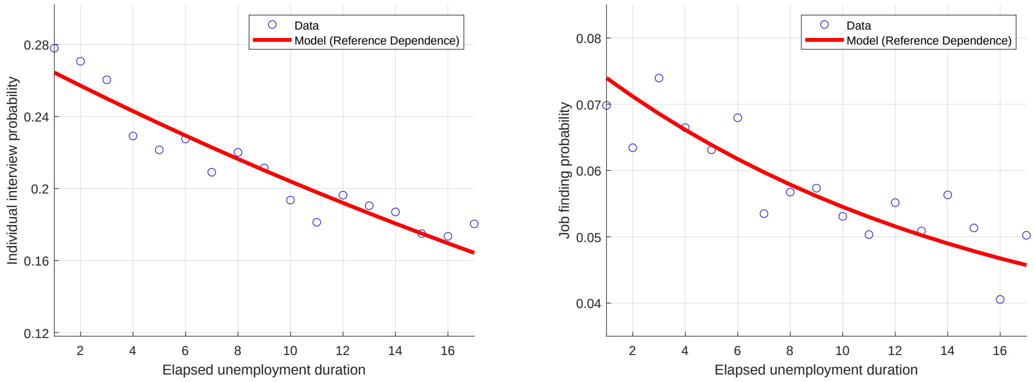
In our last model extension we allow for duration-dependent application costs by workers (DellaVigna et al., 2022). We interpret this modeling assumption as a reduced-form treatment of the depletion of personal networks. Depletion of personal networks entails that applications are increasingly more costly as the unemployment spell lengthens as job seekers run out of personal contacts and need to collect additional information on job vacancies (Beaman and Magruder, 2012; Burks et al., 2015; Hensvik and Nordström Skans, 2016).

Formally, we assume that unit application costs are duration-dependent, *i.e.* $\sigma(1;\gamma,\tau) = \frac{\psi(\gamma,\tau)}{1+\eta}$. Specifically, we posit that unit application costs are iso-elastic in search efficiency and increasing in unemployment duration: $\psi(\gamma,\tau) = \psi_0\epsilon(\gamma)^{\zeta}(1+\tau\psi_1)$.

We estimate the duration-dependent application cost model through the SMM using

Figure E5: Duration profile of interview prob. and job finding rate, reference dependence model vs data

(A) Ind. interview prob., residual (obs.) (B) Job finding prob., residual (obs.)

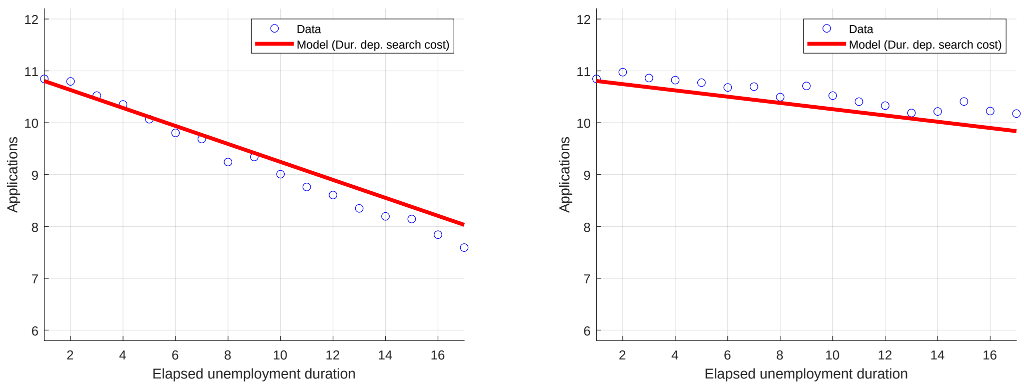


Note: This figure contrasts the duration profiles controlling for observables of the individual interview probability (Panel A) and job finding rate (Panel B) detected in the data (circles) with those implied by the estimated learning models. The duration profiles in the model are derived by estimating individual fixed effects and duration effects from a saturated regression and computing the expected values of the individual interview probability and job finding rate at any unemployment duration. Expected values are computed with respect to the ability distribution in the contemporaneous period of unemployment, *i.e.* $\mathbb{E}_\tau[c^{ind}(x,\tau)]$ and $\mathbb{E}_\tau[f(x,\tau)]$.The distribution of observables across unemployment durations is kept the same as in the first month of unemployment. The model-based duration profiles are fitted using a negative exponential function estimated via weighted least squares.

