12-2020

Decomposing duration dependence: Skill depreciation vs. statistical discrimination

Ismail BAYDUR
Singapore Management University, ismailb@smu.edu.sg

Jianhuan XU
Singapore Management University, jhxu@smu.edu.sg

Decomposing Duration Dependence: Skill Depreciation vs. Statistical Discrimination*

Ismail Baydur†
Singapore Management University

Jianhuan Xu‡
Singapore Management University

May 2020

Abstract

This paper develops a random matching model with unobserved worker heterogeneity and learning about worker types from unemployment duration. The model features negative duration dependence that stems from unobserved heterogeneity as well as statistical discrimination and skill depreciation. We estimate our model using micro-level data from Current Population Survey (CPS) and we decompose the contribution of each channel to job finding rates by duration. We find that shutting down statistical discrimination substantially increases the job finding rates of the long-term unemployed while skill depreciation mainly affects the medium-term unemployed.

*We thank Nicolas Jacquet, Thomas Sargent, Sephorah Mangin, Costas Meghir, Gueorgui Kambourov, Dirk Krueger, and Jeremy Greenwood for their comments and suggestions. We also thank seminar and conference participants at City University of Hong Kong, GRIPS, Kyushu University, National University of Singapore, Singapore Management University, Econometric Society conferences in Seoul and Auckland, and Midwest Macro Meetings in Michigan State University. Qiugu He provided valuable research assistance. This study is funded through a research grant (Fund No: C244/MSS18E001) from Singapore Management University. All errors are our own.

†Contact information: ismailb@smu.edu.sg

‡Contact information: jhxu@smu.edu.sg

Keywords: unemployment duration dependence, skill depreciation, statistical discrimination
JEL Classifications: E24,E32,J64

1 Introduction

The fraction of the unemployed who find a job within a month falls dramatically by unemployment duration. This feature of the US labor market, which is referred as negative duration dependence, has been studied by a large body of literature.[1] There are three prominent explanations for this observation. The first of these explanations asserts that negative duration dependence is simply a reflection of the "composition effect". Because workers with high job finding rates find work relatively quickly, the unemployment pool becomes populated with more individuals who have difficulty in finding a job. This dynamic selection creates leads to a decline in job finding rates by unemployment duration. The other two explanations, skill depreciation and statistical discrimination, imply negative duration dependence at an individual level (i.e. "true" duration dependence). The former argues that people lose their skills while unemployed and therefore have difficulty in finding jobs as their unemployment spell extends. The latter is based on the argument that firms use unemployment duration as a signal for a job applicant's capabilities. Individuals with less competitive skill-sets remain longer in the unemployment pool and firms factor this information into their hiring decisions.

In this paper, our goal is to empirically assess the contribution of these sources to the negative duration dependence for the US economy. To this end, we develop an equilibrium model of unemployment with matching frictions and negative unemployment duration dependence. In our model, hiring occurs in two stages: matching and interviewing. The first stage is a standard Diamond-Mortensen-Pissarides (DMP) model, where workers randomly meet with vacancies at a potentially time varying rate determined via an aggregate matching function. However, matching does not guarantee hiring. In the second stage, firms matched with a worker decide whether to go through a costly interview process and hiring only occurs if the outcome of the interview is successful. Both the cost of the interview and its outcome are random draws.

A key feature of our model is that workers are heterogeneous with respect to their

[1] More recently, Barnichon and Figura (2015) and Hall and Schulhofer-Wohl (2018) show this phenomenon in their estimation of aggregate matching function with micro-level data.

ability to turn interviews into hiring, which we call skill. There are specifically two skill types and one of these types (e.g., high-skilled) is better at turning interviews into hiring than the other type (e.g., low-skilled). Heterogeneity among the workers is the basis for the three sources of negative duration dependence. The first is due to the composition effect as high-skill workers leave the unemployment pool faster. Second, we model skill depreciation as an exogenous process whereby high-skilled workers lose their abilities and become low-skilled at a constant rate as they remain in the unemployment pool. The loss of skills also leads to negative duration dependence.[2] Lastly, the true type of a worker is not observed, but the firm can infer the probability that a worker is high-skilled from his unemployment duration. We call this probability the *resume* of the worker. Firms update these resume values based on the differences between the job finding rates of the high and low-skilled workers. Given that the interview cost is random, firms interview applicants only if the prospects are favorable relative to the cost of the interview. Because there are more low-skilled workers at longer durations, firms are less likely to interview the long-term unemployed. Interviewing the applicant selectively, which we call statistical discrimination, also generates declining job finding rates with unemployment duration. We parsimoniously parametrize the distribution of the interview costs with one parameter so that only a fraction of the firms choose not to interview the applicant. We associate this parameter with the statistical discrimination channel.[3]

We then estimate our model via maximum likelihood using individual level data from Current Population Survey (CPS) and aggregate level data from Help-Wanted-Online-Index (HWOL). Our model allows us to implement a resume updating rule and assign a resume value for each observation in our sample. Our identification of the statistical discrimination channel relies on the fact that resume values are sensitive to changes in aggregate market tightness. For example, job finding rates are

[2]In the literature, skill depreciation during unemployment is generally related to human capital loss and declining productivity. See for example Acemoglu (1995) and Ljungqvist and Sargent (1998). In our model, skill loss is defined as a reduction in a worker's abilities at the interview stage. From this perspective, skill loss may also be thought as a worker's loss of self-confidence after performing poorly at successive interviews.

[3]Barron and Bishop (1985) find that time spent by company personnel for recruiting, screening and interviewing job applicants is on average 9.87 hours per hire with a standard deviation of 17.16 hours. This finding is suggestive of varying interview costs.

low for everyone during a recession and therefore resume values are updated slowly. Consequently, individuals with identical characteristics but different aggregate labor market history have different resume values conditional on duration. We use this variation in resume values to estimate the parameter governing the statistical discrimination channel.

We assess the significance of each channel by setting the associated parameter value to zero and then calculating counter-factual job finding rates from a simulated model. We find that shutting down the statistical discrimination channel increases the job finding rate of a worker at 24 months duration relative to a recently unemployed worker by 75%. The corresponding number after shutting down skill depreciation is negligible, and the remaining difference is solely due to the composition effect. However, at its estimated value, skill depreciation makes duration dependence relationship less responsive to the changes in aggregate market tightness. When there is no skill depreciation, the duration dependence relationship becomes more responsive to the changes in aggregate labor market conditions.

Our structural estimation naturally relies on strong assumptions about the hiring process. Nonetheless, the fitted job finding rates by unemployment duration track the direct measures of the flows from unemployment to employment quite well. Moreover, we find that the duration dependence during a boom is stronger at short durations but weaker at longer durations. Abbring, Van Den Berg, and Van Ours (2001) reaches to a similar conclusion after estimating the job finding rates from the CPS data using a flexible reduced form specification.

Kroft, Lange, and Notowidigdo (2013) conduct a field experiment with fake resumes in which individuals differ only by their unemployment duration. They show that first call for an interview decreases with unemployment duration. This finding has been taken as evidence for statistical discrimination, because the experimental design addresses the concern for unobserved heterogeneity in studies estimating job finding rates from observational data. Our modeling of hiring process for statistical discrimination is motivated by their findings. In the same study, the authors also show that this negative relationship weakens during recessions, which is the basis for our identification. Because we observe the hiring decisions in CPS data, our findings are directly related to that margin although interview decisions are parallel to the hiring decisions in our model.

In a recent paper, Jarosch and Pilossoph (2019) argue that eliminating statistical discrimination does not have any significant effect on job finding rates of the long-term unemployed workers. In contrast, we find that statistical discrimination has the largest impact among other possible sources. Although the two models have several similar features and they both capture the job finding rates by unemployment duration quite well, the mechanisms that generate statistical discrimination are different. In our model, interview costs are random and long-term unemployed are called for an interview only if the interviewing cost is sufficiently small. Jarosch and Pilossoph (2019), on the other hand, assume a constant value for the cost of interviewing applicants. Instead, they assume a continuum of worker skills. Interview probability in their model still falls with unemployment duration because firms have minimum hiring requirements and they are heterogeneous with respect to these requirements. Workers at the bottom of the skills set are less likely to satisfy these requirements, and therefore they remain longer in the unemployment pool. Consequently, only a small fraction of the firms finds it worthwhile to interview the long-term unemployed. When they set the cost of interview to zero, the interview probability increases, but the job finding rate of the workers with relatively lower skills only marginally changes because they still cannot satisfy the minimum hiring requirements for most of the jobs. Whether statistical discrimination stems from variation in interview costs or in minimum hiring requirements (or both) is a modeling assumption and neither the CPS data nor the experimental data with fake resumes are informative about it. Nonetheless, we argue that the distinction matters for assessing the role of statistical discrimination in explaining the job finding rates of the long-term unemployed.

