ROCKWOOL FOUNDATION BERLIN
INSTITUTE FOR THE ECONOMY AND THE FUTURE OF WORK

CReAM Centre for Research & Analysis of Migration

DISCUSSION PAPER SERIES 15/25

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

Rafael Lalive, Aderonke Osikominu, Lorenzo Pesaresi, Jeremy Zuchuat, Josef Zweimüller

APRIL 2025

ROCKWOOL Foundation Berlin
www.rfberlin.com

Centre for Research & Analysis of Migration
www.cream-migration.org

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

Rafael Lalive
University of Lausanne

Aderonke Osikominu
University of Hohenheim

Lorenzo Pesaresi
University of Zürich

Jeremy Zuchuat
University of Lausanne

Josef Zweimüller
University of Zürich

April 24, 2025

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 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 roughly half of the decline in the job finding rate is due to structural duration dependence and the other half to 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. 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 an invitation to a job interview and, if so, whether the interview eventually resulted in a job offer. Search diaries are an important monitoring tool providing high-quality information on 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. These data allow us to dig deeper into the various steps of the job finding process and provide novel evidence on how job applications, interviews, and job offers change with the duration of unemployment. The existing literature has looked, separately, at the effect of duration on search effort and the effect of duration on employer callbacks. Our study is the first one looking at both margins simultaneously in the same data set. Moreover, our search diary data allow us to explore how the probability to obtain a job offer after an interview changes with duration. To the best of our knowledge, no previous study has yet explored the duration profile of this outcome.

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 *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, *i.e.* 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 interviews and rising job offers are due to duration dependence or dynamic selection is more complicated. Unlike job applications, which are repeatedly observed and occur throughout the unemployment spell, job interviews and job offers are rare events that are concentrated at the end of the spell. Hence, the fixed effects model does not work.[2] Instead, 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.[3] Using this statistical model, which holds the composition of *observable* characteristics constant, we find that the interview probability 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.

[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] Even if the fixed effects approach were feasible, its usefulness in capturing the recruiting firm's perspective is uncertain, as firms may themselves have imperfect information about an applicant's quality.

[3] 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.

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.[4] Finally, with 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.

As mentioned above, our empirical approach is more comprehensive than that of existing studies since we can observe the responses to duration of both job seekers *and* recruiters. This grants us more than the simple sum of the two in isolation. Specifically, we can address the pervasive concern of dynamic selection on *unobservables* when estimating duration dependence in recruiters' decisions.[5] We do this by using the individual application fixed effect as a regressor in the interview and job offer models. Because a high application fixed effect is associated with a long duration of unemployment (recall the positive dynamic selection in applications), it is indicative of low "unobserved employability", *i.e.* unobservables determining the willingness of recruiters to hire an applicant. We find that, indeed, a larger individual application fixed effect is associated with a significant (and quantitatively large) reduction in both the interview probability and the probability to obtain a job offer after an interview.

Overall, our empirical analysis suggests that both duration dependence and dynamic selection are likely to be important for the duration profile of the job finding rate. However, the fact that we can only imperfectly account for unobserved heterogeneity leaves our empirical analysis inconclusive. This points to the importance of a structural model to obtain a consistent decomposition of the falling job finding rate into duration dependence and dynamic selection.

Our second main contribution is to provide such a theoretical framework. In the model, heterogeneous job seekers, differing in ability and search efficiency, and firms,

---

[4] 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.

[5] 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.

differing in ability requirements, interact in a frictional labor market. 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 employers, which themselves depend on the job seeker's unemployment duration. Our modelling strategy is motivated by the three main empirical findings of the reduced-form analysis. First, declining applications hint at a framework in which the net benefits from search decrease with the duration of unemployment. Second, the positive dynamic selection on unobservables in applications, together with the strong negative effect of the application fixed effect on the interview probability, suggest that there are significant differences in search efficiency across job seekers.[6] Finally, the rising duration profile of job offers per interview, together with the falling profile of job offers per application, suggest that firms statistically discriminate against the long-term unemployed in their interview policy.[7]

We model search behavior in a rather general way, so that duration dependence in job search could be due to discouragement, loss-averse preferences, duration-dependent application costs, incomplete information about one's own ability, or a combination of these channels. Firms' recruiting behavior is driven by incomplete information about an applicant's ability, which is revealed after a costly interview. This boils down to a model of duration-dependent job search in which firms statistically discriminate against the long-term unemployed.[8]

Our theoretical framework delivers three important predictions. First, it predicts negative duration dependence in *job applications*. This is because of a direct effect and an equilibrium effect. The direct effect comes from a lower return from an application at a longer duration, either due to reference-point adaptation, depletion of personal network,

[6]Assuming that firms base their callback policy on the same set of observable characteristics we can control for, application fixed effects should not affect the interview probability if job seekers were equally efficient in search.

[7]A falling duration profile of job offers per application may result from negative dynamic selection on unobservables or discrimination against unemployment duration. The differential duration profile of job offers before and after an interview is inconsistent both with pure dynamic selection and with taste-based discrimination, whereas it may arise in a context of statistical discrimination. Statistical discrimination occurs when firms use unemployment duration as a signal for job seekers' (unobserved) ability when deciding whether to invite them to job interviews.

[8]The firm side of our model is a non-trivial extension of the framework of Jarosch and Pilossoph (2019). Indeed, relaxing the assumption that each job seeker sends one application per period crucially modifies firms' callback behavior. Intuitively, if higher-ability job seekers exert more (less) search effort at a given duration, a firm may decide to (not to) interview job seekers of that duration. This is because heterogeneous search effort across job seekers affects the pace of dynamic selection in job seekers' ability.

and/or learning from search. The equilibrium effect arises from applicants internalizing that firms discriminate against applicants with a longer duration. This further reduces the return from search.

Second, the model generates an equilibrium where the eventually long-term unemployed apply more at any duration because their search is, on average, less efficient. Since recruiters are less willing to interview and offer the job to long-term unemployed applicants, our model captures the empirically observed negative correlation between the application fixed effect and the interview probability.[9]

Finally, the model predicts that the job offer probability per application falls and the job offer probability per interview rises with duration. Since the average job seeker's ability decreases with duration, the duration profile of job offers per application is falling, as negative dynamic selection and duration dependence reinforce each other. At the same time, the variance of job seeker's ability decreases with duration, generating a rising duration profile of job offers per interview. Intuitively, only firms that are willing to hire low-ability workers interview job seekers with a long unemployment duration.

We estimate the structural model by targeting the observed cross-sectional properties and duration profiles of applications, interviews, and job offers. 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 duration dependence and dynamic selection in almost equal proportions.[10] According to our estimates, duration dependence accounts for 53% of the observed job finding rate decline and comes about mainly from reduced search effort by job seekers (45%), while employer behavior (interviews, job offers) is quantitatively less important (8%). Observed and unobserved heterogeneity account for the remaining 47%. Duration dependence is mainly due to job seekers' reduced search effort at longer durations. To a large extent, this is driven by workers internalizing that firms discriminate against

[9] Specifically, when the effective marginal application costs are increasing in search efficiency, *e.g.* because more efficient workers are wealthier, the model replicates negative duration dependence and positive dynamic selection in applications simultaneously.

[10] The decomposition exercise uses our empirical analysis to account for the role of observable characteristics and the theoretical analysis (along with the estimated structural parameters) to decompose the residual duration profiles after controlling for observables into dynamic selection on unobservables and duration dependence.

applicants with a longer duration.[11]

The remainder of the paper is organized as follows. In Section 2 we discuss 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 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) and Mueller and Spinnewijn (2023) find that heterogeneity is the most important driver behind the falling job finding rate, and the dynamics of labor markets more generally (Ahn, Hobijn, and Şahin, 2023).

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 paper 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 how changes in search strategies along the duration of unemployment affect

[11] According to our estimates, a 1% reduction in the job offer probability per application triggers, on average, a 3% reduction in the job finding rate. This huge amplification is mediated by the response of application effort to a declining job offer probability per application (*discouragement*). Hence, statistical discrimination by firms affects duration dependence mainly due to its indirect effect on workers' search effort.

the job finding rate (Belot, Kircher, and Muller, 2018).

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 the callback rate.

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), 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, and biased beliefs about job-finding prospects (DellaVigna, Lindner, Reizer, and Schmieder, 2017; DellaVigna, Heining, Schmieder, and Trenkle, 2022; He and Kircher, 2023). We contribute to this literature by proposing a theory of duration-dependent job search 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

The context of our analysis are job seekers in Switzerland drawing unemployment insurance (UI) benefits. Like in most unemployment insurance (UI) systems, job seekers in Switzerland who receive UI benefits are obliged to actively search for new jobs. Compliance with Swiss UI rules implies that job seekers have to document their search effort in monthly search diaries.[12] 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, search effort 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 check the correctness 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. This means that unemployed workers have a strong incentive to provide correct information in search diaries. 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.[13] Overall, search requirements and benefit sanctions provide a lower bound on the decline of job applications over an unemployment spell.

The Swiss UI system is rather generous. UI benefits are 70% of previous earnings or 80% for low income earners or job seekers with dependents. The maximum duration of UI benefits is 18 months. In what follows, we will truncate the analysis at 17 months. The main reason is the lack of statistical power at longer durations. This also means that the observed duration profiles are not determined by changes in UI benefits over time, as all job seekers are entitled to regular UI benefits when their outcomes – applications, interviews, and job offers – are observed.

The search diary data used for this study were collected between April 2012 and March 2013 in five Swiss cantons (Zürich, Bern, Vaud, Zug and St-Gallen).[14] All workers who

[12]The monthly search diary is a standardized form that job seekers have to fill out. For the design of this form, see Appendix, in Figure A1.

[13]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.

[14]Around 47% of the Swiss population live in one of these five cantons.

were unemployed in April 2012 and all who 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).[15]

We digitized more than 58,000 monthly search diaries filled out by 15,000 job seekers. These diaries document more than 600,000 job applications and their outcomes (job interview, job offer). A particular advantage is that the search diary data can be linked to the Swiss unemployment insurance register (reporting job seekers' socio-economic and demographic characteristics) and to the Swiss social security register (providing information on workers' previous and subsequent earnings- and employment history). Another advantage is that search diaries report the behavior of both job seekers and recruiters, thus allowing us to quantify the relative importance of supply and demand forces as drivers of the job finding rate.[16] We restrict our analysis samples to those job applications made in months during which a job seeker receives UI benefits. This is motivated by data reliability: only job seekers drawing unemployment 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 remain right-censored, meaning that the outcome of the job application remains unknown. However, since right-censoring in applications does not vary with unemployment duration (see Fig-

[15]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).

[16]When interpreting applications as workers' search behavior and interviews and job offers as firms' decisions, one should keep in mind the caveats to this interpretation. 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.

ure B4 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 right-censored applications would typically not have resulted in a job offer. For these reasons, we integrated right-censored job applications into the baseline analysis and code the response to right-censored job applications the same way as a rejection to the application. Our results are not sensitive to treating right-censored applications as rejections or to removing them from the pool of applications (see Figures B5 to B7 in the Appendix).[17]

The outcome of main interest is the *job finding rate* – as measured by the probability of at least one job offer from applications sent during a given month. Notice that the job finding rate is purely based on search diaries and relates the job finding event to the month when the application was made. Table 1 reports descriptive statistics on the job finding rate, and 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 about 10.5 applications on a typical search diary. Firms invite applicants to 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 make 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

[17] 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 Appendix D.

graphs are conceptually similar and based on the same population at risk.[18] 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.[19]

Our second validation of the information content of the search diaries is based on income trajectories observed in the social security data after the last job offer observed in the search diary data. Appendix Figure A2 shows that, indeed, labor earnings are close to zero during the months before the job offer and increase sharply in the 2-3 months after the last job offer. 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

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

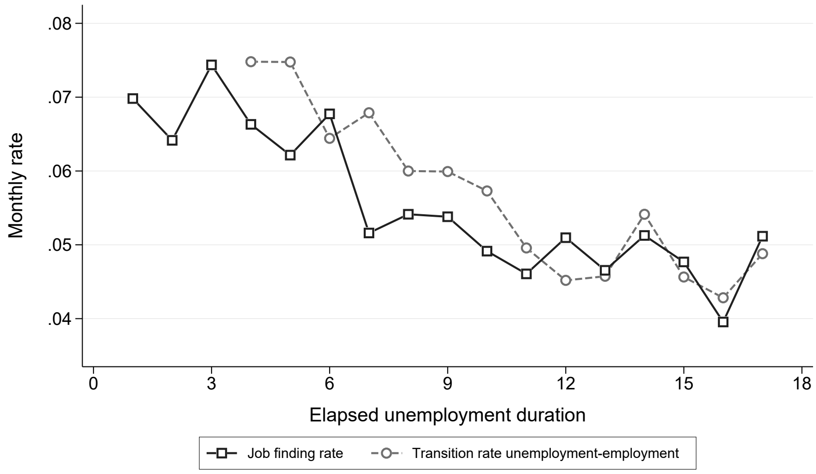
A. Person-month level (search-diary level)		
Job finding rate (per month)	0.061	(0.239)
Number of applications (per month)	10.553	(4.698)
B. Application level		
Interview probability (per application)	0.040	(0.196)
Job offer probability (per interview)	0.225	(0.418)
Job offer probability (per application)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

[18] In our definition of the job finding rate, the population at risk includes individuals 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.

[19] 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 would be identical to the latter only 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 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. 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 security data is identical to the employer who made a job offer to the job seeker.

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).

or more. Panel B shows that the probability that an application receives an invitation to 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 – *increases* with duration (Panel C). Early in the unemployment spell this probability is around 20 percent, increasing up to 30 percent at long durations.

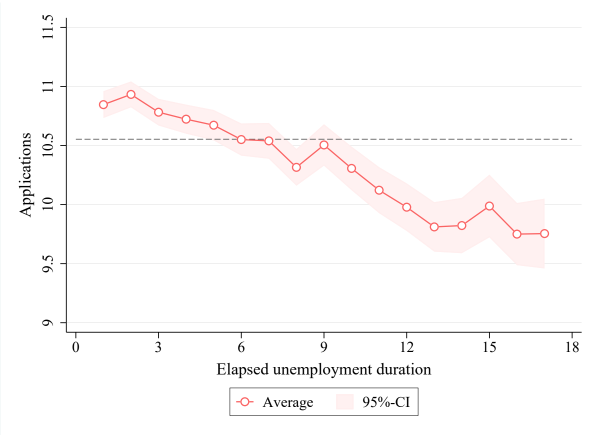
It is worth noting at this stage that, with respect to job applications and job interviews, the descriptive evidence is in the ballpark of what other studies have documented in different contexts. For instance, Faberman and Kudlyak (2019) find a decreasing profile of job applications in online job board 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.

4. Applications, interviews and job offers

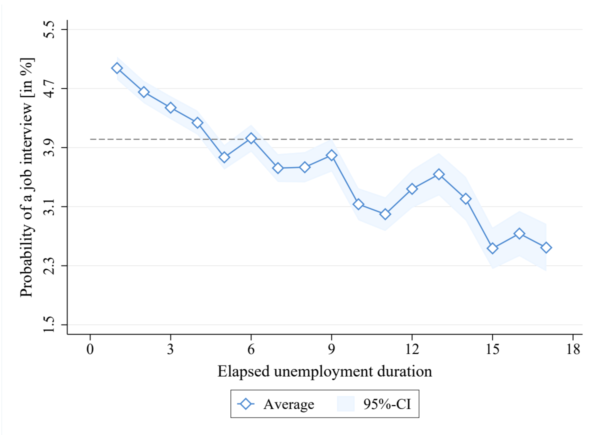
We now exploit our search diary data to study how (i) the number of job applications, (ii) the probability of an interview and (iii) the probability of a job offer (per interview) change with the duration of unemployment. These three outcomes jointly determine the job finding rate. We are not only interested in how applications, interviews and job

Figure 2: Empirical duration profiles

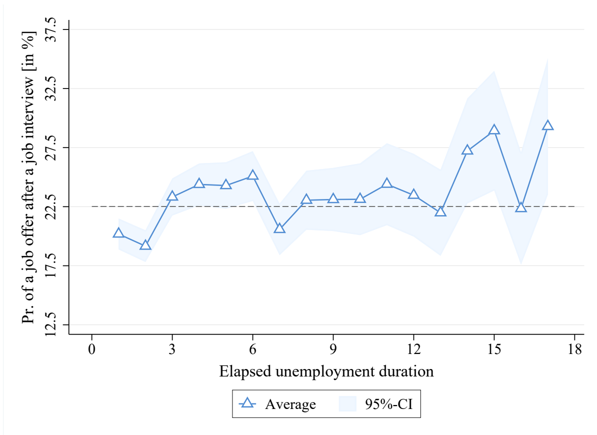
(A) Applications



(B) Interview probability



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



Note: This figure depicts the empirical duration profiles in the number of job applications made per month (Panel A), the application-level probability of a job interview (Panel B) and the application-level probability of a job offer conditional on an interview (Panel C). Dashed horizontal lines indicate sample average.

offers change with duration, but we also want to explore 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 occurs within individuals. Hence, the natural empirical approach is a statistical model which accounts for duration-invariant individual fixed effects. In the case of job applications, we can indeed apply such a fixed effects approach. Applications are both repeatedly observed and occur in positive amounts throughout the unemployment spell. Unfortunately, this is not the case for interviews and job offers. Interviews are rare events which are concentrated at the end of a spell. Job offers usually occur only once, as most job seekers accept the first job offer they get.

For this reason, the fixed effects approach does not work for interviews and job offers.

When accounting for dynamic selection in these outcomes, we are limited to the characteristics we observe in our rich administrative data. We adopt a flexible double-debiased machine learning approach (DDML) to control for these characteristics.[20] DDML consists of a two-step procedure. In the first 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 residual of the unemployment-duration equation on the residual of the outcome regression. This yields the effect of duration on the outcome after accounting for the effect of observables on both the outcome of interest and the individual's unemployment duration.[21]

Notice that, despite our rich data from UI- and social-security registers, it is unlikely that the DDML approach captures all relevant heterogeneity determining the outcomes of interest, so unobserved heterogeneity may still confound our estimates. To make progress, we take advantage of the individual fixed effect we obtain from the application regression. As we will show below, the application fixed effect is positively correlated with a job seeker's elapsed unemployment duration. To the extent that unobservables affecting application intensity are correlated with unobservables affecting the willingness of employers to hire a worker, the application fixed effect should predict interviews and job offers by acting as a proxy for unobserved employability.

Just as we, as researchers, cannot observe all characteristics affecting interview- and job-offer decisions, recruiters themselves make hiring decisions under imperfect informa-

---

[20] 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.

[21] 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 this two-step approach using a stacked DDML approach (Ahrens et al., 2024; Chernozhukov et al., 2018) that is robust to model uncertainty concerning functional form and relevance of control variables and addresses potential overfitting issues in an optimal way. We implement the stacked DDML approach by using an ensemble of base learners, each with differing strengths (*i.e.* standard parametric models, lasso and ridge regression, as well as random forest and gradient boosted trees). This enables us to perform the first-step estimations in a flexible, data-driven manner that maximizes robustness against various forms of misspecification and model selection errors (Ahrens et al., 2024; Breiman, 1996; van der Laan et al., 2007; Wolpert, 1992). Additionally, we use sample-splitting and cross-fitting techniques to make our second-step estimates robust to estimation errors in the first step and to prevent overfitting (Chernozhukov et al., 2018).

tion. After all, recruiters call back applicants to interviews precisely to figure out whether the applicant is suitable for the open position. Due to the richness of our data set, 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.

Hence, we implement our analysis in two steps. First, we estimate residual duration profiles for both job interviews and job offers controlling for the observable characteristics of the job seeker and of the application that a firm can likely infer from the application materials. Second, we add the estimated application fixed effect to explore the role of unobservable characteristics which are unknown to firms prior to a job interview.

We acknowledge that our approach may not capture all unobserved heterogeneity in interview and job offer decisions, as the application fixed effect is likely to be only an imperfect proxy for unobserved employability. Therefore, our estimates may still be biased if unobservable employability has some nonzero correlation with duration once controlling for application fixed effects. This is why, in Section 5, we will develop a theoretical framework which explicitly allows for unobserved heterogeneity across job seekers and recruiters. By imposing this structure, we will be able to provide a decomposition of the falling job finding rate into duration dependence and dynamic selection that takes unobserved and observed characteristics into account.

Job applications. Our main empirical strategy takes advantage of the fact that, for a given job seeker, the number of applications can be observed repeatedly. To disentangle duration dependence from dynamic selection in the number of applications, we adopt a fixed effects approach that accounts for time-invariant individual heterogeneity. Specifically, we model the number of applications, A_{it}, of job seeker i in month t as

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

where α_i is the individual fixed effect, X_{it} is a vector of calendar quarter $\times$ local labor market dummies, and ε^A_{it} an idiosyncratic error term.[22] The function $f^A(t)\phi^A$ captures duration dependence in the number of applications net of individual observed and unob-

[22] All the other observed variables are time-constant and therefore drop out when estimating eq. (1) with the fixed effects estimator.

served characteristics, X_{it} and α_i.[23]

We refer to the function $f^A(t)\phi^A$ as residual duration dependence because it captures the duration dependence in applications after partialling out the influence of X_{it} and α_i. 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).

Before estimating duration dependence in monthly job applications using a fixed effects model, we also examine the influence of job seekers' observed characteristics.[24] We estimate different models that rely on different specifications of the covariates: a specification that includes a hand-selected set of variables in a linear way, as well as highly flexible specifications that rely on a comprehensive dictionary of covariates that includes the original variables along with their interactions, polynomials and other transformations. We estimate this model adopting both OLS and the stacked DDML approach (see Appendix B.1 and B.2.1) to back out residual duration dependence after accounting for observed individual heterogeneity.

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-

[23]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). It delivers reliable estimates of duration dependence when the dependent variable is not directly related to exits from the sample (Zuchuat, 2023).

[24]See Appendix Table A2 for an overview of the characteristics observed in our linked register data.

selected set of individual control in an OLS regression (column 2). The stacked DDML estimates, in column (3), show that applications decline by about 0.050 per month, or around 0.46%, 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.[25]

Table 2: Duration dependence in applications, linear specification

	(1)	(2)	(3)	(4)
Dependent variable: Applications per month				
Elapsed unemployment duration	-0.078***	-0.043***	-0.050***	-0.209***
	(0.008)	(0.007)	(0.008)	(0.021)
	[-0.718%]	[-0.398%]	[-0.457%]	[-1.926%]
Individual characteristics	No	Yes	Yes	No
Policy controls	No	No	Yes	No
Local labor market conditions	No	No	Yes	Yes
DDML	No	No	Yes	No
Individual FE	No	No	No	Yes
Mean outcome 1st month	10.846	10.846	10.846	10.846
Adjusted-R^2	0.005	0.159	–	0.495
Observations	58755	58755	58755	55559
Persons	14798	14798	14798	11602

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.

In panel A of 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.[26]

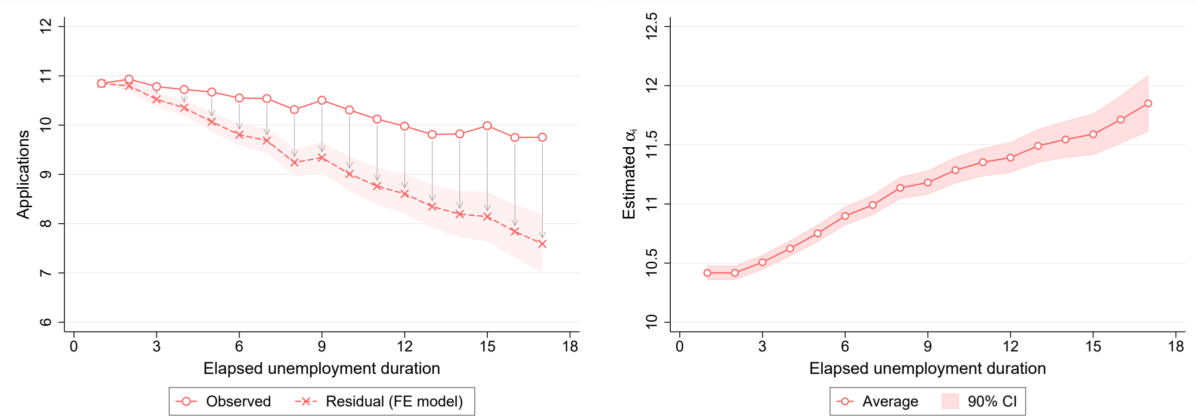
[25] In Appendix B.2.1, we explore to what extent the observed individual characteristics can predict the estimated individual fixed effects. Applying the stacked DDML approach, we find that observed characteristics explain 31% of the variation in the estimated applications fixed effects, see Table B2.