the same estimation strategy as in our baseline. Table E4 reports our estimated parameters. We estimate a duration coefficient in application costs, ψ_1, equal to 0.1% of monthly output. This implies that the unit application cost after 17 months of unemployment is 2.2% higher than in the first month.

Figure E6: Duration profile of application effort, duration-dep. app. cost model vs data

(A) Application effort, residual (FE) (B) Application effort, residual (obs.)



Note: This figure contrasts the duration profiles controlling for individual fixed effects (Panel A) and for observables (Panel B) of application effort in the data (circles) with those implied by the estimated learning model. The duration profiles in the model are derived by estimating individual fixed effects and duration effects from a saturated regression and computing the expected values of application effort at any unemployment duration. For the duration profile controlling for individual fixed effects, expected values are computed with respect to the ability distribution in the first month of unemployment, *i.e.* $\mathbb{E}_0[a(x,\tau)]$; for the duration profile controlling for observables, expected values are computed with respect to the ability distribution in the contemporaneous period of unemployment, *i.e.* $\mathbb{E}_\tau[a(x,\tau)]$. The distribution of observables across unemployment durations is kept the same as in the first month of unemployment in both specifications. The model-based duration profiles are fitted using a linear function estimated via weighted least squares.

Table E4: Parameter Estimates (Duration-dependent App. Cost Model)

Parameter	Description	Estimate	Target/Source	Data	Model
Panel A. Internally estimated structural parameters					
$B1$	1^{st} shape param. Beta distr. search eff.	0.100	$\hat{\beta}_{c^{app}(x,\tau),\alpha(x)\mid\tau}/\mathbb{E}_{\tau}[c^{app}(x,0)]$: partial effect ind. FE in app's on interview prob. per app.	-0.029	-0.029
$B2$	2^{nd} shape param. Beta distr. search eff.	0.426	$\mathbb{E}_{\tau}[a(x,0\!:\!11)]$: short-term duration profile applications, residual (obs.)	See Figure E6B	
$G1$	1^{st} shape param. Beta distr. prod.	0.195	$\mathbb{E}_{\tau}[a(x,12\!:\!16)]$: long-term duration profile applications, residual (obs.)	See Figure E6B	
$G2$	2^{nd} shape param. Beta distr. prod.	0.677	$\mathbb{E}_{\tau}[f(x,12\!:\!16)]$: long-term duration profile job finding rate, residual (obs.)	See Figure E7B	
ψ_0	Scalar search effort cost	0.013	$\sigma_{\alpha(x)}$: std. dev. ind. fixed effects in applications	4.095	4.228
ϕ	Search efficiency dispersion param.	12.42	$\mathbb{E}[f(x,\tau)]$: avg job finding rate	0.062	0.061
η	Convexity search effort cost	0.247	$\mathbb{E}_{\tau}[f(x,0\!:\!11)]$: short-term duration profile job finding rate, residual (obs.)	See Figure E7B	
q	Qualification prob. $x \geq y$	0.398	$\mathbb{E}[c^{ind}(x,\tau)]$: avg individual interview prob.	0.231	0.223
χ	App. effort elasticity search effort	0.920	$\mathbb{E}_{\tau}[c^{ind}(x,0\!:\!11)]$: short-term duration profile individual interview prob., residual (obs.)	See Figure E7A	
ζ	Search eff. elasticity app. costs	1.360	$\mathbb{E}_{\tau}[a(x,\tau)]$: duration profile applications, residual (FE)	See Figure E6A	
$\tilde{\kappa}$	Interview cost factor ($=\kappa/q$)	0.090	Avg interview cost per hire in Muehlemann and Strupler Leiser (2018)	0.129	0.125
ψ_1	Duration dependence app. costs	0.001	$\mathbb{E}_{\tau}[a(x,12:16)]$: long-term duration profile applications, residual (FE)	See Figure E6A	
Panel B. Internally estimated auxiliary parameters					
V	Measure of vacancies	0.223	$\mathbb{E}[a(x,\tau)]$: avg applications	10.69	10.39
λ	Meeting prob. per unit of search effort	0.014	$\mathbb{E}_{\tau}[c^{ind}(x,12\!:\!16)]$: long-term duration profile individual interview prob., residual (obs.)	See Figure E7A	
$L(\Theta^{\star})$	SMM loss function	0.121%			