Our model is also complementary to the directed search model with resume updating in Doppelt (2016) but it differs in two aspects. First, the two papers propose very different mechanisms to explain duration dependence although they both find that statistical discrimination plays an important role in generating heterogeneity in job finding rates. The second difference is methodological. Rather than calibrating our parameters to aggregate moments, we directly estimate our model parameters using micro-level data and therefore we also control for observables. For identification purposes, we also introduce time variation in market tightness via an aggregate shock to the cost of vacancy posting.

Finally, Alvarez, Borovičková, and Shimer (2016) decomposes duration dependence in an optimal stopping time model. Using micro-level data from Austria, they show that unobserved heterogeneity is a critical source in explaining duration dependence. We also find a substantial role for unobserved heterogeneity, but statistical discrimination is still the dominant factor in explaining the negative duration dependence.

The rest of the paper is organized as follows. The next section describes our model. We then describe our dataset and estimation procedure. In Section 4, we perform our counter-factual analysis using our model and estimated parameters. The last section concludes.

2 Model

2.1 Environment

The economy is populated by infinitely-lived risk-neutral workers with a unit measure and a large number of risk-neutral firms. Time is continuous and t denotes the calendar time. The discount rate is $\rho > 0$.

At any instant, a firm can either be active or inactive. Each active firm has one job that is either filled and producing or vacant and searching. A vacant firm attracts unemployed workers by posting a vacancy. Inactive firms can become active by creating a vacancy and start searching for a worker. The flow cost of vacancy posting, $\kappa(t)$, is common across the firms but is varying over time.

At any instant, a worker can either be employed or unemployed. An employed worker engages in production and becomes unemployed at a constant rate $\lambda > 0$. While employed, he cannot search for a job. An unemployed worker searches for a job in a frictional labor market. To be hired by a firm, an unemployed worker not only needs to meet a vacancy, but he also needs to go through a costly interview process. Hiring occurs only if the outcome of the interview is successful.

Workers are heterogeneous with respect to their ability to turn interviews into job offers. Each worker can be one of the two possible types, H and L, which we call high-skilled and low-skilled workers, respectively. A high-skilled worker is relatively more successful at the interview stage. Workers draw their skill types each time that they become unemployed from a fixed Bernoulli distribution, with r_0 denoting the probability of being high-skilled. While unemployed, a high-skilled worker loses his skills at a constant rate $\delta \geq 0$ (e.g. skill depreciation), and becomes low-skilled.

The *true* skill type of an unemployed worker is unobserved. For each unemployed worker, there is a public belief about their type, which denotes the probability of being a high-skilled worker. Let $r(\tau, t_0)$ denote this probability for an unemployed worker at duration τ of his unemployment spell who has become unemployed at calendar time $t_0 = t - \tau$. We call $r(\tau, t_0)$ the resume of worker. The resume value for all the workers entering the unemployment pool is equal to r_0 but their value changes during their unemployment spell. Because workers are heterogeneous with respect to their skills and these skills are subject to change while unemployed, the resume values can be updated over time. We explain this updating rule in greater detail below. One implication of our updating rule is that workers who enter the unemployment pool at the same time have identical resume values until they exit unemployment. Hence, $r(\tau, t_0)$ is the common resume value of workers who have become unemployed at calendar time t_0.

2.2 Match surplus

A matched firm-worker pair produces a flow value of output, y, and the flow value of unemployment is $b < y$. Both of these variables are constant over time and are independent of the worker's true skill type. Because labor markets are frictional, a firm-worker match creates a surplus. We assume that the firm gets the entire surplus value (e.g. workers have no bargaining power). Under this assumption, the wage rate is equal to b and the value of a filled position at time t, $J(t)$, satisfies the following Bellman equation:

$$\rho J(t) = y - b + \dot{J}(t) - \lambda(J(t) - V(t)), \tag{1}$$

where $V(t)$ is the value of a vacancy and "$\dot{}$"(dot) notation indicates time derivative. Intuitively, a filled position is an asset for the firm and its discounted value, $\rho J(t)$, is equal to the return on this asset given by the terms on the right-hand side. A filled position generates a flow profit of $(y - b)$, and its value potentially changes over time, $\dot{J}(t)$. In addition, the job becomes vacant at rate λ, in which case the firm loses $(J(t) - V(t))$.

The time dependence of $J(t)$ stems from the fact that $V(t)$ may change over time. Eventually we impose a free entry condition in equilibrium such that $V(t) = 0$ for all t so that the value of a filled position is also constant over time, $J(t) = \bar{J}$. After re-arranging equation (1) under the free entry condition, we obtain:

$$\bar{J} = \frac{y - b}{\lambda + \rho}. \tag{2}$$

Note that the value of a filled position is also equal to the total surplus value because the worker has no bargaining power.

The match surplus is constant over time and is the same for all the jobs. These results admittedly rest on two strong assumptions: constant productivity and constant wage rate. Nonetheless, they dramatically simplify the estimation procedure. In particular, we are able to calculate the job finding rates solely based on the market tightness and individual characteristics, which we directly observe in the data. Relaxing either of these assumptions would make the match surplus depend on aggregate state variables such as the aggregate productivity or even the entire distribution of individual states. In this case, not only would the value function calculations be tedious but the job finding rates would also depend on some aggregate state variables that we do not directly observe in the data.

2.3 Matching frictions, interview costs, and job finding rates

We model the hiring process as a two-stage process. In the first stage, unemployed workers and vacancies are randomly matched in a labor market with matching frictions. This stage is essentially the canonical DMP model. In the second stage, interview and hiring decisions are made. Each firm matched with a worker draws a random interview cost. The interview takes place only if the firm pays the interview cost and hiring occurs if the outcome of the interview is successful. Otherwise, both the firm and the worker continue to search for alternatives. If the interview cost is sufficiently high, then the firm may opt not to interview the applicant at all.

We capture the matching frictions in the first stage of the hiring process via an aggregate matching function that depends on the stock of job openings, $v(t)$, and the total mass of unemployed workers, $u(t)$. Total matches, $m(t)$, are given by the following Cobb-Douglas matching function:

$$m(t) = \eta \left[v(t)\right]^{1-\sigma} \left[u(t)\right]^{\sigma}, \tag{3}$$

where $\sigma \in (0, 1)$ is the shape parameter of the matching function and η is the matching efficiency of the unemployed workers at time t. The overall contact rate, $f(t)$, is common across the workers and is equal to:

$$f(t) = \frac{m(t)}{u(t)} = \eta \left[\theta(t)\right]^{1-\sigma}, \tag{4}$$

where we define the market tightness, $\theta(t)$, as the ratio of vacancies to unemployed: $v(t)/u(t)$.

On the firm side, each vacancy has an equal chance to contact a worker. Hence, a firm contacts workers at rate:

$$q(t) = \frac{m(t)}{v(t)} = \eta \left[\theta(t)\right]^{-\sigma}. \tag{5}$$

Not all of the matched workers are eventually hired. In the second stage, each firm matched with a worker randomly draws a cost for interviewing the applicant, c, from a fixed distribution $F(c)$. At this stage, the firm observes the resume value of the applicant. Without knowing his true skill type, the firm decides on whether to pay the cost and interview the applicant or continue searching for alternatives. If the firm eventually decides to interview the applicant, then the worker-firm pair learn the productivity of their match in the third stage. The productivity of a match is a random draw and can only take two values: 0 or y.[4] A high-skilled worker draws y with probability p^H, in which case an employment relationship is formed and the firm-worker pair starts producing. Otherwise, with probability $1 - p^H$, hiring does not occur. A low-skilled worker draws y with probability $p^L < p^H$. Therefore, when making a decision about interviewing an applicant with a resume value r, the firm compares the cost of interview with its *expected* value, $\hat{W}(r)$, which is equal to:

$$\hat{W}(r) = rp^H \bar{J} + (1 - r)p^L \bar{J}. \tag{6}$$

The expected value of the interview is linear and monotonically increasing in r. Accordingly, the interviewing decision has a cut-off property in that the firm interviews the applicant only if $c < \hat{W}(r)$. For a worker with resume value r, the probability of being interviewed, conditional on a match, is equal to $F(\hat{W}(r))$, which is also increasing r. In other words, conditional on a match, workers with high resume values are more likely to be interviewed.

We further assume that the distribution of interview cost, $F(c)$, has the following specification. With probability $(1-\alpha)$, the cost of interview is equal to zero and the firm interviews any applicant. With probability α, the interview cost comes from a uniform distribution between $\hat{W}(0)$ and $\hat{W}(r_0)$ so that the interview probability is equal to r/r_0 (i.e. recently unemployed are interviewed with certainty conditional on contacting with a vacancy). Overall, a worker with resume value r expects to be interviewed with probability $(1 - \alpha + \alpha r/r_0)$. This specification allows us to study the special case when $\alpha = 0$ (i.e. interviewing is costless for every firm) and it is at the center of our discussion about statistical discrimination below.

