

DURATION DEPENDENCE AND HETEROGENEITY: LEARNING FROM EARLY NOTICE OF LAYOFF*

DIV BHAGIA[†]

July 10, 2024

Abstract

This paper presents a novel approach to distinguish the impact of duration dependent forces and adverse selection on the exit rate from unemployment by leveraging variation in the length of layoff notices. I formulate a Mixed Hazard model in discrete time and specify the conditions under which variation in notice length enables the identification of structural duration dependence while allowing for arbitrary heterogeneity across workers. Utilizing data from the Displaced Worker Supplement (DWS), I employ the Generalized Method of Moments (GMM) to estimate the model. According to the estimates, the decline in the exit rate over the first 48 weeks of unemployment is largely due to the worsening composition of surviving jobseekers. Furthermore, I find that an individual's likelihood of exiting unemployment decreases initially, then increases until unemployment benefits run out, and remains steady thereafter. These findings are consistent with a standard search model where returns to search decline early in the spell.

Keywords: *Duration dependence, unemployment, exit rate, job-finding rate, unobserved heterogeneity, hazard models, advance notice*

JEL Codes: *C41, E24, J64, J65*

*I am grateful to Theodore Papageorgiou, Kevin Lang, Shakeeb Khan, and Arthur Lewbel for their invaluable guidance and support. I would also like to thank all participants of the Dissertation Workshop at Boston College for their helpful comments. I am solely responsible for any errors or omissions. The code and data for the paper are available at: <https://github.com/divbhagia/notice>.

[†]California State University, Fullerton; dbhagia@fullerton.edu

I INTRODUCTION

A well-established empirical regularity is that the exit rate out of unemployment decreases over the spell of unemployment, except for a spike at the time of unemployment insurance (UI) exhaustion. The decline in the exit rate may represent negative duration dependence, meaning that the longer a worker remains unemployed, the less likely they are to exit unemployment. This would be true if employers discriminate against long-term unemployed workers or if workers lose valuable skills and connections over time, which would otherwise assist them in finding employment. However, workers with different unemployment durations, who may appear similar to researchers, may actually be quite different from each other. Factors such as employability, the urgency to find a job, or the ability to secure employment may vary across individuals. Such heterogeneity across workers would imply that the observed exit rate declines even in the absence of structural duration dependence. As more employable workers leave unemployment, the remaining pool of unemployed individuals increasingly consists of those who are less likely to exit unemployment.

Understanding how the likelihood of exiting unemployment evolves over the unemployment spell and the extent of heterogeneity across workers is crucial for the design of unemployment policies.¹ Furthermore, the magnitude and direction of structural duration dependence have implications for the incidence of long-term unemployment and the speed of recovery from economic downturns (Pissarides, 1992). Given its significance, a substantial body of literature has attempted to disentangle the sources of the decline in the exit rate from unemployment. However, it has proven to be challenging to do so using observational data.

In this paper, I develop and implement a novel approach to empirically disentangle

¹See Shimer and Werning (2006), Pavoni and Violante (2007), Pavoni (2009), and Kolsrud et al. (2018).

the contributions of structural duration dependence and unobserved heterogeneity in explaining the dynamics of the exit rate from unemployment. My approach relies on leveraging variation in the length of notice workers receive from their employers before being laid off. Using data from the Displaced Worker Supplement (DWS), I compare workers with a notice period of more than two months (referred to as long notice) to workers with a notice period of one to two months (referred to as short notice). To ensure comparability across the two groups, I use inverse probability weighting (IPW) to achieve balance on a comprehensive set of observable characteristics. The analysis reveals that during the initial 12 weeks, the exit rate out of unemployment is 7 percentage points higher for long-notice workers. This difference is primarily due to a larger proportion of long-notice workers transitioning directly to their next job without experiencing a period of unemployment. However, beyond the first 12 weeks, the exit rate for workers with the longer notice is actually lower.

I argue that the lower exit rate for long-notice workers at later durations is due to the composition of this group becoming relatively worse as a larger proportion of individuals exit early in the spell. This indicates the presence of heterogeneity among workers. In the presence of heterogeneity, such as differences in employability, those who are more employable exit unemployment earlier. As more workers exit early from the long-notice group, the surviving workers from this group will have a lower proportion of highly employable workers compared to the short-notice group. Conversely, if there is no heterogeneity, a larger proportion of workers exiting early from the long-notice group will not alter the composition of long versus short-notice workers at later durations. Consequently, there would be no discernible difference in the exit rates of the two groups.² Thus, the difference in the exit rates of short and long-notice workers is indicative of the extent of underlying heterogeneity. This is the fundamental idea behind my approach, which enables me to pin down the extent of heterogeneity and

²It is also possible that receiving a longer notice directly affects a worker's exit probability, even at later durations. I discuss this possibility below while addressing the robustness of my findings.

estimate structural duration dependence.

I operationalize this intuition by formulating a Mixed Hazard (MH) model ([Lancaster, 1979](#)) in discrete time with multiple notice lengths. Within this framework, I specify the probability of an individual exiting unemployment as a product of their unobservable type, a function of observable factors, and a structural hazard that varies with the duration of unemployment and notice length. I show that structural duration dependence, characterized by how the structural hazard varies with the duration of unemployment, can be identified when two conditions hold. The first condition, commonly referred to as unconfoundedness ([Rosenbaum and Rubin, 1983](#)), requires that the length of notice is independent of the worker's unobservable type when conditioned on observable characteristics. The second condition states that the length of the notice period does not impact the structural hazard at later durations of unemployment. In other words, while a longer notice period before layoff may affect the probability of exiting at the beginning of the unemployment spell, it does not influence the probability of exiting at later durations.

Building on the identification result, I develop a method for estimating the model using the Generalized Method of Moments (GMM). The estimation method utilizes moments weighted by the inverse of propensity scores to ensure that the distribution of observable characteristics is similar across different notice lengths. Additionally, the method incorporates right-censored duration data to account for spells that are still incomplete at the time of the survey.

Relative to the existing literature on the identification and estimation of the Mixed Hazard model, my approach goes further in several dimensions. Firstly, I do not impose any functional form restrictions on the distribution of heterogeneity. This is crucial because misspecification of unobserved heterogeneity can significantly impact estimates of structural duration dependence, as demonstrated by [Heckman and Singer \(1984\)](#). Secondly, identification in my model stems from variation in a variable—the length of

notice—that is exogenous conditional on observables and is assumed to affect the structural hazard in a way that aligns with economic intuition.³ The introduction of unconfoundedness in the Mixed Hazard framework is a novel innovation. Lastly, I provide a root- n consistent estimator for the parameters of my model. My model is analogous to a Mixed Hazard model with a time-varying exogenous variable, for which [Brinch \(2007\)](#) provides a non-constructive proof of identification. The key differences here are the assumption of unconfoundedness instead of exogeneity and the exposition being in discrete time, the latter of which leads to a consistent estimator for the model’s parameters via GMM.⁴ To the best of my knowledge, [Alvarez et al. \(2021\)](#) is the only other study that utilizes moment conditions from a discrete version of the Mixed Hazard model and constructs a GMM estimator. However, their identification result and estimator pertain to multiple spell data.⁵

I estimate the model using weighted moments from the DWS data. The estimates uncover substantial heterogeneity in individual exit probabilities. I find that a quarter of the 43.6% decline in the observed exit rate over the first five months is due to the changing composition of workers over the spell of unemployment. Moreover, after the first five months, an individual worker’s exit probability increases until the time of their unemployment benefit exhaustion and remains constant after. This is in contrast to the observed exit rate, which continues to decline even after benefit exhaustion. Recently, researchers have proposed behavioral modifications to standard search theory to explain this decline ([Boone and van Ours, 2012](#); [DellaVigna et al., 2017, 2021](#)).

³Existing non-parametric identification results for the Mixed Hazard model rely on variation in an exogenous variable that enters the structural hazard multiplicatively ([Elbers and Ridder, 1982](#); [Heckman and Singer, 1984](#)). The practical implementation of these results has been limited due to the challenge of locating a variable that meets this criterion, as well as the absence of a convenient estimator. Another approach to identification is using multiple spell data ([Honoré, 1993](#)). However, this approach assumes that the unobserved characteristics of the jobseeker remain constant across repeated spells.

⁴Although the discrete-time model provides a simple, consistent estimator, it must be noted that one of its disadvantages is invariance to the definition of periods.

⁵[van den Berg \(1990\)](#) also set up a discrete-time MPH model and study identification using cohort effects; however, they use a maximum likelihood estimator on aggregate data for estimation.

However, I highlight an alternative explanation: as a substantial number of individuals exit unemployment right at the point of benefit exhaustion, the composition of the remaining unemployed workers becomes significantly worse. This compositional change contributes to the observed decline in the exit rate after benefit exhaustion. Finally, I calibrate a partial equilibrium search model with a non-stationary environment (Mortensen, 1986; van den Berg, 1990) and show that my findings can be rationalized in this framework with a decline in returns to search early in the spell.

Under the identifying assumptions specified for the Mixed Hazard model in my framework, the lower exit rate among long-notice workers after the initial 12 weeks is attributed to the presence of heterogeneity. However, two alternative explanations are possible. First, there could be unobservable differences between long- and short-notice workers. Second, a longer notice period may reduce a worker's exit probability at later durations. To address these concerns, in [Appendix E](#), I relax the assumptions of my model to allow for arbitrary differences between the two groups and for the structural hazards at later durations to vary by notice length up to a specific constant. Although I cannot show that all the parameters of this more general model are identified, I estimate the model by varying the additional parameters and find the values that minimize residuals.⁶ The estimated values that minimize the residuals suggest no mean differences in the unobservable heterogeneity distribution between the two groups nor any difference in the structural hazard beyond the first 12 weeks. This points towards the validity of the identifying assumptions employed in the analyses.

This paper contributes to the extensive literature on the dynamics of job-finding over the spell of unemployment. Previous empirical studies utilizing the Mixed Hazard model have had to make strong functional form assumptions due to challenges with estimation. Consequently, the evidence on structural duration dependence from these studies is mixed, as highlighted by [Machin and Manning \(1999\)](#) in their review.

⁶I verify that the numerical error function is locally convex in all cases.

Recently, [Alvarez et al. \(2016\)](#) revived this strand of work by estimating a Mixed Hitting-Time (MHT) model ([Abbring, 2012](#)) using Austrian social security data. They focus on a selected sample of workers with multiple unemployment spells and are able to estimate the extent of heterogeneity across workers that is fixed between spells. A relative advantage of my approach is that it captures spell-specific heterogeneity.⁷ Another related study, [Mueller et al. \(2021\)](#), utilizes variation in expectations about job-finding from survey data to pin down variation in actual job-finding rates. While both of these studies also document substantial heterogeneity across jobseekers, my estimator for structural duration dependence is flexible enough to capture changes around UI exhaustion.⁸

Given the challenges with estimating structural duration dependence, researchers have instead focused on assessing its determinants. [Kroft et al. \(2013\)](#) conduct an audit study and find that the likelihood of receiving a callback for an interview declines with the duration of unemployment. However, they note that since they cannot measure worker behavior or employers' ultimate hiring decisions, their estimates only shed light on one determinant of structural duration dependence.⁹ Several papers have also documented how search effort or reservation wages evolve over the spell of unemployment ([Krueger and Mueller, 2011](#); [Marinescu and Skandalis, 2021](#); [DellaVigna et al., 2021](#)). The evidence provided in this paper suggests that, while call-back rates or other factors affecting returns to search matter initially, a worker's optimizing behavior determines the likelihood of exiting unemployment at later durations.

⁷For instance, a worker's savings or UI eligibility may change over the months or years by the time this worker becomes unemployed again.

⁸[Alvarez et al. \(2016\)](#) utilizes an optimal-stopping model; a worker finds a job at an optimal stopping time when a Brownian motion with drift hits a barrier. Their model generates an inverse Gaussian distribution of duration for each worker. [Mueller et al. \(2021\)](#) restrict the structural hazard to be monotonic over the spell of unemployment, and their estimator yields a practically flat hazard.

⁹Using a structural model, [Jarosch and Pilossoph \(2019\)](#) argue that if employers statistically discriminate against those with longer durations, then a decline in callback rates only has a marginal effect on workers' exit rates.

Finally, a substantial body of literature highlights a spike in exit rates at UI exhaustion, where exit rates increase until benefit exhaustion and decline thereafter.¹⁰ While the initial increase is consistent with standard search theory, the subsequent decline is not. My estimates reproduce the increase in individual exit probabilities leading up to UI exhaustion but do not find evidence of a decline thereafter. [Boone and van Ours \(2012\)](#) propose storable job offers as an explanation for the spike, while [DellaVigna et al. \(2017\)](#) argue that search models incorporating reference dependence predict a decrease in search effort after benefit exhaustion. My estimates suggest that the decline in the exit rate after UI exhaustion can be attributed to a shift in the composition of surviving workers, as a significant proportion of workers exit unemployment right at benefit exhaustion. However, the individual exit probability remains constant, consistent with the predictions of standard search models.

II CONTEXT AND DATA

In this section, I describe the institutional setting and the data and document how the exit rate out of unemployment varies with the length of notice.

II.A Institutional Details

Under certain circumstances, US employers are required to give notice of layoff. The federal WARN Act mandates that employers with 100 or more full-time employees provide a 60-day advance notice for plant closings and mass layoffs. A plant closing is defined as the shutdown of a site or units within it that results in 50 or more employees losing their jobs within a 30-day period, while a mass layoff is the loss of employment for 500 or more employees during a 30-day period, or 50-499 employees if they con-

¹⁰[Katz and Meyer \(1990\)](#) first documented the spike in exit rates at benefit exhaustion in the context of the US. Some recent papers that document this pattern using administrative data are [DellaVigna et al. \(2017\)](#) (Hungary), [Ganong and Noel \(2019\)](#) (US), and [Marinescu and Skandalis \(2021\)](#) (France).

stitute one-third or more of the employer’s active workforce. The law only applies to layoffs exceeding six months, excluding discharges for cause, voluntary departures, or retirements. Some states, such as California, New York, and Illinois, have implemented their own WARN laws that expand the coverage of employment losses beyond what the federal law requires.¹¹

When it comes to unemployment insurance (UI), US workers who are terminated without cause are typically eligible to receive benefits for a limited duration. Although the UI program is a federal program, each state sets its own benefit levels and durations. Eligibility and benefits may depend on a combination of earnings, hours worked, or weeks worked during a base period, depending on the specific rules of the state’s UI program. Typically, this base period consists of the first four out of five completed calendar quarters preceding the claim filing date. In most states, the maximum period for receiving benefits is 26 weeks. Nine states have a uniform benefit duration of 26 weeks, while the benefit durations in the remaining states vary depending on the applicant’s earnings history. Additionally, a program for extended benefits has been in place since a 1970 amendment to the Federal Unemployment Tax Act (FUTA), which can be triggered by the state unemployment rate. Temporary programs have also been implemented to extend benefits during recessions.