Note: Expectations are taken with respect to the ability distribution at the duration of the respective subscript. When the subscript is not specified, the expectation is taken with respect to the ability distribution at all durations from 0 to $\bar{\tau}$. Individual fixed effects in applications are not standardized. Numeraire: cross-sectional avg monthly output.

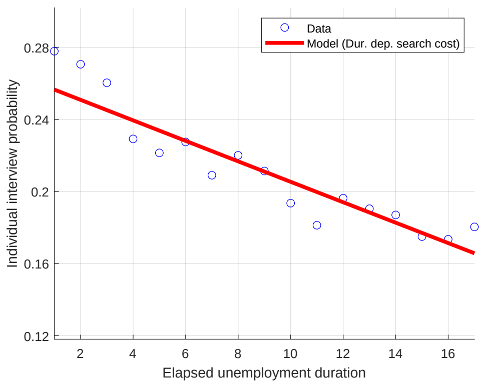
E.5 Sensitivity of decomposition shares

We finally examine how sensitive the estimated contributions of duration dependence and dynamic selection to the decline in job finding rate are to the structural modeling of the job search process. To do so, we repeat the same decomposition exercise as in equation (10) to the restricted model with exogenous search effort and the three model extensions. The results are reported in Figure E8.

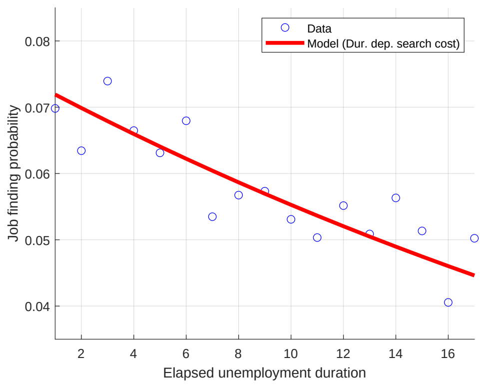
In contrast to the results of our baseline model with endogenous search effort, Figure E8A shows that dynamic selection accounts for 90% of the decline in the job finding rate in the model with exogenous search effort. Comparing Figure E8A to Figure 7, we draw two main takeaways. First, the entire share of duration dependence due to workers in our baseline model is attributed to dynamic selection on unobservables by the model with exogenous search effort. Second, the share of duration dependence due to firms is

Figure E7: Duration profile of interview prob. and job finding rate, duration-dep. app. cost model vs data

(A) Ind. interview prob., residual (obs.)



(B) Job finding prob., residual (obs.)



Note: This figure contrasts the duration profiles controlling for observables of the individual interview probability (Panel A) and job finding rate (Panel B) detected in the data (circles) with those implied by the estimated learning models. The duration profiles in the model are derived by estimating individual fixed effects and duration effects from a saturated regression and computing the expected values of the individual interview probability and job finding rate at any unemployment duration. Expected values are computed with respect to the ability distribution in the contemporaneous period of unemployment, *i.e.* $\mathbb{E}_\tau[c^{ind}(x,\tau)]$ and $\mathbb{E}_\tau[f(x,\tau)]$.The distribution of observables across unemployment durations is kept the same as in the first month of unemployment. The model-based duration profiles are fitted using a negative exponential function estimated via weighted least squares.