We can write the job finding rate for an unemployed worker at duration τ conditional on his true skill type, $s = \{H, L\}$, as follows:

$$h(\tau, t_0|s) = \eta \left[\theta(t_0 + \tau)\right]^{1-\sigma} (1 - \alpha + \alpha r(\tau, t_0)/r_0)p^s. \tag{7}$$

Because the interview probability is increasing with the resume value, workers with high resume values find jobs faster for a given skill type. Moreover, for a given resume value r, high-skilled workers can find jobs faster because they are more

[4] In principal, we could replace the outcome 0 with any value less than b.

likely to form a productive match at the interview stage.

2.4 Resume updating

Firms can make an inference about the true skill type of an unemployed worker from his unemployment duration by exploiting the difference between the job finding rates of high and low-skilled workers. We employ a Bayesian updating rule so that the job finding rates for different skill levels in equation (7) implies the following updating rule for the resume values at duration τ:

$$\frac{\dot{r}(\tau, t_0)}{r(\tau, t_0)} = -\delta - (1 - r(\tau, t_0))\, \eta\, [\theta(t_0 + \tau)]^{1-\sigma}\, (1 - \alpha + \alpha r(\tau, t_0)/r_0)\Delta_p, \quad (8)$$

where we define $\Delta_p = p^H - p^L$. We obtain this expression as a limiting case of the Bayes rule in a discrete time approximation and we defer the details to the Appendix. Given the resume value at the beginning of the unemployment spell, r_0, along with the path of $\theta(t)$, firms can calculate the evolution of resume values at every unemployment duration $\tau \in [0, t - t_0]$ using the differential equation in (8).[5]

2.5 Equilibrium

The value of a vacant position satisfies the following Bellman equation:

$$\rho V(t) = -\kappa(t) + \dot{V}(t) + q(t)(W(t) - V(t)), \quad (9)$$

where $W(t)$ is the expected value from searching conditional on matching with a worker at time t. The interpretation of equation (9) is similar to the interpretation of the Bellman equation for a filled position. The discounted value of a vacancy is equal to the expected return from searching. While vacant, firms incur the flow cost of vacancy, $\kappa(t)$, and gain or lose potentially from the change in the value of vacancy

[5]Note that if the worker finds a job at the end of his unemployment spell, then his resume value jumps to a higher value. Nonetheless, because the resume value does not affect the match surplus and its value is reset to r_0 when the worker becomes unemployed again, its exact value is redundant for calculating the equilibrium.

over time. The firms also contact a worker at the market rate $q(t)$. Conditional on a match, they expect to gain $(W(t) - V(t))$.

The value of $W(t)$ depends on the distribution of unemployed workers. Let $\Gamma(\tau, t_0)$ denote the measure of workers who lost their job at time t_0 and are currently at duration τ. Then, using the identity $t_0 = t - \tau$, we can calculate $W(t)$ as follows:

$$W(t) = \int_0^\infty \left\{(1-\alpha)\hat{W}(r(\tau, t-\tau)) + \alpha \frac{r(\tau, t-\tau)}{r_0}\left[\frac{\hat{W}(r(\tau, t-\tau)) - \hat{W}(0)}{2}\right]\right\} \frac{\Gamma(\tau, t-\tau)}{u(t)} d\tau, \tag{10}$$

where the aggregate unemployment rate is given by $u(t) = \int_0^\infty \Gamma(\tau, t-\tau)d\tau$. The interpretation of this equation is as follows. Upon meeting an unemployed worker at duration τ, with probability $(1-\alpha)$, the firm draws zero interview cost and enjoys the expected value of the interview, $\hat{W}(r(\tau, t-\tau))$, given in equation (6). Otherwise, the firm draws an interview cost and interviews the worker with probability $r(\tau, t-\tau)$. Conditional on the interview, the interview cost is a uniform random variable between $\hat{W}(0)$ and $\hat{W}(r(\tau, t-\tau))$ and its expected value is the simple average of the two. Subtracting this expected interview cost from $\hat{W}(r(\tau, t-\tau))$ yields the last term in equation (10). Integrating over all unemployment durations gives the value of $W(t)$. Using the definition of $\hat{W}(r)$, we can further simplify this expression as:

$$W(t) = \bar{J}\int_0^\infty \left\{(1-\alpha)(p^L + r(\tau, t-\tau)\Delta_p) + \alpha\frac{r(\tau, t-\tau)^2\Delta_p}{2r_0}\right\} \frac{\Gamma(\tau, t-\tau)}{u(t)} d\tau, \tag{11}$$

If we impose a free entry condition in that the value of a vacant position is always equal to zero, $V(t) = 0$ and replace $q(t)$ from equation (5), then we obtain from equation (9):

$$\kappa(t) = \eta\,[\theta(t)]^{-\sigma}\,W(t). \tag{12}$$

Note that $W(t)$ can be calculated independent of $\theta(t)$ from equation (10). Given its value, $\kappa(t)$ then pins down the equilibrium value of $\theta(t)$ from the free entry condition in equation (12).

Finally, due to Bayesian updating, the resume values truly reflect the proportion of high-skilled workers among those who has the same resume value. Accordingly, we can write the evolution of $\Gamma(\tau, t_0)$ over time as follows:

$$\frac{\dot{\Gamma}(\tau, t_0)}{\Gamma(\tau, t_0)} = -\left(r(\tau, t_0)h(\tau, t_0|H) + (1 - r(\tau, t_0))h(\tau, t_0|L)\right), \tag{13}$$

given $\Gamma(0, t_0) = \lambda(1 - u(t_0))$. Equation (13) states that the measure of the unemployed at a given duration decreases over time at a rate equal to the *average* job finding rate at that duration, which depends on the share of the high and low-skilled workers. This share is equal to the resume value at that duration.

Given these equations, we can define a dynamic equilibrium as follows:

Definition 1. A dynamic equilibrium is a sequence of $\{\theta(t), q(t), W(t), \Gamma(\tau, t_0), r(\tau, t_0)\}$ such that:

1. $q(t)$ and $\theta(t)$ are related according to equation (5).
2. $W(t)$ is given in equation (11).
3. $r(\tau, t_0)$ is updated according to equation (8) given that $r(0, t_0) = r_0$.
4. $\Gamma(\tau, t_0)$ satisfies equation (13) given $\Gamma(0, t_0) = \lambda(1 - u(t_0))$ and the hazard rates by skill type, $h(\tau, t_0|s)$, in equation (7).
5. Given $\kappa(t)$, a free entry condition, $V(t) = 0$, determines equilibrium $\theta(t)$ from equation (12).

As we discuss further below, time variation in $\theta(t)$ is necessary for estimating α and we can achieve this in the model through changes in $\kappa(t)$ over time.[6] In principle, we could specify a stochastic process for $\kappa(t)$, generate a long time series on small time intervals, simulate the distribution of workers, and calculate equilibrium $\theta(t)$ over time. Nonetheless, we are ultimately interested in how unemployment duration dependence changes when we restrict α and δ. We performed this comparison at the steady-state where we target the average market tightness in the data. To show the effects of changes in $\theta(t)$ on duration dependence, we change the value of $\kappa(t)$ to a new value associated with that steady state equilibrium value of $\theta(t)$.

[6]Alternatively, from equations (11) and (12), we could generate cyclical variation in $\theta(t)$ by assuming a time variation in $\bar{J}$.

2.6 Decomposing duration dependence

Equation (8) implies that resume values decrease with unemployment duration and, from equation (7), it generates a negative duration dependence in job finding rates. In this section, we illustrate the various channels causing negative unemployment duration dependence under specific parameter restrictions for a group of workers who have become unemployed at time t_0. Consider the following special cases.

First, assume that $\delta = \alpha = 0$. In this case, the job finding rates in equation (7) are independent of the resume value of a worker. Moreover, because there is no skill depreciation, the job finding rates are constant conditional on $\theta(t)$ and the true skill type at the beginning of the unemployment spell. In other words, there is no *true* duration dependence. However, the job finding rates are still declining with unemployment duration at the population level, because high-skilled workers leave the unemployment pool faster than the low-skilled workers. To see this point, let us write the average job finding rate for this group over time as follows:

$$h^p(\tau, t_0) = \eta \left[\theta(t_0 + \tau)\right]^{1-\sigma} (r(\tau, t_0)p^H + (1 - r(\tau, t_0)p^L)),$$

where $r(\tau, t_0)$ is the common resume value at duration τ. According to (8), $r(\tau, t_0)$ is declining with unemployment duration, and because $p^H > p^L$, the job finding rate at the population level is also declining. We call this channel the "composition effect". We note that this channel is weaker when $\theta(t)$ is lower (e.g., during recessions) because even high-skilled workers are unemployed for a longer period of time during a recession.