II.B Data Description and Sample Construction

I use data from the Displaced Worker Supplement (DWS) for the years 1996-2020. DWS is fielded biennially along with the basic monthly Current Population Survey (CPS) in January or February. The survey is administered to individuals who report having *lost or left* a job within the past three years due to a plant closure, their position being abolished, or having insufficient work at their previous employment. Apart from details

¹¹It is not possible to exploit policy variation across states, say in a differences-in-difference framework, due to confounding pre-trends; both California and New York implemented these laws in the aftermath of a national recession.

on workers' lost and current jobs, DWS also collects the length of the notice period workers received before being laid off and the length of time they took to find another job.

For the analysis, the sample is limited to individuals aged 21 to 64 years old. It excludes individuals who expected to be recalled to their previous jobs and those whose lost job was self-employment. To focus on workers who lost stable, permanent employment, the sample only includes individuals who were employed full-time for at least six months at their previous job and had health insurance benefits provided by that employer. The question on notice length in the DWS is categorical and specifies whether someone did not receive a notice or received a notice of <1 month, 1-2 months, or >2 months. For individuals who did not receive a notice, it is uncertain whether they were displaced or quit their jobs voluntarily, so they are excluded from the sample.¹² Additionally, I exclude individuals with less than a month's notice from the main analysis, as they significantly differ from the other two groups in observable characteristics, casting doubt on the validity of the unconfoundedness assumption. Nevertheless, I present results including these groups in Section V.B as a robustness check. Section C.1 in the Appendix provides additional details on data construction. Specifically, Table C1 summarizes the sample selection procedure, and Table C2 compares the characteristics of the workers in my sample to all workers in the CPS and DWS.

Throughout the remainder of the text, I use the term *long notice* interchangeably to refer to a notice period of more than two months and *short notice* to refer to a notice period of 1-2 months. Table 1 presents the summary statistics separately for workers with short and long notice in the sample. Columns (1) and (2) display the raw averages for the sample, revealing notable differences between the two groups. Workers with

¹²As noted by Farber (2017), while the DWS seeks to capture separations caused by economic difficulties within firms, it is hard to distinguish layoffs from voluntary quits. This is because financially distressed firms might reduce hours or wages rather than lay off workers, prompting some workers to voluntarily quit in search of better opportunities.

TABLE 1: DESCRIPTIVES BY NOTICE LENGTH

	Unbalanced			Balanced		
	Short (1)	Long (2)	Diff. (2)-(1)	Short (3)	Long (4)	Diff. (4)-(3)
Age	42.44 (0.24)	43.57 (0.22)	1.13*** (0.33)	43.04 (0.24)	43.01 (0.22)	-0.03 (0.33)
Female	0.45 (0.01)	0.46 (0.01)	0.02 (0.02)	0.46 (0.01)	0.46 (0.01)	-0.00 (0.02)
Married	0.59 (0.01)	0.63 (0.01)	0.04** (0.02)	0.61 (0.01)	0.61 (0.01)	-0.00 (0.02)
Black	0.10 (0.01)	0.09 (0.01)	-0.01 (0.01)	0.10 (0.01)	0.09 (0.01)	-0.00 (0.01)
College Degree	0.41 (0.01)	0.39 (0.01)	-0.03* (0.02)	0.40 (0.01)	0.40 (0.01)	-0.00 (0.02)
Plant Closure	0.46 (0.01)	0.62 (0.01)	0.16*** (0.02)	0.54 (0.01)	0.54 (0.01)	-0.00 (0.02)
Union Membership	0.15 (0.01)	0.16 (0.01)	0.01 (0.01)	0.15 (0.01)	0.15 (0.01)	0.00 (0.01)
In Metro Area	0.84 (0.01)	0.82 (0.01)	-0.01 (0.01)	0.83 (0.01)	0.83 (0.01)	0.00 (0.01)
Years of Tenure	7.12 (0.15)	9.18 (0.16)	2.06*** (0.22)	8.25 (0.16)	8.21 (0.15)	-0.04 (0.22)
Log Earnings	6.54 (0.01)	6.56 (0.01)	0.03 (0.02)	6.54 (0.01)	6.55 (0.01)	0.00 (0.02)
Observations	1959	2216		1959	2216	

Note: The sample consists of respondents from the Displaced Worker Supplement (DWS) for the years 1996-2020, who were between the ages of 21 to 64, who had worked full-time for at least six months at their previous job, received health insurance from their former employer, and did not expect to be recalled. Short notice refers to a notice period of 1-2 months, while long notice refers to a notice period exceeding two months. Columns (1) and (2) present raw averages for the sample, while columns (3) and (4) show weighted averages, where the weights correspond to the inverse of the estimated probabilities of receiving short or long notice.

longer notice tend to be older and are more likely to be married. Additionally, workers laid off during plant closures are more likely to receive longer notice, potentially due to compliance with the WARN law. Workers with longer notice also tend to have longer job tenure. However, there are no notable differences in earnings for the two groups.

To isolate the impact of notice from these additional correlates, which may affect the probability of exiting unemployment, I reweight the sample using inverse propensity score weighting. I use a logistic regression model to predict the likelihood of receiving a longer notice based on several covariates. These covariates consist of age, gender, marital status, race, education, location characteristics, the reason for displacement, year of displacement, industry and occupation of the lost job, as well as union status, tenure, and earnings at the lost job. I then utilize the propensity scores to assign weights to the observations. Specifically, individuals with the long notice are assigned a weight of $1/\hat{p}(X_i)$, where $\hat{p}(X_i)$ is the estimated probability of receiving the long notice from the regression model for an individual with covariates X_i . On the other hand, individuals who received the short notice are assigned a weight of $1/(1 - \hat{p}(X_i))$.

The summary statistics for the reweighted sample are presented in columns (3) and (4) of Table 1. After reweighting, the observable differences between the two groups disappear, indicating that the weights effectively minimize the observed disparities. Section C.2 in the Appendix provides additional details on propensity score estimation. Figure C1 demonstrates a high degree of overlap in the estimated propensity score distributions between long and short-notice workers. Additionally, Figures C2 and C3 depict the balance of the weighted sample with respect to the displacement year and industrial and occupational composition, respectively.

II.C Distribution of Unemployment Duration

In this section, I explore how a longer notice impacts the exit rate over the spell of unemployment. Workers who receive a layoff notice may start searching for a job

before separating from their previous employer. In this case, some of these workers may secure a new job during the notice period, thus avoiding any period of unemployment. In the data, 14.7% of the workers with the short layoff notice report no duration of unemployment. Since workers with longer notice periods have more time to search for a new job while still employed, we expect their chances of avoiding unemployment to be even greater.

In Table 2, panel A, I examine the relationship between receiving a long notice and reporting an unemployment duration of 0. Columns (1) and (2) present estimates from unweighted regressions, while columns (3) and (4) present weighted regression estimates using the weights described in the previous section. Additionally, columns (2) and (4) include a comprehensive set of controls identical to the ones used to generate the weights. The table shows that the impact of a lengthier notice on the exit probability is reduced after accounting for observable characteristics of the separation. The coefficient in column (2) indicates that individuals who receive a longer notice are 8 percentage points more likely to avoid unemployment. Similar estimates are observed in columns (3) and (4) as well. Notably, the inclusion of controls in column (4) does not lead to a change in the coefficient, indicating that the weighting has effectively achieved balance in terms of the covariates across the two groups. In panel B of Table 2, I present a similar regression analysis, but this time using an indicator for exiting unemployment within the first 12 weeks. The results show that the exit rate out of unemployment is around 7 percentage points higher for long-notice workers compared to the short-notice group.

To examine how the exit rate varies with the length of notice over the spell of unemployment, I bin unemployment duration into 12-week intervals.¹³ Figure 1 presents the exit rate and the survival rate separately for the long- and short-notice workers over the

¹³Exit rate is defined as the ratio of individuals who found a job during a specific interval to those who were jobless at the beginning of the interval. See Figure C5 in the Appendix for the presentation of data with alternative binning definitions.

TABLE 2: OBSERVED EXIT RATE – EARLY IN THE SPELL

	(1)	(2)	(3)	(4)
PANEL A. $\mathbb{I}\{\text{UNEMPLOYMENT DURATION} = 0 \text{ WEEKS}\}$				
> 2 month notice	0.094*** (0.012)	0.080*** (0.012)	0.077*** (0.013)	0.077*** (0.013)
PANEL B. $\mathbb{I}\{\text{UNEMPLOYMENT DURATION} \leq 12 \text{ WEEKS}\}$				
> 2 month notice	0.078*** (0.015)	0.074*** (0.016)	0.070*** (0.016)	0.070*** (0.016)
Controls	No	Yes	No	Yes
Weights	No	No	Yes	Yes
Observations	4175	4175	4175	4175

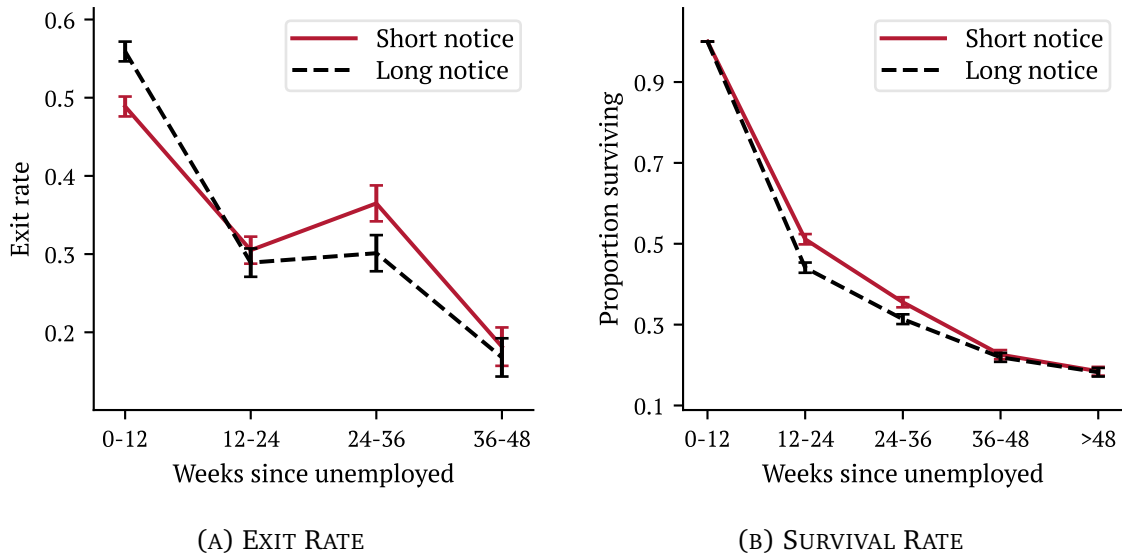
Note: The table presents estimates from linear regression models, where the main independent variable is an indicator variable that takes a value of 1 if the individual received a notice of more than 2 months and 0 if they received a notice of 1-2 months. The dependent variable is an indicator for reporting an unemployment duration of 0 weeks (Panel A) or less than 12 weeks (Panel B). The weights are generated using inverse probability weighting (IPW). Robust standard errors are reported in the parenthesis.

spell of unemployment. Note that the rates are calculated using the weighted sample to ensure that the comparison is between similar groups of workers who received different lengths of notice.¹⁴ Approximately 56.6% of individuals with a long notice exit within the first 12 weeks, while 48.7% of those with a short notice do the same. However, over the course of unemployment, individuals with shorter notice periods catch up, resulting in almost identical survival rates for both groups by the 48th week of unemployment. As shown in panel A of Figure 1, for all durations beyond 12 weeks, individuals with shorter notice periods have a higher exit rate compared to those with longer notice periods.

I interpret the higher exit rate for short-notice workers beyond the initial 12 weeks

¹⁴Appendix D presents the unweighted exit rates and corresponding estimates obtained from the Mixed Hazard model. There is little qualitative difference between unweighted and weighted quantities, suggesting that differences in other characteristics of the two groups are not driving the results.

FIGURE 1: EXIT AND SURVIVAL RATE — LATER IN THE SPELL



Note: Short notice refers to a notice of 1-2 months, and long notice refers to a notice of more than 2 months. Panel A presents the weighted proportion of individuals exiting unemployment in each interval amongst those who were still unemployed at the beginning of the interval. Panel B presents the weighted proportion of individuals who are unemployed at the beginning of each interval. Error bars represent 90% confidence intervals.

as evidence for heterogeneity across workers. When workers are heterogeneous, those with better chances of exiting unemployment do so earlier. Given that a larger proportion of long-notice workers exit earlier, the long-notice group will have a lower proportion of individuals with higher exit probabilities, which is reflected in the (average) exit rate. It is important to note that this interpretation holds under the condition that longer notice does not directly reduce the probability of exiting unemployment at later durations. In the following section, I formally outline the assumptions necessary to identify heterogeneity and duration dependence in a Mixed Hazard model, and I also discuss the plausibility of these assumptions and potential violations.

III ECONOMETRIC FRAMEWORK

This section illustrates how variation in notice length can be used to identify structural duration dependence. Specifically, I set up a Mixed Hazard model in discrete time and specify the assumptions under which the key components of this model are identified. The model and assumptions are outlined in Section III.A, while the main identification result is presented in Section III.B. The intuition behind identification is explained in Section III.C. All proofs are presented in Appendix A.

III.A Mixed Hazard Model in Discrete Time

Prior to being laid off, workers are given a notice period of length L , where L takes discrete values on some support $\mathcal{L}$. Workers are heterogeneous, each possessing an unobservable fixed type ν , characterized by the cumulative distribution $F(\cdot)$. Additionally, workers vary based on pre-notice observables, denoted by X , with the distribution of X given by $F_X(\cdot)$. These characteristics may include details about the job, the worker, or the circumstances of the layoff. Using potential outcome notation (Rubin, 1974), let D_L denote the realized unemployment duration under notice L . D_L is a random variable that takes values in $\{1, 2, 3, \dots\}$. Let's begin by assuming that the econometrician observes $D = \sum_{\ell \in \mathcal{L}} D_\ell \mathbb{I}\{L = \ell\}$, which means that the potential duration under notice ℓ is observed among those who receive notice ℓ .¹⁵ Right-censored duration data will be incorporated later.

Finally, for a set of conditioning variables Υ , define the hazard function $h(\cdot)$ as:

$$h(d|\Upsilon) = \Pr(D = d | D \geq d, \Upsilon)$$

Thus, $h(d|\nu, L, X)$ represents the probability that an individual of type ν with observed characteristics X will exit unemployment at duration d , given that the individual has

¹⁵Throughout the text, $\mathbb{I}(\cdot)$ is used to denote the indicator function.

not yet exited and had received notice L .

We are interested in understanding how individual exit probabilities $h(d|\nu, L, X)$ evolve over the duration of unemployment; this is referred to as structural duration dependence. However, these individual probabilities are not observed in the data. Instead, we can utilize the observed unemployment durations to determine the timing of each worker's exit from unemployment. Specifically, we can at most infer the exit rate $h(d|L, X)$, which represents the proportion of individuals exiting at a specific duration relative to those who have remained up to that point in the group with notice L and observable characteristics X .