[26] Figure B1 in the Appendix, repeats the exercise, showing also the duration profile obtained with OLS.

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 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.

Panel A of 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.[27] 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 make more applications at all durations. Figure 3B shows the average of the estimated individual fixed effects of those job seekers who are still unemployed in the respective unemployment month. Those still in unemployment at high elapsed durations have a higher fixed effect α_i than those who leave unemployment quickly.[28]

Recall that job seekers need to send a pre-specified number of applications in each month, which may pose a challenge for the empirical results. In Appendix B.2.2, we address this challenge by removing the number of applications that are required (usually 8 or 10 applications per month), keeping only those applications in excess of the search

[27] In Appendix Figure A4, we show direct evidence on the within-person decline in applications for job seekers who all remain unemployed for at least 3 (or 6, 9, 12, or 15) months. In all these sub-samples, applications decline much stronger than in the full population of job seekers.

[28] 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 used by around 20% of all job seekers in France, report strong within-person increases in applications just before benefit exhaustion, but application numbers are flat or somewhat decreasing otherwise.

requirement, setting the excess applications to zero for job seekers whose applications fall below the search requirement. Excess applications are likely to reflect application effort more closely than total applications. It turns out that the results are consistent across specifications (see Table B5).[29] We conclude that there is strong negative duration dependence and strong positive dynamic selection in the number job applications.

We have 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 B3 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 allow 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 to 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 an invitation to a job interview, $Y = O$ in case of a job offer per interview, and $Y = U$ in case of a job 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 callbacks 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 $f^Y(t)$ corresponds to a complete set of dummy variables, one 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 (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 $Y = U$.

[29] In addition, we estimate count-data models with individual fixed effects to verify the robustness of our results to applying linear or nonlinear models, see Table B4 in Appendix B.2.2.

We exploit our rich 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).[30] 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.3.1 for further details). Moreover, we use the stacked DDML approach to control for the estimated individual application fixed effect as a proxy for the unobserved (to us, but potentially observed by the firm during an interview) employability of a job seeker.

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 and model the effects of observed covariates non-parametrically (see Appendix B.1 and B.3.1 for further information on the implementation of the stacked DDML approach). The estimated model underlying column (4) controls in addition for the standardized estimated individual application fixed effect, $\hat{\alpha}_i$ from eq. (1), that is meant to proxy a job seeker's employability (recall that job seeker's with a higher value of $\hat{\alpha}_i$ tend to leave unemployment more slowly, see Figure 3B).

The probability of a job interview declines by about 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 the duration dependence of interviews somewhat (in absolute value) to -0.133 percent per month (column (2) in panel A).[31] The estimates based on the stacked DDML approach suggest an even smaller decline in interviews of around 0.12 percentage points per month (column (3) in panel A). Job seekers with high application fixed effects tend to remain unemployed longer than those with low application fixed effects. Column (4) in panel A presents stacked DDML estimates that control for the estimated application fixed effect, whose partial effect on the interview probability turns out to be quantitatively

[30]See Table A2 in the Appendix for an overview.

[31]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 B7 in the appendix, shows that the AUROC for the first stage model, that predicts job interviews without including duration, is 0.657.

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***	-0.119***	-0.105***
	(0.015)	(0.015)	(0.014)	(0.015)
	[-3.117%]	[-2.668%]	[-2.394%]	[-2.105%]
Employability (α_i)				-0.342***
				(0.084)
				[-6.877%]
Mean outcome 1$^{\text{st}}$ month	4.977	4.977	4.977	4.977
Area under ROC	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***	0.310***	0.342***
	(0.099)	(0.096)	(0.102)	(0.103)
	[1.736%]	[1.769%]	[1.537%]	[1.695%]
Employability ($\hat{\alpha}_i$)				-0.988**
				(0.460)
				[-4.894%]
Mean outcome 1$^{\text{st}}$ month	20.187	20.187	20.187	20.187
Area under ROC	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***	-0.011*
	(0.006)	(0.006)	(0.006)	(0.006)
	[-2.339%]	[-1.642%]	[-1.814%]	[-1.104%]
Employability (α_i)				-0.174***
				(0.037)
				[-16.935%]
Mean outcome 1$^{\text{st}}$ month	1.027	1.027	1.027	1.027
Area under ROC	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. Standard errors (in parentheses) are clustered at the individual level. Stars indicate the following significance levels: * 0.1, ** 0.05 and *** 0.01.

important. Estimates in column (4) of panel A suggest that a one-standard-deviation increase in the estimated application fixed effect (*i.e.*, about four more applications per month) reduces the probability to be invited to an interview by 0.342 percentage points, which is around 6.8% of the baseline interview probability. The chances of an interview decline by 0.11 percentage points per month, or 2.1%, holding constant job seeker application fixed effects. The residual decline in interview chances is substantial, with 10 months of unemployment decreasing interview chances by around 21%. The results in column (4) of panel A suggest that the job seeker's application behavior captures an important element of a job seeker's employability, possibly because job seekers with low application fixed effects target applications better, or write applications of higher quality. In fact, this significant effect points to labor market frictions on the side of job seekers (*e.g.* search inefficiencies) that are not directly observed by firms, but whose implications are relevant for the job seeker's callback chances. Moreover, 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.

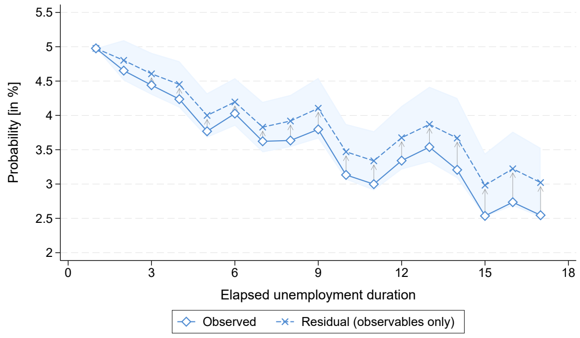
Unlike the interview probability, the probability of a job offer per interview *increases* with unemployment duration, by 0.31 percentage points per month, which is about 1.54% of the baseline job offer probability (panel B of Table 3, column (3)). Taken at face value, these estimates suggest that firms are more likely to offer a job to job seekers interviewed later in the spell – job interviews appear more targeted.[32] A comparison of the estimates in columns (1) through (3) of panel B reveals that the estimated residual duration dependence only slightly declines, from 0.35 to 0.31 percentage points per month, as the conditioning on control variables becomes more comprehensive and as the impact of observed characteristics is modeled non-parametrically using stacked DDML. Like for interviews, a one-standard deviation higher application fixed effect reduces job offer chances by around 0.99 percentage points per month, or about 4.99% of the baseline job offer probability (column (4) in panel B). Job seekers with a high application fixed effect have both a reduced chance of being called to an interview, and a reduced chance of receiving a job offer. Hence, the application fixed effect is indeed a proxy for employability of job seekers, reducing both interview and job offer chances: job seekers with a high application fixed effect have low chances of leaving unemployment and are over-represented among the

[32]We have explored alternative explanations, *e.g.* perhaps job seekers interviewed late in the spell were already doing their second or third interview, and are therefore more experienced. There is little support for this idea in the data: the duration dependence in job offers is positive and significant for the first interview as well as the second or higher order interview of job seekers. Also, the duration dependence in job offers remains positive and significant when we remove the first two months, or the last two months of data. Results available upon request.

long-term unemployed (Figure 3A). Moreover, comparing columns (1) and (3) of panel B, we see that controlling for observed characteristics reduces duration dependence somewhat, but adding application fixed effects increases duration dependence again. Dynamic selection does not play an important role for job offers per interview.

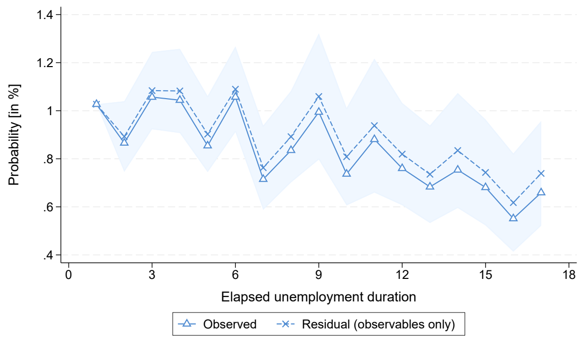
Figure 4: Empirical and residual duration profiles of job interviews and job offers

(A) Job interviews



(B) Job offers (per interview)

(C) Job offers (per application)



Note: In each panel, the solid line depicts the empirical duration profile of the interview probability (panel A), the job offer probability per interview (panel B), and the job offer probability per application (panel C) and 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 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 decreases from 5 percent in month one to 2.5 percent in month 15 of unemployment. In contrast, the decline in the interview probability adjusted for observed heterogeneity is substantially lower and decreases from 5 to only 3 percent, *i.e* by 0.13 percentage points every 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.[33] Note, however, that the falling interview probability need not be driven by duration dependence originating from the job seeker's behavior. Rather it may reflect the firm's reaction to 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 application fixed effects – a proxy for unobserved employability – 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 B7 and B8 in Appendix B.3.1). 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 application fixed effects 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, while the probability of a job offer per interview increases with duration, so it is not clear which effect dominates when considering the chances that a job application results in a job offer. Here, we report estimates of the probability that an application results in a job offer, considering all applications, including those that did not lead to a job interview,

[33] Eriksson and Rooth (2014) report a callback rate of 25 percent, for fictitious job applicants in Sweden. The callback rate is high because applications were sent to high skilled jobs, and applications had excellent fit for the job.

adopting the same empirical strategy as for 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 1.8%, 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 B 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 application fixed effect to the model reduces the impact of unemployment duration by over 40%, bringing it down to 1.1 percentage points (1.1%). A one standard deviation increase in the application fixed effect significantly decreases the probability of receiving a job offer across all applications by 17.4 percentage points, or 16.9%. Again, there is negative dynamic selection of job seekers, *i.e.* job seeker characteristics unobserved (to us, but potentially observed by the firm during an interview) plays a key role in driving the negative duration dependence in 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 application fixed effect 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.

Implications for job search and recruitment. We have used descriptive and reduced-form evidence to document job seekers' application behavior and recruiters' interview and job offer decisions. Here we discuss three main results and their implications for models of search and recruitment, which will guide our theoretical analysis below.

Fact 1 (Drop in individual applications)**.** *Application effort displays negative duration dependence.*

The drop in applications with unemployment duration can be rationalized by a number of different mechanisms. The most promising in our context are discouragement, learning from search, depletion of personal network, and reference-dependent preferences.[34] Discouragement is the response to a declining job offer probability per application over the unemployment spell, that is, the equilibrium effect of firms' discrimination against unemployment duration on the optimal number of applications (Pissarides, 2000). Learning from search occurs in models of incomplete information about own job prospects, where negative dynamic selection makes unemployment duration an informative signal for job seekers about their actual job prospects (Burdett and Vishwanath, 1988; Gonzalez and Shi, 2010; He and Kircher, 2023). Hence, job seekers revise their beliefs about the own job offer probability per application downward as the unemployment spell lengthens, and scale down their applications as a result. Depletion of personal network 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). 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).[35]

[34] Another popular explanation for negative duration dependence in applications is stock-flow matching (Ebrahimy and Shimer, 2010). Stock-flow matching posits that job seekers apply to the stock of suitable vacancies in the first period(s), to then apply to the (smaller) inflow of new vacancies subsequently. This mechanism entails a non-gradual decline in applications with elapsed unemployment duration. We find that applications decreases gradually and linearly over time, which is not in line with the stock-flow matching hypothesis.

[35] Reference dependence has the distinctive prediction that job seekers scale up applications as UI benefits

Fact 2 (Heterogeneity in search strategies)**.** *Application effort displays positive dynamic selection on unobservables.*

While application effort declines strongly within individuals, it remains quite flat on average across observationally equivalent job seekers. Hence, there is strong and positive dynamic selection of job seekers, with individuals making more applications being over-represented at longer durations. It follows that application fixed effects represent a proxy for the job seeker's unobserved *employability*.[36] Controlling for employability in the interview probability equation (Table 3 column (4)) allows us to uncover the potential reasons why individuals differ so systematically in their search strategies. If the chance that each application is considered by a firm ("search efficiency") were constant across workers and search effort were linear in applications (as in standard sequential search models), application fixed effects should not affect the interview probability in any way, since by construction job seekers are observationally equivalent to firms. On the contrary, we detect a negative and strongly significant partial effect for given unemployment duration. Hence, our data unveil heterogeneous search efficiency across job seekers and/or decreasing returns in applications (Gregory, Menzio, and Wiczer, 2021; Lafuente, 2023).

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

For job seekers who are observationally equivalent to a firm, unemployment duration depresses job offer chances on average, but not once they have landed an interview: the duration profile for job applicants is different from that of interviewees.

Fact 3 suggests that dynamic selection is not the only driver of the observed duration profile of the job offer probability per application. Indeed, pure dynamic selection (of the more able workers into jobs) would imply that both the job offer probability per

approach exhaustion, to then decrease again. Since we lack statistical power to document search behavior around UI benefit exhaustion reliably, we are unable to provide a definitive test for reference dependence.

[36] Fact 2 can result from higher-employability job seekers having higher (marginal-utility-weighted) unitary application costs (Lentz and Tranæs, 2005) or applications being strategic substitutes for the job offer probability per application (Galenianos and Kircher, 2009; Mukoyama, Patterson, and Şahin, 2018). A joint reading of Facts 1 and 2 makes the latter explanation less plausible: by the same logic of strategic substitution, job seekers should scale *up* their applications over the unemployment spell as their job offer probability per application reduces, which is inconsistent with Fact 1 (negative duration dependence in applications).

application and per interview are downward sloping controlling for observable characteristics (and flat once controlling for unobservable characteristics, as well). Moreover, the differential duration profile of job offers per application and per interview also rules out basic models of taste-based discrimination (Blanchard and Diamond, 1994), since duration should not play any role once a job seeker is invited to an interview. However, Fact 3 is in line with statistical discrimination against long-term unemployed. On the one hand, the job offer probability per application decreases with duration as the latter is used as a signal of job seeker's ability when the firm decides to call her back to an interview. On the other hand, we prove in Proposition 1 that the job offer probability per interview may increase because interviews become more targeted as unemployment duration lengthens.

These three empirical findings are consistent with models of job search where unemployed workers face duration-dependent costs and/or returns from applications (Fact 1); where job seekers are heterogeneous with respect to their search efficiency (Fact 2); and where firms statistically discriminate against long-term unemployed applicants (Fact 3).

5. A theory of job search under statistical discrimination

In this section, we develop a model of job search when firms statistically discriminate against applicants with longer unemployment durations. The goal of the model is twofold. First, we want to rationalize the three facts of the last section within a consistent theoretical framework. Second, estimating this structural model provides us with a precise decomposition of the empirically observed duration profile of the job finding rate into duration dependence and dynamic selection. Moreover, it also allows us to highlight the extent to which duration dependence is due to job seekers' and recruiters' choices.

On the worker side, we build a new job search model where heterogeneous job seekers, differing in search efficiency, make endogenous application decisions subject to loss-averse preferences, duration-dependent application costs, and incomplete information about individual ability.[37] On the firm side, we consider an extended version of Jarosch and Pilossoph (2019)'s model of statistical discrimination with endogenous job creation.

[37] While the learning model triggered by incomplete information is an original contribution of our paper, our treatment of reference-dependent preferences and duration-dependent application costs follows the existing literature (DellaVigna, Heining, Schmieder, and Trenkle, 2022). Yet, we are the first to study the interaction between such drivers of workers' search behavior and statistical discrimination by firms.

Our theoretical model encompasses various mechanisms that lead workers to search less over an unemployment spell. In particular, we show that learning from search, reference-point adaptation, and depletion of personal network are observationally equivalent mechanisms generating negative duration dependence in applications of low-employability workers.[38] Because our data do not allow us to identify them separately, we estimate three variants of the general model – each featuring just one such mechanisms. The idea is that, if all the mechanisms were at play in reality, the outcome would be a convex combination of the three model variants.[39] Our data are also silent as to whether the main source of statistical discrimination is negative dynamic selection on unobservables (across-individual heterogeneity) or ability depreciation during unemployment (within-individual heterogeneity). Yet, we uncover sizable cross-sectional heterogeneity in applications and the job offer probability per application at each unemployment duration (Figure B11). It follows that across-individual heterogeneity is necessary to replicate the empirical evidence. We therefore model firms' statistical discrimination as driven by negative dynamic selection on (time-constant) unobservables.[40]

Environment. We consider a discrete-time economy populated by a unit mass of workers, who differ in their search efficiency type $\epsilon \sim \mathcal{L}(\epsilon),\ \epsilon \in (\underline{\epsilon}, \overline{\epsilon})$, and a continuum of firms differing in their productivity $y \sim G(y),\ y \in \left(\underline{y}, \overline{y}\right)$.

Search efficiency is an unobservable characteristic measuring how effective a worker is in overcoming meeting frictions: the higher the search efficiency, the fewer applications are needed to meet a vacancy with a given probability. Every time a worker of type

[38]If the drop in applications were mainly motivated by a declining job offer probability per application, high-employability job seekers should scale down their applications the most. Indeed, low-employability job seekers have low chances of being offered a job since the beginning of their spell, so they would be barely affected by firms' statistical discrimination (Jarosch and Pilossoph, 2019). On the contrary, our data show that *low-employability* job seekers scale down their applications the most (Table B12). Hence, the drop in applications cannot be solely driven by discouragement from statistical discrimination.

[39]We think that building a general model of workers' search under statistical discrimination is an important theoretical contribution of our paper. Indeed, with data on job seekers' expectations on job finding prospects, the model would allow future researchers to discipline the learning process. Likewise, monitoring applications before and after UI benefit exhaustion would allow identifying reference dependence.

[40]Direct evidence on skill decpreciation is scarce, mainly because panel data tracking individuals' skills during unemployment are usually not available. An exception is the recent paper by Cohen et al. (2023), who use a unique data set with information on skills for a large sample of German workers at the onset of unemployment and three additional times thereafter. They do not find a decline in cognitive and noncognitive skills, even though indicators of depression and loneliness rise substantially.

ϵ separates from a job, nature draws a new ability x from an exogenous distribution $\mathcal{H}(x|\epsilon, \tau = 0)$, where $\tau \in \mathbb{N}$ stands for elapsed unemployment duration, and $\partial \mathbb{E}[x|\epsilon]/\partial \epsilon > 0$.[41] Information is incomplete: workers do not observe their own ability draw. However, the underlying distribution $\mathcal{H}(x|\epsilon, \tau = 0)$ is common knowledge.

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 rate δ_H and the endogenously determined job finding rate $f(\epsilon, \tau, x)$. The exogenous separation rate 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.* $s(\epsilon, \tau) \equiv \epsilon a(\epsilon, \tau)^\chi$, with $\chi > 0$.[42] A job seeker's job finding chances are higher either if she makes more applications (higher a) or *better* applications (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 type-specific and duration-dependent application cost function $\sigma(a; \epsilon, \tau)$, $\sigma_a > 0$, $\sigma_{aa} > 0$, $\sigma_\epsilon > 0$, $\sigma_\tau > 0$ (Pissarides, 2000; DellaVigna et al., 2022). Type-specific costs capture invariant heterogeneity with unemployment duration (*e.g.* in permanent income or value of leisure), while duration-dependent application costs catch depletion of personal network in reduced-form.[43]

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

[41] 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.

[42] 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.

[43] Type-specific application costs are needed to generate positive dynamic selection in applications (Fact 2). In Appendix C.1 we provide two potential microfoundations for application costs being increasing in search efficiency, based on a positive correlation between search efficiency and permanent income (*wealth effect*) or the value of social leisure (*time allocation effect*).

with probability $s\lambda(\theta)$, where $\lambda(\theta) \equiv \frac{\mathcal{M}(\mathcal{S},\mathcal{V})}{S} = \mathcal{M}(1,\theta)$ and $\theta \equiv \frac{V}{S}$ represents labor market tightness. Upon meeting, the only relevant information released to firms 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 true ability x and decides whether to offer her a job.

Match output is governed by a production technology $p(x, y)$ characterized by positive assortative matching, *i.e.* the most productive firms are the most selective in terms of workers' ability:[44]

$$p(x,y) = \begin{cases} x+y & \text{if } x \geq y, \\ 0 & \text{else.} \end{cases} \tag{3}$$

A worker is thus qualified for a job if her ability x exceeds the firm's productivity y, meaning that higher-ability job seekers enjoy a higher job offer probability per unit of search effort. For any (x, y) pair, let $\mathcal{Q}$ be a qualification indicator such that $\mathcal{Q}(x, y) = \mathbb{1}\{x \geq y\}$. 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.[45]

Workers. Workers have linear and reference-dependent preferences over consumption represented by the following utility function:

$$u(c_t; r_t) = \begin{cases} c_t + \Upsilon(c_t - r_t) & \text{if } c_t < r_t, \\ c_t & \text{if } c_t \geq r_t, \end{cases} \tag{4}$$

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

[44]We adopt the modified Albrecht and Vroman (2002)'s production function proposed by Jarosch and Pilossoph (2019) as it grants an intuitive notion of a worker's qualification for a job, on top of being consistent with the production function estimation of Lise and Robin (2017). Our results extend to any alternative specification giving rise to positive assortative matching.

[45]The assumption of rigid wages allows us to focus on sources of duration dependence unrelated to changes in the individual reservation wage, as well as to simplify the model significantly. However, our results would go through more sophisticated wage setting protocols giving rise to compressed wage structures, *i.e.* as long as reservation wages do not adjust so much that firms are indifferent between workers of different abilities, or prefer lower-ability ones.

point be the average consumption over the $\tilde{\tau}$ previous periods, *i.e.* $r_t = \frac{1}{\tilde{\tau}} \sum_{j=1}^{\tilde{\tau}} c_{t-j}$.[46]

Workers 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:

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

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

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

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 expected job finding rate weighted by the discounted capital gain upon employment:

$$a(\epsilon,\tau) : \frac{\partial\sigma(a;\epsilon,\tau)}{\partial a} = \beta\;\frac{\partial s(a,\epsilon)}{\partial a}\hat{o}(\epsilon,\tau)\;\Big[W(\epsilon) - U(\epsilon,\tau+1)\Big]. \tag{5}$$

Notice that job seekers with higher search efficiency have both higher marginal benefit (because search effort is super-modular in search efficiency and application effort, and the expected job offer probability per unit of search effort is increasing in search efficiency via ability) and higher marginal cost of exerting application effort (because each unit of application effort is more costly).

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 $\lambda(\theta)/\theta$. The value of a filled job is given by the present discounted value of flow profits, *i.e.* $J(x,y) = \frac{p(x,y)-\omega}{1-\beta(1-\delta_H)}$.

[46] 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.

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 $\mathcal{C}(y, \tau) = \mathbb{1}\left\{\int \max\left\{J(x, y), 0\right\} \mu(x|\tau)\ dx \geq \kappa\right\}$, 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: $c(\tau) = \lambda(\theta) \int \mathcal{C}(y, \tau)\ dG(y)$. It follows that the *interview rate*, the probability that a job seeker exerting optimal search effort receives an interview at duration τ, equals $c(\epsilon, \tau, x) = c(\epsilon, \tau) = s(\epsilon, \tau)\ c(\tau)$.