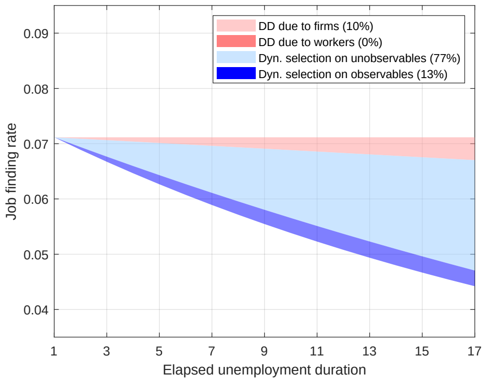
also lower in the model with exogenous search effort than in our baseline (10% versus 26%). Hence, we conclude that modeling endogenous search effort by workers is key to study duration dependence in the equilibrium job finding process.

For completeness, we also estimated a model with exogenous search effort and coordination frictions. Allowing for coordination frictions improves the model fit to the data only marginally, but the estimated share of duration dependence due to firms increases significantly up to 30%. This suggests that models with exogenous search effort *and* coordination frictions may provide a good approximation of the share of duration dependence due to firms.

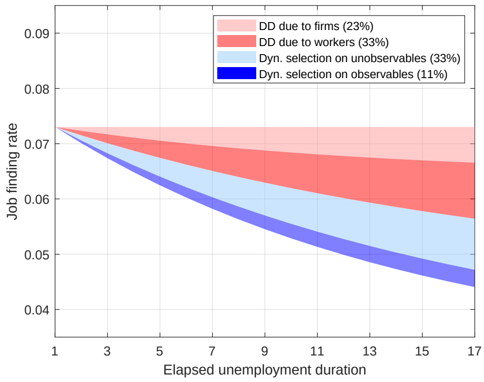
Allowing for other drivers of duration dependence in application effort than firms' statistical discrimination, Figure E8B-Figure E8D show that the contribution of duration dependence to the decline of the job finding probability ranges between 52% (duration-dependent application costs) and 57% (reference-dependent preferences). Therefore, our conclusion that duration dependence is the main driver of the decline in the job finding probability is robust to the structural modeling of the job search process.

Figure E8: Duration profile of the job finding rate, decomposition across model variants