It is worth noting that, according to the definition of the hazard function, we have:

$$h(d|L, X) = \mathbb{E}[h(d|\nu, L, X)|D \geq d, L, X]$$

The above reformulation clarifies that the exit rate $h(d|L, X)$ captures the average hazard for individuals who have survived until duration d rather than all individuals. This is what makes identifying structural duration dependence from observed exit rates challenging. To make the problem more manageable, the Mixed Hazard model introduces additional structure by assuming that the individual exit probabilities $h(d|\nu, L, X)$ can be expressed as the product of the individual's unobserved type ν and a structural component $\psi_L(d, X)$.¹⁶

Assumption 1. (*Mixed Hazard*) An individual's exit probability at duration d is given by:

$$h(d|\nu, L, X) = \psi_L(d, X)\nu$$

where the structural hazard $\psi_L(d, X) \in (0, \infty)$ and worker's type $\nu \in (0, \bar{\nu}]$ with $\bar{\nu} =$

¹⁶Lancaster (1979) expanded the proportional hazard model (Cox, 1972) to incorporate unobserved heterogeneity. His Mixed Proportional Hazard (MPH) model represented the hazard rate as a product of a regression function, a structural hazard that varies with duration, and the worker's unobserved type. The Mixed Hazard model formulated here is similar to Lancaster's MPH model but permits non-proportional effects of observable characteristics and distinguishes the length of notice from other observed variables.

$$1/\max_{d,L,X}\{\psi_L(d,X)\}.$$

The structural hazard $\psi_L(d,X)$ is common to all individuals with observable characteristics X and a notice period of length L , but it varies with the duration of unemployment. The restrictions on the structural hazard and the support of ν in Assumption 1 guarantee that individual exit probabilities lie between 0 and 1. Under Assumption 1, workers with lower values of type ν have lower exit probabilities at all durations. Consequently, low-type workers tend to remain unemployed for longer durations. Given that the observed exit rate reflects the average exit probability of surviving workers, it will decline more steeply than individual exit probabilities over the unemployment spell. The following proposition formally states this result.

Proposition 1. *Under Assumption 1, the exit rate $h(d|L,X)$ can be expressed as follows:*

$$h(d|L,X) = \psi_L(d,X)\mathbb{E}(\nu|D \geq d, L, X)$$

Moreover, the average type of workers who survive until d , $\mathbb{E}(\nu|D \geq d, L, X)$, decreases with the unemployment duration d .

This proposition clarifies why identifying the structural hazard $\psi_L(d,X)$ is challenging. The exit rate $h(d|L,X)$ is impacted by both the structural duration dependence $\psi_L(d,X)$ and the changing worker composition over the unemployment spell captured by $\mathbb{E}(\nu|D \geq d, L, X)$. If the observed exit rate $h(d|L,X)$ declines over the spell of unemployment, it is not possible to distinguish between the scenario where there is no structural duration dependence but significant worker heterogeneity causes the average type of workers and the observed exit rate to decline, and the scenario where there is no worker heterogeneity, but the structural hazard declines over the spell of unemployment. Both of these scenarios would be consistent with the observed decline in $h(d|L,X)$. Hence, structural duration dependence is not identified from exit rates in the model formulated so far.

I now introduce two additional assumptions under which variation in notice length

leads to the identification of structural duration dependence. The first assumption is conditional independence, which states that the length of notice is independent of the worker's unobservable type given observable characteristics. In other words, for workers with similar observable characteristics, there is no systematic difference in the length of notice given to workers with different unobservable types.¹⁷

Assumption 2. (*Conditional Independence*) *The length of notice L is independent of the worker's unobservable type v , given observable characteristics X , i.e., $L \perp v|X$.*

This assumption ensures that differences in the exit rates $h(d|L, X)$ for various notice lengths are only due to the direct effect of the notice on individual exit probabilities rather than due to differences in the types of workers who receive different lengths of notice. In other words, the assumption implies that for any notice ℓ , $h(d|\ell, X) = \mathbb{E}[h(d|v, \ell, X)|D_\ell \geq d, X]$. The second assumption, referred to as stationarity, states that the length of notice does not affect an individual's exit probability after the first period.

Assumption 3. (*Stationarity*) *For all L, X , and $d > 1$,*

$$\psi_L(d, X) = \psi(d, X)$$

The rationale for Assumption 3 is that workers with longer notice periods have more time to search for a new job before separating from their previous employer, potentially increasing their likelihood of finding a job at the beginning of their unemployment spell. However, suppose duration dependence in job-finding is caused by factors such as human capital depreciation due to prolonged unemployment or employers discriminating against long-term unemployed workers. In that case, a worker's exit probability later in the spell should only vary with the unemployment duration and not with the length

¹⁷In [Appendix E](#), I provide estimates from a model that permits the underlying type distribution to differ across various notice lengths instead of assuming conditional independence.

of notice received at the onset of the spell. Given that I bin unemployment duration in 12-week intervals, this assumption translates to the length of notice only impacting the probability of exit within the first 12 weeks and not after that.

More generally, Assumption 3 implies that individual exit probabilities vary only with the duration of unemployment and not with time elapsed since the start of the job search.¹⁸ This assumption would be violated if time spent searching increases or decreases an individual's likelihood of exiting unemployment. For instance, if workers learn while searching and become better at job search (Burdett and Vishwanath, 1988; Gonzalez and Shi, 2010), then those with longer notice would have a higher hazard even beyond the initial period. On the other hand, time spent searching may decrease the exit probability if workers first apply to all jobs in stock but subsequently only apply to newly posted jobs (Coles and Smith, 1998).¹⁹ In Appendix E, I investigate whether Assumption 3 is violated by estimating a more general model where structural hazards are allowed to vary with the length of notice beyond the initial period, and I do not find evidence against it.

III.B Identification Results

Theorem 1. *Under Assumptions 1–3, for any ℓ, ℓ' with $\psi_\ell(1, X) \neq \psi_{\ell'}(1, X)$ and some integer $\bar{D}$, the structural hazards $\{\psi_\ell(1, X), \psi_{\ell'}(1, X), \{\psi(d, X)\}_{d=2}^{\bar{D}}\}$ and the conditional moments of the type distribution $\{\mathbb{E}(v^k | X)\}_{k=1}^{\bar{D}}$ are identified up to a scale from the conditional exit rates $\{h(d | \ell, X), h(d | \ell', X)\}_{d=1}^{\bar{D}}$.*

¹⁸This assumption aligns with a large class of search models, including those that involve non-stationarity. For instance, the model proposed by Lentz and Tranæs (2005), in which workers start searching harder over time as their savings run down, would be consistent with this assumption as savings only start depleting once unemployed.

¹⁹Another possibility for why individual exit probability may decline with time spent searching could be that workers get discouraged over time and stop trying. However, this seems unlikely since 85% of individuals in the sample eventually find employment after displacement, and among those who do not, none report being out of the labor force.

The above theorem establishes that if the first-period hazard varies for two different notice lengths, we can identify the structural hazards up to $\bar{D}$ and the first $\bar{D}$ moments of ν conditional on X using the conditional exit rates for both notice lengths up to $\bar{D}$. A direct implication of this result is that if X does not enter the structural hazard and we assume independence instead of conditional independence, we can identify the model using exit rates unconditional on X . The following corollary presents this result formally.

Corollary 1. *Assume the following conditions hold:*

- (i) $h(d|\nu, L) = \psi_L(d)\nu$, where $\psi_L(d)$ and ν are bounded to ensure $h(d|\nu, L) \in (0, 1)$
- (ii) L is independent of ν
- (iii) $\psi_L(d) = \psi(d)$ for $d > 1$

Then for any ℓ and ℓ' , with $\psi_\ell(1) \neq \psi_{\ell'}(1)$ and some integer $\bar{D}$, the structural hazards $\{\psi_\ell(1), \psi_{\ell'}(1), \{\psi(d)\}_{d=2}^{\bar{D}}\}$ and the moments of the type distribution $\{\mathbb{E}(\nu^k)\}_{k=1}^{\bar{D}}$ are identified up to a scale from the exit rates $\{h(d|\ell), \{h(d|\ell')\}_{d=1}^{\bar{D}}\}$

Neither of the two results mentioned above is ideal for application to the data. The first result has a limitation in that $h(d|L, X)$ is only well-defined for discrete values of X , and even then, it may be imprecisely estimated if each bin size is not large enough. On the other hand, the second result allows us to use duration distributions that are unconditional on X , but it imposes a stronger restriction of unconditional independence, which may not hold in the data.

To address these limitations, I present an additional result below, which allows controlling for observables more flexibly. Specifically, suppose observable characteristics enter the structural hazard proportionally, as in the MPH model. In that case, the model's parameters are identified under conditional independence using inverse propensity score weighted (IPW) exit rates, denoted by $h^w(d|L)$.²⁰ Specifically, $h^w(d|L)$ is defined as follows:

²⁰Note that the length of notice still enters the structural hazard non-proportionally.

$$h^w(d|L) = \mathbb{E} \left[\frac{\Pr(D = d | \nu, L, X)}{p_L(X)} \middle| L \right] / \mathbb{E} \left[\frac{\Pr(D \geq d | \nu, L, X)}{p_L(X)} \middle| L \right]$$

where $p_\ell(X) = \Pr(L = \ell | X)$.

In words, the IPW exit rate $h^w(d|L)$ is computed by using weighted averages of individual probabilities, where the weights are equal to the probability of the individual receiving a specific notice length based on their observable characteristics X . As shown in the context of treatment effects (Rosenbaum, 1987), inverse propensity score weighting leads to the elimination of bias due to selection into treatment based on observable characteristics. This is because it gives less weight to individuals with characteristics that lead to a higher treatment probability, thereby readjusting the sample to reflect the general population more accurately. Therefore, the variation in $h^w(d|L)$ by notice length will capture the direct effect of notice on exit rates without being confounded with the effects of observable characteristics that correlate with notice length.

Put another way, for any notice length ℓ , while the unweighted exit rate $h(d|\ell)$ reflects $\mathbb{E}[h(d|\nu, \ell, X) | D_\ell \geq d, \ell]$, the IPW adjusted exit rate captures $\mathbb{E}[h(d|\nu, \ell, X) | D_\ell \geq d]$. This fact is formally shown in the proof of the subsequent proposition, which presents the identification result using weighted exit rates. It is important to reiterate that, in this context, we must assume that X enters the structural hazard proportionally, such that $\psi_L(d, X) = \psi_L(d)\phi(X)$. Consequently, it is useful to define $\theta(X, \nu) = \phi(X)\nu$, which now represents the combined type of workers based on both observed and unobserved characteristics.

Proposition 2. *Assume that Assumptions 1–3 are satisfied, and also that the additional conditions $\psi_L(d, X) = \psi_L(d)\phi(X)$ and $0 < p_L(X) < 1$ for all L, X are met. Then for any ℓ and ℓ' , with $\psi_\ell(1) \neq \psi_{\ell'}(1)$ and some integer $\bar{D}$, the structural hazards $\{\psi_\ell(1), \psi_{\ell'}(1), \{\psi(d)\}_{d=2}^{\bar{D}}\}$ and the moments $\{\mathbb{E}(\theta(X, \nu)^k)\}_{k=1}^{\bar{D}}$ are identified up to a scale from the weighted exit rates $\{h^w(d|\ell), \{h^w(d|\ell')\}_{d=1}^{\bar{D}}\}$*

The identification results discussed thus far rely on observing completed unemploy-

ment durations. However, as is typical in many datasets, some individuals are still unemployed at the time of the DWS survey. For the unemployed individuals, we observe how long they have been unemployed, but we do not know if and when they will find a job. I now extend Proposition 2 to incorporate right-censored unemployment durations.

Let D^c denote the censoring time, which is the time elapsed from when an individual becomes unemployed to the time of the survey. For individuals who have already exited unemployment at the time of the survey, we observe their completed unemployment duration D in the data. However, we only observe the censoring time D^c for currently unemployed individuals. Specifically, for each individual, we observe $\tilde{D} = \min\{D, D^c\}$ along with an indicator variable for whether the individual was censored or not, denoted by $C = \mathbb{I}\{D^c < D\}$.

Intuitively, the exit rate at duration d can be calculated from observed durations as the proportion of individuals who find a job at d , indicated by $\tilde{D} = d$ and $C = 0$, among those whose observed duration $\tilde{D}$ is at least equal to d . Since we are limiting ourselves to individuals whose observed duration is at least d , it means that everyone in this group has remained unemployed up to duration d . However, this approach excludes individuals who are censored before d , among whom some may have also remained unemployed until d . Therefore, the exit rate calculated in this manner captures the exit rate for those who are censored after d . Mathematically,

$$\tilde{h}(d|\Upsilon) = \Pr(\tilde{D} = d, C = 0 | \tilde{D} \geq d, \Upsilon) = \Pr(D = d | D \geq d, D^c \geq d, \Upsilon)$$

The second equality in the above expression follows from the definition of $\tilde{D}$ and C .

To incorporate inverse propensity score weighting while calculating exit rates based on observed durations and the censoring indicator, we can proceed as before and define the IPW exit rate using observed durations as follows:

$$\tilde{h}^w(d|L) = \mathbb{E} \left[\frac{\Pr(\tilde{D} = d, C = 0 | v, L, X)}{p_L(X)} \middle| L \right] / \mathbb{E} \left[\frac{\Pr(\tilde{D} \geq d | v, L, X)}{p_L(X)} \middle| L \right]$$

The proposition below states that, under the assumptions specified in Proposition 2, the structural hazards are identified from $\tilde{h}^w(d|L)$ with the additional assumption $D^c \perp L \mid X$. Additionally, if we also assume that D^c is independent of both X and ν , moments of $\theta(X, \nu)$ are also identified. In my application, these assumptions are generally innocuous, as the censoring time is determined by when an individual was surveyed in DWS, which should not correlate with individual outcomes.

Proposition 3. *Assume that the assumptions stated in Proposition 2 hold and additionally $D^c \perp L \mid X$. Then for any ℓ and ℓ' , with $\psi_\ell(1) \neq \psi_{\ell'}(1)$ and some integer $\bar{D}$, the structural hazards $\{\psi_\ell(1), \psi_{\ell'}(1), \{\psi(d)\}_{d=2}^{\bar{D}}\}$ are identified up to a scale from the weighted exit rates $\{\tilde{h}^w(d|\ell), \{\tilde{h}^w(d|\ell')\}_{d=1}^{\bar{D}}\}$. Furthermore, if $D^c \perp \nu$ and $D^c \perp X$, then the moments $\{\mathbb{E}(\theta(X, \nu)^k)\}_{k=1}^{\bar{D}}$ are also identified.*

Proposition 3 is the main proposition utilized in the application in the paper. Section IV builds a GMM estimator based on this identification result.