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 (3), regardless of unemployment duration. For any (x, y, τ) triple, denote a job offer indicator as $\mathcal{O}(x, y, \tau) = \mathcal{C}(y, \tau)\mathcal{Q}(x, y)$. In words, a firm of productivity y which meets a job seeker of ability x at duration τ offers her a job if her unemployment duration makes it profitable to interview her *and* she is qualified for its production technology. On the job seeker's side, this implies that the job offer probability per interview equals: $o|c(x, \tau) = \frac{\int \mathcal{O}(x,y,\tau)\ dG(y)}{\int \mathcal{C}(y,\tau)\ dG(y)}$. In turn, the job offer probability per unit of search effort for a job seeker of ability x at duration τ is given by:[47]

$$o(x, \tau) \equiv c(\tau)\ o|c(x, \tau) = \lambda(\theta) \int \mathcal{O}(x, y, \tau)\ dG(y). \tag{6}$$

Finally, the individual *job finding rate* at duration τ reads:

$$f(\epsilon, \tau, x) = s(\epsilon, \tau)\ o(x, \tau). \tag{7}$$

Stationary equilibrium. Closing the model requires to specify the equilibrium conditions for the measure of unemployed of each type and duration. To do so, we solve

[47]Absent statistical discrimination, *i.e.* if $\kappa = 0$, the job offer probability per unit of search effort would read $o^{ND}(x) = \lambda(\theta) \int \mathcal{Q}(x, y)\ dG(y)$. Contrasting it with equation (6), we notice that statistical discrimination by firm y affects a job seeker's job offer probability per unit of search effort if and only if $\mathcal{Q}(x, y) = 1$, that is, if the job seeker is denied an interview for a job she would have been qualified for.

the model in stationary equilibrium by imposing balance of flows:[48]

$$u(\epsilon,\tau) = \begin{cases} \delta_L\big(1-\sum_{t=0}^{\infty} u(\epsilon,t)\big) & \text{if } \tau = 0, \\ u(\epsilon,\tau-1)\big[1-f(\epsilon,\tau-1)\big] & \text{if } \tau > 0, \end{cases} \tag{8}$$

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

The key equilibrium objects of the model are the belief function about job seeker's ability, $\hat{h}(x|\epsilon,\tau) = \hat{\mathcal{H}}'(x|\epsilon,\tau)$, which drives job seekers' application decisions, and the search-effort-weighted density of job seekers' ability at each duration, $\mu(x|\tau)$, which drives firms' callback decisions. For given $\hat{h}(x|\epsilon,0) = h(x|\epsilon,0)$, the belief function about job seeker's ability evolves according to Bayesian updating:

$$\hat{h}(x|\epsilon,\tau) = \frac{\big(1-f(\epsilon,\tau,x)\big)\,\hat{h}(x|\epsilon,\tau-1)}{\int \big(1-f(\epsilon,\tau,x)\big)\; d\hat{\mathcal{H}}(x|\epsilon,\tau-1)} \quad \forall \tau > 0. \tag{9}$$

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. In equilibrium, job seekers' belief function about own ability equals the type-specific ability distribution at each duration, *i.e.* $\mathcal{H}(x|\epsilon,\tau) = \hat{\mathcal{H}}(x|\epsilon,\tau)$. The search-effort-weighted density of job seeker's ability at each duration reads:

$$\mu(x|\tau) = \frac{\int s(\epsilon,\tau)\; u(\epsilon,\tau)\; h(x|\epsilon,\tau)\; d\mathcal{L}(\epsilon)}{\int s(\epsilon,\tau)\; u(\epsilon,\tau)\; d\mathcal{L}(\epsilon)}. \tag{10}$$

Definition 1. *A stationary equilibrium of the economy is a tuple* $\big\{a(\epsilon,\tau), o(x,\tau), \hat{h}(x|\epsilon,\tau), u(\epsilon,\tau), \theta\big\}$, *where application effort satisfies equation (5), the job offer probability per unit of search effort satisfies equation (6), the belief function satisfies equation (9) for given* $\mathcal{H}(x|\epsilon,0)$, *the unemployment rate satisfies equation (8), and the labor market tightness is pinned down by equation (C.1) (free entry condition).*

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.

Upon meeting a job seeker with unemployment duration τ, firms form an expectation

[48]Intuitively, the stationary measure of unemployed at $\tau = 0$ equals the measure of employed that separate from their employer. In turn, the stationary measure of unemployed at longer durations equals the share of unemployed who have not found a job in the previous period.

about her ability based on $\mu(x|\tau)$. Since job seekers with high ability x match more easily according to the production technology (3), 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.[49] 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 screening device when choosing whether to call back a job seeker for an interview, some job seekers are denied interviews by firms they would 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, thus making firms' callbacks more targeted. This induces positive duration dependence in the job offer probability per interview.[50] To see why this is the case, suppose that the pool of job seekers resembles the population ability distribution at short unemployment duration. All firms call back any job seeker at that duration and reject the unqualified ones at the interview stage. Then, the job offer probability per interview approaches the probability that a worker is qualified for a job. Now suppose that the pool of job seekers is composed almost entirely by low-ability workers at long

[49]Notice that the right model counterpart of the job offer probability per application analyzed in Figure 4C is per unit of *application* effort, *i.e.* $\epsilon a(\epsilon,\tau)^{\chi-1}o(x,\tau)$. If $\chi \to 1$, the individual duration profiles of the two variables coincide.

[50]Even if the job offer probability per interview exhibits positive duration dependence, the average job offer probability per interview does not necessarily increases with duration. This is because negative dynamic selection in ability tilts the composition of the pool of job seekers towards those who have lower job offer probability per interview to start with (because they are qualified for fewer firms). Hence, our model replicates Fact 3 when positive duration dependence in the job offer probability per interview, due to better interview targeting at longer durations, outweighs the compositional change in the pool of job seekers, due to negative dynamic selection in ability.

unemployment duration. Most of the firms find it unprofitable to call back a job seeker at such duration. As a result, only firms that are willing to hire a low-ability worker call back job seekers at that duration. Therefore, the job offer probability per interview approaches 1. In Appendix Table B11 we provide empirical validation to this mechanism by documenting that the positive duration dependence in the job offer probability per interview is driven by low-employability job seekers.

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.2.

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.

In equilibrium, 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 according to equation (5). Proposition 2 reports a sufficient condition on the path of the individual expected job offer probability per unit of search effort for application effort to exhibit negative duration dependence under statistical discrimination.

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

Proposition 2 develops a general condition that holds independently of the specific worker's search behavior.[51] Indeed, our model rationalizes Fact 1 through four – complementary

[51]The sufficient condition in Proposition 2 makes sure that the total reduction in the expected job offer probability per unit of search effort happens smoothly over the unemployment spell. Indeed, if it followed a step-like process, job seekers would find it optimal to scale *up* their application effort during

or alternative – channels. First, for given belief about own ability, firms' statistical discrimination induces negative duration dependence in the *true* job offer probability per unit of search effort (*discouragement*). Second, for given job offer probability per unit of search effort per ability level, negative dynamic selection induces job seekers to revise their beliefs about own ability downward, which results in a lower expected job offer probability per unit of search effort due to changing probability weights attached to each ability level (*learning*). Third, marginal application costs increase with duration (*depletion of personal network*). Finally, 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*).

Since job seekers with higher search efficiency are, on average, of higher ability as well, average search efficiency decreases with duration. Recall from Equation (5) that whether or not job seekers with higher search efficiency exert higher application effort is a priori ambiguous. Proposition 3 provides a sufficient condition for application effort to be decreasing in search efficiency, which entails that application effort displays positive dynamic selection.

Proposition 3. *Let ζ be the elasticity of application costs with respect to search efficiency. If expected ability declines with duration, i.e. $d\mathbb{E}[x|\tau]/d\tau < 0$, and the elasticity is high enough such that $\zeta > 1 + \frac{\partial \ln(\hat{o}(\epsilon,\tau)[W(\epsilon)-U(\epsilon,\tau+1)])}{\partial \ln(\epsilon)}$ $\forall \tau, \epsilon$, then application effort exhibits positive dynamic selection, i.e. $\mathbb{E}_0[a(\epsilon,\tau)] < \mathbb{E}_\tau[a(\epsilon,\tau)]$ $\forall \tau > 0$, where $\mathbb{E}_t[.]$ denotes the expectation with respect to the distribution of search efficiency ϵ at duration t.*

Hence, our model rationalizes Fact 2 through higher-ability job seekers having higher marginal application costs, *e.g.* because they are wealthier.

6. Quantitative analysis

To make the model amenable for quantification, we enrich the framework outlined in the previous section with two additional components. First, following Blanchard and Diamond (1994) and Shimer (2005a), we allow for coordination frictions in the form of multiple job seekers per vacancy. Coordination frictions are a standard assumption in

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.

the existing literature as, in their presence, firms need to sort potentially multiple job seekers. As long as firms sort job seekers by unemployment duration (interviewing those with shorter duration first), coordination frictions induce negative duration dependence in the interview probability.[52] We introduce coordination frictions to smooth out the duration profile of the job offer probability per unit of search effort for given ability, which makes sure that application effort is monotonically decreasing in unemployment duration as per Proposition 2. Second, we assume that qualified job seekers get offered a job after an interview with probability $q \in (0, 1)$. This assumption catches idiosyncratic matching frictions as in models of stochastic match quality (Pissarides, 2000; Menzio and Shi, 2011; Wright, Kircher, Julien, and Guerrieri, 2021) and allows us to replicate the scale of the job offer probability per interview observed in the data. Appendix C.3 develops the extended model.

Functional forms. We assume that worker ability x and firm productivity y lie in the unit interval, *i.e.* $\text{supp}(x) = \text{supp}(y) = [0, 1]$. Worker search efficiency and firm productivity follow flexible (shifted) Beta distributions. Formally, $\epsilon \sim \mathcal{L}(\epsilon) = 1 + \phi\text{Beta}(B_1, B_2)$, where $\text{supp}(\epsilon) = [1, 1 + \phi]$, and $y \sim G(y) = \text{Beta}(G_1, G_2)$.

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

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. Following the literature, we adopt an iso-elastic application cost function, *i.e.* $\sigma(a; \epsilon, \tau) = \psi(\epsilon, \tau)\frac{a^{1+\eta}}{1+\eta}$, that is increasing and convex ($\eta > 0$). In turn, we assume that the intercept of the application cost function is iso-elastic

[52] Unlike in models of taste-based discrimination such as Blanchard and Diamond (1994), in our model job seekers with different unemployment duration are on average not equally productive for firms, due to negative dynamic selection in job seekers' ability. As a result, coordination frictions do not give rise to an additional source of duration dependence in the interview probability but simply amplify the impact of statistical discrimination.

in search efficiency and increasing in unemployment duration: $\psi(\epsilon,\tau) = \epsilon^{\zeta}\psi_0\left(1+\tau\psi_1\right)$.

Structural estimation. We estimate the structural model at monthly frequency for unemployment duration $\tau = 0,\ldots,\tilde{\tau}$. We set the grid size 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 from external sources. Then, we estimate the remaining moments internally via indirect inference. Since our empirical findings do not allow us to separately identify learning, reference dependence and duration-dependent application costs, we estimate three variants of the general model, each featuring just one such mechanisms.

Table 4: Externally chosen parameters

Parameter	Description	Value	Target/Source
β	Discount factor	0.996	5% annual interest rate in Davis and von Wachter (2011)
δ_L	Separation rate (workers)	0.009	Monthly EU rate
δ_H	Separation rate (firms)	0.019	Monthly EE+EU rate
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)
κ	Interview cost per hire	0.090	Interview costs in Muehlemann and Strupler Leiser (2018)

Numeraire: cross-sectional avg monthly output.

Table 4 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 rates from the EU rate and EE rate 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. To calibrate interview costs, we resort to administrative and representative Swiss survey data at the establishment-level, breaking down firms' hiring costs into their specific components (Muehlemann and Strupler Leiser, 2018). According to the data, average interview costs per hire amounts to to 8.97% of average monthly output.[53]

[53] The figure results from average monetary interview costs per hire of 519 CHF in the face of an average weekly wage of $1,311.21$ CHF.

Table 5: Estimated parameters

Parameter	Description	Target	Value		
			Learning	Loss aversion	DD costs
B_1	1^{st} shape param. Beta distr. search eff.	$\hat{\beta}_{\ln c(\epsilon,\tau,x),\tau}$: duration effect interview rate, residual (obs.)	0.113	0.095	0.086
B_2	2^{nd} shape param. Beta distr. search eff.	$\mathbb{E}_\tau[c(\epsilon,\tilde{\tau},x))]$: long-term avg interview rate	0.498	0.467	0.424
G_1	1^{st} shape param. Beta distr. prod.	$\hat{\beta}_{\ln f(\epsilon,\tau,x),\tau}$: duration effect job finding rate, residual (obs.)	0.192	0.159	0.162
G_2	2^{nd} shape param. Beta distr. prod.	$\mathbb{E}_\tau[f(\epsilon,\tilde{\tau},x))]$: long-term avg job finding rate	0.550	0.687	0.710
ξ	Subst. param. meeting function	$\mathbb{E}_\tau[c(\epsilon,\tau,x))]$: avg interview rate	0.190	0.179	0.164
η	Convexity app. effort cost	$\hat{\beta}_{\ln a(\epsilon,\tau),\tau\mid\epsilon}$: duration effect applications, residual (FE)	0.239	0.268	0.257
ψ_0	Scalar app. effort cost	$\mathbb{E}_\tau[a(\epsilon,\tau)]$: avg applications	0.019	0.014	0.015
ϕ	Search efficiency dispersion param.	$\sigma(\alpha(\epsilon))$: std. dev. application fixed effects	10.46	11.42	12.85
κ_v	Vacancy posting cost	$\mathbb{E}_\tau[f(\epsilon,\tau,x))]$: avg job finding rate	0.007	0.004	0.002
q	Cond. job offer prob. qualified job seeker	$\mathbb{E}_\tau[a(\epsilon,\tilde{\tau})]$: long-term avg applications	0.398	0.380	0.364
χ	App. effort elasticity search effort	$\hat{\beta}_{\ln[c(\epsilon,\tau,x)/a(\epsilon,\tau)],\alpha(\epsilon)\mid\tau}$: partial effect app FE on interview prob.	0.999	0.954	0.986
ζ	Search eff. elasticity app. costs	$\hat{\beta}_{\ln a(\epsilon,\tau),\tau}$: duration effect applications, residual (obs.)	1.143	1.328	1.268
ρ	Equally ranked ability-eff. correlation	$\hat{\beta}_{\ln a(\epsilon,\tau),\tau\mid\epsilon,\alpha(\epsilon)\geq \text{med}[\alpha]}$: duration effect applications (app FEs above median), residual (FE)	0.530	1.000	1.000
Υ	Loss aversion coefficient		0.000	0.400	0.000
ψ_1	Duration dependence app. costs		0.000	0.000	0.006
Θ	Loss function		0.037	0.034	0.026

Note: All duration effects are computed from a linear model and expressed as semi-elasticities, *i.e.* the duration coefficient is normalized by the average variable at $\tau = 0$. All averages are computed with respect to the distribution of observables at $\tau = 0$. Application fixed effects are not standardized. Numeraire: cross-sectional avg monthly output.

We then estimate the remaining set of parameters via indirect inference through the simulated method of moments. Each such parameters conceptually relates to some moment in the data through the equilibrium conditions of the model. Formally, let Θ be the vector of parameters still to be determined: $\Theta = \Big\{B_1, B_2, G_1, G_2, \xi, \eta, \psi_0, \phi, \kappa_v, q, \chi, \zeta, \rho, \Upsilon, \psi_1\Big\}$. We choose parameter values that minimize the sum of weighted squared percentage deviations between a set of empirical moments (μ) and model-generated moments ($\hat{\mu}$):

$$\Theta^* = \arg\min_{\Theta\in\mathcal{P}} \sum_{m\in\mathcal{M}} w_m \left(\frac{\hat{\mu}_m(\Theta) - \mu_m}{\mu_m}\right)^2,$$

where $\mathcal{P}$ denotes the parameter space, $\mathcal{M}$ the set of targeted moments, and w some weighting factor. Table 5 reports the internally chosen parameters, along with the respective targeted moments. We make use of the cross-sectional properties and duration profiles of the interview rate, job finding rate and application effort from our search diary data to identify the model parameters. Since job seekers of low ability are barely affected by firms' statistical discrimination, we let the duration profile of application effort of the low-employability job seekers (with application fixed effects above median) inform the distinctive parameter of the three model variants – namely, the correlation between ability and search efficiency ρ, loss aversion Υ, and duration dependence of application costs ψ_1. In Appendix C.5 we explain the rationale behind our choice of the targeted moments and comment our estimation results.

Model fit. All the three variants of the estimated model are able to closely replicate the duration profiles of the outcome variables and the observed workers' flows.

On the workers' side, the estimated models generate duration profiles of applications in line with the empirical ones. Figure 5A compares the duration profile of applications controlling for observables in the data with that of average application effort in the three model variants, $\mathbb{E}_\tau[a(\epsilon, \tau)]$. Figure 5B compares the duration profile of applications controlling for individual fixed effects in the data with that predicted by the three model variants when the composition of the unemployment pool is kept constant, $\mathbb{E}_0[a(\epsilon, \tau)]$. Comparison across panels reveals that, quantitatively, the divergence between the two duration profiles is slightly lower in the models than in the data, though the discrepancy is small.[54] The model with duration-dependent application costs (dash-dotted green line) provides the best approximation to the duration profiles of applications. Importantly, all the model variants deliver the positive dynamic selection in job applications highlighted in Figure 3B (see Figure C6 for the model counterpart). In the model, job seekers with higher search efficiency exert lower application effort in equilibrium. Since ability and search efficiency are positively correlated, job seekers who exert higher application effort at every duration are therefore more likely to experience longer unemployment spells.

On the firms' side, the empirical duration profiles of the interview rate and job finding rate (controlling for observables) are replicated accurately by all the model variants. Figure 6A displays the average interview rate, $\mathbb{E}_\tau[c(\epsilon, \tau, x)]$, while Figure 6B shows the average job finding rate, $\mathbb{E}_\tau[f(\epsilon, \tau, x)]$. The learning and reference-dependence model (solid red line and dashed blue line) are equally able to match the empirical duration profiles accurately, whereas the model with duration-dependent application costs (dot-dashed green line) misses the convex behavior of the interview rate. Notice that the duration profile of the interview rate is somewhat steeper than that of the job finding rate. It follows that the average individual job offer probability per interview increases slightly with duration, as observed in our data (Figure 4B). This means that, quantitatively, negative dynamic selection on unobservables does not fully offset the positive duration

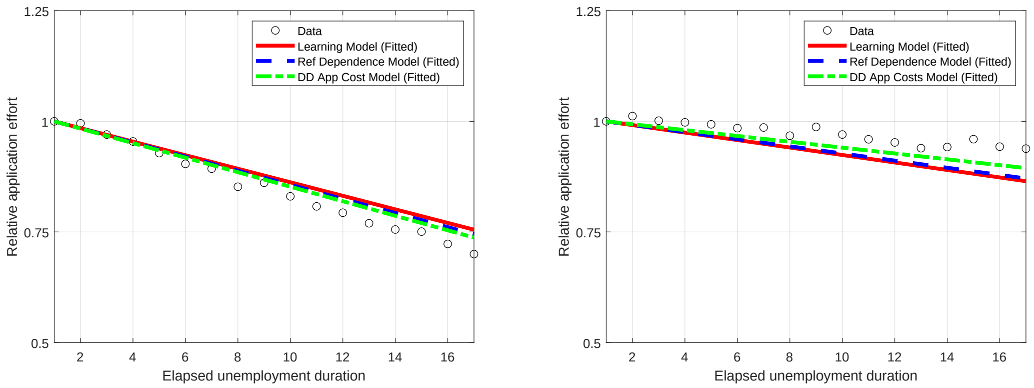
[54] The small discrepancy between the empirical and model-implied duration profiles of application effort is likely a product of the time lag between application records and the corresponding job finding dates.

dependence – established in Proposition 1 – in the individual job offer probability per interview (see Figure C4 for the quantitative fit).

Notably, our estimated models are able to replicate all the duration profiles not only in relative terms but also in levels (see Table C1-Table C3). It follows that the pace of dynamic selection – the driver of statistical discrimination, learning and compositional changes – is virtually the same in the model and in the data, as governed by the observed job finding rate and separation rate.

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

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



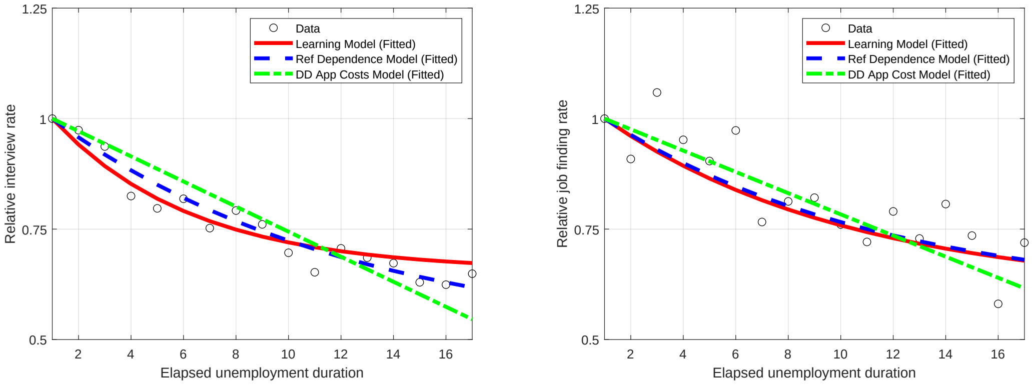
Note: This figure contrasts the duration profiles controlling for individual fixed effects (Panel A) and for observables (Panel B) of application effort in the data (circles) with those implied by the estimated models. The learning model is depicted in solid red, the reference dependence model in dashed blue and the duration-dependent application costs model in dotted-dashed green. The duration profiles in the model are derived by estimating individual fixed effects and duration effects from a saturated regression, computing the expected values of application effort at any unemployment duration, and normalizing them with respect to the first month of unemployment. For the duration profile controlling for individual fixed effects, expected values are computed with respect to workers' search efficiency distribution in the first month of unemployment, *i.e.* $\mathbb{E}_0[a(\epsilon, \tau)]$; for the duration profile controlling for observables, expected values are computed with respect to workers' search efficiency distribution in the contemporaneous period of unemployment, *i.e.* $\mathbb{E}_\tau[a(\epsilon, \tau)]$. The distribution of observables across unemployment durations is kept the same as in the first month of unemployment in both specifications. Finally, the duration profiles are fitted by a linear function.

Duration dependence versus dynamic selection. We now use our model as an accounting framework to break down the decrease of the observed job finding rate into duration dependence and dynamic selection. In turn, our model allows us to separate duration dependence due to workers from that due to firms.

On the workers' side, the estimated models allow us to map observed application effort into the relevant notion of search effort for the sake of job finding, *i.e.* $s(\epsilon, \tau) = \epsilon a(\epsilon, \tau)^\chi$. This has two important implications. First, the negative dynamic selection in search efficiency we estimate turns out to outweigh the positive dynamic selection in application effort we observe in the data. Hence, search effort inherits negative duration dependence from application effort, but the direction of dynamic selection is flipped with respect to the latter (see Figure C5). Second, our estimates of the decreasing returns in application effort are consistently small across model variants. Hence, negative duration dependence

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

(A) Interview rate, residual (obs.) (B) Job finding rate, residual (obs.)



Note: This figure contrasts the duration profiles controlling for observables of the interview rate (Panel A) and job finding rate (Panel B) detected in the data (circles) with those implied by the estimated models. The learning model is depicted in solid red, the reference dependence model in dashed blue and the duration-dependent application costs model in dotted-dashed green. The duration profiles in the model are derived by estimating individual fixed effects and duration effects from a saturated regression, computing the expected values of the interview rate and job finding rate at any unemployment duration, and normalizing them with respect to the first month of unemployment. Expected values are computed with respect to the joint distribution of workers' search efficiency and ability in the contemporaneous period of unemployment, *i.e.* $\mathbb{E}_\tau[c(\epsilon,\tau,x)]$ and $\mathbb{E}_\tau[f(\epsilon,\tau,x)]$.The distribution of observables across unemployment durations is kept the same as in the first month of unemployment. Finally, the model-implied duration profiles are fitted by a negative exponential function estimated via weighted nonlinear least squares.

in application effort gets transmitted to search effort with an almost unitary elasticity (specifically, the estimated concavity parameter χ ranges between 0.95 and 1 across model variants).