Next, consider the case when $\delta = 0$ but $\alpha > 0$. In this case, firms care about the resume value of the worker and the resume values explicitly enter the job finding rates described in equation (7). The existence of interview costs generates the "statistical discrimination" channel. Because resume values decline with unemployment duration, workers find it harder to receive interviews because they remain longer in the unemployment pool. This channel is different than the composition effect in that even a truly high-skilled worker faces declining job finding rates by duration. Nonetheless, both the statistical discrimination channel and the composition effect imply that resume updating is slower when $\theta(t)$ is lower (e.g., during recessions).

Finally, consider the case when $\delta > 0$ but $\alpha = 0$. The job finding rates are independent of the resume value of the worker, but a high-skilled worker may become low-skilled while unemployed. Therefore, in addition to the composition effect above, we have an additional source of duration dependence, which we call "skill depreciation" channel. Note that this channel generates a true duration dependence in that the job finding rate for a high-skilled worker, conditional on $\theta(t)$, drops from p^H to p^L at the instant when he exogenously loses his skills. The *average* job finding rate of the unemployed workers who enter the unemployment pool as high-skilled is related to τ as follows:

$$\bar{h}^H(\tau, t_0) = \eta \left[\theta(t_0 + \tau)\right]^{1-\sigma} (\exp\{-\delta\tau\}p^H + (1 - \exp\{-\delta\tau\})p^L).$$

Because skill depreciation is an exogenous process, it is independent of the aggregate market tightness once we control for the true skill type at the beginning of the job spell and $\theta(t)$. This feature of the skill depreciation channel is different than both the composition effect and the statistical discrimination channels. We exploit the business cycle variation in market tightness to identify these channels in the data.

3 Estimation

3.1 Data Description and Sample Restrictions

We use micro-level data from the CPS obtained from the Integrated Public Use Microdata Series (IPUMS-CPS) database (Flood et al. (2020)). CPS uses rotation groups where each individual is interviewed for four consecutive months, not interviewed for eight months, and then reinterviewed for another four months before leaving the sample. The rotating panel design of CPS allows us to observe the employment status of each individual next month. In addition, the unemployed individuals are asked to report their duration of unemployment, which allows us to create a market tightness history during their current unemployment spell.

We also collect information about age, gender, education, region, occupation, industry, class of worker, race, marital status, and reason for unemployment. The CPS monthly files contain inconsistent information, therefore we drop the individuals whose gender or race changes across different surveys. We also drop individuals when their age decreases, or when their age increases by two or more years. We perform our analysis with white males only.

A worker's age changes while unemployed, but the reference month for the reported age is the age during the interview and not the age of the worker when he entered the unemployment pool. For example, if the worker remains in the unemployment pool two years, then his age is substantially different than his age on entry to the unemployment pool. To address this concern, we randomly assigned a month component to the reported age in years and we indicate each individual's age at the time of the interview in months. We restrict our sample to those individuals who were at least 24 years old when their unemployment spell started. Consequently, we first normalized the age variable so that an individual who is 24 years old and is randomly assigned to January is one month old. Then, we excluded unemployment spells if an individual's adjusted age is negative when he started his current unemployment spell. We also dropped individuals whose adjusted age is greater than 384 months, or 55 years, because interview and hiring decisions might be substantially different for those near retirement (e.g., due to short expected tenure at the time of the hiring).

We create four occupation categories: (a) professional and managerial, (b) personal services,(c) sales and office, and (d) production. Industry variable indicates whether the individual is in the goods or services producing sector. We define three education categories: high school and less than high school, some college, and at least a 4-year college degree. The marital status variable indicates whether the respondent is married. We excluded family workers and self-employed individuals from our sample. We indicate government jobs by the class of worker variable.

IPUMS-CPS provides six categories for reason of unemployment: (a) job loser or on layoff (b) other job loser, (c) end of temporary job, (d) job leaver, (e) new-entrant, and (f) re-entrant. Fujita and Moscarini (2017) report that many temporary lay off workers are called back by their previous employer and their duration dependence is different than other workers. Therefore, we exclude temporary layoff or end of temporary jobs. Moreover, new-entrants do not report an occupation or industry. Consequently, we do not have any observation in this category.

Apart from the micro-level CPS data, we obtained market tightness data from Barnichon (2010) available monthly and updated until December 2016. This dataset is constructed using vacancies data HWOL and HWI after correcting for the bias due to the secular decline in paper based job ads.

Our final sample consists of 106,493 observations covering the period from September 1995 to December 2016. Because the market tightness data is seasonally adjusted, we also include monthly seasonal dummies in our regression. Finally, CPS uses a highly stratified sampling scheme. To have a representative sample of the population, we use the longitudinal sampling weights for the two consecutive months available at IPUMS-CPS database.

3.2 Maximum Likelihood Framework

Let x_{it} be a vector of regressors including a constant in month t for individual i who is recorded as unemployed in the same month. This vector includes individual characteristics that are constant throughout the unemployment spell, such as education, and also market tightness as a time-varying regressor. Let y_{it} be equal to 1 if an individual is recorded as employed in the following month.[7] In addition to the regressors and employment status data, we utilized other information, such as calendar time indicators for t_0 and unemployment duration τ, to calculate individual resume values denoted by $r_i(\tau, t_0)$.[8]

We define the hazard function for individual i at duration τ conditional on observables and his skill as follows:

$$h_i(\tau, t_0|x_{it}, L) = \quad \exp\left(x_{it}\beta\right)\left(1-\alpha+\alpha r_i(\tau,t_0)/r_0\right)\frac{1-\Delta_p}{2} \qquad (14)$$

$$h_i(\tau, t_0|x_{it}, H) = \quad \exp\left(x_{it}\beta\right)\left(1-\alpha+\alpha r_i(\tau,t_0)/r_0\right)\frac{1+\Delta_p}{2}, \qquad (15)$$

[7] We treat switching to not-in-the-labor force status as an unsuccessful search.

[8] In the original dataset, unemployment duration is recorded in weeks. To facilitate the calculation of resume values, we convert these quantities to months assuming that the interview takes place in the middle of the month.

where α and Δ_p are defined as in the model and to be estimated from the data. Note that we implicitly normalize the hiring probability for high and low-skilled workers, which we discuss in the identification section later on. Moreover, x_{it} includes a constant and natural log of market tightness and this specification is consistent with the Cobb-Douglas form that we employed in the model section for the matching function. The coefficient estimates for the constant and the market tightness refer to $\log(\eta)$ and $(1-\sigma)$ in the model, respectively. Moreover, we introduced other regressors proportionally to the job finding rate, similar to the market tightness, and this specification can be rationalized under some mild assumptions about the matching function.

Our data is monthly and we assume that time-varying regressors are constant throughout the month. Similarly, we assume that $r_i(\tau, t_0)$ is constant during the month and use the following discrete version of the resume updating rule:

$$r_i(\tau, t_0) = \frac{r_i(\tau, t_0)\exp(-\delta - h_i(\tau, t_0|x_{it}, H))}{r_i(\tau, t_0)\exp(-h_i(\tau, t_0|x_{it}, H)) + (1 - r_i(\tau, t_0))\exp(-h_i(\tau, t_0|x_{it}, L))} \tag{16}$$

Given the initial resume value r_0, we can think of $r_i(\tau, t_0)$ as function of data and the model parameters. Let us define Ω_{it} as the set of all the information for individual i and ξ as the set of all the model parameters and write $r_i(\tau, t_0)$ compactly as $r_{it}(\xi)$. Then, the likelihood contribution of individual i can be written as:

$$\begin{aligned} L_i(\xi, r_{it}(\xi)|\Omega_{it}) \quad &= y_{it}\left(r_{it}(\xi)(1-\exp(-h_{it}^H)) + (1-r_{it}(\xi))(1-\exp(-h_{it}^L))\right) \\ &\quad +(1-y_{it})\left(r_{it}(\xi)\exp(-h_{it}^H) + (1-r_{it}(\xi))\exp(-h_{it}^L)\right). \end{aligned} \tag{17}$$

where we use h_{it}^L and h_{it}^H as a short-hand notation for the expressions in equations (14) and (15), respectively. The log-likelihood function can be written as:

$$\mathcal{L}(\xi) = \sum_{i=1}^{N} \omega_i \log\left(L_i(\xi, r_{it}(\xi)|\Omega_{it})\right), \tag{18}$$

where ω_i is the sampling weight for individual i. Because we calculate the resume values conditional on the data and the model parameters, the resume value of individual i at time t is a function of the model parameters, individual characteristics at time t, and the entire market tightness history for his current unemployment spell. When we maximize the log-likelihood function with respect to ξ, we take the dependence of the resume values on other model parameters into account and we employ the chain rule to calculate the gradient of the likelihood.