III.C Intuition for Identification

In this section, I elucidate the intuition behind the identification result. To simplify the explanation, I focus on the case without observable characteristics, as incorporating them does not provide any additional insights regarding identification.²¹ In this model, an individual worker's exit probability is given by $h(d|\ell, \nu) = \psi_\ell(d)\nu$ and ν is independent of L . Note that independence implies $f(\nu|L) = f(\nu)$. For brevity, let us denote the first and second moments of ν by $\mu_1 = \mathbb{E}(\nu)$ and $\mu_2 = \mathbb{E}(\nu^2)$, respectively. It is worth noting that the variance of ν , given by $\text{var}(\nu) = \mu_2 - \mu_1^2$, captures the extent of heterogeneity across workers.

To see why the identification result holds, note that the exit rate in the first period is

²¹A similar exposition can also be found in van den Berg and van Ours (1996), where the authors discuss identification utilizing calendar-time effects.

given by $h(1|\ell) = \psi_\ell(1)\mu_1$ and the exit rate at $d = 2$ is given by:

$$h(2|\ell) = \frac{Pr(D = 2|\ell)}{Pr(D \geq 2|\ell)} = \psi(2) \left(\frac{\mu_1 - \psi_\ell(1)\mu_2}{1 - \psi_\ell(1)\mu_1} \right) = \psi(2)\mu_1 \left(\frac{1 - h(1|\ell)(\mu_2/\mu_1^2)}{1 - h(1|\ell)} \right)$$

The third equality in the above equation follows from $\psi_\ell(1) = h(1|\ell)/\mu_1$. In the presence of heterogeneity, the variance of ν is greater than zero, which means that $\mu_2/\mu_1^2 > 1$. Therefore, based on the expressions for $h(1|\ell)$ and $h(2|\ell)$, we can observe that $h(2|\ell)/h(1|\ell)$ will always be smaller than $\psi(2)/\psi_\ell(1)$. Furthermore, the greater the variance of ν (i.e., the more heterogeneity across workers), the larger μ_2/μ_1^2 will be, and the more distant $h(2|\ell)/h(1|\ell)$ will be from $\psi(2)/\psi_\ell(1)$. This occurs because greater heterogeneity across workers implies that the composition of workers from the first to the second period changes more drastically. For instance, in the absence of heterogeneity across workers where $\mu_2/\mu_1^2 = 1$, the composition across both periods is unchanged, and thus $h(2|\ell)/h(1|\ell) = \psi(2)/\psi_\ell(1)$.

If we knew the extent of heterogeneity across workers as captured by μ_2/μ_1^2 , we could determine how the composition changes from the first to the second period and estimate the structural duration dependence $\psi(2)/\psi_\ell(1)$ from the observed duration dependence $h(2|\ell)/h(1|\ell)$. The variation in notice lengths allows us to learn about the underlying heterogeneity and estimate structural duration dependence. To understand why this is the case, note that for two lengths of notice ℓ and ℓ' , the following expression holds:

$$\frac{h(2|\ell)}{h(2|\ell')} = \left(\frac{1 - h(1|\ell)(\mu_2/\mu_1^2)}{1 - h(1|\ell)} \right) / \left(\frac{1 - h(1|\ell')(\mu_2/\mu_1^2)}{1 - h(1|\ell')} \right)$$

Assuming without loss of generality that $h(1|\ell') > h(1|\ell)$, we can see from the above expression that then $h(2|\ell)/h(2|\ell') \geq 1$. This is because more individuals with notice length ℓ' leave in the first period, leading to a worse composition for that group in the second period. Furthermore, when the variance across workers is higher, $h(2|\ell)$ will be further above $h(2|\ell')$. Thus, the difference in exit rates among workers with different notice lengths provides information about the degree of heterogeneity, and we can use

the above expression to compute μ_2/μ_1^2 . Once we know μ_2/μ_1^2 , we can plug that back into the expression for $h(2|\ell)/h(1|\ell)$ and estimate the structural duration dependence $\psi(2)/\psi_\ell(1)$. In summary, the difference in exit rates at duration $d = 2$ across notice lengths reflects differences in the composition of remaining workers. Therefore, comparing exit rates of workers with different notice lengths can provide insights into the extent to which underlying heterogeneity impacts exit rates. A similar argument applies to identifying structural hazards beyond the second period.²²

IV ESTIMATION

Generalized Method of Moments (GMM). Using the identification result presented in Proposition 3, we can use the Generalized Method of Moments (GMM) to construct a consistent estimator for the structural hazards and moments of the heterogeneity distribution. Since the model is identified only up to scale, I normalize the first moment by setting $\mathbb{E}[\theta(X, \nu)] = 1$. With J possible notice lengths, the vector of unknown parameters is given by $\Theta = \{\{\psi_\ell(1)\}_{\ell=1}^J, \{\psi(d)\}_{d=2}^{\bar{D}}, \{\mathbb{E}[\theta(X, \nu)^k]\}_{k=2}^{\bar{D}}\}$ and has a total of $2(\bar{D} - 1) + J$ unknown parameters.

Now, for each individual i , define the following moment condition:

$$m_i(\ell, d; \Theta) = \mathbb{I}\{L_i = \ell\} \cdot \left[\frac{\mathbb{I}\{\tilde{D}_i = d\} \mathbb{I}\{C_i = 0\}}{p_\ell(X_i)} - \tilde{h}^w(d|\ell; \Theta) \cdot \frac{\mathbb{I}\{\tilde{D}_i \geq d\}}{p_\ell(X_i)} \right]$$

Given the definition of $\tilde{h}^w(d|\ell; \Theta)$, it follows that $\mathbb{E}[m_i(\ell, d; \Theta)] = 0$, as reasoned in Section A.7. We can now stack moment conditions pertaining to different notice lengths and durations in one vector, denoted by $\mathbf{m}_i(\Theta) = \{\{m_i(\ell, d, \Theta)\}_{d=1}^{\bar{D}}\}_{\ell=1}^J$. Observe that

²²To understand why higher moments determine the hazard at later durations, we can consider how the composition of workers changes from $d = 2$ to $d = 3$. This change depends on the level of heterogeneity across workers at the start of $d = 2$. If the distribution of heterogeneity has a positive skew, the variance among individuals who survive to $d = 2$ would be lower than that among individuals at the start of $d = 1$. This is because the few individuals with a high likelihood of exiting unemployment would have already left, reducing the variance among surviving workers.

$\mathbf{m}_i(\Theta)$ contains $J \times \bar{D}$ moment conditions. As shown in Proposition 3, our parameters of interest are identified from these moment conditions as long as $J > 1$. This condition also ensures that the number of moments is equal to or greater than the number of parameters.

The corresponding sample average to $\mathbb{E}[m_i(\ell, d; \Theta)]$ can be written as:

$$\hat{m}(\ell, d; \Theta) = \frac{1}{n} \sum_{i=1}^n m_i(\ell, d; \Theta) = \hat{h}^{num}(d|\ell) - \tilde{h}^w(d|\ell; \Theta) \cdot \hat{h}^{den}(d|\ell)$$

Here, n is the sample size, and $\hat{h}^{num}(d|\ell)$ and $\hat{h}^{den}(d|\ell)$ are defined as follows:

$$\hat{h}^{num}(d|\ell) = \frac{1}{n} \sum_{i=1}^n \frac{\mathbb{I}\{L_i = \ell\} \mathbb{I}\{\tilde{D}_i = d\} \mathbb{I}\{C_i = 0\}}{p_\ell(X_i)}, \quad \hat{h}^{den}(d|\ell) = \frac{1}{n} \sum_{i=1}^n \frac{\mathbb{I}\{L_i = \ell\} \mathbb{I}\{\tilde{D}_i \geq d\}}{p_\ell(X_i)}$$

Note that $\hat{h}(d|\ell) = \hat{h}^{num}(d|\ell) / \hat{h}^{den}(d|\ell)$ represents the sample counterpart to $\tilde{h}^w(d|\ell; \Theta)$.

As before, stack the sample moments in a single vector $\hat{\mathbf{m}}(\Theta)$, such that $\hat{\mathbf{m}}(\Theta) = \sum_{i=1}^n \mathbf{m}_i(\Theta) / n$.

The GMM estimator $\hat{\Theta}$ is then given by: $\hat{\Theta} = \arg \max_{\Theta} \hat{\mathbf{m}}(\Theta)' \hat{W} \hat{\mathbf{m}}(\Theta)$. When the model is just-identified, $\hat{W}$ is given by the identity matrix. In the case of over-identification, the efficient weighting matrix is given by $\hat{W} = \hat{\Omega}^{-1}$, where $\hat{\Omega} = [\frac{1}{n} \sum_{i=1}^n \mathbf{m}_i(\hat{\Theta}) \mathbf{m}_i(\hat{\Theta})']$. Using the two-step estimation process, we can compute $\hat{\Theta}$. The asymptotic distribution of this estimator is given by $\sqrt{n}(\hat{\Theta} - \Theta) \rightarrow N(0, (\hat{M}' \hat{\Omega}^{-1} \hat{M})^{-1})$, where $\hat{M} = \partial \hat{\mathbf{m}}(\hat{\Theta}) / \partial \Theta$.

Functional Form for Structural Hazard. Even though the model is identified non-parametrically, given small sample sizes, to minimize the number of estimated parameters, I assume that the structural hazard $\psi(d)$ for $d > 1$ has a log-logistic form as follows

$$\psi(d) = \frac{(\alpha_2 / \alpha_1)(d / \alpha_1)^{\alpha_2 - 1}}{1 + (d / \alpha_1)^{\alpha_2}} \quad (1)$$

where $\alpha_1 > 0, \alpha_2 > 0$. The hazard function in equation (1) is monotonically decreasing when $\alpha_2 \leq 1$ and is unimodal, initially increasing and subsequently decreasing when $\alpha_2 > 1$. The mode or the turning point is $\alpha_1(\alpha_2 - 1)^{1/\alpha_2}$.²³

²³This provides a flexible parametrization for the structural hazard relative to other commonly used

V DURATION DEPENDENCE AND HETEROGENEITY

VA *Baseline Estimates*

Table 3 presents the main estimates from the Mixed Hazard model. Since I normalized the mean of worker type to equal 1, the estimated structural hazards corresponding to the first period for short and long-notice individuals coincide with their corresponding observed exit rates in the data. The last two lines in panel A of Table 3 show the estimated parameters for the log-logistic function specified in equation (1) used to model structural dependence.

The structural hazards implied by these parameters are presented in panel B of Table 3 and panel A of Figure 2. Additionally, panel A of Figure 2 shows the observed exit rate from the data, averaged across workers with short and long notice, alongside the estimated hazard. This figure shows that the estimated hazard consistently exceeds the observed hazard throughout the unemployment spell, indicating the role of underlying heterogeneity. While the observed hazard in the data declines by 43.6% over the first 24 weeks, the estimated structural hazard only decreases by 33.6% during the same period. Hence, after accounting for heterogeneity, almost three-quarters of the observed decline in the first 24 weeks can be attributed to structural duration dependence.

However, the estimated structural hazard increases by 75% from 12-24 to 24-36 weeks, a much more pronounced increase than the observed hazard. This pattern possibly reflects individuals approaching the exhaustion of their unemployment insurance (UI) benefits.²⁴ As previously noted, there is variation across individuals in the eligible duration of UI receipt. However, a significant proportion of individuals are eligible for UI benefits that last for 26 weeks, which coincides with the third interval. Figure C4

parametrization, such as Weibull or Gompertz, as it allows the structural hazard to be non-monotonic. However, I also present non-parametric estimates in [Appendix D](#).

²⁴Since the sample consists of displaced workers, a substantial portion of these individuals should be eligible for UI benefits. Table C4 in the Appendix shows that more than 80% of individuals unemployed for longer than 12 weeks report receiving UI benefits in the sample.

TABLE 3: ESTIMATION RESULTS

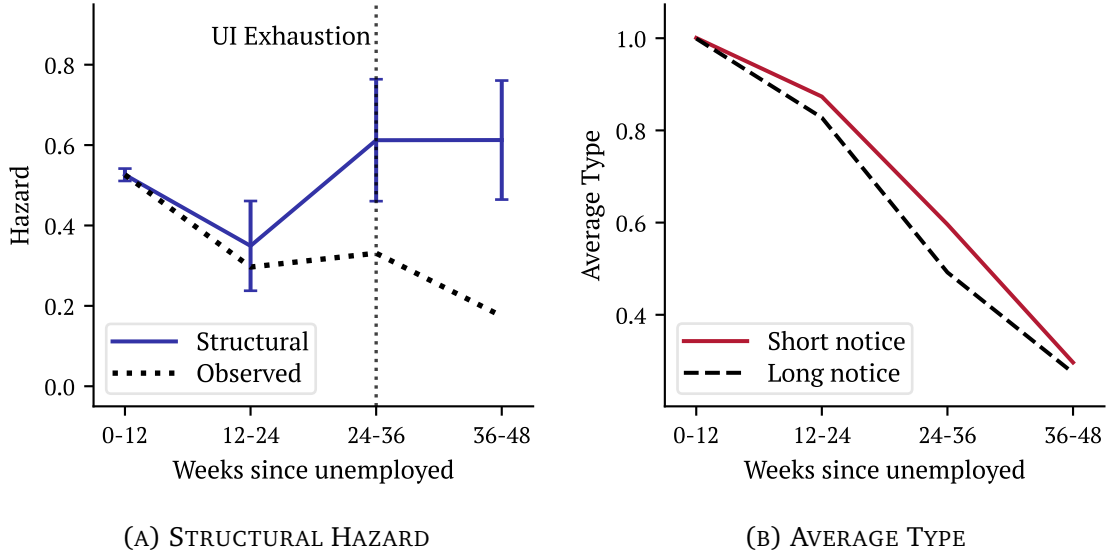
Parameter	Explanation	Estimate	SE
<i>Panel A: Estimated Parameters</i>			
$\psi_S(1)$	Structural hazard 0-12 weeks: Short notice	0.49	0.01
$\psi_L(1)$	Structural hazard 0-12 weeks: Long notice	0.56	0.01
α_1	Scale parameter for $\psi(d)$	2.06	0.17
α_2	Shape parameter for $\psi(d)$	2.54	0.27
<i>Panel B: Duration Dependence</i>			
$\bar{\psi}(1)$	Structural hazard: 0-12 weeks	0.53	0.01
$\psi(2)$	Structural hazard: 12-24 weeks	0.35	0.07
$\psi(3)$	Structural hazard: 24-36 weeks	0.61	0.09
$\psi(4)$	Structural hazard: 36-48 weeks	0.61	0.09
<i>Hansen-Sargan Test</i>			
Test statistic: 0.01		Critical value, $df = 1, \chi^2_{0.05}$: 3.84	

Note: The table presents estimates from the Mixed Hazard model. The first moment is normalized to one, and structural duration dependence is specified by equation (1). Panel A shows the estimated parameters from the model, and panel B presents structural hazards implied by the estimated parameters. The standard errors for the structural hazards are calculated using the delta method.

in the Appendix shows that the proportion of individuals reporting exhausting their UI benefits jumps up significantly in the third interval. This finding of increasing structural hazard leading up to benefit exhaustion is consistent with individuals intensifying their job search efforts or lowering their expectations to secure employment before depleting their benefits. Finally, while the observed hazard continues declining even after 36 weeks, the structural hazard remains constant.