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

$$\begin{aligned}\underbrace{\mathbb{E}_\tau\left[f(\epsilon,\tau,x)\right]-\mathbb{E}_0[f(\epsilon,0,x)]}_{\text{Duration profile controlling for obs.}} = {} & \underbrace{\mathbb{E}_\tau\left[s(\epsilon,0)\ \big(o(x,\tau)-o(x,0)\big)\right]}_{\text{DD due to firms}}+\underbrace{\mathbb{E}_\tau\left[\big(s(\epsilon,\tau)-s(\epsilon,0)\big)\ o(x,\tau)\right]}_{\text{DD due to workers}}\\ & +\underbrace{\mathbb{E}_\tau\left[s(\epsilon,0)\ o(x,0)\right]-\mathbb{E}_0\left[s(\epsilon,0)\ o(x,0)\right]}_{\text{Dynamic selection on unobservables}}, \qquad (11)\end{aligned}$$

where $\mathbb{E}_t[.]$ denotes the expectation with respect to the distribution of workers' unobservable characteristics, *i.e.* type ϵ and ability 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 a 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.

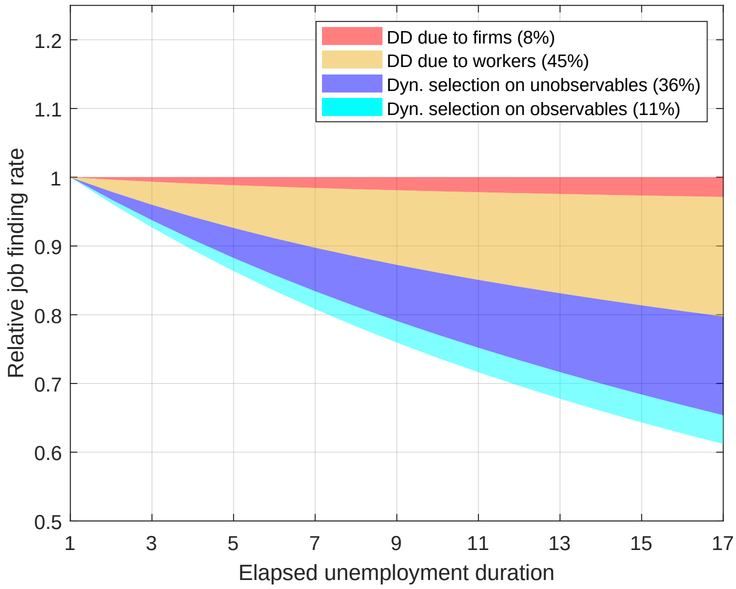
We notice that our model assumes that workers are homogeneous in terms of observable characteristics in a given labor market. Accordingly, when estimating the model, we target the duration profile of the job finding rate controlling for observables. This amounts to positing that the distribution of observables at any unemployment duration is the same as in the first month of unemployment. Let X be a vector of observable characteristics. Hence, $\mathbb{E}_t[f(\epsilon,\tau,x)] \equiv \tilde{\mathbb{E}}_0\left[\mathbb{E}_t[f(\epsilon,\tau,x)|X]\right]$, 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 Figure B10. The observed duration profile of the job finding rate can be decomposed as follows:

$$\begin{aligned}\underbrace{\hat{\mathbb{E}}_\tau\left[f(\epsilon,\tau,x,X)\right]-\hat{\mathbb{E}}_0\left[f(\epsilon,0,x,X)\right]}_{\text{Observed duration profile}} = & \underbrace{\mathbb{E}_\tau\left[f(\epsilon,\tau,x)\right]-\mathbb{E}_0\left[f(\epsilon,0,x)\right]}_{\text{Duration profile controlling for obs.}} \\ & +\underbrace{\hat{\mathbb{E}}_\tau\left[f(\epsilon,\tau,x,X)\right]-\mathbb{E}_\tau\left[f(\epsilon,\tau,x)\right]}_{\text{Dynamic selection on observables}},\end{aligned} \tag{12}$$

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

We perform the decomposition for each model variant in Appendix Figure C7. Figure 7 reports the weighted average decomposition across variants graphically. The weight attached to each model variant equals the corresponding inverse loss function in the SMM estimation – a measure of goodness of model fit to the data moments. On average, our estimated model attributes 53% of the observed decline of the job finding rate to duration

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 (11) and equation (12). For each model variant, the duration profiles of the components of the job finding rate reported in equation (11) are derived by estimating individual fixed effects and duration effects from a saturated regression, computing the expected values of each component at any unemployment duration, and normalizing them with respect to the first month of unemployment. Expected values are computed with respect to the joint distribution of workers' search efficiency and ability 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 (12), 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 B10). 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. Finally, the reported shares equal the weighted average shares across the three model variants, where the weight attached to each model variant equals the inverse loss function in the SMM estimation (last row of Table 5).

dependence and 47% to dynamic selection.[55] More specifically, duration dependence is mainly driven by workers' search behavior, which accounts on average for 45% of the observed decline of the job finding rate, and largely outweighs the role of firms' hiring behavior (8%). Dynamic selection happens primarily on unobservables, which accounts for 36% of the observed decline, while the role of observables is more muted (11%).

Our results show that duration dependence is at least as important as dynamic selection in explaining the observed decline in the job finding rate. This is different from the estimates of Mueller, Spinnewijn, and Topa (2021) and Mueller and Spinnewijn (2023), who attribute most of the falling job finding rate to a change in the composition of the unemployment pool. Our results are more in line with the estimates of Kroft et al. (2016) and Alvarez, Borovičková, and Shimer (2023), who find that structural duration dependence is of first-order importance.[56]

[55] The estimated duration dependence share ranges from 45% in the model with duration-dependent application costs to 59% in the reference-dependence model.

[56] Mueller and Spinnewijn (2023) estimate in Swedish administrative data that dynamic selection on observables explains (at least) 49% of the observed decline of the 6-month job finding rate in the first

Breaking down the sources of duration dependence, we find that duration dependence is by and largely driven by workers' sizable reduction in application effort over an unemployment spell. On the one hand, our model-based assessment of the (limited) importance of firms' statistical discrimination is in line with Jarosch and Pilossoph (2019)'s. On the other hand, we claim that ignoring endogenous search effort by workers leads to underestimate the role of statistical discrimination in driving duration dependence in the equilibrium job finding process. Indeed, we estimate that the general-equilibrium elasticity of the job finding rate with respect to the job offer probability per unit of search effort is about 3, owing to the induced workers' discouragement (with exogenous search effort it would be just 1).[57] It follows that statistical discrimination mainly determines duration dependence in the job finding rate via its indirect effect on workers' application effort.[58] These results shows the importance of jointly analyzing workers' search behavior and firms' hiring choices to explain the duration dependence in the job finding rate.

7. 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 increase with duration. We propose a new theoretical framework with duration-dependent search by workers and statistical discrimination against the long-term unemployed by firms, which matches these findings closely.

Our data set is unique in allowing us to track how applications, interviews and job offers change with duration. Together with our theoretical framework, our analysis generates interesting new insights on how workers' search behavior and employers' recruitment policies interact – and its implications for the relative importance of duration dependence and dynamic selection.

6 months and 36% in the successive 6 months. However, note that Mueller and Spinnewijn (2023) focuses on the cumulative job finding rate over a 6-month horizon, while we provide a decomposition of the monthly job finding rate, so the two empirical approaches differ substantially with respect to time aggregation. Alvarez, Borovičková, and Shimer (2023) estimates, on Austrian social security data, that dynamic selection (duration dependence) is the main driver of the duration profile of the job finding rate in (after) the first 20 weeks of unemployment.

[57]See Appendix C.6 for details on the computation of the general-equilibrium elasticity.

[58]We validate this statement in Appendix C.6 by running the counterfactual exercise of setting the interview cost to zero, while adjusting the vacancy posting costs to keep the mass of vacancies fixed.

On the one hand, we find that duration dependence is mainly driven by a strong within-individual decline in job applications. As we highlight in our theoretical framework, one potentially important mechanism is job seekers' reaction to statistical discrimination by firms against long-term unemployed applicants. The corresponding lower return from search discourages applications at longer durations. Our quantitative analysis shows that this discouragement effect is quantitatively important.

On the other hand, our empirical analysis reveals interesting patterns of dynamic selection. As duration progresses, the unemployment pool does not only consist of less employable workers, but also of workers who send disproportionately more applications at any duration. This negative correlation points to the potential importance of heterogeneity across job seekers in search efficiency, in line with recent business cycle research (Gregory, Menzio, and Wiczer, 2021; Lafuente, 2023). Indeed, our quantitative analysis shows that accounting for heterogeneity in search efficiency is important to match the duration patterns we observe in our Swiss search diary data.

An important caveat of our analysis is that it abstracts from depreciation of human capital during an unemployment spell. Lack of relevant information in our search diary data prevents us from exploring this channel in more detail. In our theoretical analysis, any heterogeneity in unobserved ability is across individuals, while ability is not allowed to change within individuals over time. In the light of the vast cross-sectional heterogeneity that is needed to match our data, we expect that within-individual ability depreciation during unemployment would not change our results significantly.[59]

While our analysis points to search behavior as an important driver of duration dependence, the precise reason why job seekers decrease applications at longer durations is less clear. Our analysis shows that discouragement, reference-dependent search behavior, learning from unsuccessful search outcomes, and lower net benefits from search (due to depletion of one's personal network) are all consistent with the data. We cannot study this further because we lack the necessary information in our data. However, we think that exploring the relative importance of these channels in more detail is an interesting direction for future research.

[59]Consistent with this conjecture, adding ability depreciation has little bearing on interview choices of employers in the model of Jarosch and Pilossoph (2019).

References

Abbring, J. H. and Van den Berg, G. J. The nonparametric identification of treatment effects in duration models. *Econometrica*, 71(5):1491–1517, 2003.

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.

Albrecht, J. and Vroman, S. A matching model with endogenous skill requirements. *International Economic Review*, 43(1):283–305, 2002.

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`.

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.

Barron, J., Berger, M., and Black, D. How well do we measure training? *Journal of Labor Economics*, 15(3):507–28, 1997.

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.

Behncke, S., Frölich, M., and Lechner, M. A caseworker like me - Does the similarity between the unemployed and their caseworkers increase job placements? *The Economic Journal*, 120(549):1430–1459, 2010.

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.

Birinci, S., See, K., and Wee, S. L. Job Applications and Labour Market Flows. Staff Working Papers 21-49, Bank of Canada, October 2023.

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.

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`.

Choudhury, J. B. Non-parametric confidence interval estimation for competing risks analysis: application to contraceptive data. *Statistics in medicine*, 21(8):1129–1144, 2002.

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.

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.

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.

Elbers, C. and Ridder, G. True and spurious duration dependence: The identifiability of the proportional hazard model. *The Review of Economic Studies*, 49(3):403–409, 1982.

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., Huffman, D., and Sunde, U. Self-confidence and search. IZA Discussion Papers 2525, Institute of Labor Economics (IZA), 2006.

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.

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.

Gavazza, A., Mongey, S., and Violante, G. L. Aggregate recruiting intensity. *American Economic Review*, 108(8):2088–2127, August 2018.

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.

Gourieroux, C., Monfort, A., and Trognon, A. Pseudo maximum likelihood methods: Applications to poisson models. *Econometrica: Journal of the Econometric Society*, pages 701–720, 1984.

Gregory, V., Menzio, G., and Wiczer, D. G. The alpha beta gamma of the labor market. Working Paper 28663, National Bureau of Economic Research, April 2021.

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.

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.

Kircher, P. Efficiency of Simultaneous Search. *Journal of Political Economy*, 117(5): 861–913, October 2009.

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., 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.

Lafuente, C. Hysteresis for the young: Search capital and unemployment. *Working Paper*, 2023.

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.

Leduc, S. and Liu, Z. The weak job recovery in a macro model of search and recruiting intensity. *American Economic Journal: Macroeconomics*, 12(1):310–43, January 2020.

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

Lentz, R. and Tranæs, T. Job search and savings: Wealth effects and duration dependence. *Journal of Labor Economics*, 23(3):467–489, 2005. ISSN 0734306X, 15375307. URL `http://www.jstor.org/stable/10.1086/430284`.

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

Lise, J. and Robin, J.-M. The macrodynamics of sorting between workers and firms. *American Economic Review*, 107(4):1104–35, April 2017.

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`.

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

Matzkin, R. L. Nonparametric estimation of nonadditive random functions. *Econometrica*, 71(5):1339–1375, 2003.

McCall, J. J. Economics of information and job search. *The Quarterly Journal of Economics*, pages 113–126, 1970.

Menzio, G. and Shi, S. Efficient Search on the Job and the Business Cycle. *Journal of Political Economy*, 119(3):468–510, 2011.

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.

Mukoyama, T., Patterson, C., and Şahin, A. Job search behavior over the business cycle. *American Economic Journal: Macroeconomics*, 10(1):190–215, January 2018.

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.

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`.

Salop, S. C. Systematic job search and unemployment. *The Review of Economic Studies*, 40(2):191–201, 1973.

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.

Silva, J. and Toledo, M. Labor turnover costs and the cyclical behavior of vacancies and unemployment. *Macroeconomic Dynamics*, 13:76–96, 05 2009.

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. Nonstationarity in job search theory. *The Review of Economic Studies*, 57(2):255–277, 1990.

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`.

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`.

Wolthoff, R. Applications and Interviews: Firms' Recruiting Decisions in a Frictional Labour Market. *Review of Economic Studies*, 85(2):1314–1351, 2018.

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. *Working paper*, 2023.

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*.

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

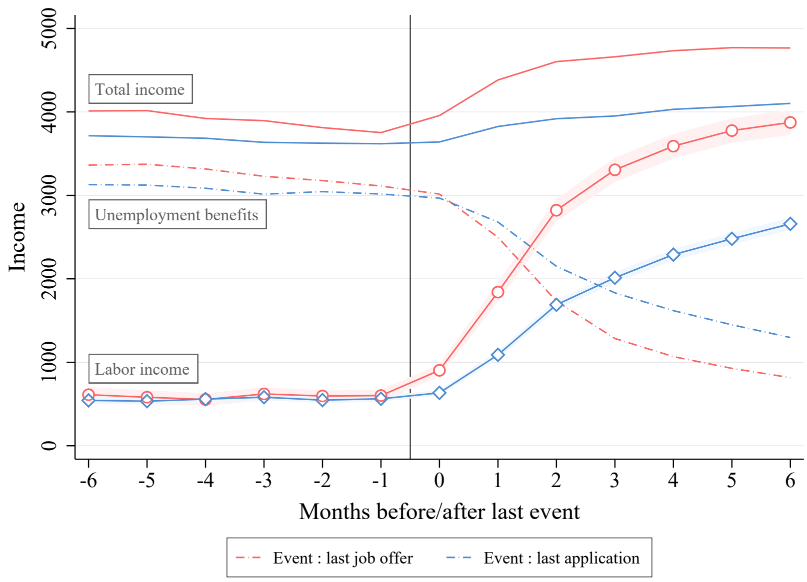
Nom et prénoms	No AVS	Mois et année

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

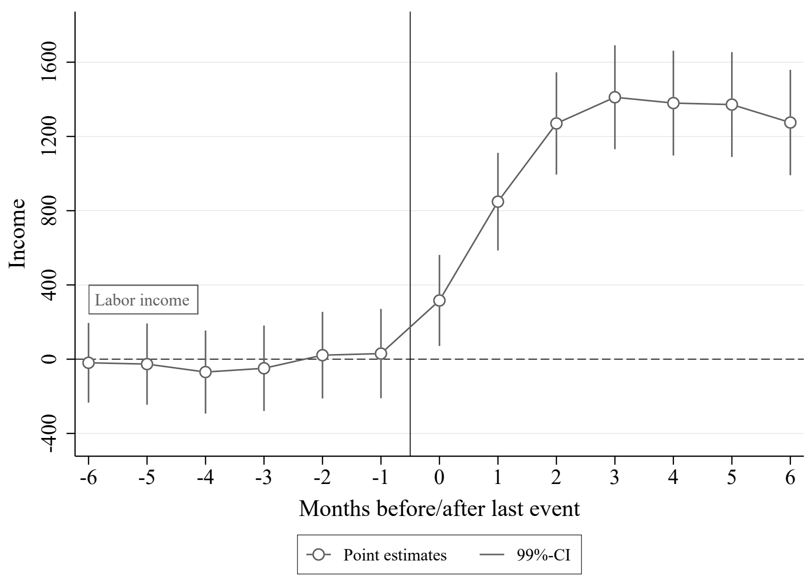
Note: This figure presents the job search diary form that job seekers have to use to document their search activities.

Figure A2: Job offers and income trajectories

(A) Observed average income trajectories

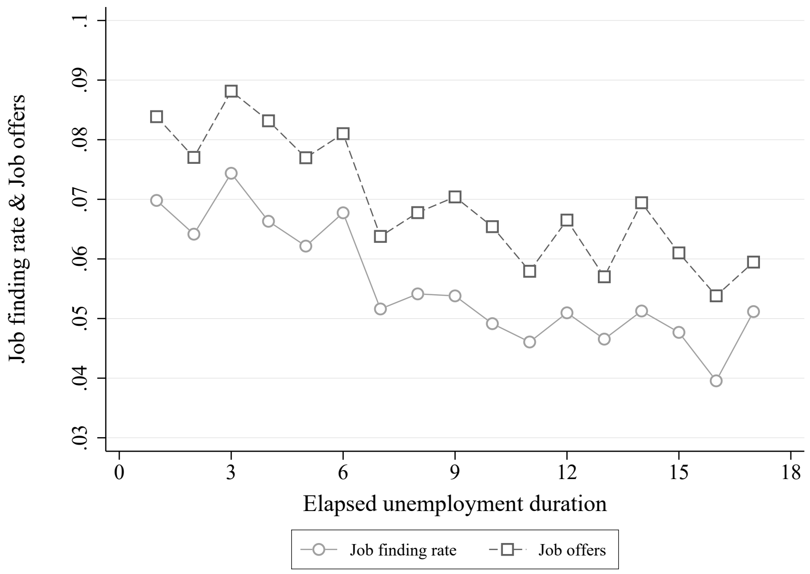


(B) Δ in labor income trajectories (accounting for heterogeneity)



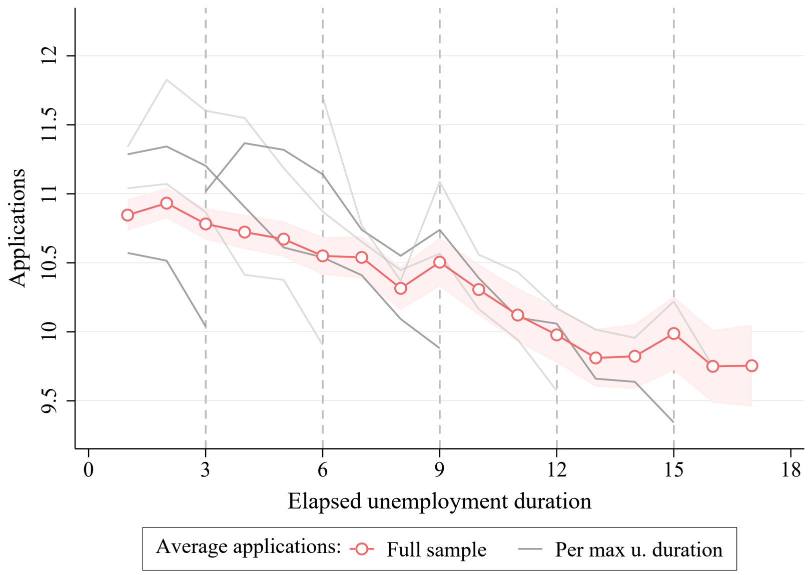
Note: This figure presents an event-study analysis, contrasting information from the search diary data and the social security data. It highlights the informational content of the search diaries. Panel A shows the average evolution of total income, labor income and unemployment benefits in months before 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 elapsed duration of unemployment (in grey). The first grey line, which ends at 3 months, is for job seekers who were unemployed for at least three months, the second line for job seekers unemployed for 6 months, etc. and the last grey line is for job seekers unemployed for at least 15 months.

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 $Y = O$.[60] 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).[61] Stacking does not require the researcher to pre-select one particu-

[60] 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.

[61] 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).[62]

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.

[62] 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 as follows. We regress $\check{Y}_{\iota t}$ 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). We use the DDML approach and the stacking approach to predict also the conditional expectation of $\hat{\alpha}_{ijt}$ with respect to X_{ijt} in step one. This gives us the residuals $\check{\alpha}_{ijt}$ which we then use in step two to run the regression of $\check{Y}_{ijt}$ on $\check{t}_{ijt}$ and $\check{\alpha}_{ijt}$, see column (4) of Table 3.

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 35.5% (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.

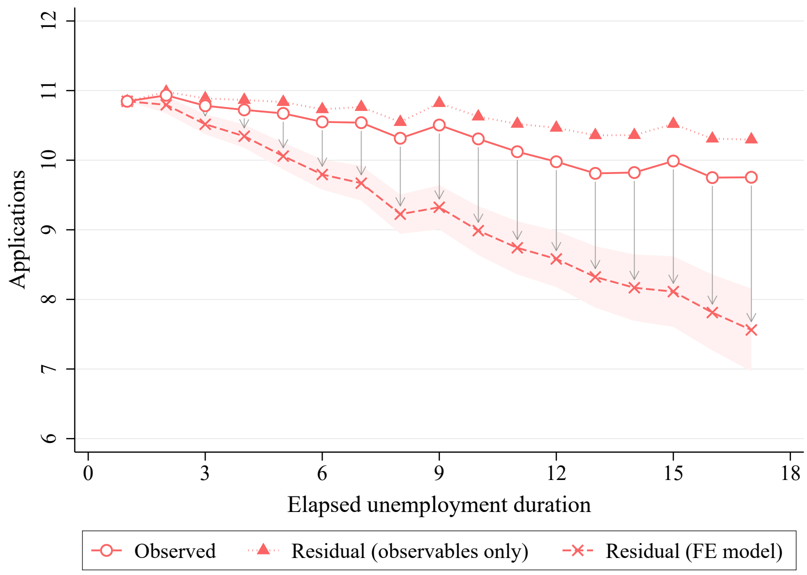
Table B1: Predictive performance and stacking weights, eq. (B.1) for applications

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.

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.

To better understand what characterizes job seekers with higher values of the individual fixed effect, we predict the estimated individual fixed 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 a weighted average 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 for predicting the estimated individual fixed effect. According to Table B2 gradient boosted trees is the best performing learner. It achieves an out-of-sample R^2 of 30.4%, 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 29.3%. The R^2 of the ensemble learner is 31.2%. Table B3 shows that variables capturing the employment prospects, *e.g.* desired occupation, and local labor market policy, *i.e.*, fixed effects for the local labor market offices, explain most of the variation in the estimated individual fixed effects when an OLS regression is used, column (1). In contrast, when gradient boosted trees are fit, local policy effects again stand out as the most important group of predictor variables, whereas the other groups of predictors are nearly equally important.

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

	(1)	(2)
	Stacking weights	R squared
Hand-curated	0.000	0.272
Very flexible	0.310	0.289
Lasso	0.066	0.293
Ridge	0.000	0.273
Random forest	0.017	0.265
Boosting	0.608	0.304
Stacked	1.000	0.312

Note: The estimated 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 ensemble learner. 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). Total sample size is 11602 observations.

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

	(1)	(2)	(3)
	OLS	Random Forest	Grad. Boosted Trees
Demographic characteristics	0.090	0.199	0.103
Employment history	0.011	0.429	0.112
Employment prospects	0.200	0.129	0.131
Local policy	0.635	0.181	0.553
Calendar time at start	0.065	0.062	0.101

Note: The estimated individual fixed effect in eq. (1), $\hat{\alpha}_i$ is predicted based on an OLS regression, 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 the estimated application fixed effect. Column (1) shows 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.