3.3 Identification

Consider our hazard rate specification in equations (15) and (14) and, for illustration purposes, suppose that $r_i(\tau, t_0)$ is a function of unemployment duration alone. In that case, we can write the hazard rate at duration τ for a given skill type as follows:

$$h(\tau|x_{it}, s) = \exp(x_{it}\beta)h_0(\tau)p^s. \tag{19}$$

This specification is a proportional hazard model with random frailty.[9] In this specification, $h_0(\tau)$ is called the baseline hazard function and it can be estimated parametrically or non-parametrically. p^s is called the random frailty, which is unobserved but is known to come from a known distribution. We discuss the identification of our model parameters by comparing our model to the proportional hazard model given above. Unlike the estimation of the hazard models, we construct our likelihood function in equation (17) as a complementary log-log function. We employ this specification because our sample does not come from a duration data but we rather observe a binary outcome that is conditional on unemployment duration.

When estimating the frailty terms, p^s, typical choices in the literature are (log)normal and gamma distributions, both of which have two parameters. If the estimation includes a constant or a set of dummy variables for each (discrete) duration, the mean of the frailty distribution is not identified and is therefore normalized, typically to 1. The variance is then estimated from the data. In our case, the distribution of the frailty terms, p^s, have a particular meaning and a distributional specification. They can take only two values and must lie between 0 and 1. Based on parallel arguments to estimate a proportional hazard model, we normalize the weighted average of p^H and p^L of the recently unemployed to 0.5. Note that this value not only depends on

[9] For detailed discussion about the non-parametric identification of this model, see Van den Berg (2001).

p^H and p^L but also on the proportion of the true high-skill types among the recently unemployed workers. We further impose that $r_0 = 0.5$ so that p^H and p^L are symmetric around 0.5 and estimate their difference, Δ_p. Given the normalization and symmetry assumptions, the job finding rate of the long-term unemployed identifies Δ_p, whereas the overall job finding rate identifies η.

Next, suppose that $\alpha = 0$. Then, following on our arguments in Section 2.6, the true duration dependence stems only from skill depreciation. Through the lenses of the proportional hazard model in equation (19), the baseline hazard would be a (parametric) function of δ. Hence, any portion of the duration dependence that is not captured by the compositional effect due to unobserved heterogeneity identify δ.

Finally, even if α is positive but the market tightness is constant for individuals with identical x_{it}, then, by equation (16), they would all have the same resume value. In the hazard equations (14) and (15), the interviewing probability would be a constant for a given x_{it}. In that case, α would not be identified. When individuals with the same individual characteristics at same unemployment duration experience a different market tightness history, they also have different resume values. The variation in the resume values then identifies α. Roughly speaking, we can think of the effect of α as reflecting the interaction of market tightness with unemployment duration. In a model similar to that in equation (19), such an effect could be captured by an interaction term for market tightness and unemployment duration.

3.4 Estimation Results

We present our estimation results in Tables 1 and 2. In Table 1, we report our estimation results for the parameters mainly governing the duration dependence: δ, α, and Δ_p. The first column of Table 1 shows our point estimate. Because the parameters have certain bound restrictions, we performed a Lagrange multiplier test to assess the statistical significance of our estimates. We report the χ^2 test statistics when the parameter of interest is set to zero in column 2. Column 3 reports the p-values associated with the χ^2 test statistic. Note that when we restrict the model so that $\Delta_p = 0$, neither δ nor α is identified. Therefore, in the third row where we test

Variable	Estimate	χ^2	p-value
δ: Skill Depreciation	0.168	15.513	0.000
α: Statistical Discrimination	0.535	149.444	0.000
Δ_p: Interview success probability (difference)	0.265	2113.105	0.000

Table 1: Maximum Likelihood Estimation Results for the parameters governing duration dependence.

Variable	Estimate	Standard Error	p-value
$\log(\eta)$: Scale	0.100	0.052	0.056
$1-\sigma$: Elasticity	0.366	0.014	0.000

Table 2: Maximum Likelihood Estimation Results for Matching Function.

for significance of Δ_p, we restrict both of the other parameters to be equal to zero. Under these restrictions all of the duration dependence is attributed to variation in observed characteristics. We report the estimates of the matching function in Table 2. In this Table, the second and the third columns report the standard error and the p-value, respectively, which are evaluated at the point estimate of the structural parameters in Table 1. An exhaustive list for the remaining parameter estimates is available on our Online Appendix.[10]

Our point estimates indicate that a high-skilled worker turns about 63% interview opportunities into hiring while the success probability of a low-skilled worker is only 37%. Every month about 17% of the unemployed workers lose skills and their success probability at the interview stage drops by 26.5 percentage points. Roughly half of the firms interviews all of the applicants for their vacancy, while the remaining half decides after observing the interview cost. We estimate the log of the matching efficiency as 0.100. At our sample mean, this number implies that each worker contacts a vacancy at a monthly rate of 0.914. Moreover, the elasticity of the contact rate with respect to market tightness, $(1-\sigma)$, is 0.366. In our sample, the standard deviation of the log of market tightness is 0.456. Putting these two numbers together, we conclude that one standard deviation increase in the log

[10] We note that our estimates for individual characteristics mimic the findings in Barnichon and Figura (2017).

of market tightness from its sample mean increases vacancy contact rate by 16.7%. Overall, our estimates are statistically significant at conventional levels.

To conduct an overall assessment of the fit our model, we compare the monthly job finding probabilities fitted from our model against those directly measured from the data. As in the literature, we calculate a direct measure of job finding rate as the share of the unemployed who find a job within a month. In Figure 1, we plot the both of these measures by unemployment duration on. Overall, our model tracks the job finding rates by duration very closely even at longer durations where we have fewer observations.

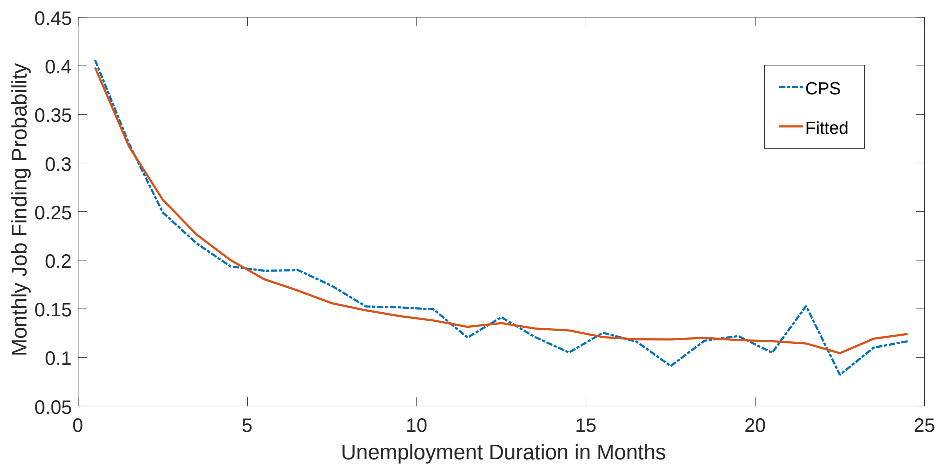

Figure 1: Model Fit by Unemployment Duration

Motivated by the random matching assumption, we used aggregate market tightness in our regression. However, in practice the search could be directed and market tightness could differ across different labor market segments. As a robustness check, we ran our regression using market tightness by region and occupation using disaggregated HWOL data and we then used this series in place of the aggregate market tightness.[11] We obtain very similar estimates with one exception when we use disaggregated market tightness data: our estimate for matching function elasticity with respect to market tightness decreases from 0.366 to 0.205. The decline in this parameter estimate echoes the result in Barnichon and Figura (2015) and our

[11] We follow Barnichon and Figura (2015) to define occupation classes and regions.

point estimates are remarkably similar. We present the other parameter estimates in the Appendix.

4 Counterfactual Analysis

In this section, we analyze the role of each channel that causes duration dependence comparing the job finding rates by unemployment duration under certain parameter restrictions. Because we are ultimately interested in job finding rates by unemployment duration, we perform these comparison at a steady state equilibrium by fixing the value of the cost of vacancy posting: that is, $\kappa(t) = \bar{\kappa}$. However, we also analyze the effect of changes in market tightness on job finding rates by changing the value of $\bar{\kappa}$.

4.1 Calibrated Parameters

We do not have an estimate for three parameters of our model: λ, $\bar{\kappa}$, and $\bar{J}$. First, the job separation rate, λ, does not enter into our likelihood function. For the counterfactual analysis in this section, we set its value to 0.017 so that the steady state unemployment is 6%. Second, the steady state equilibrium value of the market tightness, $\bar{\theta}$, depends on the ratio of $\bar{\kappa}$ to $\bar{J}$ from equations (11) and (12) given the distribution of the resume values over the unemployed workers. In the following, we normalize the value of $\bar{J}$ to 1 and set $\bar{\kappa}$ to 0.346. This value implies that the log of the equilibrium market tightness is equal to -0.663, which corresponds to the mean value we observe in the data.