Panel B of Figure 2 displays the average type implied by the model for individuals with short and long notice periods throughout the unemployment spell. For both groups, the average type deteriorates over the course of unemployment. However, for individuals with longer notice periods, the composition worsens more between 0-12 and 12-24 weeks, indicating a higher exit rate in the initial period for this group. By the end of 36 weeks, when a significant number of individuals have already left unemployment,

FIGURE 2: BASELINE ESTIMATES



Note: The solid blue line in panel A presents estimates for structural hazards as implied by the estimated parameters in panel A of Table 3. The dotted black line in panel A presents the observed exit rate from the data, averaged across workers with short and long notice. Panel B presents the implied average type at each duration for those with short and long notice. Error bars represent 90% confidence intervals.

there is little difference in the average type between the two groups.

Overall, the estimated pattern of structural duration dependence aligns with existing evidence from audit studies on call-back rates and with the predictions of search theory. I find that individual exit probabilities decline during the first five months, which can be attributed to duration-based employer discrimination. Further, I find that an individual's exit probability increases leading up to benefit exhaustion and remains constant after. This is consistent with search theory, which predicts that individuals increase their search effort or lower their reservation wages until they reach benefit exhaustion. After that point, if there are no further changes in workers' incentives, their probability of exiting unemployment should remain constant. Interestingly, in their audit study, [Kroft et al. \(2013\)](#) find a decrease in callback rates only during the first six months of unemployment (refer to Figure 2 in their paper). In Section VI, I formally illustrate that my findings are consistent with a search model incorporating heterogeneous workers

and falling callback rates early in the unemployment spell.

In recent studies, researchers have introduced behavioral modifications to the standard search theory in order to explain the observed decline in exit rates after UI exhaustion, which deviates from the predictions of the standard search model. Most notably, [DellaVigna et al. \(2021\)](#) introduces reference dependence in utility to account for this decline. However, after adjusting for compositional effects, I do not find evidence of a decline in individual exit probabilities after UI exhaustion. Hence, I show that the data can be reconciled with the standard model by incorporating heterogeneous workers.

VB Robustness Checks

In this section, I examine the robustness of the main results by presenting non-parametric estimates, results using different moments for estimation, and findings from an extension of the MH model that relaxes the identifying assumptions.

[Appendix D](#) presents results from several robustness checks. [Figure D2](#) displays the data and estimates using the unweighted sample. The results reveal that the exit rates and estimates for the unweighted sample are very similar to those obtained using the weighted sample, suggesting that the observable characteristics play a limited role in explaining the differences in exit rates between notice groups. [Figure D1](#) compares the non-parametric estimate for the structural hazard with the baseline log-logistic estimate. The non-parametric estimates are practically equivalent to the baseline but have larger standard errors, particularly for the last data point where the standard error increases significantly.

I also test the robustness of my findings by including additional notice length categories to the original estimation sample. First, I add individuals with no notice as a third group, and then I repeat the process by including individuals with <1 month notice as the third group. Details for these analyses are provided in [Section D.3](#) in the Appendix, and estimates are presented in [Figure D3](#). The results show that the estimated struc-

tural hazard is qualitatively similar to the baseline estimates in both cases. However, when adding the <1 month notice group, the increase in the structural hazard leading up to benefit exhaustion is less pronounced. Note that the model with three notice categories is overidentified with 4 degrees of freedom. The Sargan-Hansen J-statistic for testing overidentifying restrictions is reported in both cases and indicates no evidence against the null hypothesis of a correctly specified model.

For the main estimates, unemployment durations were grouped into 12-week intervals. I now assess the robustness of these findings to different bin sizes through two exercises. Firstly, I present estimates from the MH model using unemployment duration data binned into 9-week intervals. The estimated structural hazard, shown in Figure D4, mirrors the pattern reported in the main results. Specifically, it decreases from the first to the second interval, followed by an increase in the third interval spanning 18-27 weeks, corresponding to UI exhaustion at 26 weeks.

Secondly, I conduct a simulation where I calculate exit rates using various bin sizes under a specified data-generating process. Using these differently binned exit rates, I estimate the MH model to examine how the estimates are affected by the choice of bin size. Additional details for this simulation are presented in Section D.4 in the Appendix, and the results are illustrated in Figures D5 and D6. The key takeaway is that the binned estimates reflect the cumulative structural hazard within an interval, representing the probability of an individual exiting unemployment at some point during that interval. Given this, while we may miss some intra-interval changes, the binned estimates still capture a meaningful quantity and are informative about the underlying duration dependence.

Finally, in Appendix E, I consider the unconditional model as in Corollary 1 and show identification under more general conditions than independence and stationarity. Specifically, with $h(d|\nu, L) = \psi_L(d)\nu$ and considering two notice lengths, ℓ and ℓ' , I define the following two quantities:

$$\kappa_d = E(v^d|\ell') - E(v^d|\ell), \quad \gamma_d = \frac{\psi_{\ell'}(d)}{\psi_{\ell}(d)}$$

κ_d is the difference between the d^{th} moment of v for the two lengths of notice, while γ_d is the ratio of structural hazards at duration d for those lengths. In other words, these parameters determine how the distribution of heterogeneity and structural hazard vary by notice length. I then show that if κ_d is known for all d and γ_d for $d > 1$, the remaining parameters of the model are identified. The result can be straightforwardly extended to the model in Proposition 3. In fact, in the applications of this result described below, I utilize the IPW exit rates constructed using observed durations.

I utilize this result to test the independence and stationarity assumptions as follows: First, I relax the independence assumption by allowing the mean of the heterogeneity distribution to vary by notice length ($\kappa_1 \neq 0$) while assuming stationarity ($\gamma_d = 1$ for $d > 1$).²⁵ Specifically, I reestimate the model for varying values of κ_1 to see where the residuals are minimized, which tells us which value of κ_1 is most consistent with the data under the given assumptions. Second, while maintaining independence (setting all $\kappa_d = 0$), I relax stationarity by assuming $\gamma_d = \gamma$ for all $d > 1$. Once again, I reestimate the model, varying γ values to find the residual-minimizing point. The results for these exercises are presented in Figures E1 and E2. In both cases, the residuals plotted as a function of the parameters resemble a convex function with a clear local minimum. Moreover, the residuals are minimized when $\kappa_1 \approx 0$ and $\gamma \approx 1$, supporting the identifying assumptions. Finally, as a last exercise, I relax both of these assumptions simultaneously by reestimating the model for a grid of κ_1 and γ values. The results, presented in Figure E3, still show the residual-minimizing values close to 0 and 1, respectively. Overall, this provides evidence in favor of the identifying assumptions

²⁵In this exercise, I allow only the mean of the heterogeneity distribution to vary while keeping the shape of the distribution, as determined by the central moments, consistent across the two groups. Note that κ_d for $d > 1$ will still be non-zero because non-central moments depend on scale, so I let them vary accordingly to ensure that the central moments, apart from the mean, are the same for both groups.

employed in the paper.

VI A MODEL OF JOB-SEARCH

The estimates obtained from the Mixed Hazard model suggest a decline in an individual worker's probability of exiting unemployment during the initial five months. Additionally, I find evidence that an individual's likelihood of exiting unemployment increases as they approach the exhaustion of unemployment insurance (UI) benefits and remains constant after that. The latter is in contrast to the observed exit rate, which continues to decline even after benefit exhaustion. Researchers have tried to explain the decline in the observed rate after exhaustion using behavioral explanations such as storable offers ([Boone and van Ours, 2012](#)) or reference dependence in utility ([DellaVigna et al., 2021](#)). In this section, I show that my findings align with standard search theory, incorporating heterogeneous workers, and are consistent with evidence from the audit study conducted by [Kroft et al. \(2013\)](#), which documents an initial decline in callback rates during the unemployment spell.

In particular, I set up a search model with heterogeneous workers. Within this model, workers choose search effort to maximize their expected utility. The likelihood of finding a job depends on the offer arrival rate and a worker's search effort. Moreover, the offer arrival rate varies by the duration of unemployment and the type of worker. I calibrate the model to match the implied structural dependence to my estimate from the Mixed Hazard model and also match the exit rate implied by the model to the data. I then examine the trajectory of the offer arrival rate and search effort. This exercise also allows me to discern the impact on exit probabilities arising from two sources: the actions of optimizing agents in response to changing incentives and external factors directly influencing a worker's employment prospects.

VI.A Model Setup

I consider a stylized model of job search where a worker's search environment is non-stationary (Mortensen, 1986; van den Berg, 1990) and workers are heterogeneous. At every duration d , workers choose how much search effort s to exert to maximize their discounted expected utility.²⁶ Costs of search effort are given by the function $c(s)$, which is increasing, convex, and twice continuously differentiable, with $c(0) = 0$ and $c'(0) = 0$. The probability that a worker finds a job $\lambda(d, \nu, s)$ depends on the time elapsed since unemployed d , their search effort s , and their type ν as follows: $\lambda(d, \nu, s) = \delta(d)\nu s$. Here, $\delta(d)\nu$ is the offer arrival rate, which varies over the duration of unemployment and across workers of different types. Once workers find a job, they remain employed forever. A worker receives unemployment insurance (UI) benefits $b(d)$ when unemployed and wages w when employed. The function $u(\cdot)$ gives the flow utility from consumption. Then the value function for a worker of type ν unemployed at duration d is given by:

$$V_u(d, \nu) = \max_s u(b(d)) - c(s) + \beta [\lambda(s, d, \nu)V_e + (1 - \lambda(s, d, \nu))V_u(d + 1, \nu)]$$

Here, β is the discount rate, and V_e is the value of employment given by $V_e = u(w) + \beta V_e$. The UI benefits $b(d)$ are equal to b for $d \leq D_B$ and equal to 0 otherwise. I also assume that after some time $D_T \geq D_B$ the job-finding function $\lambda(d, s, \nu)$ does not depend on the duration of unemployment d , such that for $d > D_T$, $\delta(d) = \delta_T$. This ensures that after D_T , jobseekers face a stationary environment, and hence, we can solve for the optimal search strategy of each worker in each period using backward induction. Finally, I consider the case of two types of workers: a high type H and a low type L with $\nu_H > \nu_L$, with π denoting the share of workers with the higher arrival rate.

²⁶Alternatively, the model could feature a reservation wage choice, and all conclusions about search effort would instead be regarding reservation wages.

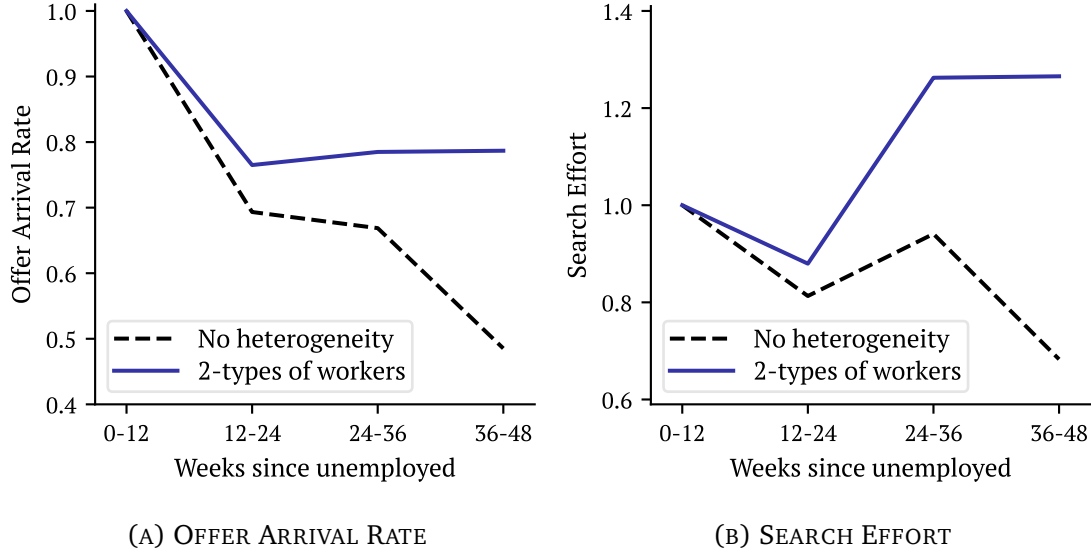
VI.B Numerical Analysis

I now calibrate the model specified in the previous section. Let $s(d, \nu)$ denote a worker's optimal search effort at duration d . The probability that this worker finds a job $h(d|\nu)$ is given by $\delta(d)s(d, \nu)\nu$. So, a worker's exit rate evolves over the spell of unemployment due to changes in the offer arrival rate $\delta(d)$ and the worker's search effort. However, just as before, the observed exit rate $h(d) = \mathbb{E}[h(d|\nu)|D \geq d]$ also changes due to changes in composition over the spell of unemployment. I use my estimate of the structural hazard from the Mixed Hazard model to target structural duration dependence $\mathbb{E}[h(d|\nu)]$ from the model.²⁷ Additionally, I match the exit rate implied by the model $h(d)$ to the data. In order to compare the predictions from this model to a model with no heterogeneity, I also calibrate the model assuming just one type of worker. In this case, I match the structural duration dependence or the exit rate implied by the model to the exit rate in the data. Further details for the calibration are provided in [Appendix B](#).

Figure [B1](#) shows that both the model with and without heterogeneity fit the data almost perfectly. Figure [3](#) presents the search effort and the offer arrival rate implied by the two calibration exercises. The offer arrival rate implied by the model with heterogeneity declines during the first five months and is constant after that, which is consistent with evidence from [Kroft et al. \(2013\)](#). Conversely, in the model with only one type of worker, the offer arrival rate continues to decline throughout the spell of unemployment. Finally, the model calibration implies that an individual's search effort decreases slightly during the first five months but then increases up to UI exhaustion and remains stable after that. In summary, the data and my findings can be rationalized with conventional search theory, without any behavioral adjustments, but by

²⁷Note that the search model does not correspond precisely to the econometric framework since it does not imply that $s(d, \nu)$ evolves in the same manner for each type of worker. However, in [Appendix F](#), I simulate data from the search model with notice periods and show that my estimator does reasonably well in capturing movements in $\mathbb{E}[h(d|\nu)]$.

FIGURE 3: CALIBRATION OF THE SEARCH MODEL



Notes: The figure presents the search effort and the offer arrival rate from the calibration of the search model, assuming no heterogeneity (dashed black line) and assuming two types of workers (solid blue line). The search effort is averaged over two types of workers.

incorporating heterogeneity among workers and declining returns to search early in the unemployment spell.