B.2.2 Robustness Checks

We perform several robustness checks to assess the validity of our finding that unemployment duration affects the number of applications per month negatively.

First, we consider an alternative model specification. Given the count data nature of the dependent variable, we estimate a Poisson pseudo maximum likelihood model with fixed effects. The corresponding results are reported in Table B4 and are very close to our baseline OLS estimates, both qualitatively and quantitatively. Specifically, accounting for unobserved heterogeneity through fixed effects consistently leads to a marked steepening in the estimated effect of duration (from a semi-elasticity of -0.9% to -2.1%).

Second, we consider alternative measures of the number of applications made in a month. This robustness check is motivated by the following two observations. 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. As a result, we observe in the search diary data some bunching at common values for minimum search

Table B4: Duration dependence in applications, Poisson pseudo maximum likelihood

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: Applications						
Elapsed unemployment duration	-0.009***	-0.006***	-0.004***	-0.004***	-0.020***	-0.021***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
	[-0.097]	[-0.069]	[-0.048]	[-0.050]	[-0.226]	[-0.230]
Individual controls	No	Yes	Yes	Yes	No	Yes
Policy controls	No	No	Yes	Yes	No	Yes
Local labor market conditions	No	No	No	Yes	No	Yes
Individual FE	No	No	No	No	Yes	Yes
Mean outcome 1$^{\text{st}}$ month	11.107	11.107	11.107	11.107	11.107	11.107
Pseudo-R^2	0.002	0.014	0.066	0.071	0.201	0.206
Observations	55559	55559	55559	55559	55559	55559

Note: This table reports estimates of duration dependence after controlling for observed and unobserved individual heterogeneity using a Poisson pseudo maximum likelihood estimator with individual fixed effects, where the duration function $f^A(t)\phi^A$ is specified linearly. All models are estimated on the restricted sample that excludes individuals with only a single unemployment month recorded in the data. Each column sequentially adds a set of controls or fixed effects. Standard errors clustered at the individual level are reported in parentheses. Absolute coefficients (measuring the monthly decrease in application effort) are indicated in square brackets and are directly comparable to the OLS estimates. Stars indicate the following significance levels: * 0.1, ** 0.05 and *** 0.01.

requirements, *i.e.*, $\underline{A} = 8, 10$. In addition, not all applications directly result from the job seeker's own initiative, but some occur in response to a suggestion by the caseworker. For instance, caseworkers may refer job seekers to apply to jobs Therefore, one might argue that the total number of applications made in a given month, A_{it}, does not capture application effort accurately enough. As a robustness check, we re-estimate our model using alternative search effort measures as dependent variables: In one specification, we use the excess application effort defined as the number of applications exceeding the standard minimum search requirement, $\bar{A}_{it} = \max(0, A_{it} - \underline{A})$, where $\underline{A} = 8, 10$ (see Figure B2 for descriptive evidence). In another specification, we consider the monthly number of applications that are not a response to a referral. The corresponding estimates are reported in Table B5 and are very much in line with our baseline findings.

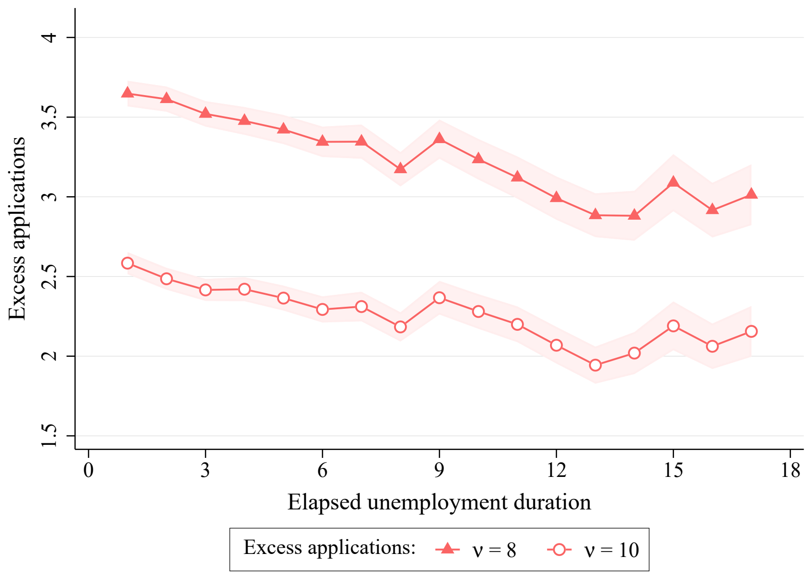
Third, we discuss the existence of a potential within-estimation bias of duration effects in our baseline estimates. As shown in Zuchuat (2023), using fixed effects models to estimate duration dependence profiles from data subject to attrition might entail a strong bias in the estimated duration effects. This is notably the case if the dependent variable is closely related to the attrition mechanism, as this mechanically generates a correlation between the within-variation of the regressor and the error term. In our context, applications are observed repeatedly within an unemployment spell and do not directly translate into exits from unemployment. As an additional check, we re-estimate our baseline specification on a subsample that excludes the last observation of each non-right-censored spell, *i.e.*, using only those observations at the person-month level that are not contem-

Table B5: Duration dependence in applications, alternative application measures

Dependent variables:	Excess applications				Applications on own initiative	
	$\underline{A} = 8$		$\underline{A} = 10$			
	(1)	(2)	(3)	(4)	(5)	(6)
A. OLS						
Elapsed unemployment duration	-0.069***	-0.201***	-0.058***	-0.179***	-0.099***	-0.202***
	(0.008)	(0.022)	(0.007)	(0.022)	(0.009)	(0.022)
Individual controls	No	Yes	No	Yes	No	Yes
Policy controls	No	Yes	No	Yes	No	Yes
Local labor market conditions	No	Yes	No	Yes	No	Yes
Individual FE	No	Yes	No	Yes	No	Yes
Mean outcome 1$^{\text{st}}$ month	4.274	4.274	3.316	3.316	11.055	11.055
Adjusted-R^2	0.005	0.393	0.004	0.338	0.008	0.468
Observations	45901	45901	39563	39563	51305	51305
B. Poisson						
Elapsed unemployment duration	-0.019***	-0.057***	-0.022***	-0.070***	-0.010***	-0.020***
	(0.002)	(0.006)	(0.003)	(0.008)	(0.001)	(0.002)
	[-0.082]	[-0.245]	[-0.072]	[-0.232]	[-0.107]	[-0.217]
Individual controls	No	Yes	No	Yes	No	Yes
Policy controls	No	Yes	No	Yes	No	Yes
Local labor market conditions	No	Yes	No	Yes	No	Yes
Individual FE	No	Yes	No	Yes	No	Yes
Mean outcome 1$^{\text{st}}$ month	4.274	4.274	3.316	3.316	11.055	11.055
Pseudo-R^2	0.004	0.328	0.004	0.334	0.003	0.200
Observations	45901	45901	39563	39563	51305	51305

Note: This table reports estimates of equation (1) for our alternative measures of applications (excess applications and applications on own initiative), where the duration function $f^A(t)\phi^A$ is specified linearly. Models are estimated using OLS (panel A) or Poisson pseudo maximum likelihood (panel B). 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.

Figure B2: Empirical duration dependence in excess applications



Note: This figure reports the empirical duration profiles of excess applications using two different values for the minimum search requirement.

poraneous to an unemployment exit. The corresponding estimation results are reported in Table B6 and turn out to be highly similar to our baseline estimates.

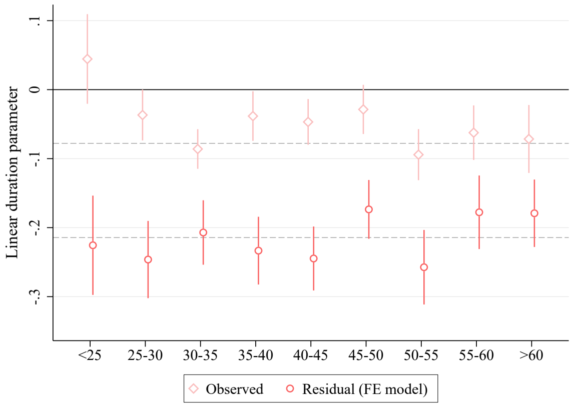
Table B6: Duration dependence in applications, dropping exit months

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: Applications						
Elapsed unemployment duration	-0.082***	-0.056***	-0.037***	-0.041***	-0.190***	-0.215***
	(0.008)	(0.008)	(0.007)	(0.007)	(0.010)	(0.021)
	[-0.750%]	[-0.518%]	[-0.343%]	[-0.378%]	[-1.747%]	[-1.975%]
Individual controls	No	Yes	Yes	Yes	No	Yes
Policy controls	No	No	Yes	Yes	No	Yes
Local labor market conditions	No	No	No	Yes	No	Yes
Individual FE	No	No	No	No	Yes	Yes
Mean outcome 1$^{\text{st}}$ month	10.846	10.846	10.846	10.846	10.846	10.846
Adjusted-R^2	0.006	0.035	0.179	0.193	0.495	0.502
Observations	56646	56646	56646	56646	56646	56646

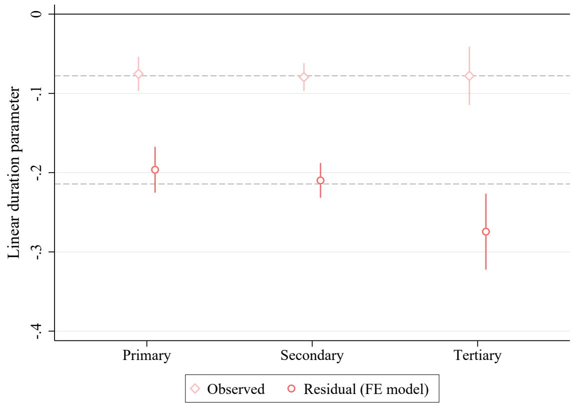
Note: This table reports estimates of equation (1), where the duration function $f^A(t)\phi^A$ is specified linearly. Models are estimated on a restricted sample, that discards those observations at the person-month level in which an unemployment exit is observed. Each column sequentially adds a set of controls or fixed effects. 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 B3: Heterogeneity in the effect of duration on applications

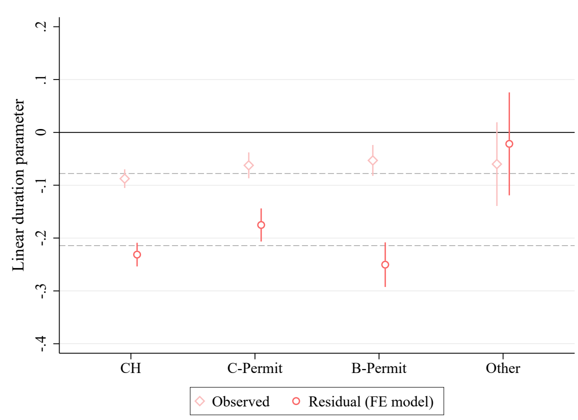
(A) Age



(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 90% confidence intervals.

B.3 Job interviews and job offers

B.3.1 Additional estimation results

Tables B7 through B9 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 $Y = C, O, U$ (interviews/callbacks, 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 B7, 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 B8. 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 B8. Both in Table B7 and in Table B8, 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 B9 in columns (3) through (6) are by definition identical to those for interviews in Table B7. According to column (2) of Table B9 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 B8.

Tables B10 and B11 show the average partial effects of elapsed unemployment duration evaluated at three different percentiles of the empirical distribution of the employability index measured by the standardized individual fixed effect from the application equation (eq. (1)). Specifically, the average partial effects are based on the following modified version of equation (2):

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

with $Y = C, O$.

Table B7: 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	0.108	0.643	0.123	0.047	0.809	0.345
Very flexible	0.000	0.472	0.000	0.077	0.191	0.246
Lasso	0.000	0.621	0.214	0.079	0.000	0.248
Ridge	0.261	0.649	0.000	0.078	0.000	0.247
Random forest	0.147	0.620	0.202	0.077	0.000	0.208
Boosting	0.484	0.654	0.461	0.092	0.000	0.253
Stacked	1.000	0.657	1.000	0.099	1.000	0.351

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 B8: 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	0.273	0.596	0.186	0.023	0.280	0.187
Very flexible	0.000	0.568	0.000	0.039	0.110	0.191
Lasso	0.000	0.588	0.037	0.040	0.225	0.192
Ridge	0.288	0.612	0.214	0.039	0.000	0.179
Random forest	0.184	0.601	0.335	0.042	0.061	0.164
Boosting	0.254	0.619	0.228	0.042	0.325	0.191
Stacked	1.000	0.627	1.000	0.053	1.000	0.214

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.

Table B10 displays estimates of the average partial effect of unemployment duration on the probability that an application leads to an interview for a model that includes in addition the estimated individual fixed effect from the application equation, eq. (1), and its interaction with elapsed unemployment duration are included as regressors. The estimation results in Table B10 do not point to a clear pattern in the way duration dependence in job interviews interacts with the individual fixed effects from the application equation. According to the DDML estimates in column (2) duration dependence in job interviews is nearly the same for individuals at the 25th, the 50th, and the 75th percentile of the distribution of the estimated individual fixed effect, *i.e.*, job seekers with high, medium, and low unobserved employability. In contrast, the results for job offers per interview in Table B11 suggest that job seekers with a lower employability exhibit a higher probability to receive a job offer interview has taken place.

Table B9: 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	0.513	0.638	0.123	0.047	0.809	0.345
Very flexible	0.000	0.493	0.000	0.077	0.191	0.246
Lasso	0.000	0.494	0.214	0.079	0.000	0.248
Ridge	0.242	0.646	0.000	0.078	0.000	0.247
Random forest	0.182	0.617	0.202	0.077	0.000	0.208
Boosting	0.064	0.663	0.461	0.092	0.000	0.253
Stacked	1.000	0.651	1.000	0.099	1.000	0.351

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.

Table B10: Duration dependence in interviews

	(1)	(2)
Average partial effect of elapsed unemployment duration		
At P25 (=-.542) of employability (α_i)	-0.134***	-0.099***
	(0.021)	(0.022)
	[-2.684%]	[-1.985%]
At P50 (=.109) of employability (α_i)	-0.120***	-0.103***
	(0.015)	(0.016)
	[-2.406%]	[-2.075%]
At P75 (=.671) of employability (α_i)	-0.109***	-0.108***
	(0.015)	(0.014)
	[-2.183%]	[-2.169%]
Individual controls	Yes	Yes
Policy controls	Yes	Yes
Local labor market conditions	Yes	Yes
DDML	No	Yes
Mean outcome 1$^{\text{st}}$ month	4.977	4.977
Observations	600323	600323

Note: This table reports estimates of duration effects on the probability of a job offer per interview according to equation (B.4). Column (1) corresponds to a standard logit regression with $g^I(X_{ijt}) = X_{ijt}\beta^I$, whereas column (2) 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). Standard errors (in parentheses) are clustered at the individual level. Stars indicate the following significance levels: * 0.1, ** 0.05 and *** 0.01.

Table B11: Duration dependence in job offers per interview

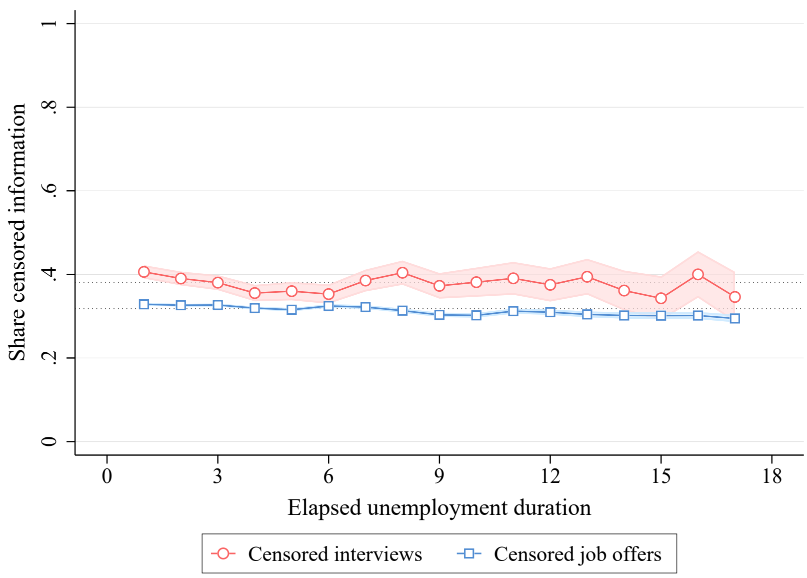
	(1)	(2)
Average partial effect of elapsed unemployment duration		
At P25 (=-.498) of employability (α_i)	0.347***	0.282**
	(0.122)	(0.133)
	[1.718%]	[1.396%]
At P50 (=.015) of employability (α_i)	0.384***	0.323***
	(0.100)	(0.108)
	[1.900%]	[1.600%]
At P75 (=.637) of employability (α_i)	0.426***	0.370***
	(0.103)	(0.107)
	[2.111%]	[1.833%]
Individual controls	Yes	Yes
Policy controls	Yes	Yes
Local labor market conditions	Yes	Yes
DDML	No	Yes
Mean outcome 1^{st} month	20.187	20.187
Observations	22422	22422

Note: This table reports estimates of duration effects on the probability of a job offer per interview according to equation (B.4). Column (1) corresponds to a standard logit regression with $g^O(X_{ijt}) = X_{ijt}\beta^O$, whereas column (2) models $g^O(X_{ijt})$ 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). Standard errors (in parentheses) are clustered at the individual level. Stars indicate the following significance levels: * 0.1, ** 0.05 and *** 0.01.

B.4 Robustness of results to right censoring

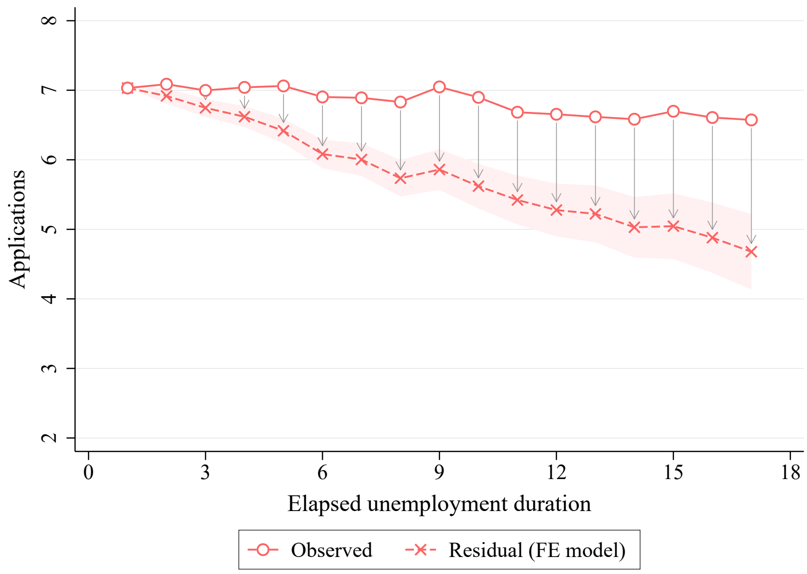
In this subsection, we provide information on the incidence of right censoring by elapsed unemployment duration and show versions of our main empirical results estimated from non-right-censored data. In sum, the incidence of right censoring does not change with unemployment duration (Figure B4), and the duration profiles of applications, job interviews and job offers computed from non-right-censored data shown in Figures B5 to B7 look qualitatively very similar to those computed from the censored data shown in the main text (Figures 3A, 4A and 4B).

Figure B4: Incidence of right censoring by elapsed unemployment duration



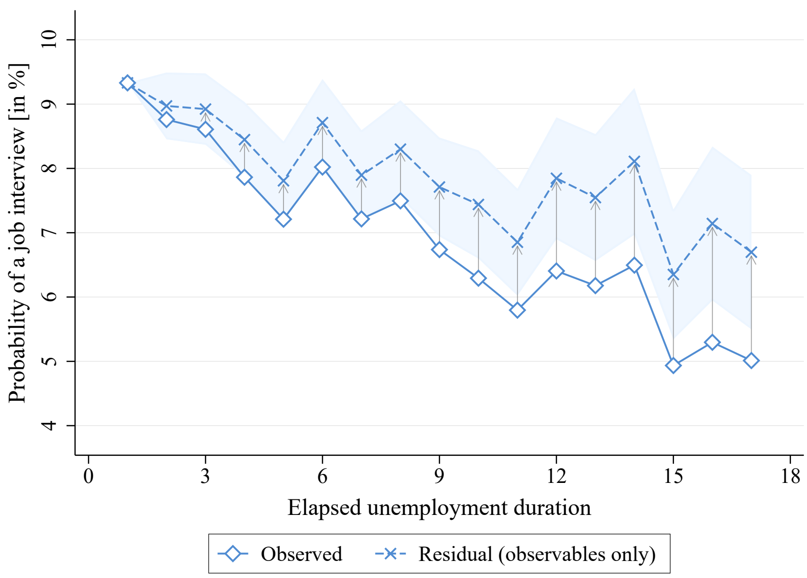
Note: This figure depicts the incidence of right censoring of (i) interviews (out of all applications) and (ii) job offers (out of all interviews) by elapsed duration of unemployment. 95% confidence intervals for the (conditional) means are reported.

Figure B5: Duration profile of applications, non-censored applications only



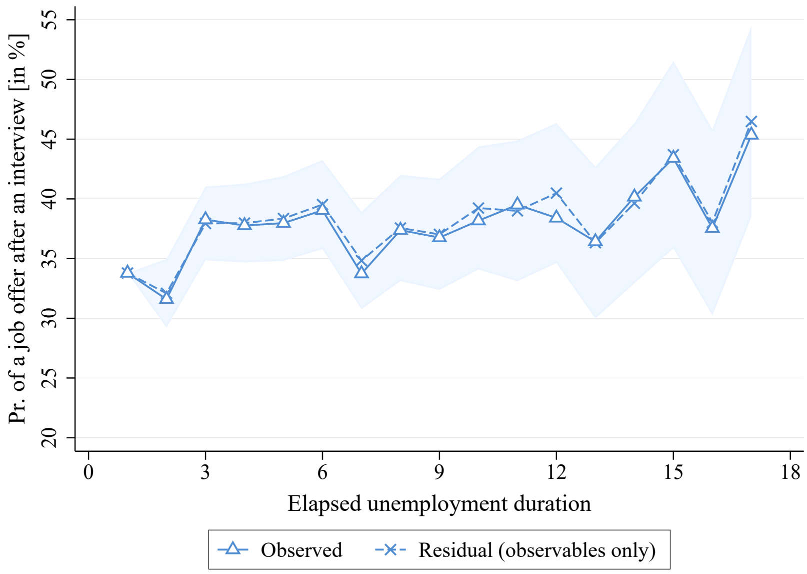
Note: This Figure depicts the empirical duration dependence in the number of job applications (solid line) and the estimated duration dependence obtained after controlling for observable heterogeneity and individual fixed effects (dashed line), with function $f^A(t)\phi^A$ in equation (1) modeled as 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 B6: Duration profile of job interviews, applications with known interview outcome only



Note: This figure depicts the empirical duration dependence in the probability of a job 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 90% confidence interval. Only non-censored applications are considered.

Figure B7: Duration profile of job offers, applications with known job offer outcome only



Note: This figure depicts the empirical duration dependence in the job offer probability (solid line) and the estimated duration dependence obtained after controlling for observable heterogeneity (dashed line), with function $f^O(t)\phi^O$ 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 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 job interview rate and job finding rate at the person-month level that are used as targets in the structural estimation (see Section C.5). For this purpose, we aggregate the application-level information to the person$\times$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.