There are observed individual heterogeneities in the data, which we assumed away in the theoretical model. Our estimate for the scale parameter of the matching function, η, is precisely for an individual with observed characteristics represented with a zero vector. Therefore, we adjusted the value of η so that the average monthly job finding probability matches its corresponding value in the data, $\bar{f}$, which is equal to 0.231.[12] The implied value of $\log(\eta)$ is 0.115. Table 3 summarizes our calibration.

[12]In the data, the maximum unemployment duration is top-coded at 104 weeks, or 24 months. In our simulation, we allowed for unemployment duration longer than this to calculate equilibrium

Parameter	Value	Target
λ: Separation rate	0.017	$u = 0.060$
$\bar{\kappa}$: Cost of vacancy posting	0.346	$\log(\bar{\theta}) = -0.663$
$\bar{J}$: Match surplus value	1.000	Normalization
$\log(\eta)$: Matching function Scale	0.115	$\bar{f} = 0.231$

Table 3: Calibrated Parameters.

For the other parameters, we used our point estimates from the previous section.

In Figure 2, we plot the job finding probability within a month by unemployment duration from our simulated model against our fitted job finding rates from the CPS data. This comparison gives us an idea about the contribution of observed individual heterogeneity to unemployment duration dependence. From the simulated data, we have calculated the job finding rates at short time intervals whereas the unemployment duration measured in CPS represents a bracket. For example, the first bracket includes the individuals who have been unemployed for less than a month. For comparability, we plot job finding probabilities from our simulated model starting from unemployment duration at the middle of the first month.

An inspection of Figure 2 reveals that observed individual heterogeneity plays a significant role in explaining the negative duration dependence. As we move from the shortest duration to the longest duration in the CPS data, the job finding probability within a month is four times lower. In our simulated data, it only goes down by half. These numbers suggest that about half of the duration dependence is attributable to observed differences in job finding rates. Our following analysis focuses on the remaining part.

4.2 Skill Decay, Statistical Discrimination, and Unobserved Heterogeneity

In this section, we calculate job finding rates by duration by setting the values of δ and α to zero in turn to evaluate their impact on duration dependence together

market tightness more precisely. For comparability, we match the average job finding rate among those who are unemployed less than 24 months.

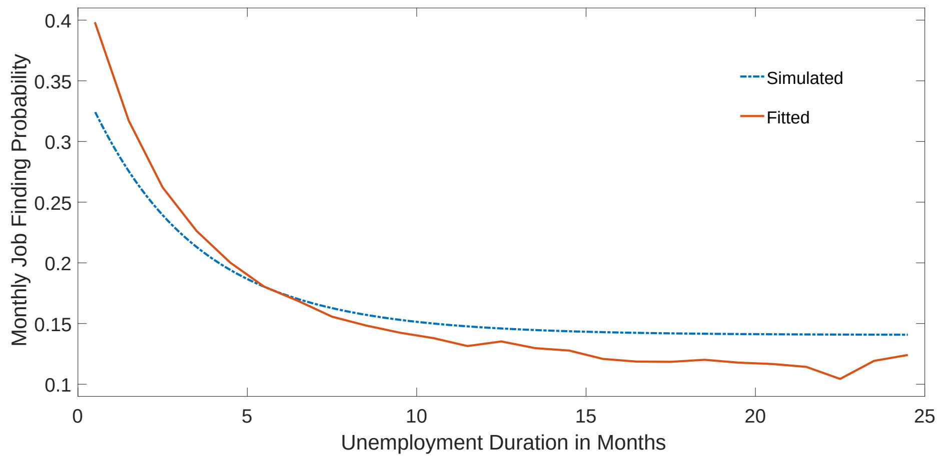

Figure 2: Monthly Job Finding Probabilities by Unemployment Duration: Simulated vs. Fitted

with unobserved heterogeneity. Changing the model parameters has two effects. First, there is direct effect on individual job finding rates through changes in resume values. Second, changes in the distribution of resume values affects the equilibrium market tightness. To better understand the role of direct and equilibrium effects, we perform two different analysis. In the first, we change the cost of vacancy posting in a way in which the equilibrium market tightness remains the same and we assess only the direct effect on duration dependence. In the second analysis, we keep the cost of posting a vacancy constant and we calculate the new equilibrium market tightness.

Figure 3 shows the decomposition of duration dependence under the assumption that market tightness remains the same. Our results suggest that the statistical discrimination channel is the dominant factor in explaining the negative duration dependence. When we eliminate statistical discrimination by setting α to zero, the black solid line in the figure, the job finding probability for the long-term unemployed jumps from 0.15 to 0.30 compared to 0.35 for the recently unemployed. On the other hand, eliminating skill depreciation, the dashed red line, improves job finding rates at short term duration while leaving that of long-term unemployed marginally changed. When both parameters are set equal to zero, the dotted green line, we have only the composition effect due to unobserved heterogeneity gener-

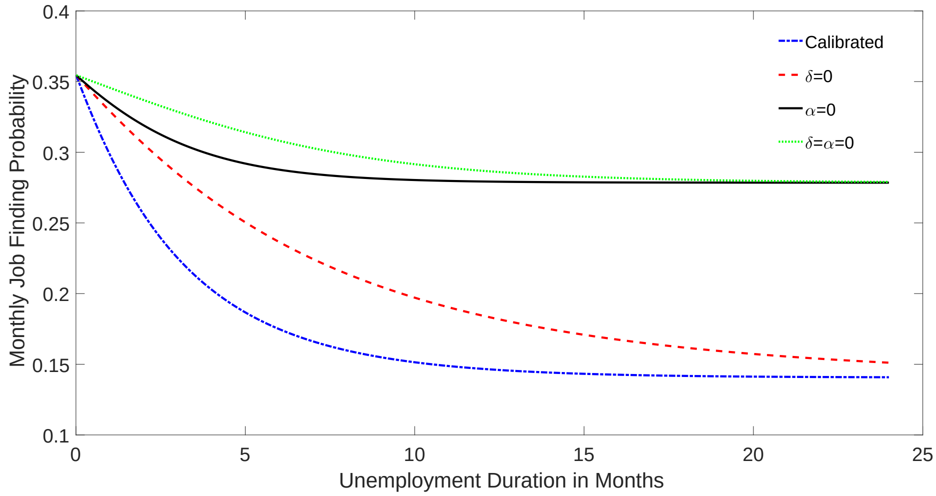

Figure 3: Counterfactual Monthly Job Finding Probabilities by Unemployment Duration with Constant $\bar{\theta}$.

ating negative duration dependence. Taken altogether, the statistical discrimination channel explains 75% of the decline in job finding probability moving from short term to long term unemployment duration while the remaining is mainly attributable to composition effect.

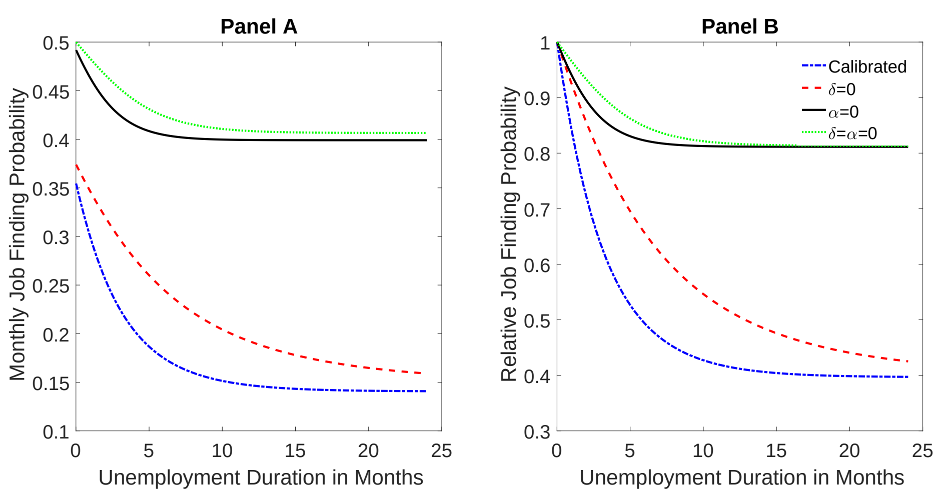

Figure 4: Counterfactual Monthly Job Finding Probabilities by Unemployment Duration with Constant $\bar{\kappa}$.

Outcome	Calibrated	Interview Subsidy
$\bar{u}$: Unemployment rate	0.060	0.027
$\log(\bar{\theta})$: Market tightness	-0.663	0.580
$\bar{f}$: Average monthly job finding probability	0.231	0.454
Cost of vacancy posting per hire	0.688	1.000
Average interview cost per hire	0.312	0.521

Table 4: Subsidizing Interview Costs.