VII CONCLUSION

In this paper, I use a novel source of variation to disentangle the role of structural duration dependence from heterogeneity in the dynamics of the observed exit rate. I document that workers who receive a longer notice before being laid off are more likely to exit unemployment early in the spell. However, the observed exit rate is lower for long-notice workers at later durations. This points towards the presence of heterogeneity across workers. As a higher proportion of the more employable workers from the long-notice group exit early, the composition of surviving long-notice workers at later durations is worse. I utilize these reduced-form moments and estimate a Mixed Hazard

model.

The estimates from the hazard model uncover substantial heterogeneity in individual exit probabilities. The observed exit rate declines by about 43.6% over the first five months. In contrast, the estimated individual hazard only declines by 33.6% over this period. Moreover, I find that after the first five months, none of the depreciation in the observed exit rate is due to structural duration dependence. Instead, an individual's exit probability increases up to UI exhaustion and remains constant after. The observed exit rate continues to decline after exhaustion as well, which has led researchers to suggest behavioral explanations for this pattern. I show that the alternative explanation of heterogeneity across individuals effectively accounts for this pattern. Specifically, as shown in the paper, my estimates can be rationalized within a standard search model with heterogeneous workers. These findings underscore the importance of incorporating heterogeneity when estimating and calibrating search models.

APPENDIX A PROOFS AND DERIVATIONS

This section provides proofs of the results discussed in the main text. Before proceeding, let us introduce some additional notation. Denote the probability distribution function of observed duration by $g(\cdot)$. Additionally, define the corresponding survival function $S(\cdot)$ to represent the probability that unemployment duration exceeds a specific value.

At this point, it will be useful to examine the expressions for the quantities defined above under Assumption 1, which specifies that $h(d|\nu, L, X) = \psi_L(d, X)\nu$. First, note that we can express the survival function $S(d|\nu, L, X)$ as follows:

$$S(d|\nu, L, X) = \Pr(D > d|\nu, L, X) = \prod_{s=1}^d [1 - \psi_L(s, X)\nu]$$

Given the above expression, the conditional distribution of observed duration is given by:

$$\begin{aligned} g(d|\nu, L, X) &= \Pr(D = d|\nu, L, X) \\ &= S(d-1|\nu, L, X) - S(d|\nu, L, X) \\ &= \psi_L(d, X)\nu S(d-1|\nu, L, X) \end{aligned} \tag{2}$$

with $S(0|\nu, L, X) = 1$.

Finally, note that according to the definition of the hazard function in the text, for a set of conditioning variables Υ , we have:

$$h(d|\Upsilon) = \Pr(D = d|D \geq d, \Upsilon) = \frac{\Pr(D = d|\Upsilon)}{\Pr(D \geq d|\Upsilon)} = \frac{g(d|\Upsilon)}{S(d-1|\Upsilon)}$$

The above expression illustrates that, given $h(d|\Upsilon)$ for $d = 1, 2, \dots, \bar{D}$, we can compute $g(d|\Upsilon)$ and $S(d|\Upsilon)$ for the same durations, and vice versa. This means that knowing any one of these three quantities enables the calculation of the other two. Therefore,

although the statement of results might specify a set of parameters being identified from $h(d|\Upsilon)$, it does not matter if the proof utilizes derived quantities $g(d|\Upsilon)$ or $S(d|\Upsilon)$.

A.1 Proof of Proposition 1

Proof. Given equation (2), we can write $g(d|L, X) = \psi_L(d, X)\mathbb{E}[\nu S(d-1|\nu, L, X)|L, X]$. Plugging this in the definition of $h(d|L, X)$, we get:

$$h(d|L, X) = \psi_L(d, X)\mathbb{E}\left(\nu \cdot \frac{S(d-1|\nu, L, X)}{S(d-1|L, X)} \middle| L, X\right)$$

To see that the second term in the above expression is the average type $\mathbb{E}(\nu|D \geq d, L, X)$ amongst surviving workers at the beginning of d , note that

$$f(\nu|D \geq d, L, X) = \frac{\Pr(D > d-1|\nu, L, X)f(\nu|L, X)}{\Pr(D > d-1|L, X)} = \frac{S(d-1|\nu, L, X)f(\nu|L, X)}{S(d-1|L, X)}$$

where the first inequality follows from the Bayes rule.

To see that the average type declines with duration, note that for any d and $\nu_H > \nu_L$,

$$\frac{S(d|\nu_H, L, X)}{S(d-1|\nu_H, L, X)} < \frac{S(d|\nu_L, L, X)}{S(d-1|\nu_L, L, X)}$$

The above equation implies that,

$$\frac{f(\nu_H|D \geq d+1, L, X)}{f(\nu_L|D \geq d+1, L, X)} < \frac{f(\nu_H|D \geq d, L, X)}{f(\nu_L|D \geq d, L, X)}$$

In which case, $f(\nu|D \geq d, L, X)$ first-order stochastically dominates $f(\nu|D \geq d+1, L, X)$ which implies that $\mathbb{E}(\nu|D \geq d, L, X) \geq \mathbb{E}(\nu|D \geq d+1, L, X)$. $\square$

A.2 Statement and Proof of Lemma 1

The following lemma states that the identification of structural hazards implies the identification of higher moments of the unobserved type distribution.

Lemma 1. Under Assumption 1, if $\psi_L(d, X)$ is known for $d = 1, \dots, \bar{D}$, then we can identify the first $\bar{D}$ conditional moments of v , given by $\{\mathbb{E}(v^k|L, X)\}_{k=1}^{\bar{D}}$, from the conditional unemployment distribution $g(d|L, X)$ for $d = 1, \dots, \bar{D}$.

Proof. Expanding equation (2) for $d = 1, 2, 3, \dots$, we can write:

$$\begin{aligned} g(1|v, L, X) &= \psi_L(1, X)v \\ g(2|v, L, X) &= \psi_L(2, X)[v - \psi_L(1, X)v^2] \\ g(3|v, L, X) &= \psi_L(3, X)[v - [\psi_L(1, X) + \psi_L(2, X)]v^2 + \psi_L(1, X)\psi_L(2, X)v^3] \\ &\vdots \end{aligned}$$

Or, more compactly,

$$g(d|v, L, X) = \psi_L(d, X) \sum_{k=1}^{\bar{D}} c_k(d, \boldsymbol{\psi}_{L,X}) v^k \quad (3)$$

where $\boldsymbol{\psi}_{L,X} = \{\psi_L(d, X)\}_{d=1}^{\bar{D}}$ and

$$c_k(d, \boldsymbol{\psi}_{L,X}) = \begin{cases} 1 & \text{for } k = 1 \\ c_k(d-1, \boldsymbol{\psi}_{L,X}) - \psi_L(d-1, X)c_{k-1}(d-1, \boldsymbol{\psi}_{L,X}) & \text{for } 1 \leq k \leq d \\ 0 & \text{for } k > d \end{cases}$$

Taking the expectation conditional on L and X of the above expression, we can write:

$$g(d|L, X) = \psi_L(d, X) \sum_{k=1}^{\bar{D}} c_k(d, \boldsymbol{\psi}_{L,X}) \mathbb{E}(v^k|L, X)$$

Let us denote $\mathbf{g}_{L,X} = \{g(d|L, X)\}_{d=1}^{\bar{D}}$ and $\boldsymbol{\mu}_{L,X} = \{\mathbb{E}(v^k|L, X)\}_{k=1}^{\bar{D}}$. Then we can write $\mathbf{g}_{L,X} = C(\boldsymbol{\psi}_{L,X})\boldsymbol{\mu}_{L,X}$ where $C(\boldsymbol{\psi}_{L,X})$ is the $\bar{D} \times \bar{D}$ upper triangular matrix with $C_{s,k}(\boldsymbol{\psi}_{L,X}) = \psi_L(s, X)c_k(s, \boldsymbol{\psi}_{L,X})$. In addition, the diagonal elements of $C(\boldsymbol{\psi}_{L,X})$ are non-zero. To see this note that, $C_{d,d}(\boldsymbol{\psi}_{L,X}) = (-1)^{d-1} \prod_{s=1}^d \psi_L(s, X)$ and each $\psi_L(s, X) > 0$. Hence,

$C(\psi_{L,X})$ is invertible and we can plug in $\psi_{L,X}$ in $g_{L,X} = C(\psi_{L,X})\mu_{L,X}$ to solve for $\mu_{L,X}$. $\square$

A.3 Proof of Theorem 1

Proof. Assumption 3 states that $\psi_L(d,X) = \psi(d,X)$ for $d > 1$. For this reason, define $\check{S}(d|\nu,X)$ as:

$$\check{S}(d|\nu,X) = \prod_{s=2}^d [1 - \psi(s,X)\nu]$$

Then by Assumption 3, for $d > 1$, we can write:

$$S(d|\nu,L,X) = [1 - \psi_L(1,X)\nu]\check{S}(d|\nu,X) \quad (4)$$

Next, by substituting the expression for $S(d-1|\nu,L,X)$ implied by equation (4) into the expression for $g(d|\nu,L,X)$ given in equation (2), we obtain:

$$g(d|\nu,L,X) = \psi_L(d,X) [\nu\check{S}(d-1|\nu,X) - \psi_L(1,X)\nu^2\check{S}(d-1|\nu,X)] \quad (5)$$

Now note that since Assumption 2 states that ν is independent of L given X , it follows that for any $L = \ell$:

$$q(d|\ell,X) = \mathbb{E}[q(d|\nu,\ell,X)|\ell,X] = \mathbb{E}[q(d|\nu,\ell,X)|X] \quad \text{for } q = S, g$$

Therefore, by taking the conditional expectation of the expressions implied by equations (4) and (5) under Assumptions 2 and 3, we get:

$$S(d-1|L,X) = \mathbb{E}[\check{S}(d-1|\nu,X)|X] - \psi_L(1,X)\mathbb{E}[\nu\check{S}(d-1|\nu,X)|X] \quad (6)$$

$$g(d|L,X) = \psi(d,X)[\mathbb{E}[\nu\check{S}(d-1|\nu,X)|X] - \psi_L(1,X)\mathbb{E}[\nu^2\check{S}(d-1|\nu,X)|X]] \quad (7)$$

Next, consider two notice lengths ℓ and ℓ' . Substituting $L = \ell, \ell'$ in equations (6) and (7) and taking the difference between expressions for ℓ' and ℓ , we get:

$$S(d-1|\ell', X) - S(d-1|\ell, X) = [\psi_\ell(1, X) - \psi_{\ell'}(1, X)]\mathbb{E}[\nu\check{S}(d-1|\nu, X)|X] \quad (8)$$

$$g(d|\ell', X) - g(d|\ell, X) = \psi(d, X)[\psi_\ell(1, X) - \psi_{\ell'}(1, X)]\mathbb{E}[\nu^2\check{S}(d-1|\nu, X)|X] \quad (9)$$

From equation (8), we can write:

$$\mathbb{E}[\nu\check{S}(d-1|\nu, X)|X] = \frac{S(d-1|\ell', X) - S(d-1|\ell, X)}{\psi_\ell(1, X) - \psi_{\ell'}(1, X)} \quad (10)$$

Similarly, from equation (9), we obtain:

$$\mathbb{E}[\nu^2\check{S}(d-1|\nu, X)|X] = \frac{g(d|\ell', X) - g(d|\ell, X)}{\psi(d, X)(\psi_\ell(1, X) - \psi_{\ell'}(1, X))} \quad (11)$$

Plugging terms from equations (10) and (11) in the expression for $g(d|\ell, X)$ given by equation (7), implies that for $d > 1$:

$$\psi(d, X) = \frac{g(d|\ell', X)\psi_\ell(1, X) - g(d|\ell, X)\psi_{\ell'}(1, X)}{S(d-1|\ell', X) - S(d-1|\ell, X)}$$

Here, the denominator is not equal to zero as we assumed $\psi_{\ell'}(1, X) \neq \psi_\ell(1, X)$.

Now note that for $d = 1$, $g(1|\ell, X) = \psi_\ell(1, X)\mathbb{E}[\nu|X]$. So plugging in $\psi_\ell(1, X) = g(1|\ell, X)/\mathbb{E}[\nu|X]$ in the expression for $\psi(d, X)$ above, we can write:

$$\psi(d, X)\mathbb{E}[\nu|X] = \frac{g(d|\ell', X)g(1|\ell, X) - g(d|\ell, X)g(1|\ell', X)}{S(d-1|\ell', X) - S(d-1|\ell, X)}$$

Since all the quantities on the right-hand side of the above equation can be derived from the conditional exit rates, it follows that $\{\psi_\ell(1, X), \psi_{\ell'}(1, X), \{\psi(d, X)\}_{d=2}^{\bar{D}}\}$ are identified up to a multiplicative constant from $\{h(d|\ell, X), h(d|\ell', X)\}_{d=1}^{\bar{D}}$. Identification of conditional moments follows from Lemma 1 by noting that $\mathbb{E}(\nu^k|L, X) = \mathbb{E}(\nu^k|L)$ under conditional independence. $\square$

A.4 Proof of Corollary 1

Proof. Given that $h(d|\nu, L) = \psi_L(d)\nu$, we can write:

$$S(d|\nu, L) = \prod_{s=1}^d [1 - \psi_L(s)\nu], \quad g(d|\nu, L) = \psi_L(d)\nu S(d-1|\nu, L)$$

Moreover, $L \perp \nu$ implies that for any $L = \ell$, $q(d|\ell) = \mathbb{E}[q(d|\nu, \ell) | \ell] = \mathbb{E}[q(d|\nu, \ell)]$ for $q = S, g$. Since we assumed that $\psi_L(d) = \psi(d)$ for $d > 1$, let us define as before: $\check{S}(d|\nu) = \prod_{s=2}^d [1 - \psi(s)\nu]$ for $d > 1$. We then obtain the following expressions:

$$\begin{aligned} S(d-1|L) &= \mathbb{E}[\check{S}(d-1|\nu)] - \psi_L(1)\mathbb{E}[\nu\check{S}(d-1|\nu)] \\ g(d|L) &= \psi(d) \left[\mathbb{E}[\nu\check{S}(d-1|\nu)] - \psi_L(1)\mathbb{E}[\nu^2\check{S}(d-1|\nu)] \right] \end{aligned}$$

Note that the above equations take a similar form to equations (6) and (7), with the only difference being that the expressions are free of X . Thus, by considering $L = \ell, \ell'$ and replicating the steps following equation (7) in the proof of Theorem 1, we can derive that $\psi_L(1) = g(1|L)/\mathbb{E}(\nu)$ for $L \in \{\ell, \ell'\}$, and

$$\psi(d)\mathbb{E}(\nu) = \frac{g(d|\ell')g(1|\ell) - g(d|\ell)g(1|\ell')}{S(d-1|\ell') - S(d-1|\ell)} \text{ for } d > 1.$$

This concludes the proof for identifying structural hazards up to a scale. For the identification of the unconditional moments of the type distribution, we can follow the steps from Lemma 1. In particular, $\mathbf{g}_L = [g(d|L)]_{d=1}^{\bar{D}}$ can be expressed using as a system of $\bar{D}$ equations, in terms of the structural hazards $\boldsymbol{\psi}_L = [\psi_L(d)]_{d=1}^{\bar{D}}$ and the moments of the type distribution $\boldsymbol{\mu} = [\mathbb{E}(\nu^k)]_{k=1}^{\bar{D}}$. Specifically, $\mathbf{g}_L = C(\boldsymbol{\psi}_L)\boldsymbol{\mu}$, where $C(\boldsymbol{\psi}_L)$ is an invertible matrix, as outlined in Lemma 1. Hence, we can plug in the identified structural hazards to derive the moments □

A.5 Proof of Proposition 2

Proof. Define $g^w(\cdot)$ as the weighted distribution and $S^w(\cdot)$ as the weighted survival function. Specifically, $q^w(d|L) = \mathbb{E}[q(d|\nu, L, X)/p_L(X)|L]$ for $q = g, S$. According to the definition in the main text, $h^w(d|L) = g^w(d|L)/S^w(d-1|L)$. Therefore, given $h^w(d|L)$ for $d = 1, 2, \dots, \bar{D}$, we can compute $g^w(d|L)$ and $S^w(d|L)$ for the same durations as well.