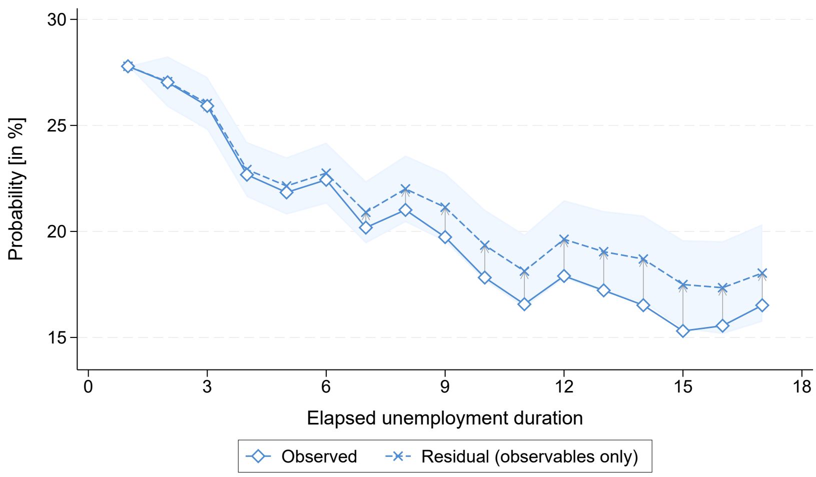
In Figure B8 and Figure B10, the blue line refers to the empirical duration profile, while the red line refers to the duration profile obtained from a regression that controls for those job seeker characteristics that are observable to the recruiting firm at the time when the application is made. Specifically, we estimate two versions, one for the outcome job interview ($Y = I$) and one for job finding ($Y = F$), of the following model for individual i in month t of unemployment, of the following type:

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

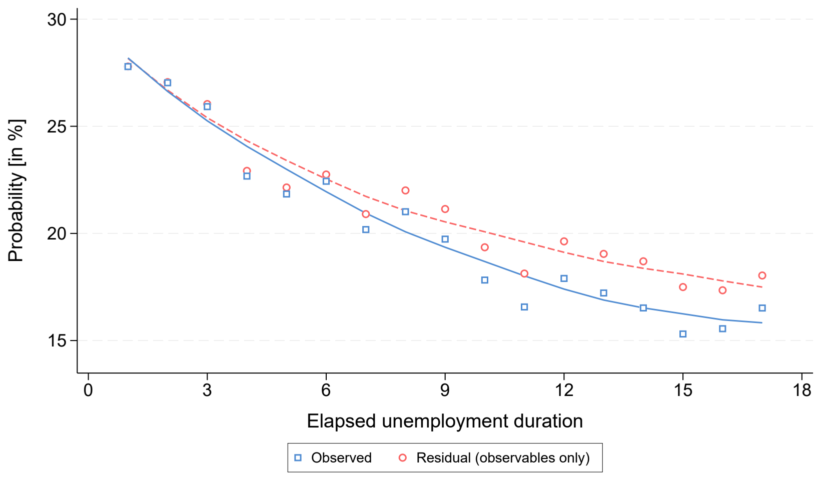
where the function $H(\cdot)$ is a link function. The term $f^Y(t)\phi^Y$ denotes a linear (in ϕ^Y) specification of duration dependence and $g^Y(X_{it})$ a potentially nonparametric function of a rich set of observed covariates X_{it}, including job seeker characteristics as well as calendar quarter 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 per interview.

Figure B8: Interview rate, observed and controlling for observables

(A) Base version

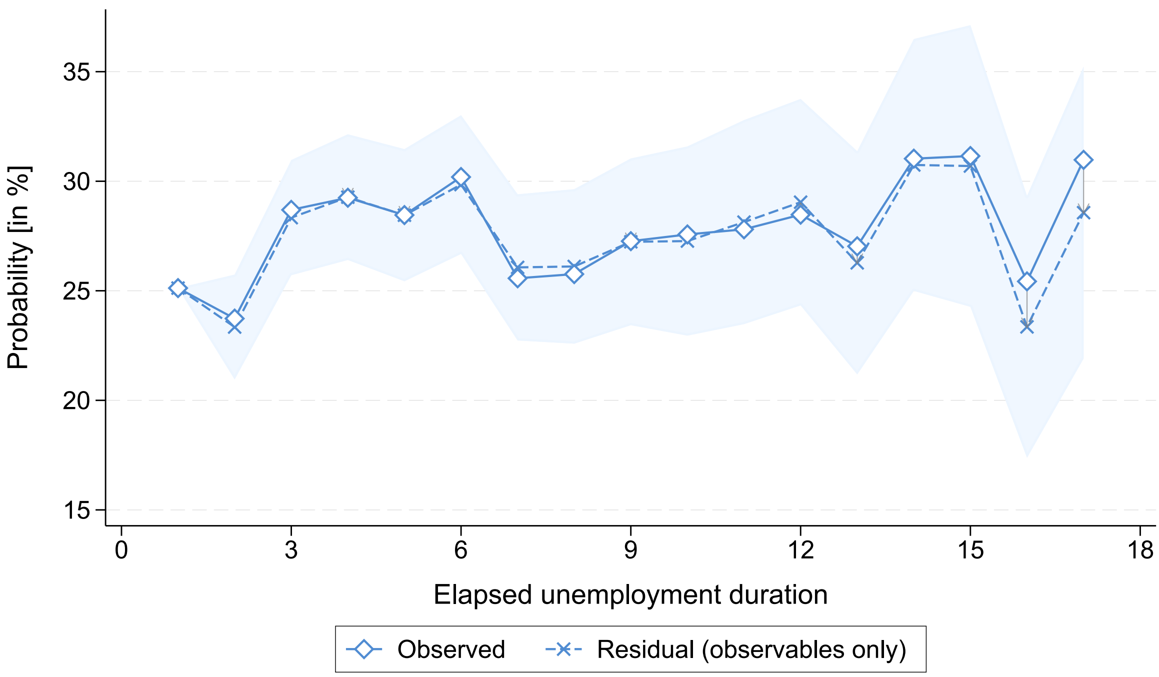


(B) Smoothed version



Note: This figure reports the empirical duration dependence in the job interview rate (probability of at least one job interview in a month) and the estimated duration dependence obtained after controlling for observable heterogeneity. In the latter case, the control variables enter directly in the regression equation. Panel (A) reports the point estimates, while panel (B) reports a smoothed version.

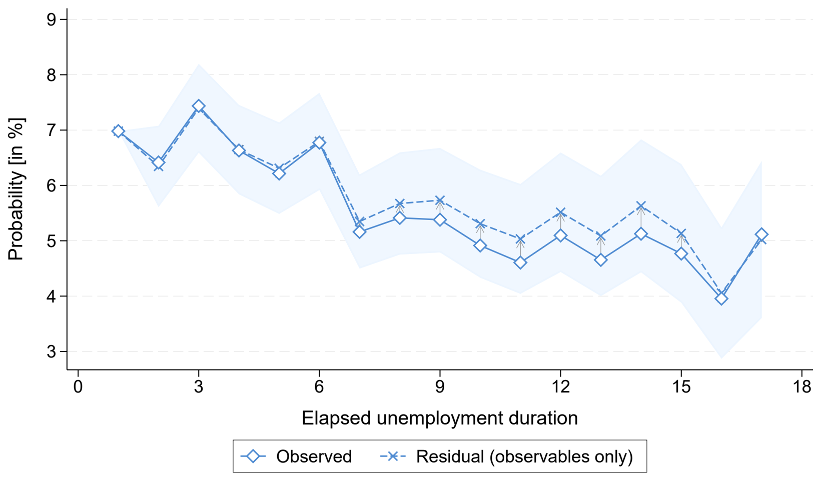
Figure B9: Empirical and residual duration profiles of job offers per interview (at the person-month level)



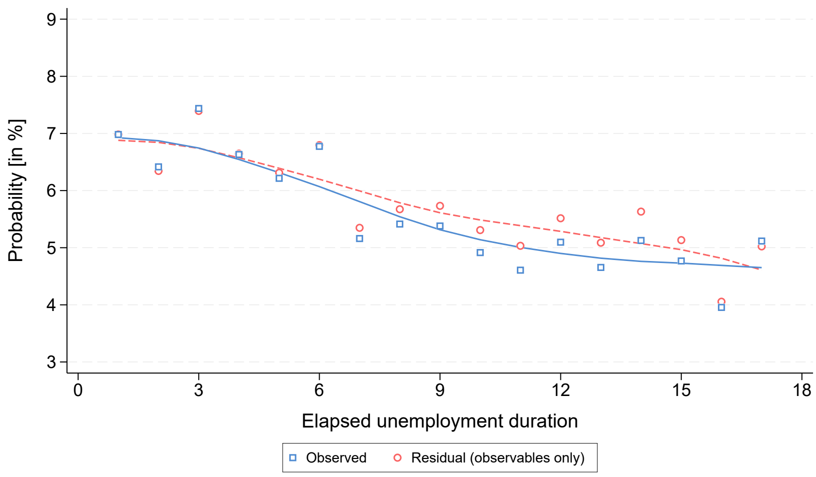
Note: The Figure depicts the empirical duration profile in the job offer probability per interview (solid line) and the estimated duration dependence obtained after controlling for observable heterogeneity (dashed line) using double/debiased machine learning with stacking, with duration dependence modeled 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 90% confidence interval.

Figure B10: Job finding rate, observed and controlling for observables

(A) Base version



(B) Smoothed version



Note: This figure reports the empirical duration dependence in the job finding rate (probability of at least one job offer in a month) and the estimated duration dependence obtained after controlling for observable heterogeneity. In the latter case, the control variables enter directly in the regression equation. Panel (A) reports the point estimates, while panel (B) reports a smoothed version.

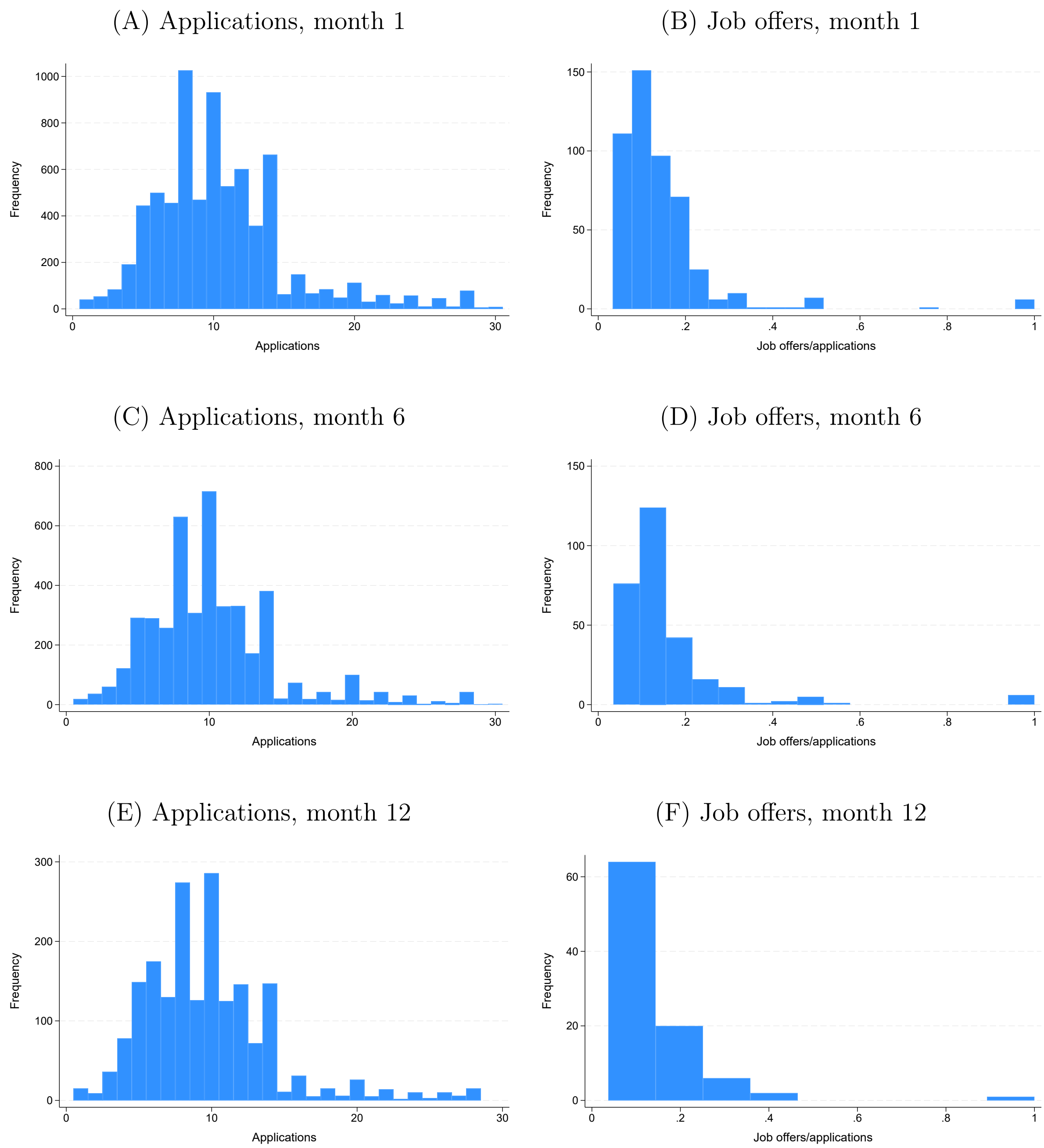
B.6 Evidence to inform modelling choices

Table B12: Duration dependence in applications by employability subgroup

	(1) All	(2) $\geq$ 50th p.	(3) $\geq$ 75th p.	(4) $\geq$ 90th p.	(5) $\geq$ 95th p.
(a) Controls: Individual fixed effects (FE) only					
Elapsed unemp. duration	-0.214	-0.303	-0.448	-0.643	-0.914
	(0.010)	(0.016)	(0.026)	(0.057)	(0.096)
	[-1.976%]	[-2.399%]	[-3.453%]	[-5.258%]	[-8.040%]
(b) Controls: Individual FE plus time-varying local labor market FE					
Elapsed unemp. duration	-0.209	-0.262	-0.349	-0.456	-0.566
	(0.021)	(0.035)	(0.059)	(0.118)	(0.180)
	[-1.926%]	[-2.072%]	[-2.690%]	[-3.726%]	[-4.980%]
Observations	55559	27198	13280	4867	2239

Note: This table reports Standard errors (in parentheses) are clustered at the individual level. Coefficients in relative terms (relative to the average in the first month of unemployment) are reported in square brackets.

Figure B11: Distribution of applications and job offers by elapsed unemployment duration



Note: Panel A, C, E depict the distribution of the number of applications per month in unemployment months 1, 6, and 12, respectively. Panel B, D, and F show the ratio of job offers to applications in unemployment month 1,6, and 12, respectively. In panels B, D, and E, for the vast majority of job seekers this ratio is equal to zero. It is omitted from the histogram.

C. Details of the structural model

Free entry into the labor market dictates that, in equilibrium, the labor market tightness adjusts to arbitrage out any pure profit from vacancy creation:

$$-\kappa_v + \beta \, \frac{\lambda(\theta)}{\theta} \int \Pi(y) dG(y) = 0, \tag{C.1}$$

where $\Pi(y)$ denotes expected profits of a firm with productivity y upon meeting a job seeker.

As a result of the two-stage recruitment process, expected profits of a firm with productivity y upon meeting a job seeker equal:

$$\Pi(y) = \sum_{\tau=0}^{\infty} r(\tau) \left(\int J(x,y) \mathcal{Q}(x,y) \mu(x|\tau) dx - \kappa \right) \mathcal{C}(y,\tau),$$

where $r(\tau)$ equals the probability of meeting a job seeker with unemployment duration τ.

C.1 Microfoundation for application cost function

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

Permanent income increasing in search efficiency (wealth effect) We set out by proposing an interpretation of our baseline model as a reduced-form for an extended model with risk-averse agents and incomplete asset markets. Accordingly, the (positive) dependence of application costs on search efficiency is the result of the positive correlation between permanent income and search efficiency, *i.e.*, a wealth effect.[63]

We consider the same economy studied in Section 5 up to two tweaks. First, we assume that workers are risk-averse and can save through incomplete asset markets. Asset supply is perfectly elastic and represented by an exogenous interest rate R. Second, application costs are independent of search efficiency. Hence, workers make their consumption-savings

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

decisions by solving the following utility maximization problem:[64]

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

where n and u_τ are dummies for the respective labor market state. The solution to the utility maximization problem is represented by the familiar Euler equation, $c^{-\gamma} = \beta R \mathbb{E}[(c')^{-\gamma}]$, and transversality condition. Since ability and search efficiency are positively correlated, lifetime consumption is increasing in ϵ, and so is current consumption conditional on each labor market state $\{n, u_0, u_1, \ldots, u_\tau, \ldots\}$.

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

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

Hence, the value of unemployment and employment are the same as in the baseline model, up to the fact the flow utility is weighted by the current marginal utility of consumption. Importantly, the capital gain upon employment, $\tilde{W}(\epsilon) - \tilde{U}(\epsilon, \tau+1)$, is lower than in the baseline model – the more so, the higher the job seeker's lifetime consumption and, therefore, the higher search efficiency.[65] Let $\alpha(\epsilon, \tau) \equiv \frac{\tilde{W}(\epsilon) - \tilde{U}(\epsilon, \tau+1)}{W(\epsilon) - U(\epsilon, \tau+1)}$. Under regularity condition, $\frac{\partial \alpha(\epsilon, \tau)}{\partial \epsilon} < 0$ (Lentz and Tranæs, 2005). Optimal application effort solves:

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

Comparing this optimality condition to Equation (5), we conclude that the application cost function increasing in search efficiency of the baseline model can be interpreted as catching a wealth effect in reduced-form, which stems from a positive correlation

[64] For simplicity, we abstract from reference dependence. No results are affected by its exclusion.

[65] For given application costs, the difference in the capital gain upon employment between the extended model and the baseline equals $[\tilde{W}(\epsilon) - \tilde{U}(\epsilon, \tau+1)] - [W(\epsilon) - U(\epsilon, \tau+1)] = \left(\frac{\omega}{c(n,\epsilon)^\gamma} - \frac{b}{c(u_\tau,\epsilon)^\gamma}\right) - (\omega - b)$, where $c(n, \epsilon) < \omega$, $c(u_\tau, \epsilon) > b \ \forall \tau$ and $c(n, \epsilon) > c(u_0, \epsilon) > c(u_1, \epsilon) > \cdots > c(u_\tau, \epsilon)$.

between permanent income and search efficiency when workers are risk averse. Formally, $\sigma(a;\epsilon,\tau) \equiv \frac{\tilde{\sigma}(a;\tau)}{\alpha(\epsilon,\tau)}$.

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,

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

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.* $\nu(\ell) = \Lambda\ell$, 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 modelling 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, 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, Λ. It follows that workers of character Λ have search efficiency $\epsilon = g(\Lambda)$, where $g' > 0$.

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

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

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:

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

Following the same argument as before, we assume convex and iso-elastic application effort hours function, *i.e.* $h(a) = \psi \frac{a^{1+\eta}}{1+\eta}$. 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 $\sigma(a;\epsilon) = g^{-1}(\epsilon)h(a)$.

C.2 Proofs

Proof Proposition 1. Consider a job seeker of ability x, whose job offer probability per unit of search effort is given by equation (6). 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,

$$y^*(\tau) : \int \max\{J\left(x, y^*(\tau)\right), 0\}\mu(x|\tau)dx = \kappa$$

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:

$$\frac{do|c(x,\tau)}{d\tau} = \frac{1}{G\left(y^*(\tau)\right)}\left[\frac{dG\left(\min\{x, y^*(\tau)\}\right)}{d\tau} - \frac{G\left(\min\{x, y^*(\tau-1)\}\right)}{G\left(y^*(\tau-1)\right)}\frac{dG\left(y^*(\tau)\right)}{d\tau}\right] \quad \text{(C.2)}$$

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.2), we distinguish two cases.

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

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

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

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

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

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

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

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

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 $\tilde{\tau} = 2$ and abstract from reference dependence and duration-dependent application costs.[66] Optimal application effort, $a(\epsilon, \tau)$, solves:

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

Henceforth, we lighten notation by letting $o(\epsilon,\tau) = \hat{o}(\epsilon,\tau)$. 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:

$$W(\epsilon) = \omega + \beta[W(\epsilon) - \delta_L(W(\epsilon) - U(\epsilon,0))], \tag{C.3}$$

$$U(\epsilon,0) = b - \sigma(a(\epsilon,0);\epsilon) + \beta[U(\epsilon,1) + s(\epsilon,0)o(\epsilon,0)(W(\epsilon) - U(\epsilon,1))], \tag{C.4}$$

$$U(\epsilon,1) = b - \sigma(a(\epsilon,1);\epsilon) + \beta[U(\epsilon,2) + s(\epsilon,1)o(\epsilon,1)(W(\epsilon) - U(\epsilon,2))], \tag{C.5}$$

$$U(\epsilon,2) = b - \sigma(a(\epsilon,2);\epsilon) + \beta[U(\epsilon,2) + s(\epsilon,2)o(\epsilon,2)(W(\epsilon) - U(\epsilon,2))]. \tag{C.6}$$

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

[66] 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.

Let $\Delta^u_\tau(\epsilon) \equiv \omega - (b - \sigma(a(\epsilon,\tau)))$ be the excess flow value of employment over unemployment at duration τ. By subtracting Equation (C.3) from Equation (C.4)-Equation (C.6), we are left with a 3-equation system in $(\Delta_0(\epsilon), \Delta_1(\epsilon), \Delta_2(\epsilon))$:

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

Solving the system yields:

$$\Delta_2(\epsilon) = \tfrac{1}{1-\beta(1-f(\epsilon,2))}\left[\Delta^u_2(\epsilon) - \tfrac{\beta\delta_L}{1+\beta\delta_L}\Delta^u_0 - \tfrac{\beta\delta_L}{1+\beta\delta_L}\beta(1-f(\epsilon,0))\Delta_1(\epsilon)\right], \tag{C.7}$$

$$\Delta_1(\epsilon) = \tfrac{1}{1+\frac{\beta\delta_L}{1+\beta\delta_L}\beta(1-f(\epsilon,0))\frac{1-\beta(f(\epsilon,1)-f(\epsilon,2))}{1-\beta(1-f(\epsilon,2))}}\left[\Delta^u_1(\epsilon) - \tfrac{\beta\delta_L}{1+\beta\delta_L}\tfrac{1-\beta(f(\epsilon,1)-f(\epsilon,2))}{1-\beta(1-f(\epsilon,2))}\Delta^u_0(\epsilon) + \tfrac{\beta(1-f(\epsilon,1))}{1-\beta(1-f(\epsilon,2))}\Delta^u_2(\epsilon)\right] \tag{C.8}$$

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

Suppose by contradiction that $a(\epsilon,0) < a(\epsilon,1)$. From Equation (C.7), it follows that:

$$\Delta_1(\epsilon) < \tfrac{\tilde{\alpha}(\epsilon,1)(1+\beta\delta_L)}{(1+\beta\delta_L)(1-\beta(1-f(\epsilon,2)))+\tilde{\alpha}(\epsilon)\beta\delta_L\beta(1-f(\epsilon,0))}\left(\Delta^u_2(\epsilon) - \tfrac{\beta\delta_L}{1+\beta\delta_L}\Delta^u_0(\epsilon)\right), \tag{C.9}$$

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

Let $\tilde{\Delta}^u_{ij}(\epsilon) \equiv \Delta^u_i(\epsilon) - \frac{\beta\delta_L}{1+\beta\delta_L}\Delta^u_j(\epsilon)$. Rearranging Equation (C.8) yields:

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

Plugging this expression into Equation (C.9) yields:

$$\tilde{\Delta}^u_{10} < \mathcal{M}(\alpha(\epsilon,1))\tilde{\Delta}^u_{20}, \tag{C.10}$$

$$\begin{aligned}\mathcal{M}(\alpha(\epsilon,1)) \equiv \tfrac{1}{1-\beta(1-\underline{f}(\epsilon))}\Bigg[&\tilde{\alpha}(\alpha(\epsilon,1))\left(1+\beta\delta_L\beta(1-\bar{f}(\epsilon))\tfrac{1-\beta(\varphi(\alpha(\epsilon,1))\tilde{\alpha}(\alpha(\epsilon,1))\bar{f}(\epsilon)-\underline{f}(\epsilon))-\tilde{\alpha}(\alpha(\epsilon,1))}{(1+\beta\delta_L)(1-\beta(1-\underline{f}(\epsilon)))+\tilde{\alpha}(\alpha(\epsilon,1))\beta\delta_L\beta(1-\bar{f}(\epsilon))}\right) \\ &-\beta(1-\varphi(\alpha(\epsilon,1))\tilde{\alpha}(\alpha(\epsilon,1))\bar{f}(\epsilon))\Bigg],\end{aligned} \tag{C.11}$$

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

Since $\tilde{\Delta}^u_{10} > \tilde{\Delta}^u_{30}$, if $\mathcal{M}(\alpha(\epsilon)) < 1$, then Equation (C.9) never holds. We now study how the multiplier (C.11) varies with $\alpha(\epsilon)$. We start by analyzing its limiting behavior as $\alpha(\epsilon)$ approaches 1 and 0.