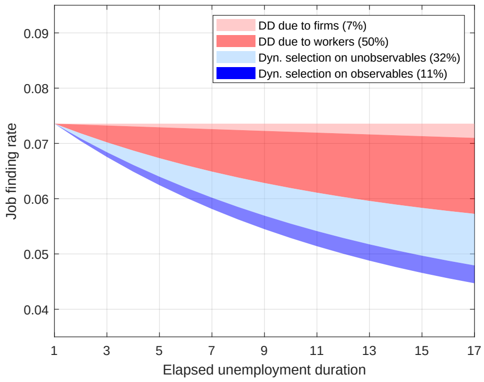
(A) Exogenous search effort



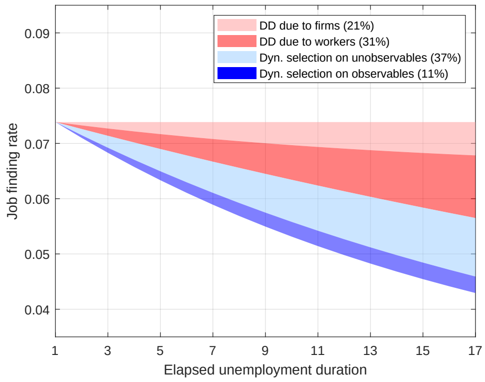
(B) Learning



(C) Reference dependence



(D) Duration-dep. app. cost



Note: These figures report the decomposition of the duration profile of the job finding rate into the different sources of duration dependence and dynamic selection derived in equation (9) and equation (10) across model variants. The duration profiles of the components of the job finding rate reported in equation (9) are derived by estimating individual fixed effects and duration effects from a saturated regression and computing the expected values of each component at any unemployment duration. Expected values are computed with respect to the ability distribution in the contemporaneous period of unemployment. The distribution of observables across unemployment durations is kept the same as in the first month of unemployment. According to equation (10), the duration profiles of the component due to dynamic selection on observables is computed as the difference between the observed duration profile of the job finding rate and the duration profile controlling for observables (see Figure B9C). Then, all duration profiles are fitted by a negative exponential function estimated via weighted nonlinear least squares. The shares of each component are computed as the frequency-weighted average shares of the respective raw components over the entire unemployment spell.

F. The role of statistical discrimination

In our model, statistical discrimination by firms – showing up as negative duration dependence in the job offer probability per unit of search effort – affects the duration dependence in the job finding rate in two ways. First, negative duration dependence in the job offer probability per unit of search effort reflects in negative duration dependence in the job finding rate with unitary elasticity (*direct effect*). Second, statistical discrimination reduces the return from workers' search as the unemployment spell lengthens, thus inducing negative duration dependence in application effort due to discouragement (*indirect effect*). In this section, we aim to assess how statistical discrimination affects the job finding rate in general equilibrium.

Local approach. We start our analysis by providing a simple analytical formula to estimate the local general-equilibrium elasticity of the job finding rate to the job offer probability per unit of search effort. Intuitively, the general-equilibrium elasticity measures how much the duration profile of the job finding rate would flatten out by reducing statistical discrimination against unemployed of a given duration.

Taking log of the job finding rate definition (6) and differentiating it with respect to duration yields:

$$\frac{d \ln f(x, \gamma, \tau)}{d\tau} = \chi \frac{d \ln a(\boldsymbol{o}(x); \gamma)}{d\tau} + \frac{d \ln o(x, \tau)}{d\tau}, \tag{F.1}$$

where $\boldsymbol{o}(x)$ denotes the full sequence of job offer probability per unit of search effort for a worker of ability x at every duration. If application effort were exogenous, the elasticity of the job finding rate to the job offer probability per unit of search effort would be simply 1. Still, in our model, the job offer probability per unit of search effort affects optimal application effort through its effect on $\boldsymbol{o}(x)$.

To compute the local general-equilibrium elasticity, we make use of equation (3), as well as our functional form assumptions, to express optimal application effort as a function of the job offer probability per unit of search effort:

$$a(x, \tau) = \left[\frac{\beta\chi}{\psi(x)} \epsilon(x) o(x, \tau)(W(x) - U(x, \tau + 1)) \right]^{\frac{1}{1+\eta-\chi}}.$$

Hence, for given capital gain upon employment, optimal application effort depends on the current job offer probability per unit of search effort with elasticity $\frac{1}{1+\eta-\chi}$. Substituting for optimal application effort into the duration profile of the job finding rate (F.1) yields:

$$\frac{d \ln f(x, \tau)}{d\tau} = \left(1 + \frac{\chi}{1 + \eta - \chi}\right) \frac{d \ln o(x, \tau)}{d\tau} + \frac{\chi}{1 + \eta - \chi} \frac{d \ln[(W(x) - U(x, \tau))/\psi(x)]}{d\tau}.$$

The local general-equilibrium elasticity of the job finding rate to the job offer probability per unit of search effort equals $1 + \frac{\chi}{1+\eta-\chi}$. According to our estimated parameters, this amounts to 3.67. It follows that the indirect effect of statistical discrimination outweighs its direct effect.