When we allow the market tightness to change, we obtain similar results. Figure 4 shows our results. In panel A, we draw the actual job finding probability under each parameter restriction and in panel B we plot job finding probability relative to the recently unemployed. Shutting down statistical discrimination channel alone causes a larger increase in equilibrium market tightness than eliminating skill depreciation. Nonetheless, the relative changes in job finding probabilities by duration mimic our findings in Figure 3 with constant market tightness.

From a policy standpoint, our decomposition analysis suggests that eliminating statistical discrimination would improve job finding rates significantly. To get an idea about the cost of such a policy, assume that firms still face interview costs but are fully subsidized so that they indiscriminately interview every applicant. Table 4 shows the effect of such an interview subsidy on the equilibrium. In the first column we report the selected equilibrium outcomes at the calibrated parameter values. The cost of vacancy posting and interviewing are equal to 0.688 and 0.312 per hire, respectively.[13] Note that they add up to 1, the match surplus value $\bar{J}$, from the free entry condition. In the second column, we report equilibrium outcomes with hiring subsidy. In this case, the unemployment rate, $\bar{u}$, falls from 6% to 2.7%, the log of market tightness, $\log(\bar{\theta})$ increases from -0.663 to 0.580 and average job finding rate, $\bar{f}$, increases from 0.231 to 0.454. Because firms effectively only pay for posting a vacancy, the free entry condition implies that its value per hire is equal

[13] Jarosch and Pilossoph (2019) choose their interview cost parameter as 10% of the flow value of output, y. Moreover, the surplus value, $\bar{J}$, is 65% of y in their equilibrium. If we also assume that same ratio between $\bar{J}$ and y, then the average interview costs incurred in our model is about 20% of y, which is twice as much compared to Jarosch and Pilossoph (2019).

to $\bar{J} = 1$. However, interview cost per hire, which are now subsidized, increases by about 67% to 0.521. Overall, the interview costs that firms face are large and deter many of them from interviewing long term unemployed workers.

4.3 Duration Dependence and Market Tightness

We have not yet discussed how duration dependence changes with market tightness. In Figure 5, we plot job finding probabilities relative to the recently unemployed for low and high market tightness values under various parametrizations. Low and high market tightness values correspond to two standard deviation from the mean value of log market tightness in the data. For each calculation, the underlying assumption is that κ changes in a way that the prevailing steady state market tightness is equal to either low or high market tightness.

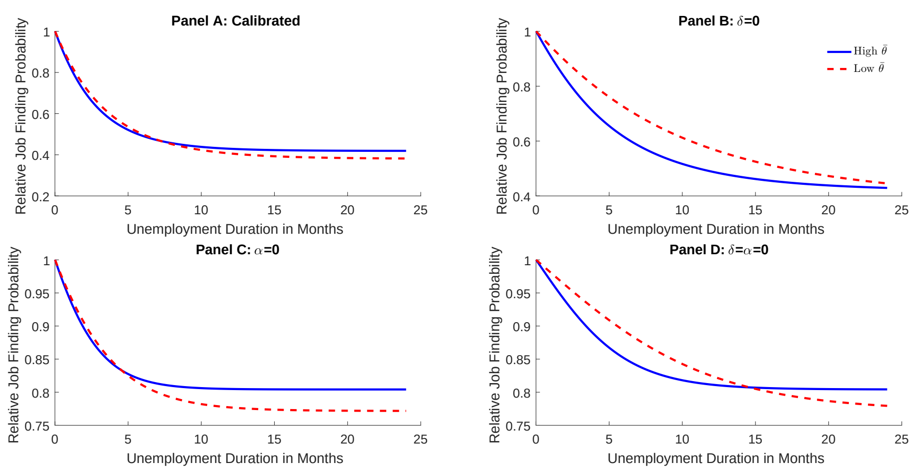

Figure 5: Relative Monthly Job Finding Probabilities by Unemployment Duration for Different $\bar{\theta}$.

Panel A in Figure 5 depicts unemployment duration dependence for low and high $\bar{\theta}$ when δ, α, and Δ_p are set to their calibrated values. We see that duration dependence is stronger when market tightness is higher for short durations, but the relationship reverses after the seventh month. Abbring, Van Den Berg, and Van Ours (2001) find a similar result in their study where they analyze the contribution of incidence of unemployment and unemployment duration to the fluctuations in aggregate unemployment rate.[14] In their analysis, they employ a flexibly parametrized reduced form model to estimate job finding rates from the monthly CPS data covering the period from 1968 to 1992. They capture the response of the job finding rates at different durations to business cycle fluctuations by a full set of interaction variables of duration and calendar time variables along with Help Wanted Index (HWI), which was an earlier version of HWOL that was in use before the Internet. They propose that a model with unobserved heterogeneity, or sorting in their terminology, and ranking of unemployed workers by duration as in Blanchard and Diamond (1994) can explain this asymmetry in duration dependence.

Our model provides an alternative explanation. To understand the dynamics behind the asymmetric response of duration dependence to market tightness, we set both δ and α to zero and we plotted the relative job finding probabilities only in the presence of unobserved heterogeneity in panel D of Figure 5. Compared to panel A, the difference in duration dependence under low and high θ is now starker and the relationship reverses at a longer duration. When market tightness is high, the overall job finding rate is also high and high-skilled workers leave the pool relatively quickly compared to the market equilibrium with low-market tightness. Because the unemployment pool deteriorates relatively quickly, the duration dependence is steeper at short durations. However, after some point, the unemployment pool becomes relatively more homogeneous and heavily populated with low-skilled workers. The crossing of the curves in panel D of Figure 5 implies that job finding probability increases proportionally more for the low-skilled workers.

Statistical discrimination makes the difference between duration dependence at high- and low-market tightness even bigger. We show this in panel B of Figure 5, where we set α to its estimated value while keeping $\delta = 0$. There are two forces behind this result. First, when market tightness is low, resume values decline more slowly with unemployment duration and probability of receiving an interview does not fall too quickly with unemployment duration. Second, when $\alpha > 0$, even high-skilled workers face declining job finding rates and hence they remain longer in the

[14] Our plots in Figure 5 corresponds to Figure 10 in their working paper version, Abbring, Van Den Berg, and Van Ours (1999), with the exception that they plot relative job finding probabilities on a log scale.

unemployment pool. Compared to the case when $\alpha = 0$, the unemployment pool has relatively more skilled workers at longer durations. Consequently, the proportionally stronger response of low-skilled workers at long duration does not show up in panel B of Figure 5. To illustrate our point, we plot the resume values, which also show the proportion of high-skilled workers, by duration in Figure 6 for each case in Figure 5.

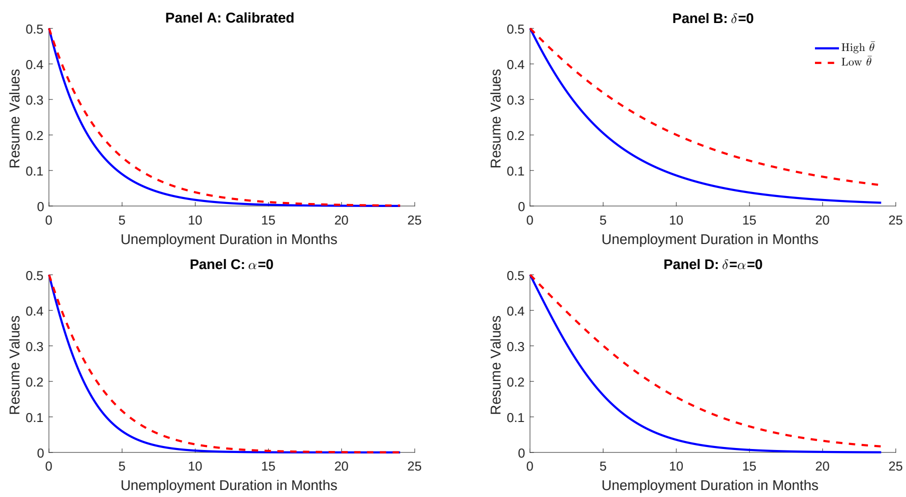

Figure 6: Resume Values by Unemployment Duration for Different $\bar{\theta}$.

Panel A of Figure 5 suggests that the duration dependence does not change much over the business cycle. Both the composition effect and the statistical discrimination channels imply a stronger response to the business cycle fluctuations. We obtain muted response of duration dependence to changes in market tightness in panel C of Figure 5, where we set δ to its calibrated value and α to zero. In this case, the duration dependence at short durations is indistinguishably similar at low- and high-market tightness equilibria. We can explain the underlying mechanism with the resume values in panel C of Figure 6. When there is skill depreciation, regardless of the market tightness, the resume values and the share of high-skilled workers decline relatively quickly. By five months of unemployment duration, the share of high-skilled workers is below 10% even when the market tightness is low. Consequently, duration dependence flattens out in either case. We still see the asymmetry of duration dependence at short and long durations because the job finding probability of low-skilled workers increases proportionally more when the market tightness increases.