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

Proof. Evaluating Equation (C.11) at $\tilde{\alpha}(1) = 1$ yields:

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

It follows that:

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

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

Evaluating Equation (C.11) at $\tilde{\alpha}(0) = \frac{\underline{o}(\epsilon)}{\bar{o}(\epsilon)}, \varphi(0) = \frac{s(\epsilon,2)}{s(\epsilon,2)}$ yields:

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

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 ($\alpha(\epsilon, 1) = 1$), then application effort may be non-monotonic, *i.e.* job seekers may exert more application effort at duration $\tau = 1$ than at duration $\tau = 0$.[67] On the contrary, if the drop in the job offer probability per unit of search effort happens fully between duration 0 and 1 ($\alpha(\epsilon) = 0$), then application effort is always monotonic, *i.e.* application effort displays negative duration dependence.

Finally, we study how the multiplier C.11 behaves within the two bounds.

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

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

Proof. Taking the derivative of Equation (C.11) with respect to $\tilde{\alpha}(\epsilon)$ yields:

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

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

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

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

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

which similarly always holds. □

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

□

Proof Proposition 3. Since $\partial\mathbb{E}[x|\epsilon]/\partial\epsilon > 0$ by assumption, negative dynamic selection in ability determines negative dynamic selection in search efficiency, *i.e.* $d\mathbb{E}[\epsilon|\tau]/d\tau < 0$. From Equation (5), optimal application effort at duration τ is increasing (decreasing) in search efficiency if the elasticity of the marginal cost exceeds (falls short of) the elasticity of the marginal benefit with respect to search efficiency itself. Hence, if $\zeta > 1 + \frac{\partial \ln(\hat{o}(\epsilon,\tau)[W(\epsilon)-U(\epsilon,\tau+1)])}{\partial \ln(\epsilon)}$ $\forall \tau, \epsilon$, then job seekers with higher search efficiency exert less application effort at any given duration. Since average search efficiency decreases with duration, average application effort at duration τ, $\mathbb{E}_0[a(\epsilon,\tau)]$, exceeds the counterfactual average application effort if the search efficiency distribution were constant over the unemployment spell, $\mathbb{E}_0[a(\epsilon,\tau)]$. □

C.3 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 along two dimensions. First, job seekers and firms/vacancies are assumed to get together through an urn-ball meeting process generating coordination frictions.[68] Second, qualified job seekers get offered a job after an interview with probability $q \in (0, 1)$.

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 κq; (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 is qualified, 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:[69]

$$\int \max \Big\{ J(x, y) - J(\hat{x}, y), 0 \Big\} \; \mu(x|\tau) \; dx \geq \kappa \tag{C.12}$$

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.4), the interview probability writes:

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

[68] 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.

[69] 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.

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.12).

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.4), the job offer probability per interview writes:

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

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:

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

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

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

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.12) at cost κq each. Upon selecting the highest-ability job seeker among them (as long as she is qualified, *i.e.*

$\mathcal{Q}(\bar{\mathbf{x}}_{1,n}, y) = 1$), the firm offers her a job with probability q – the job seekers' selection process being therefore independent of the latter.

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 (C.1):

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

C.4 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 $S \equiv \sum_{\tau=0}^{\infty} \int s(\epsilon, \tau)u(\epsilon, \tau) \; d\mathcal{L}(\epsilon)$ 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 callback stage reads:

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

where $\bar{\tau}(x', y)$ denotes the highest duration τ such that equation (C.12) 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:

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

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).

C.5 Details of structural estimation

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

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 to select individual-level targets – rather than application-level ones – from search diaries (see Table A1 for the respective descriptive statistics). Even though all the parameters are estimated jointly, in what follows we discuss how the empirical moments we select relate to the identification of each parameter.

The Beta shape parameters of the search efficiency and productivity distributions, $[B_1, B_2, G_1, G_2]$, govern the variance and skewness of job seekers' ability and firms' productivity, respectively. As in Jarosch and Pilossoph (2019), the higher the variance in ability and productivity, the higher the scope for negative dynamic selection, which determines the steepness of the duration profiles of the interview rate and job finding rate. In turn, the higher the skewness in ability and productivity, the faster dynamic selection occurs, which determines the convexity of such duration profiles and, as a result, the levels at which the interview rate and job finding rate eventually plateau. We therefore target the duration profiles controlling for observables of the interview rate (Figure B8) and job finding rate (Figure B10), as well as their long-term averages, to identify such parameters.

The substitution parameter of the meeting function, ξ, controls the job seekers' meeting probability per unit of search effort for given labor market tightness, thus being identified by the average interview rate. The convexity of the application effort cost function, η, is the reciprocal of the elasticity of application effort to the expected job offer probability per unit of search effort, which makes the duration profile of application effort controlling for individual fixed effects (Figure 3A) its natural target. The scalar of the application effort cost function, ψ_0, determines the level of application effort, thereby being identified by average application effort. The dispersion parameter of search efficiency, ϕ, governs the cross-sectional variance in application effort for given unemployment duration. Thus, we identify it by targeting the standard deviation of the application fixed effects. As standard in the literature, the vacancy posting cost, κ_v, is identified by the average job finding rate, given that it determines the labor market tightness. The job offer probability per interview of qualified job seekers, q, relates to the long-term job offer probability per unit of search effort and, as such, informs the application decisions of long-term unemployed. We therefore target long-term average application effort to identify it. The elasticity of search effort with respect to application effort, χ, governs how much the (application-level) interview probability is affected by the amount of application effort exerted. For given distribution of search efficiency, this parameter is identified by the partial effect of application fixed effects on the interview probability for given duration (Table 3 Column 4). As formalized in Proposition 3, the elasticity of application costs with respect to search efficiency ζ governs the correlation between dynamic selection in application effort and search efficiency. Hence, we identify it by targeting the duration profile of application effort net of observable heterogeneity (see Figure B1).

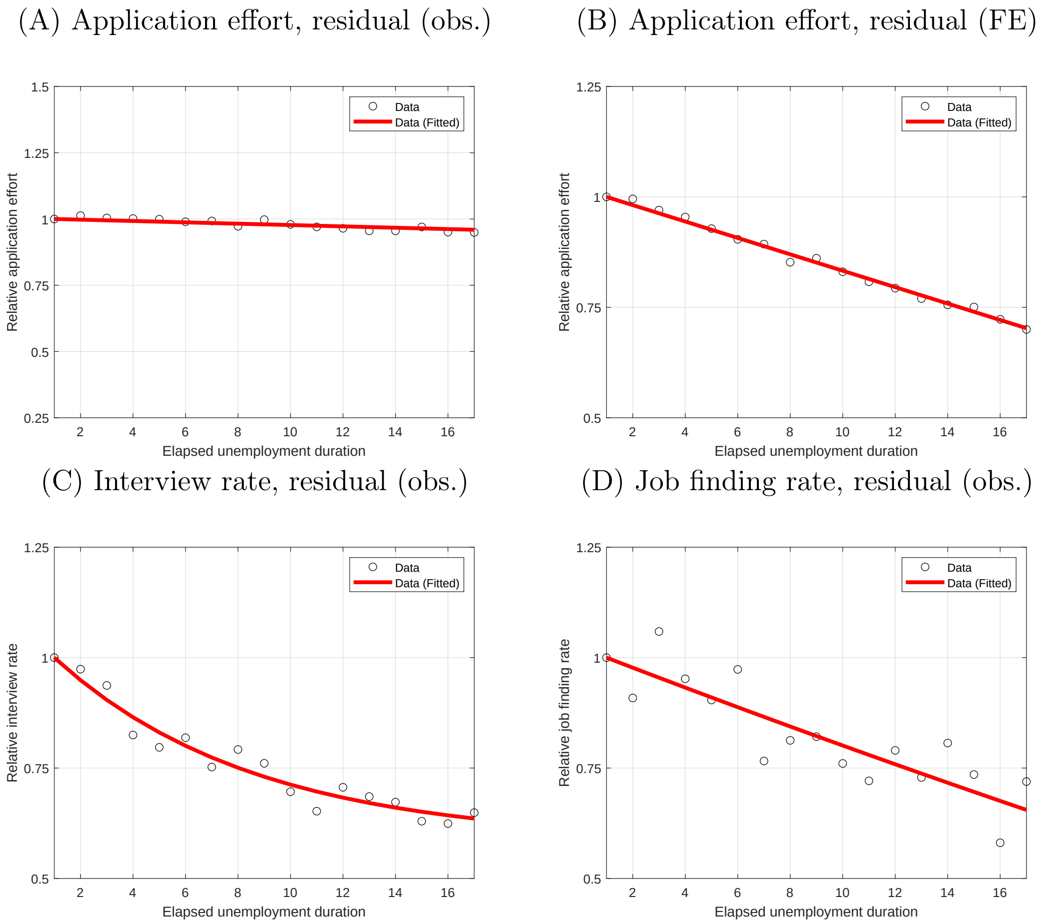
The distinctive parameters of the three model variants are naturally related to the duration profile of application effort by low-ability job seekers, who are barely affected by firms' statistical discrimination. Therefore, we identify them by targeting the duration profile (controlling for fixed effects) of application effort of the job seekers with application fixed effects above median (see Table B12). Specifically, the correlation between equally-ranked ability and search efficiency grid points, ρ, pins down the strength of learning from search, the loss aversion parameter, Υ, determines the magnitude of the application effort response to reference-point adaptation, and the duration dependence coefficient of application costs, ψ_1, directly parametrizes the reduction in optimal application effort with duration.

Details of estimation strategy. Our treatment of the duration profiles follows closely Jarosch and Pilossoph (2019)'s. In particular, we first make functional form assumptions on the duration profile of each variable normalized with respect to the first period of unemployment. As in Jarosch and Pilossoph (2019) and Kroft, Lange, and Notowidigdo (2013), we estimate a negative exponential relationship for the duration profiles controlling for observables of the interview rate and job finding rate via weighted nonlinear least squares. Guided by our empirical results, we then estimate a linear relationship for the duration profiles controlling for observables and for individual fixed effects of application effort. Figure C1 reports the fitted duration profiles along with the raw data. For the sake of our indirect inference exercise, we treat the duration profiles implied by the model exactly as those in the data, by repeating the same steps outlined above. Following Jarosch and Pilossoph (2019), we choose as targets the entire duration profiles of each normalized variable rather than just its linear trend. In practice, each duration-related target is a vector of equally-weighted values for each duration $\tau = 1, \ldots, 17$. Given our main focus on duration dependence, we assign weight $w_1 = 5$ to the 4 duration-related moments and weight $w_2 = 1$ to the remaining 9 moments.

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 about half of the mass is concentrated in every model variant.[70] This is perfectly in line with Jarosch and Pilossoph (2019), which finds the same spike with density ranging from 40% to 64% across different model specifica-

[70]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 $y = 0$.

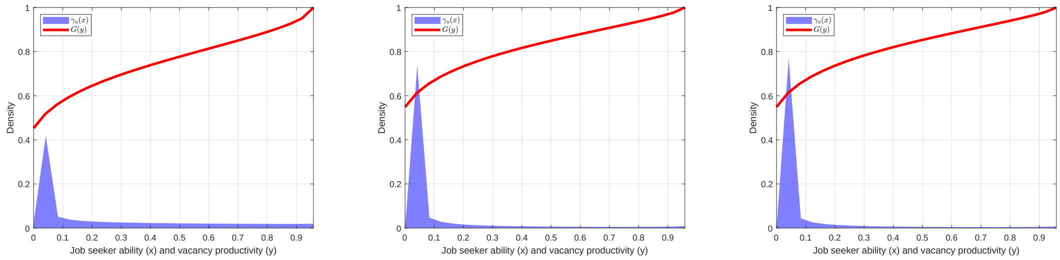
Figure C1: Goodness of fit, functional forms



Note: This figure reports the fitted and raw duration profiles controlling for observables of application effort (Panel A), interview rate (Panel C), and job finding rate (Panel D), as well as the duration profile controlling for individual fixed effects of application effort (Panel B). The fitted duration profiles of applications are estimated through a linear relationship, those of the interview rate and job finding rate through a negative exponential relationship via weighted nonlinear least squares.

Figure C2: Matching probability, job seekers

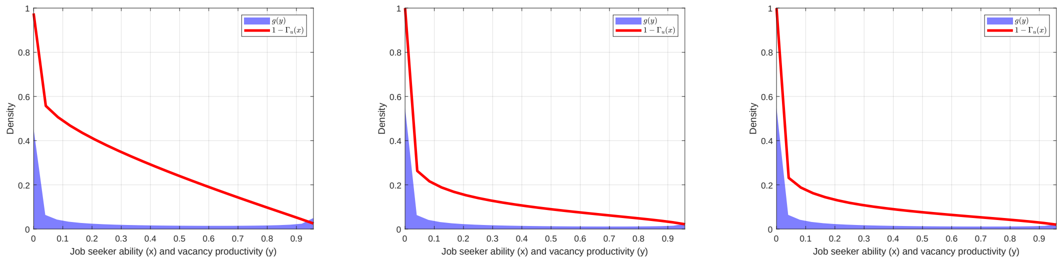
(A) Learning (B) Reference dependence (C) Duration-dep. app. cost



Note: The figures report the after-meeting matching probability faced by job seekers across the ability distribution in the three model variants. The red solid line represents the probability that a worker meets a firm she is qualified for conditional on meeting one; the blue area displays the density of the job seeker ability distribution $\gamma_u(x)$.

Figure C3: Matching probability, vacancies

(A) Learning (B) Reference dependence (C) Duration-dep. app. cost



Note: The figures report the after-meeting matching probability faced by firms across the vacancy productivity distribution in the three model variants. The red solid line represents the probability that a firm meets a qualified worker conditional on meeting one; the blue area displays the density of the vacancy productivity distribution $g(y)$.

tions. Instead of the uniform pattern imposed by Jarosch and Pilossoph (2019), we allow the rest of the productivity distribution to assume any shape consistent with a Beta distribution. According to our estimates, the rest of the productivity distribution exhibits a decreasing density with a small final increase in correspondence to the highest productivity level.[71] Similarly, the equilibrium job seeker ability distribution displays a spike at the lowest positive grid point accounting for 40 to 75% of the total mass. The rest of the distribution is instead relatively close to uniform. The relative shape of the ability and productivity distribution is informative of the extent of matching frictions faced by searching agents. Figure C2 plots the matching probability faced by job seekers across the ability distribution, *i.e.* the probability of meeting a firm they are qualified for (conditional on meeting one). As a result of the production technology (3), such matching probability is increasing in ability. Figure C3 reports the same graph under the firms'

[71] 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 rate, being high-productivity firms the most prone to statistical discrimination.

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 between 0.16 and 0.19, 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, between 1.4 and 1.9% of average monthly output and a mild convexity ($\eta \in [0.24, 0.27]$), which implies a sizable elasticity of application effort to the expected job offer probability per unit of search effort of 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).[72] The search efficiency dispersion parameter ϕ is estimated to be around 11.5, meaning that the highest-efficiency workers are about 8 times more likely to get a callback than lowest-efficiency 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 application fixed effects for the sake of identification. The estimated vacancy posting cost, κ_v, equals just 0.2 to 0.7% of average monthly output, consistently with the reasonable notion that most of hiring costs arises from interview costs rather than entry costs. The job offer probability per interview of qualified applicants equals 36 to 40%, supporting an important role of idiosyncratic matching frictions (unrelated to workers' qualification) in the hiring process. The elasticity of search effort with respect to application effort, χ, is estimated between 0.95 and 1, meaning that decreasing returns in applications are at most mild. The elasticity of application costs with respect to search efficiency, ζ, lies between 1.14 and

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

1.33. This means that the marginal application cost of job seekers with the highest search efficiency is about 13 times larger than that of job seekers with the lowest search efficiency, reflecting *e.g.* wealth differences between the two groups. Finally, we turn to the distinctive parameters of the three model variants. We estimate a correlation between equally-ranked ability and search efficiency grid points, ρ, of 53%, meaning that short-term unemployed are expected to place on average half of the probability on their true ability – the rest being equally split across other ability levels by construction. We estimate a loss aversion coefficient, υ, of 0.4, meaning that workers suffer a utility loss of 40% the gap between current consumption and reference point (provided that the former is lower than the latter).[73] We estimate a duration coefficient in application costs, ψ_1, equal to 0.6% of monthly output. This implies that the marginal application cost after 17 months of unemployment is 10% higher than in the first month.

Table C1: Estimated parameters (learning model)

Parameter	Description	Value	Target	Data	Model
B_1	1^{st} shape param. Beta distr. search eff.	0.113	$\hat{\beta}_{\ln c(\epsilon,\tau,x),\tau}$: duration effect interview rate, residual (obs.)	-0.022	-0.019
B_2	2^{nd} shape param. Beta distr. search eff.	0.498	$\mathbb{E}[c(\epsilon,\tilde{\tau},x)]$: long-term avg interview rate	0.177	0.173
G_1	1^{st} shape param. Beta distr. prod.	0.192	$\hat{\beta}_{\ln f(\epsilon,\tau,x)),\tau}$: duration effect job finding rate, residual (obs.)	-0.018	-0.020
G_2	2^{nd} shape param. Beta distr. prod.	0.550	$\mathbb{E}[f(\epsilon,\tilde{\tau},x))]$: long-term avg job finding rate	0.046	0.050
ξ	Subst. param. meeting function	0.190	$\mathbb{E}[c(\epsilon,\tau,x))]$: avg interview rate	0.230	0.214
η	Convexity app. effort cost	0.239	$\hat{\beta}_{\ln a(\epsilon,\tau),\tau\|\epsilon}$: duration effect applications, residual (FE)	-0.019	-0.014
ψ_0	Scalar app. effort cost	0.019	$\mathbb{E}[a(\epsilon,\tau)]$: avg applications	10.65	11.06
ϕ	Search efficiency dispersion param.	10.46	$\sigma(\epsilon)$: std. dev. application fixed effects	3.966	4.095
κ_v	Vacancy posting cost	0.007	$\mathbb{E}[f(\epsilon,\tau,x))]$: avg job finding rate	0.062	0.062
q	Cond. job offer prob. qualified job seeker	0.398	$\mathbb{E}[a(\epsilon,\tilde{\tau})]$: long-term avg applications	10.24	10.33
χ	App. effort elasticity search effort	0.999	$\hat{\beta}_{\ln[c(\epsilon,\tau,x)/a(\epsilon,\tau)],\alpha(\epsilon)\|\tau}$: partial effect app FE on interview prob.	-0.018	-0.018
ζ	Search eff. elasticity app. costs	1.143	$\hat{\beta}_{\ln a(\epsilon,\tau),\tau}$: duration effect applications, residual (obs.)	-0.005	-0.007
ρ	Equally ranked ability-eff. correlation	0.530	$\hat{\beta}_{\ln a(\epsilon,\tau),\tau\|\epsilon,\alpha(\epsilon)\geq \text{med}[\alpha]}$: duration effect applications (app FEs above median), residual (FE)	-0.021	-0.013

Note: All duration effects are computed from a linear model and expressed as semi-elasticities, *i.e.* the duration coefficient is normalized by the average variable at $\tau = 0$. All averages are computed with respect to the distribution of observables at $\tau = 0$. Application fixed effects are not standardized. Numeraire: cross-sectional avg monthly output.

[73] 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 C2: Estimated parameters (reference dependence model)

Parameter	Description	Value	Target	Data	Model
B_1	1^{st} shape param. Beta distr. search eff.	0.095	$\hat{\beta}_{\ln c(\epsilon,\tau,x),\tau}$: duration effect interview rate, residual (obs.)	-0.022	-0.024
B_2	2^{nd} shape param. Beta distr. search eff.	0.467	$\mathbb{E}[c(\epsilon,\tilde{\tau},x)]$: long-term avg interview rate	0.177	0.173
G_1	1^{st} shape param. Beta distr. prod.	0.159	$\hat{\beta}_{\ln f(\epsilon,\tau,x)),\tau}$: duration effect job finding rate, residual (obs.)	-0.018	-0.020
G_2	2^{nd} shape param. Beta distr. prod.	0.687	$\mathbb{E}[f(\epsilon,\tilde{\tau},x))]$: long-term avg job finding rate	0.046	0.052
ξ	Subst. param. meeting function	0.179	$\mathbb{E}[c(\epsilon,\tau,x))]$: avg interview rate	0.230	0.226
η	Convexity app. effort cost	0.268	$\hat{\beta}_{\ln a(\epsilon,\tau),\tau\mid\epsilon}$: duration effect applications, residual (FE)	-0.019	-0.014
ψ_0	Scalar app. effort cost	0.014	$\mathbb{E}[a(\epsilon,\tau)]$: avg applications	10.65	11.09
ϕ	Search efficiency dispersion param.	11.42	$\sigma(\epsilon)$: std. dev. application fixed effects	3.966	4.166
κ_v	Vacancy posting cost	0.004	$\mathbb{E}[f(\epsilon,\tau,x))]$: avg job finding rate	0.062	0.064
q	Cond. job offer prob. qualified job seeker	0.380	$\mathbb{E}[a(\epsilon,\tilde{\tau})]$: long-term avg applications	10.24	10.57
χ	App. effort elasticity search effort	0.954	$\hat{\beta}_{c(\epsilon,\tau,x)/a(\epsilon,\tau),\alpha(\epsilon)\mid\tau}$: partial effect app FE on interview prob.	-0.018	-0.015
ζ	Search eff. elasticity app. costs	1.328	$\hat{\beta}_{\ln a(\epsilon,\tau),\tau}$: duration effect applications, residual (obs.)	-0.005	-0.007
Υ	Loss aversion coefficient	0.400	$\hat{\beta}_{\ln a(\epsilon,\tau),\tau\mid\epsilon,\alpha(\epsilon)\geq\text{med}[\alpha]}$: duration effect applications (app FEs above median), residual (FE)	-0.021	-0.013

Note: All duration effects are computed from a linear model and expressed as semi-elasticities, *i.e.* the duration coefficient is normalized by the average variable at $\tau = 0$. All averages are computed with respect to the distribution of observables at $\tau = 0$. Application fixed effects are not standardized. Numeraire: cross-sectional avg monthly output.