Global approach. The general-equilibrium elasticity of the job finding rate to the job offer probability per unit of search effort derived in the previous paragraph provides a local estimate of the sensitivity of the job finding rate to changes in the job offer probability per unit of search effort *at a given duration*. A natural question to ask is how the stationary equilibrium of the economy would look like if firms did not discriminate against unemployment duration *at all*, *i.e.* if the job offer probability per unit of search effort were flat over the entire unemployment spell. Indeed, since optimal application effort depends on the entire path of job offer probability per unit of search effort, changes in the latter are expected to reflect both in the level and in the duration dependence of application effort. We analyze this counterfactual scenario by setting interview costs equal zero, *i.e.* $\kappa \to 0$, and mandating firms to interview all applicants.[56] Since the presence of interview costs is the structural determinant of statistical discrimination, getting rid of the former allows removing the latter. However, interview costs are also a source of search costs for firms, so removing them would induce more vacancy posting. To single out the effect of removing statistical discrimination from that of stimulating vacancy posting, we adjust the vacancy posting cost κ_v to keep the mass of vacancies fixed at its baseline level. In this way, we guarantee the absence of pure profits from job creation, which allows us to compute heterogeneous welfare effects across workers without taking a stance on how profits are rebated. Overall, changes in the stationary equilibrium of our estimated models are to be entirely attributed to the complete flattening of the individual duration profile of the job offer probability per unit of search effort and the induced response of application effort.

Table F1 shows that job seekers react to a flat profile of the job offer probability per unit of search effort by reducing their average application effort (which is constant over the unemployment spell for each job seeker). This allows them to save on application costs. Overall, we estimate that removing statistical discrimination would reduce the long-term unemployment rate by 0.7pp and increase aggregate welfare by 1% consumption-equivalent units.[57]

The aggregate welfare gain masks some heterogeneity across worker types. Figure F1

[56]In our quantitative model, we allow for coordination frictions in the form of multiple applicants per vacancy. Absent interview costs, firms are indifferent about the ranking among multiple applicants. Mandating firms to interview all applicants rule out ranking schemes based on unemployment duration.

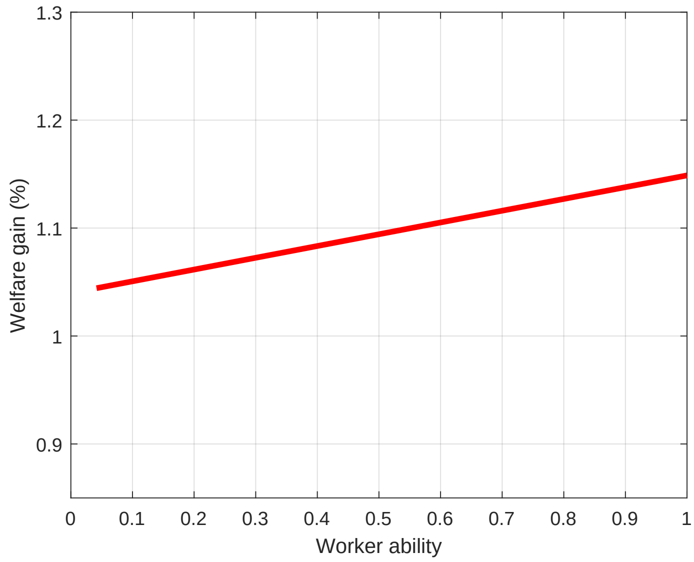
[57]With costly application effort, the unemployment rate is no longer a sufficient statistic for welfare.

Table F1: Baseline vs No statistical discrimination

	Baseline	No statistical discrimination
short-term avg job offer prob. per app. (%)	0.59	0.63
long-term avg job offer prob. per app. (%)	0.48	0.55
short-term avg application effort	10.39	9.57
long-term avg application effort	9.76	10.23
short-term avg interview prob. (%)	22.27	23.99
long-term avg interview prob. (%)	15.57	23.29
short-term avg job finding prob. (%)	6.08	5.98
long-term avg job finding prob. (%)	4.64	5.60
unemployment rate (%)	13.97	13.27
long-term unemployment rate (%)	7.62	6.70
welfare	100	100.99

Note: Averages are taken with respect to the ability distribution at all durations from 0 to $\bar{\tau}$. Long-term averages are taken with respect to the ability distribution at duration $\bar{\tau}$.