We can show that for any $L = \ell$, $q^w(d|\ell) = \mathbb{E}[q(d|\nu, \ell, X)]/\pi_\ell$ for $q = g, S$, under conditional independence, by following the steps outlined below:

$$\begin{aligned} \mathbb{E}\left[\frac{q(d|\nu, \ell, X)}{p_\ell(X)}\middle|\ell\right] &= \mathbb{E}\left[\mathbb{E}\left[\frac{q(d|\nu, \ell, X)}{p_\ell(X)}\middle|\ell, X\right]\middle|\ell\right] & (i) \\ &= \mathbb{E}\left[\mathbb{E}\left[\frac{q(d|\nu, \ell, X)}{p_\ell(X)}\middle|X\right]\middle|\ell\right] & (ii) \\ &= \int \frac{\mathbb{E}[q(d|\nu, \ell, x)|x]}{p_\ell(x)} \cdot f_{X|\ell}(x) \partial x & (iii) \\ &= \frac{1}{\pi_\ell} \cdot \mathbb{E}[q(d|\nu, \ell, x)] & (iv) \end{aligned}$$

Here, $\pi_\ell = \Pr(L = \ell)$ and $f_{X|L}(\cdot)$ represents the distribution of X given L . Step (i) utilizes the law of iterated expectations, step (ii) follows from $L \perp \nu|X$, step (iii) uses the definition of conditional expectation, and step (iv) follows from the Bayes' rule, which implies that $f_{X|\ell}(x) = p_\ell(x)f_X(x)/\pi_\ell$.

As an aside, note that this implies that:

$$h^w(d|\ell) = \frac{\mathbb{E}[g(d|\nu, \ell, X)]}{\mathbb{E}[S(d-1|\nu, \ell, X)]} = \frac{\mathbb{E}[\Pr(D_\ell = d|\nu, X)]}{\mathbb{E}[\Pr(D_\ell \geq d|\nu, X)]} = \frac{\Pr(D_\ell = d)}{\Pr(D_\ell \geq d)} = \Pr(D_\ell = d|D_\ell \geq d)$$

Having shown that $\pi_\ell q^w(d|\ell) = \mathbb{E}[q(d|\nu, \ell, X)]$ for $q = g, S$, if we now take the expectation of the expressions implied by equations (4) and (5) and substitute $\psi_L(d, X) = \psi_L(d)\phi(X)$ with $\psi_L(d) = \psi(d)$ for $d > 1$, we obtain:

$$\begin{aligned} \pi_L S^w(d-1|L) &= \mathbb{E}[\check{S}(d-1|\nu, X)] - \psi_L(1)\mathbb{E}[\theta(X, \nu)\check{S}(d-1|\nu, X)] \\ \pi_L g^w(d|L) &= \psi(d)\left[\mathbb{E}[\theta(X, \nu)\check{S}(d-1|\nu, X)] - \psi_L(1)\mathbb{E}[\theta(X, \nu)^2\check{S}(d-1|\nu, X)]\right] \end{aligned}$$

where $\theta(X, \nu) = \phi(X)\nu$.

Note that the above equations take a similar form to equations (6) and (7). Thus, by considering $L = \ell, \ell'$ and replicating the steps following equation (7) in the proof of Theorem 1, we can derive that $\psi_L(1) = \pi_L g^w(1|L)/\mathbb{E}[\theta(X, \nu)]$ for $L \in \{\ell, \ell'\}$, and

$$\psi(d)\mathbb{E}[\theta(X, \nu)] = \frac{g^w(d|\ell')g^w(1|\ell) - g^w(d|\ell)g^w(1|\ell')}{S^w(d-1|\ell')/\pi_{\ell'} - S^w(d-1|\ell)/\pi_{\ell}}, \quad \text{for } d > 1$$

To see that the moments are identified as well, first note that by plugging in $\psi_L(d, X) = \psi_L(d)\phi(X)$, we can rewrite equation (3) as:

$$g(d|\nu, L, X) = \psi_L(d) \sum_{k=1}^{\bar{D}} c_k(d, \psi_L) \phi(X)^k \nu^k \quad (12)$$

where $\psi_L = \{\psi_L(d)\}_{d=1}^{\bar{D}}$ and $c_k(d, \psi_L)$ is defined as in Lemma 1. Since we established that $\pi_{\ell} g^w(d|\ell) = \mathbb{E}[g(d|\nu, \ell, X)]$, we can write:

$$\pi_L g^w(d|L) = \psi_L(d) \sum_{k=1}^{\bar{D}} c_k(d, \psi_L) \mathbb{E}(\phi(X)^k \nu^k)$$

The rest of the proof follows as in the proof for Lemma 1. Denote $\mathbf{g}_L^w = \{\pi_L g^w(d|L)\}_{d=1}^{\bar{D}}$ and $\boldsymbol{\mu}^w = \{\mathbb{E}[\theta(X, \nu)^k]\}_{k=1}^{\bar{D}}$. Then we can write $\mathbf{g}_L^w = C(\psi_L)\boldsymbol{\mu}^w$ where $C(\psi_L)$ is an invertible matrix, as outlined in Lemma 1. Hence, we can plug in the identified structural hazards to find $\boldsymbol{\mu}^w$. $\square$

A.6 Proof of Proposition 3

Proof. First note that, by the definition of $\tilde{D}$ and C , we can write $\tilde{h}^w(d|L)$ as follows:

$$\tilde{h}^w(d|L) = \mathbb{E} \left[\frac{\Pr(D = d, D^c \geq d|\nu, L, X)}{p_L(X)} \middle| L \right] / \mathbb{E} \left[\frac{\Pr(D \geq d, D^c \geq d|\nu, L, X)}{p_L(X)} \middle| L \right]$$

Now note that,

$$\begin{aligned}
\mathbb{E} \left[\frac{\Pr(D \in \Delta, D^c \geq d | \nu, L, X)}{p_L(X)} \middle| L \right] &= \mathbb{E} \left[\mathbb{E} \left[\frac{\Pr(D \in \Delta, D^c \geq d | \nu, L, X)}{p_L(X)} \middle| L, X \right] \middle| L \right] \quad (i) \\
&= \mathbb{E} \left[\mathbb{E} \left[\frac{\Pr(D \in \Delta, D^c \geq d | \nu, L, X)}{p_L(X)} \middle| X \right] \middle| L \right] \quad (ii) \\
&= \int \frac{\mathbb{E}[\Pr(D_L \in \Delta, D^c \geq d | \nu, x) | x]}{p_L(x)} \cdot f_{X|L}(x) \quad (iii) \\
&= \mathbb{E}[\Pr(D_L \in \Delta, D^c \geq d | \nu, X)] \quad (iv) \\
&= \Pr(D_L \in \Delta, D^c \geq d)
\end{aligned}$$

Here, $\pi_\ell = \Pr(L = \ell)$ and $f_{X|L}(\cdot)$ represents the distribution of X given L . Step (i) utilizes the law of iterated expectations, step (ii) follows from $L \perp \nu | X$ and $D^c \perp \nu | X$, step (iii) uses the definition of conditional expectation and swaps D with potential duration D_L , and step (iv) follows from the Bayes' rule, which implies that $f_{X|L}(x) = p_L(x)f_X(x)/\pi_L$.

Note that this implies that:

$$\tilde{h}^w(d|L) = \frac{\Pr(D_L = d, D^c \geq d)}{\Pr(D_L \geq d, D^c \geq d)} = \frac{\Pr(D_L = d | D^c \geq d)}{\Pr(D_L \geq d | D^c \geq d)} = \Pr(D_L = d | D_L \geq d, D^c \geq d)$$

As before, if we are given the hazard rate, we can also calculate the corresponding density and survival rate. In what follows, the identification of parameters is shown using the density and survival rates $\Pr(D_L = d | D^c \geq d)$ and $\Pr(D_L \geq d | D^c \geq d)$. In particular, note that $\Pr(D_L = d | D^c \geq d) = \mathbb{E}[g(d | \nu, L, X) | D^c \geq d]$ and $\Pr(D_L \geq d | D^c \geq d) = \mathbb{E}[S(d | \nu, L, X) | D^c \geq d]$. Therefore, taking the conditional expectation of the expressions implied by equations (4) and (5) and substituting $\psi_L(d, X) = \psi_L(d)\phi(X)$ with $\psi_L(d) = \psi(d)$ for $d > 1$, we get:

$$\begin{aligned}
\Pr(D_L \geq d | D^c \geq d) &= \mathbb{E}_d[\check{S}(d-1 | \nu, X)] - \psi_L(1)\mathbb{E}_d[\theta(X, \nu)\check{S}(d-1 | \nu, X)] \\
\Pr(D_L = d | D^c \geq d) &= \psi(d) \left[\mathbb{E}_d[\theta(X, \nu)\check{S}(d-1 | \nu, X)] - \psi_L(1)\mathbb{E}_d[\theta(X, \nu)^2\check{S}(d-1 | \nu, X)] \right]
\end{aligned}$$

Here, $\theta(X, \nu) = \phi(X)\nu$, and I use $\mathbb{E}_d$ to denote expectations conditional on $D^c \geq d$ for

brevity; specifically, $\mathbb{E}_d(\cdot) = \mathbb{E}(\cdot | D^c \geq d)$. Except for the expectations being conditional on $D^c \geq d$, the expressions on the right-hand side in the equations are identical to those in Proposition 2. Therefore, one can follow the same steps as in Proposition 2 to show the identification of structural hazards. However, it should be noted that the moments are not yet identified. This is because taking the expectation of equation (12) conditional on $D^c \geq d$ results in different moments entering different period densities. For instance, we will have $\mathbb{E}[\theta(X, \nu) | D^c \geq 1]$ in the first-period density, but the second-period density will have $\mathbb{E}[\theta(X, \nu) | D^c \geq 2]$. However, if we further assume that $D^c \perp X$ and $D^c \perp \nu$, then $\mathbb{E}[\theta(X, \nu)^k | D^c \geq d] = \mathbb{E}[\theta(X, \nu)^k]$, and identification of moments follows from Proposition 2. $\square$

A.7 Proof for $\mathbb{E}[m_i(\ell, d; \Theta)] = 0$

For brevity, I omit the i subscript from the terms in the following exposition. Additionally, let us denote the indicators for exit and survival used in the individual moments as follows: $I^E(d) = \mathbb{I}\{\tilde{D} = d, C = 0\}$ and $I^S(d) = \mathbb{I}\{\tilde{D} \geq d\}$. In this case, we can rewrite the expression for individual moments as follows:

$$m(\ell, d; \Theta) = \mathbb{I}\{L = \ell\} \left[\frac{I^E(d)}{p_\ell(X)} - \tilde{h}^w(d|\ell; \Theta) \cdot \frac{I^S(d)}{p_\ell(X)} \right]$$

Proof. By taking the expectation of the $m(\ell, d; \Theta)$ expression above and applying the law of iterated expectations, we get:

$$\mathbb{E}[m(\ell, d; \Theta)] = \pi_\ell \left[\mathbb{E} \left[\frac{I^E(d)}{p_\ell(X)} \middle| L = \ell \right] - \tilde{h}^w(d|\ell; \Theta) \cdot \mathbb{E} \left[\frac{I^S(d)}{p_\ell(X)} \middle| L = \ell \right] \right]$$

Applying the law of iterated expectations again, we can write:

$$\frac{\mathbb{E}[m(\ell, d; \Theta)]}{\pi_\ell} = \mathbb{E} \left[\frac{\mathbb{E} [I^E(d) | L = \ell, X]}{p_\ell(X)} \middle| L = \ell \right] - \tilde{h}^w(d|\ell; \Theta) \cdot \mathbb{E} \left[\frac{\mathbb{E} [I^S(d) | L = \ell, X]}{p_\ell(X)} \middle| L = \ell \right]$$

Finally, by substituting $\mathbb{E}[I^E(d) | L, X] = \Pr(\tilde{D} = d, C = 0 | \nu, L, X)$ and $\mathbb{E}[I^S(d) | L, X] =$

$\Pr(\tilde{D} \geq d | \nu, L, X)$ into the above expression, and applying the definition of $\tilde{h}^w(d|\ell; \Theta)$, it becomes clear that the right-hand side equals zero. $\square$

APPENDIX B SEARCH MODEL CALIBRATION

I calibrate the model under standard values for model parameters. To maintain consistency with the econometric model, each period is assumed to be 12 weeks long. Corresponding to a 5 percent annual interest rate, the discount factor β is set equal to 0.985. I normalize the wage to 1 and set the replacement rate for unemployment benefits at 0.5. In addition, I assume individuals receive an annuity payment of 0.1 times their wages in each period, regardless of their employment status. This can be interpreted as the income of a secondary earner. Utility from consumption is given by the constant relative risk aversion (CRRA) utility function, $u(c) = c^{1-\sigma}/(1-\sigma)$ with $\sigma = 1.75$. I follow [DellaVigna et al. \(2017\)](#) and [Marinescu and Skandalis \(2021\)](#) in assuming that costs of job search are given by $c(s) = \theta s^{1+\rho}/(1+\rho)$. I set $\rho = 1$ and $\theta = 50$.²⁸ Table [B1](#) summarizes the calibration parameters and Figure [B1](#) displays the fit of the calibrated model.