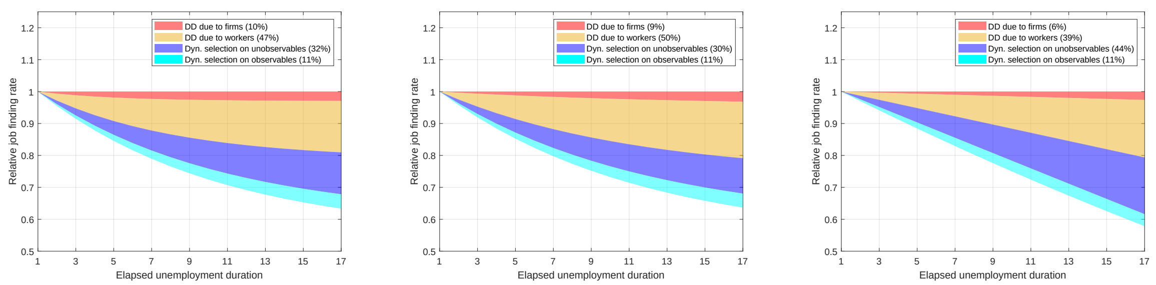
Table C3: Estimated parameters (duration-dependent application cost model)

Parameter	Description	Value	Target	Data	Model
B_1	1^{st} shape param. Beta distr. search eff.	0.086	$\hat{\beta}_{\ln c(\epsilon,\tau,x),\tau}$: duration effect interview rate, residual (obs.)	-0.022	-0.030
B_2	2^{nd} shape param. Beta distr. search eff.	0.424	$\mathbb{E}[c(\epsilon,\tilde{\tau},x)]$: long-term avg interview rate	0.177	0.161
G_1	1^{st} shape param. Beta distr. prod.	0.162	$\hat{\beta}_{\ln f(\epsilon,\tau,x)),\tau}$: duration effect job finding rate, residual (obs.)	-0.018	-0.024
G_2	2^{nd} shape param. Beta distr. prod.	0.710	$\mathbb{E}[f(\epsilon,\tilde{\tau},x))]$: long-term avg job finding rate	0.046	0.050
ξ	Subst. param. meeting function	0.164	$\mathbb{E}[c(\epsilon,\tau,x))]$: avg interview rate	0.230	0.251
η	Convexity app. effort cost	0.257	$\hat{\beta}_{\ln a(\epsilon,\tau),\tau\mid\epsilon}$: duration effect applications, residual (FE)	-0.019	-0.017
ψ_0	Scalar app. effort cost	0.015	$\mathbb{E}[a(\epsilon,\tau)]$: avg applications	10.65	11.50
ϕ	Search efficiency dispersion param.	12.85	$\sigma(\epsilon)$: std. dev. application fixed effects	3.966	3.982
κ_v	Vacancy posting cost	0.002	$\mathbb{E}[f(\epsilon,\tau,x))]$: avg job finding rate	0.062	0.067
q	Cond. job offer prob. qualified job seeker	0.364	$\mathbb{E}[a(\epsilon,\tilde{\tau})]$: long-term avg applications	10.24	10.39
χ	App. effort elasticity search effort	0.986	$\hat{\beta}_{c(\epsilon,\tau,x)/a(\epsilon,\tau),\alpha(\epsilon)\mid\tau}$: partial effect app FE on interview prob.	-0.018	-0.016
ζ	Search eff. elasticity app. costs	1.268	$\hat{\beta}_{\ln a(\epsilon,\tau),\tau}$: duration effect applications, residual (obs.)	-0.005	-0.008
ψ_1	Duration dependence app. costs	0.006	$\hat{\beta}_{\ln a(\epsilon,\tau),\tau\mid\epsilon,\alpha(\epsilon)\geq\text{med}[\alpha]}$: duration effect applications (app FEs above median), residual (FE)	-0.021	-0.016

Note: All duration effects are computed from a linear model and expressed as semi-elasticities, *i.e.* the duration coefficient is normalized by the average variable at $\tau = 0$. All averages are computed with respect to the distribution of observables at $\tau = 0$. Application fixed effects are not standardized. Numeraire: cross-sectional avg monthly output.

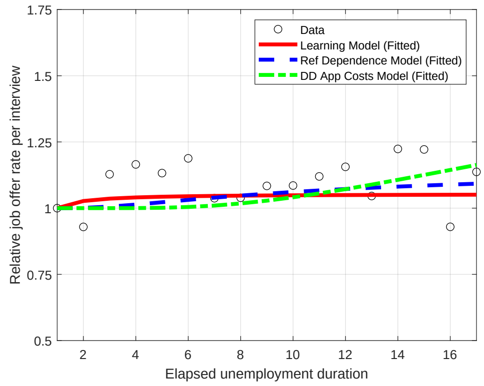
Figure C7: Duration profile of the job finding rate, decomposition across models

(A) Learning (B) Reference dependence (C) Duration-dep. app. cost



Note: This figure reports the fitted duration profiles controlling for observables (in solid red) and fixed effects (in dashed blue) of search effort in the learning model (Panel A), reference dependence model (Panel B), and duration-dependent application costs (Panel C). The fitted duration profiles are estimated through a negative exponential relationship via weighted nonlinear least squares.

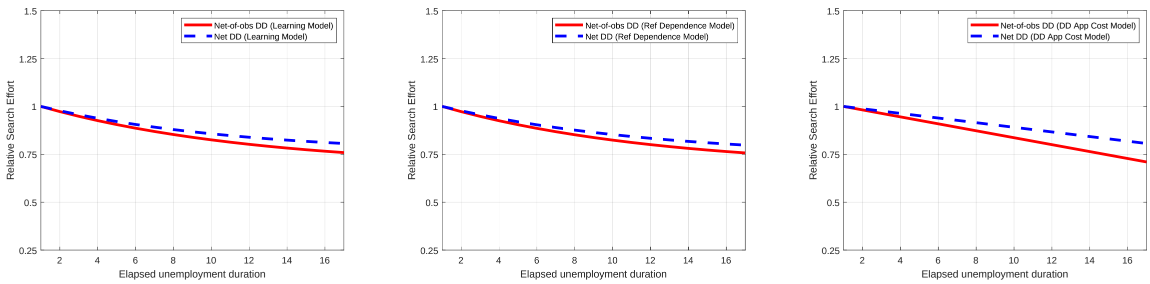
Figure C4: Duration profile of job offer probability per interview, model vs data



Note: This figure contrasts the duration profiles controlling for observables of the individual-level job offer probability per interview in the data (circles) with those implied by the estimated models. The learning model is depicted in solid red, the reference dependence model in dashed blue, and the duration-dependent application cost model in dotted-dashed green. The duration profiles in the model are derived by estimating individual fixed effects and duration effects from a saturated regression, computing the expected values of application effort at any unemployment duration, and normalizing them with respect to the first month of unemployment. For the duration profile controlling for observables, expected values are computed with respect to workers' search efficiency distribution in the contemporaneous period of unemployment, *i.e.* $\mathbb{E}_{\tau}[\hat{a}(\epsilon, \tau)]$. The distribution of observables across unemployment durations is kept the same as in the first month of unemployment in both specifications. Finally, the model-implied duration profiles are fitted by an inverse negative exponential function via nonlinear least squares.

Figure C5: Duration profile of search effort

(A) Learning (B) Reference dependence (C) Duration-dep. app. cost

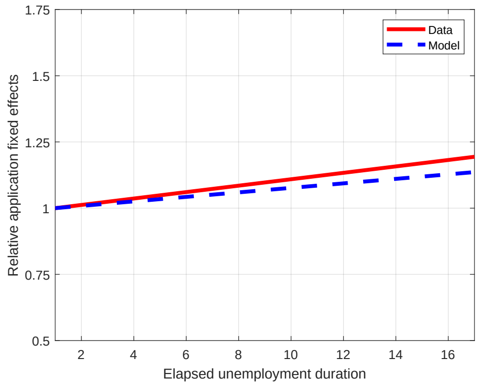


Note: This figure reports the fitted duration profiles controlling for observables (in solid red) and fixed effects (in dashed blue) of search effort in the learning model (Panel A), reference dependence model (Panel B), and duration-dependent application costs (Panel C). The fitted duration profiles are estimated through a negative exponential relationship via weighted nonlinear least squares.

C.6 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*). However, the decomposition of the duration profile of the job finding rate carried out in Figure 7 attributes to statistical discrimination by firms only the direct

Figure C6: Application fixed effects, residual (obs.)



Note: This figure contrasts the duration profiles controlling for individual fixed effects (Panel A) and for observables (Panel B) of application effort fixed effects in the data (solid red) with those implied by the estimated model (dashed blue). Both the duration profiles in the data and in the model are derived by estimating individual fixed effects and duration effects from a saturated regression, computing the expected values of application effort fixed effects at any unemployment duration, and normalizing them with respect to the first month of unemployment. Expected values are computed with respect to workers' search efficiency distribution in the first month of unemployment, *i.e.* $\mathbb{E}_0[\hat{\alpha}(\epsilon,\tau)]$. The distribution of observables across unemployment durations is kept the same as in the first month of unemployment. Finally, both the data- and the model-implied duration profiles are fitted by a linear function.

effect. In this section, we aim to assess the total (direct and indirect) effect of statistical discrimination on duration dependence in the job finding rate.

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.

Taking log of the job finding rate definition (7) and differentiating it with respect to duration yields:

$$\frac{d\ln f(\epsilon,\tau,x)}{d\tau} = \chi\frac{d\ln a(\hat{\boldsymbol{o}}(\epsilon,\tau);\epsilon)}{d\tau} + \frac{d\ln o(x,\tau)}{d\tau}, \tag{C.13}$$

where $\hat{\boldsymbol{o}}(\epsilon,\tau)$ denotes the sequence of expected job offer probability per unit of search effort. 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 $\hat{\boldsymbol{o}}(\epsilon,\tau)$.

To compute the local general-equilibrium elasticity, we make use of Equation (5), as well as our functional form assumptions, to express optimal application effort as a function of the expected job offer probability per unit of search effort:

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

Hence, for given capital gain upon employment, optimal application effort depends on the expected 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 (C.13) yields:

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

If information about own ability is perfect, the elasticity of the expected job offer probability per unit of search effort to the individual one, *i.e.* $\frac{\partial\ln(\hat{o}(\epsilon,\tau))}{\partial\ln(o(x,\tau))}$, is precisely 1, while with incomplete information the same holds as an approximation. Hence, the local general-equilibrium elasticity of the job finding rate to the job offer probability per unit of search effort equals $1+\frac{\chi}{1+\eta-\chi}\frac{\partial\ln(\hat{o}(\epsilon,\tau))}{\partial\ln(o(x,\tau))}\approx 1+\frac{\chi}{1+\eta-\chi}$. According to our estimated parameters, the general-equilibrium elasticity ranges between 4.04 and 5.17 across model variants. It follows that the indirect effect of statistical discrimination largely 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 duration dependence in the job finding rate 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.[74] 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.* $\kappa\to 0$. Since the presence of interview costs is the structural determinant of

[74] 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. Hence, choosing a ranking by duration makes the job offer probability per unit of search effort still exhibit negative duration dependence due to taste-based discrimination.

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 absence of pure profits from job creation, which allows us to compute heterogeneous welfare effects across workers without taking a stance on how profits get rebated. Overall, changes in the stationary equilibrium of our estimated models are to be entirely attributed to the flattening of the net duration profile of the job offer probability per unit of search effort and the induced response of application effort.

Table C4: Baseline vs No statistical discrimination

	Learning		*Loss aversion*		*DD costs*	
	Baseline	No stat. discr.	Baseline	No stat. discr.	Baseline	No stat. discr.
avg job offer rate (%)	0.56	0.57	0.57	0.58	0.59	0.58
long-term avg job offer rate (%)	0.49	0.49	0.49	0.52	0.48	0.52
avg application effort	11.06	10.87	11.09	10.82	11.50	11.11
long-term avg application effort	10.33	10.50	10.57	10.57	10.39	10.16
avg interview rate (%)	21.45	23.20	22.64	24.12	25.08	25.85
long-term avg interview rate (%)	17.33	20.45	17.26	21.84	16.09	21.89
avg job finding rate (%)	6.23	6.20	6.37	6.25	6.74	6.45
long-term avg job finding rate (%)	5.03	5.20	5.21	5.48	4.95	5.28
welfare gain wrt baseline (%)	0.00	0.17	0.00	0.16	0.00	0.19

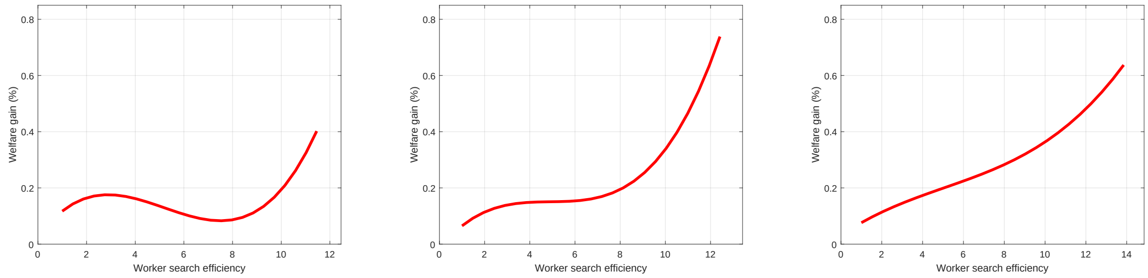
Table C5: Baseline vs No statistical discrimination, decomposition of net DD in job finding rate

	Baseline (%)	No interview cost (%)	Change (pp)	Elasticity
		Learning		
$\frac{d\ln o(x,\tau)}{d\tau}$	-0.21	-0.05	0.16	1.00
$\frac{d\ln a(\epsilon,\tau)}{d\tau}$	-1.38	-0.99	0.39	2.50
$\frac{d\ln f(\epsilon,\tau,x)}{d\tau}$	-1.59	-1.03	0.56	3.50
		Loss aversion		
$\frac{d\ln o(x,\tau)}{d\tau}$	-0.14	-0.00	0.14	1.00
$\frac{d\ln a(\epsilon,\tau)}{d\tau}$	-1.53	-1.24	0.27	2.02
$\frac{d\ln f(\epsilon,\tau,x)}{d\tau}$	-1.66	-1.19	0.46	3.02
		Duration-dependent app. cost		
$\frac{d\ln o(x,\tau)}{d\tau}$	-0.09	-0.00	0.09	1.00
$\frac{d\ln a(\epsilon,\tau)}{d\tau}$	-1.69	-1.57	0.12	1.28
$\frac{d\ln f(\epsilon,\tau,x)}{d\tau}$	-1.82	-1.55	0.27	2.28

Table C4 shows that job seekers react to a flat profile of the job offer probability per unit of search effort both by (i) reducing application effort at any duration and (ii)

Figure C8: Welfare gains across workers

(A) Learning (B) Reference dependence (C) Duration-dep. app. cost



Note: This figure reports the fitted percentage increase in welfare enjoyed by workers across the search efficiency distribution. Welfare gains are fitted via a 5-th order polynomial.

by scaling down application effort by less over the unemployment spell. Jointly, these two margins of adjustment imply that the local general-equilibrium elasticity is likely to over-estimate the *global* general-equilibrium elasticity of the job finding rate to the job offer probability per unit of search effort.[75] In Table C5 we revisit the relative importance of the indirect effect of statistical discrimination by leveraging Equation (C.13) as an accounting identity. As expected, the indirect effect of statistical discrimination gets dampened in magnitude compared to the local estimate. Specifically, the global general-equilibrium elasticity hovers around 3. However, the main message goes through: Statistical discrimination mainly affects the duration dependence in the job finding rate via its indirect effect on application effort.

Overall, we estimate that removing statistical discrimination would increase aggregate welfare by 0.2% consumption-equivalent units.[76] The muted aggregate welfare gain masks significant cross-sectional heterogeneity. Indeed, as shown in Figure C8, removing statistical discrimination mainly benefit the most productive workers (in terms of ability and search efficiency), who are relieved from the risk of being denied interviews for jobs they would have been qualified for. On the other hand, welfare of low productive workers is barely affected by statistical discrimination.

[75] Intuitively, both the intercept and the slope of the application effort function adjust downward. Instead, the local general-equilibrium elasticity attributes the entire adjustment to the slope.

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

D. Complementary explanations

Our structural model encompasses the the mechanisms most likely to drive our empirical findings, as discussed in Figure 4. In this section we provide some additional evidence on potential further mechanisms that are not directly captured in our structural model.

Human capital depreciation. Human capital depreciation over the course of unemployment as, for instance, in Ljungqvist and Sargent (1998) represents the main alternative (or complementary) explanation for negative dynamic selection on ability and negative duration dependence in application effort. On the one hand, even absent any cross-sectional heterogeneity in ability at the start of the unemployment spell, firms would statistically discriminate against long unemployment durations to the extent that a job seeker's ability deteriorates with the duration of unemployment. Note, however, that in our data substantial heterogeneity in the job offer probability per application (and application effort) remains after controlling for observed characteristics. This suggests that ability depreciation is unlikely to be the only determinant of dynamic selection.[77] On the other hand, job seekers would scale down application effort over the unemployment spell due to both firms' statistical discrimination for given individual ability and progressive ability downgrading. Hence, human capital depreciation generates negative duration dependence in application effort in a qualitatively similar way to our model of learning from search (how quickly the job offer probability per application declines would depend on the exogenous depreciation process rather than the endogenous job finding process). Our takeaway is that, while our model generates a duration profile of application effort that is able to replicate the empirical evidence, the precise reasons why job seekers get increasingly discouraged are still unclear. We conclude that the extent to which human capital depreciation drives workers' search behavior remains an open question. Also the recent empirical literature remains inconclusive on this issue (Cohen, Johnston, and Lindner, 2023; Dinerstein, Megalokonomou, and Yannelis, 2022; Arellano-Bover, 2022).

Changes in application quality. The observed decline in the interview probability may relate to changes in application quality over time: (part of) the downward-sloping

[77] Jarosch and Pilossoph (2019) estimates a model putting together both ability depreciation and cross-sectional heterogeneity as sources of dynamic selection and finds that ability should depreciate very slowly to be consistent with the observed decline in the interview rate.

duration profile in the interview probability could reflect a gradual downgrading of job application characteristics. In our context, we observe an important dimension of application quality: the channel used to contact the firm (Beaman and Magruder, 2012; Burks, Cowgill, Hoffman, and Housman, 2015; Hensvik and Nordström Skans, 2016). As we show in Figure D1, the application channel is strongly predictive of an application's success at the interview stage, with personal applications being more successful than written resumes or phone applications.

However, changes in application quality as captured by the application channel seem unlikely to represent the main driver of the effect of our duration results. On the one hand, the relative importance of each channel remains constant with duration, even after controlling for individual heterogeneity (see Figure D1). On the other hand, we still find evidence of a marked decline in the interview probability after controlling for the application channel in our regressions.[78]

Increasing the search radius. Another potential explanation for the decline in the interview probability lies in application targeting (Galenianos and Kircher, 2009; Wright, Kircher, Julien, and Guerrieri, 2021; Lehmann, 2023). Initially, job seekers might target a specific occupation, before starting to search more broadly and to apply to a wider set of job ads as unemployment duration increases. This may reduce callback chances, as job seekers are potentially less suited to the positions they newly apply to. If this mechanism is at play, we should observe adjustments in job search targets over time. We assess this point by studying how occupational targeting changes over time in the *auxiliary sample*, for which we have information on the occupation of the vacancies reported in the search diary. Specifically, we construct two measures that characterize the types of occupations job seekers target: a binary variable indicating whether the targeted occupation is the same as the occupation desired by the job seeker, and a measure of net cognitive requirements of targeted occupations.[79] The range of occupations for which

[78] If anything, we would expect (unobserved determinants of) application quality to be actually increasing over time, as job seekers learn how to make better applications over time. Such omitted determinants would induce an upward bias in the estimated duration profile that controls for observable characteristics, implying that the true duration dependence in the interview probability would actually be more negative.

[79] In our data, occupations are categorized according to the *Swiss Standard Classification of Occupations 2000* (SSCO 2000). This job nomenclature follows a hierarchical structure, and presents 5 different levels of occupational groups. The binary indicator for occupational similarity between the desired

job seekers apply, as measured by the same-occupation indicator (Figure D2A), hardly changes with unemployment duration, regardless of whether we control for job seeker fixed effects or not. At short unemployment durations, job seekers apply to occupations that have on average higher cognitive requirements than physical requirements, whereas job applications later in the spell target less cognitively intense occupations. However, the decline in cognitive intensity of target occupations is strongly attenuated once job seeker fixed effects are added (Figure D2B). This suggests that the decline in cognitive intensity is largely driven by the changing composition of the pool of unemployed rather than by a change application targeting within individuals. Altogether, these pieces of evidence point towards a limited role of application targeting in the decline of the interview probability.

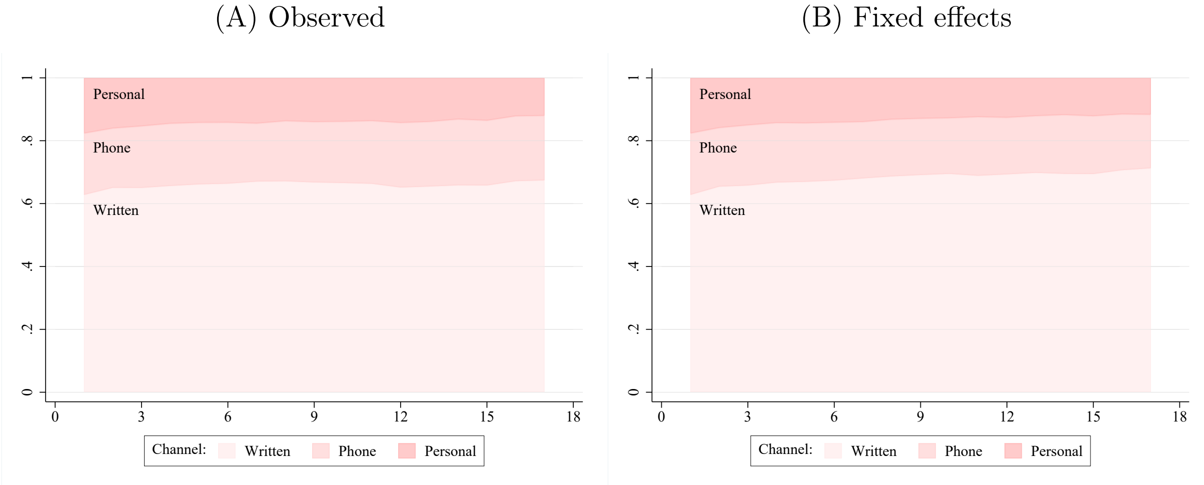
Table D1: Duration profile of applications by channel

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.

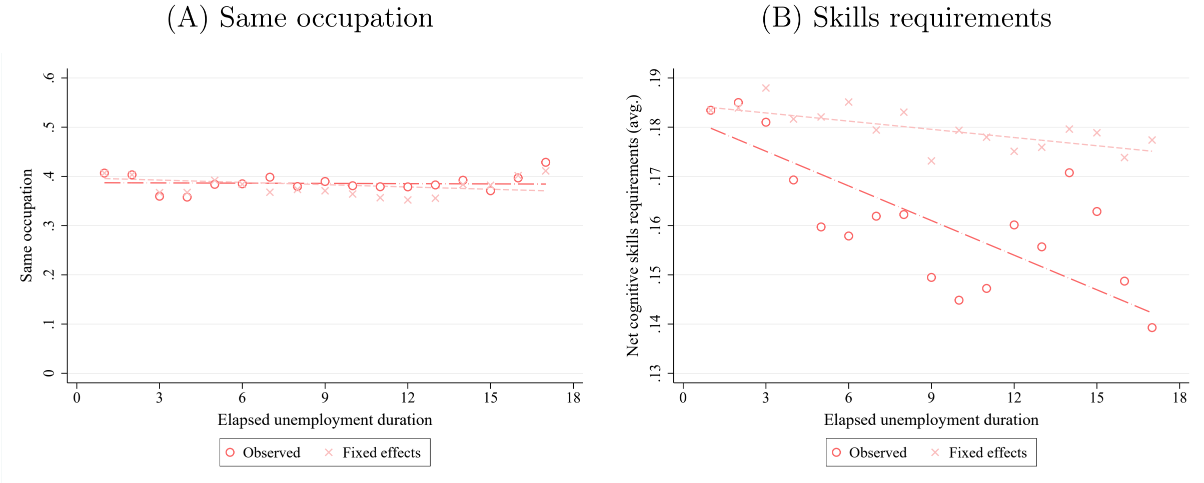
occupation (at the spell level) and targeted occupation (at the application level) can be constructed for the different levels of the SSCO 2000. As for the net cognitive requirements measure, we use *O*Net* skill and ability requirements for each occupation. *O*Net* provides 52 abilities and skills, grouped into cognitive and physical. Our net cognitive measure is based on the difference between weighted importance of cognitive skill requirements and physical requirements.

Figure D1: Changes in the shares of application channels



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.

Figure D2: Changes in application targeting



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).

Table D2: 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.