Figure F1: Welfare gains across workers



Note: This figure reports the fitted percentage increase in welfare enjoyed by workers across the search efficiency distribution. Welfare gains are fitted via a linear relationship.

shows that removing statistical discrimination mainly benefits the highest-ability workers (in terms of ability and search efficiency), who are relieved from the risk of being denied interviews for jobs they may have been qualified for because of their unemployment duration.

G. Complementary explanations

Table G1: Duration profile of applications by channel

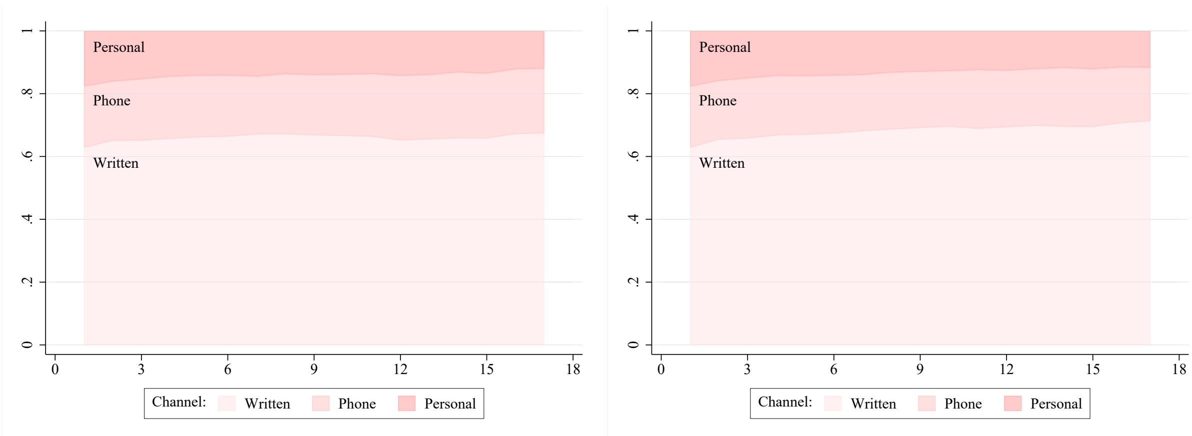
Dependent variables:	In writing		By phone		In person	
	(1)	(2)	(3)	(4)	(5)	(6)
Elapsed unemployment duration	-0.037***	-0.132***	-0.003	-0.046***	-0.038***	-0.075***
	(0.009)	(0.020)	(0.006)	(0.012)	(0.005)	(0.012)
	[-0.535%]	[-1.899%]	[-0.146%]	[-2.278%]	[-2.034%]	[-4.034%]
Constant	7.224***		2.025***		1.771***	
	(0.071)		(0.039)		(0.039)	
Individual controls	No	Yes	No	Yes	No	Yes
Policy controls	No	Yes	No	Yes	No	Yes
LLMC	No	Yes	No	Yes	No	Yes
Individual FE	No	Yes	No	Yes	No	Yes
Mean outcome 1st month	6.962	6.962	2.035	2.035	1.849	1.849
adj.-R^2	0.001	0.631	0.000	0.614	0.003	0.615
Observations	58755	58755	58755	58755	58755	58755

Note: This table reports empirical estimates of equation (1) using OLS, where the duration function $f^A(t)\phi^A$ is specified linearly. The dependent variables are the number of applications made in writing (columns 1-2), by phone (columns 3-4) and in person (columns 5-6). For each dependent variable, we report estimation results from a simple binary regression (on duration only) and from the full specification described in equation (1). Standard errors are clustered at the individual level and reported in parentheses. Coefficients in relative terms (with respect to the average in the first month of unemployment) are indicated in square brackets. Stars indicate the following significance levels: * 0.1, ** 0.05 and *** 0.01.