We conclude that shutting down statistical discrimination has a greater impact on job finding probability of long-term unemployed workers, but skill depreciation channel is more important to capture the cyclical behavior of duration dependence.

5 Conclusion

We study unemployment duration dependence in a random matching model with worker heterogeneity in terms of their skill to turn interviews into hires. Apart from the composition effect, the model displays negative duration dependence due to skill depreciation and statistical depreciation. We estimate our model using micro-level data from CPS via maximum likelihood. Our identification relies on the fact that the statistical discrimination is weaker when the market tightness is low whereas skill depreciation is independent of the aggregate labor market conditions.

We find that shutting down statistical discrimination alone improves the job finding rates of the long-term unemployed relative to a recently unemployed by 75%. The corresponding number for skill depreciation is negligible. We also find that the duration dependence relationship changes only slightly in response to a change in the market tightness. Our counter-factual exercise suggests that this result is mainly driven by skill depreciation in that average resume values and the share of high-skilled workers quickly become relatively small, regardless of the market tightness.

Our results also suggest that the interview costs that firms face are large and this deters many of them from interviewing long-term unemployed workers in equilibrium. If firms are fully subsidized for their interview costs, then the unemployment rate declines from 6% to 2.7% as a result of increased job finding rates but this outcome comes at a large cost. The total interview costs the economy incurs increases by 67%.

Our conclusion with respect to the role of statistical discrimination in explaining the negative duration dependence is in contrast to the findings in Jarosch and Pilossoph (2019). Although these two models have several features in common and they closely capture the duration dependence relationship in the data, they have different mechanisms to generate statistical discrimination. We motivate the idea of statistical discrimination by the variation in interview costs, whereas statistical discrimination is due to the variation in minimum hiring requirements in Jarosch and Pilossoph (2019). Examining the quantitative role of each type of variation in accounting for statistical discrimination is an important future research topic.

References

Abbring, J. H., G. J. Van Den Berg, and J. C. Van Ours (1999). Business Cycles and Compositional Variation in US Unemployment. Discussion Paper 9965.

Abbring, J. H., G. J. Van Den Berg, and J. C. Van Ours (2001). Business Cycles and Compositional Variation in US Unemployment. *Journal of Business & Economic Statistics 19*(4), 436–448.

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

Alvarez, F. E., K. Borovičková, and R. Shimer (2016). Decomposing Duration Dependence in a Stopping Time Model. Working Paper.

Barnichon, R. (2010). Building a Composite Help-Wanted Index. *Economics Letters 109*(3), 175–178.

Barnichon, R. and A. Figura (2015). Labor Market Heterogeneity and the Aggregate Matching Function. *American Economic Journal: Macroeconomics 7*(4), 222–49.

Barron, J. M. and J. Bishop (1985). Extensive Search, Intensive Search, and Hiring Costs: New Evidence on Employer Hiring Activity. *Economic Inquiry 23*(3), 363–382.

Blanchard, O. J. and P. Diamond (1994). Ranking, Unemployment Duration, and Wages. *The Review of Economic Studies 61*(3), 417–434.

Doppelt, R. (2016). The Hazards of Unemployment: A Macroeconomic Model of Job Search and Résumé Dynamics. Working Paper.

Flood, S., M. King, R. Rodgers, S. Ruggles, and J. R. Warren (2020). Integrated Public Use Microdata Series, Current Population Survey: Version 7.0 [dataset]. Minneapolis, MN: IPUMS, 2020. https://doi.org/10.18128/D030.V7.0.

Fujita, S. and G. Moscarini (2017). Recall and Unemployment. *American Economic Review 107*(12), 3875–3916.

Hall, R. E. and S. Schulhofer-Wohl (2018). Measuring Job-Finding Rates and Matching Efficiency with Heterogeneous Job-Seekers. *American Economic Journal: Macroeconomics 10*(1), 1–32.

Jarosch, G. and L. Pilossoph (2019). Statistical Discrimination and Duration Dependence in the Job Finding Rate. *The Review of Economic Studies 86*(4), 1631–1665.

Kroft, K., F. Lange, and M. J. Notowidigdo (2013). Duration Dependence and Labor Market Conditions: Evidence from a Field Experiment. *The Quarterly Journal of Economics 128*(3), 1123–1167.

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

Van den Berg, G. J. (2001). Duration Models: Specification, Identification and Multiple Durations. In *Handbook of Econometrics*, Volume 5, pp. 3381–3460. Elsevier.

6 Appendix

6.1 Derivation of Resume Updating Rule

Consider the Bayesian updating rule in discrete time for an unemployed worker in a short time interval $[\tau, \tau{+}d\tau]$. To simplify the notation, let us suppress the dependence of the hazard rates on calendar time and write the Bayesian updating rule as:

$$r(\tau+d\tau) = \frac{r(\tau)\,\text{exp}(-\gamma d\tau - h(\tau|H)d\tau)}{r(\tau)\,\text{exp}(-h(\tau|H)d\tau) + (1-r(\tau))\,\text{exp}(-h(\tau|L)d\tau)}.$$

For small $d\tau$, we can write this equation approximately as:

$$r(\tau+d\tau) = \frac{r(\tau)(1-\gamma d\tau)(1-h(\tau|H)d\tau)}{r(\tau)(1-h(\tau|H)d\tau) + (1-r(\tau))(1-h(\tau|L)d\tau)}.$$

Subtracting $r(\tau)$ from both sides and after re-arranging, we obtain:

$$r(\tau+d\tau) - r(\tau) = \frac{r(\tau)d\tau(-\gamma - (1-r(\tau))(h(\tau|H) - h(\tau|L)) - \gamma h(\tau|H)d\tau)}{r(\tau)(1-h(\tau|H)d\tau) + (1-r(\tau))(1-h(\tau|L)d\tau)}.$$

Dividing both sides by $d\tau$ and taking the limit as $d\tau \to 0$, we obtain:

$$\frac{dr(\tau)}{d\tau} = -r(\tau)(\gamma + (1-r(\tau))(h(\tau|H) - h(\tau|L))).$$

Using the expression for job finding rates in equation (7), we obtain the resume updating rule in equation (8).

6.2 Market Tightness by Occupation and Region

Tables 5 and 6 show our estimation results using market tightness data calculated separately for occupation classes and regions. Following Barnichon and Figura (2015), we define four occupation classes and nine regions. The occupation classes are SOC 2010 high-level occupation groups: professional and managerial, personal services, sales and office support, and production. For regions, we use nine census divisions. Vacancy data comes from HWOL database and is seasonally adjusted. We calculated monthly series for the total number of unemployed for each occupation and region group from basic monthly CPS files. We use Census X-ARIMA suit to seasonally adjust these series. Then, we calculate market tightness as the ratio of total vacancies to total unemployed.

Our estimation results are quite similar to main text except for the elasticity of matching function, $(1-\sigma)$. Because the mean value of market tightness is not

Variable	Estimate	χ^2	p-value
δ: Skill Depreciation	0.183	12.083	0.000
α: Statistical Discrimination	0.593	111.147	0.000
Δ_p: Interview success probability (difference)	0.255	1468.952	0.000

Table 5: Maximum Likelihood Estimation Results for the parameters governing duration dependence. Market tightness is calculated separately for occupation and region groups.

Variable	Estimate	Standard Error	p-value
$\log(\eta)$: Scale	-0.135	0.078	0.083
$1 - \sigma$: Elasticity	0.205	0.027	0.000

Table 6: Maximum Likelihood Estimation Results for Matching Function. Market tightness is calculated separately for occupation and region groups.

zero, the estimate of the scale parameter, $\log(\eta)$, is also different. Barnichon and Figura (2015) argues that the decline in elasticity of matching function implies that heterogeneities in the labor market is pro-cyclical.

Variable	Estimate	χ^2	p-value
δ: Skill Depreciation	0.232	11.321	0.000
α: Statistical Discrimination	0.587	98.419	0.000
Δ_p: Interview success probability (difference)	0.189	1234.916	0.000

Table 7: Maximum Likelihood Estimation Results for the parameters governing duration dependence. We use aggregate market tightness, but our sample covers the period from May 2005 to December 2016 for comparability.

Variable	Estimate	Standard Error	p-value
$\log(\eta)$: Scale	-0.050	0.052	0.056
$1 - \sigma$: Elasticity	0.304	0.014	0.000

Table 8: Maximum Likelihood Estimation Results for Matching Function.

Note also that our disaggregated data for HWOL starts in May 2005. For comparability, we also provide estimation results using aggregate market tightness for the same period in Tables 7 and 8. The estimate is of $(1-\sigma)$ is similar to the main text, which implies that the change in the estimate of this parameter is not due to the time frame we study. We also find that α is closer to our main estimate in Table 5, although δ and Δ_p are slightly different.