Figure G1: Changes in the shares of application channels

(A) Observed

(B) Fixed effects



Note: This figure represents the share of applications sent out through the written, phone and personal channels, per month of elapsed unemployment. Panel A corresponds to the patterns in the raw data, without accounting for changes in the pool of applicants. Panel B corresponds to the results of a fixed effects regression, that accounts for the evolution of the pool of applicants.

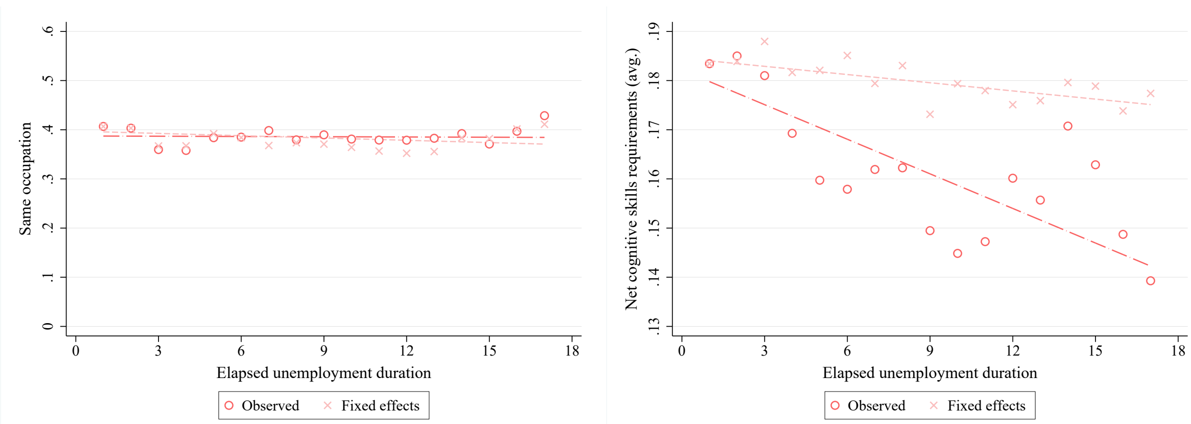
Table G2: Job search effort provision and application channels' shares

	Written channel		**Phone channel**		**Personal channel**	
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: estimated α_i						
Application channel's share	0.569***	0.857***	-0.863***	-0.662***	-0.251	-0.973***
	(0.123)	(0.126)	(0.169)	(0.165)	(0.184)	(0.186)
Individual controls	No	Yes	No	Yes	No	Yes
Mean outcome	10.224	10.224	10.224	10.224	10.224	10.224
Adjusted R^2	0.002	0.156	0.002	0.153	0.000	0.154
Observations	14798	14798	14798	14798	14798	14798

Note: This table reports evidence of the correlation between job search effort provision and the use of application channels. Each column reports the partial correlation between the estimated α_i from equation (1) and the share of each channel (written, phone, personal) in all applications sent by job seeker i (aggregated at the individual level). Odd columns correspond to bi-variate regressions, whereas even columns additionally control for job seekers' characteristics. Stars indicate the following significance levels: * 0.1, ** 0.05 and *** 0.01.

Figure G2: Changes in application targeting

(A) Same occupation (B) Skills requirements



Note: This figure describes the evolution of application characteristics with respect to elapsed unemployment duration. The two panels are based on the *Auxiliary sample*. Panel A shows results for the share of targeted positions that are the same as occupations desired by the job seekers. Panel B reports evidence for the net-cognitive skill requirements of targeted occupations. Both panels show evidence based on the raw data (circle) and evidence controlling for individual heterogeneity, through individual fixed effects (x-cross).