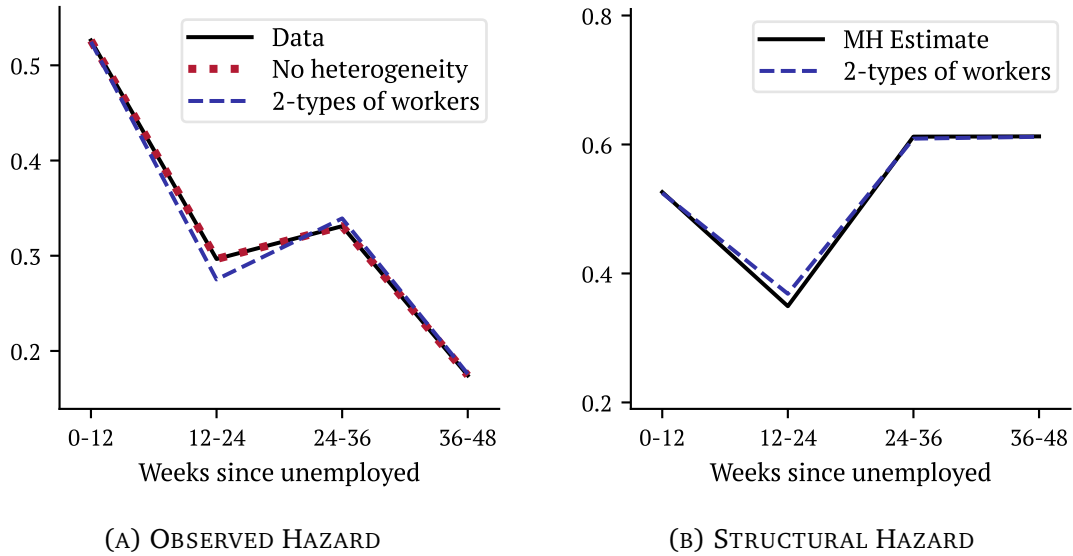
²⁸Different parameters for the cost function do not change qualitative predictions of my exercise but do lead to changes in the scale of the search effort.

TABLE B1: CALIBRATION PARAMETERS FOR THE SEARCH MODEL

Parameter	Value
Length of each period	12 Weeks
Discount factor β	0.985
Relative risk aversion σ	1.75
Per period wages w	1
Annuity Payments	0.1
Unemployment benefits	0.5
Benefit exhaustion D_B	3
Search cost parameter ρ	1
Search cost parameter θ	50
First period arrival rate $\delta(1)$	1

Note: The table presents the parameters used for calibrating the search model in Section VI.

FIGURE B1: SEARCH MODEL CALIBRATION: FIT



Note: The figure displays the fit of the search model for the two calibration exercises described in the text. Panel A shows the observed exit rate in the data (solid black line) alongside the corresponding fitted values obtained from calibrating the search model without heterogeneity (dotted red line) and with two types of workers (dashed blue line). Panel B displays the estimated structural hazard from the Mixed Hazard (MH) model (solid black line) and the fitted structural hazard from calibrating the search model with two types of workers (dashed blue line).

APPENDIX C DATA

This section provides more details about data construction and sample selection. It also elaborates on the estimation of propensity scores and checks the balance of the weighted sample on additional measures. Supplementary data descriptives are also presented here.

C.1 Data Construction and Sample Selection

The Displaced Worker Supplement (DWS) was introduced in 1984, but the variable on the length of notice was not included in the first two samples. Furthermore, the definition of displaced workers has changed over time.²⁹ Before 1998, self-employed individuals or those who expected to be recalled to their lost job within six months were also included in the survey. However, the information on whether a worker expected to be recalled is only available for the years 1994 and 1996. In addition, the data on the length of time individuals took to find their next job is miscoded and largely missing for the year 1994. For these reasons, my analysis begins from 1996. Moreover, to maintain consistency across years, I exclude self-employed individuals or those who expected to be recalled from the 1996 sample.

The final sample only includes individuals with non-missing information on all the variables utilized in the analyses. Additionally, to capture individuals who had lost a job representing stable, non-temporary employment, the sample is limited to those who worked full-time for at least six months at their previous job and had health insurance provided through that employment. Finally, to minimize retrospective bias, I exclude individuals who report switching more than two jobs since losing their previous job. The sample selection procedure is summarized in Table C1.

The duration of unemployment for individuals who have secured a job by the time

²⁹The recall window was 5 years instead of 3 before 1994.

TABLE C1: SAMPLE SELECTION

Sample condition	Observations
DWS 1996-2020, 21-64 year old respondents	44707
No recall expectation	44537
Lost job was not self-employment	43471
Non-missing values for variables used	29443
Worked full-time at lost job	25293
Employed for at least 6 months at lost job	22991
Had health insurance at lost job	14825
Held less than 3 jobs since lost job	13784
Got a notice before job loss	5898
Got a notice of >1 month	4175

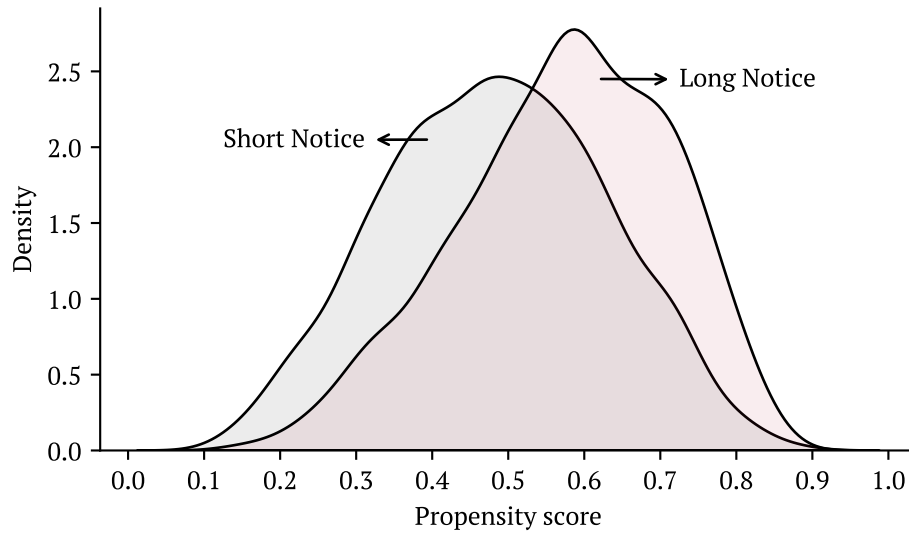
Note: The table shows the number of observations remaining at each step of sample selection.

TABLE C2: COMPARISON OF THE ANALYTICAL SAMPLE TO ALL INDIVIDUALS IN THE DISPLACED WORKER SUPPLEMENT (DWS) AND THE CURRENT POPULATION SURVEY (CPS)

	Sample (1)	DWS (2)	CPS (3)
Age	43.04	40.61	42.22
Female	0.45	0.44	0.52
Married	0.61	0.54	0.60
Black	0.10	0.11	0.10
High School	0.30	0.42	0.41
Some College	0.30	0.32	0.29
College Degree	0.40	0.26	0.30
Employed	0.80	0.67	0.75
Unemployed	0.18	0.21	0.04
NILF	0.02	0.12	0.21
Observations	4175	44707	964225

Note: All samples are restricted to individuals between the ages of 21 to 64 and pertain to years 1996-2020. Column (1) includes individuals from the DWS who worked full-time for at least six months and were provided health insurance at their lost job, did not expect to be recalled, and received a layoff notice of 1-2 months or greater than 2 months. Columns (2) and (3) include all individuals in the DWS and the monthly CPS, respectively.

FIGURE C1: ASSESSING OVERLAP OF PROPENSITY SCORE DISTRIBUTIONS



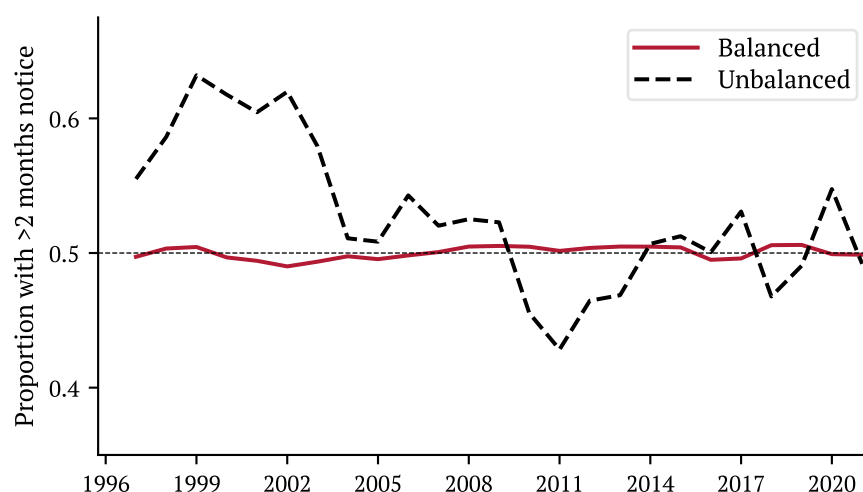
Note: The figure presents the density of estimated propensity scores for individuals with short and long notice separately.

of the survey is given by the *dwwksun* variable, which measures the number of weeks the person was unemployed between leaving or losing one job and starting another. For those who report not holding another job since their lost job, censored duration is obtained using the *durunemp* variable from the CPS. Since 2012, tenure at the lost job was top-coded at 24 years. To maintain consistency across samples, I also implement a top code of 24 years for all years prior to 2012. Earnings are reported in 1999 dollars. Table C2 presents the descriptive statistics for my analytical sample relative to all individuals aged 21 to 64 in the DWS and the CPS over the sample period.

C.2 Propensity Score Weighting

To ensure individuals with long and short notice are comparable, I reweight the sample using inverse propensity score weighting. The weight for each individual is calculated as the inverse of the likelihood of receiving the reported notice length. To estimate the propensity scores, I utilize a logistic regression where the odds of receiving a longer

FIGURE C2: LENGTH OF NOTICE OVER TIME

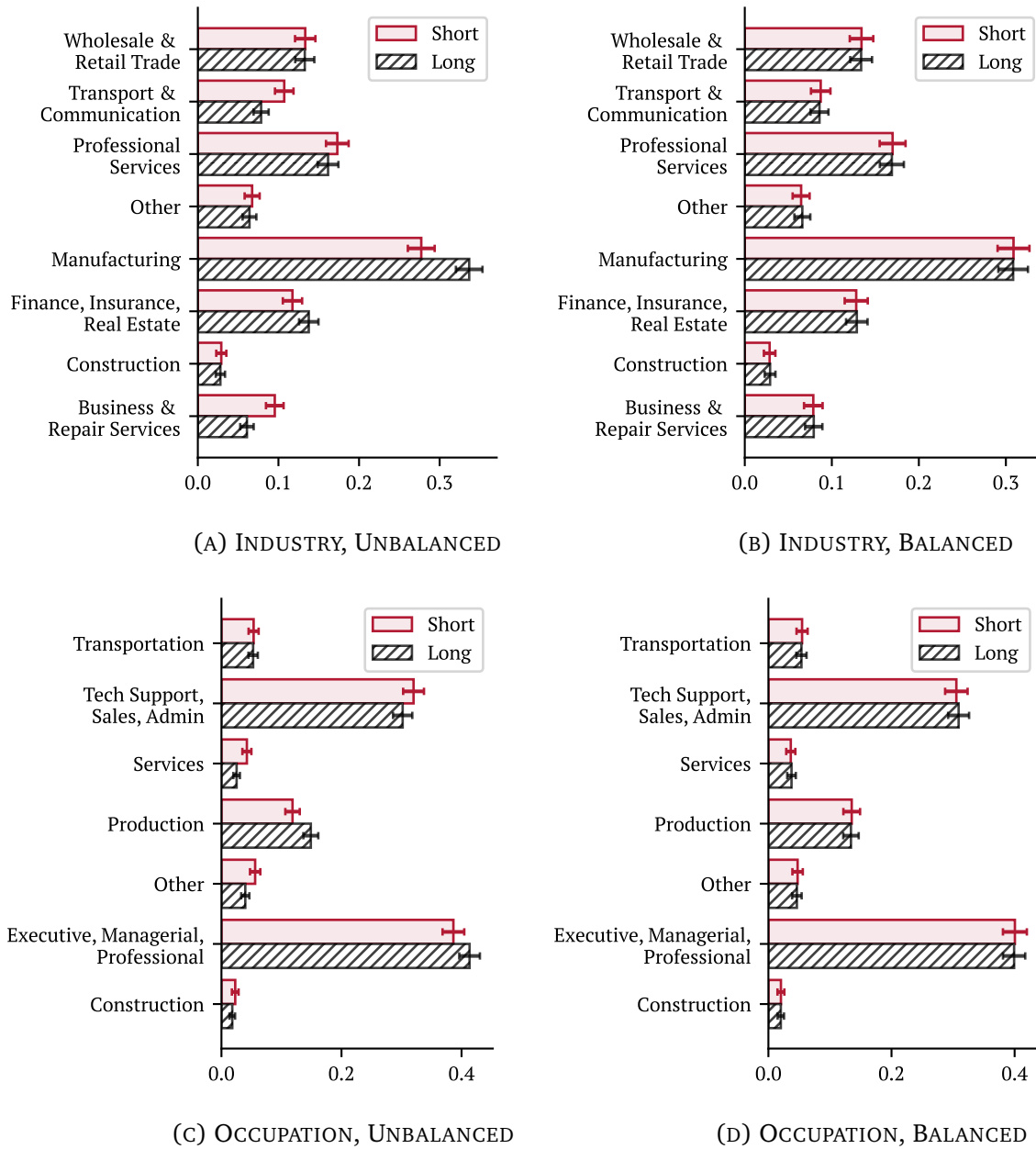


Note: The figure plots a 3-year moving average of the proportion of individuals who received a notice of more than 2 months amongst all individuals in the sample who were displaced in a given year.

notice are modeled as a function of several variables. These variables include age, gender, marital status, race (specifically, an indicator for Black), education level, reason for layoff, union membership, residence in a metropolitan area, tenure, and earnings at the lost job. Additionally, fixed effects for occupation and industry of the lost job, state of residence, and year of displacement are included in the model. The density of estimated propensity scores for short and long-notice individuals is displayed in Figure C1. The figure indicates a significant overlap between the two distributions, with all estimated propensity scores falling in the 0.1 to 0.9 range, thereby rendering further data trimming unnecessary.

Table 1 in the main text provides evidence that the reweighting achieves balance across certain observable variables. Figure C2 demonstrates that reweighting leads to balance with respect to the year of displacement. In addition, Figure C3 presents occupation and industry distributions for short and long-notice workers in both the balanced and unbalanced samples. Notably, the weighted sample exhibits more similarity in the industrial and occupational composition of short and long-notice workers.

FIGURE C3: INDUSTRY AND OCCUPATION OF THE LOST JOB



Note: The figure presents the proportions of individuals whose displaced jobs were in specific industries (panels A and B) and occupations (panels C and D) among long-notice and short-notice workers in both the unbalanced and balanced samples. The error bars represent the 90% confidence intervals.

TABLE C3: EARNINGS AT THE SUBSEQUENT JOB

	Weekly Log Earnings			
	(1)	(2)	(3)	(4)
> 2 month notice	0.091** (0.037)	0.066** (0.032)	0.093** (0.038)	0.063* (0.032)
Controls	No	Yes	No	Yes
Weights	No	No	Yes	Yes
Observations	2657	2657	2657	2657

Note: The table shows results from linear regressions of log weekly wages at the subsequent job on an indicator for receiving a notice of more than 2 months. The sample used is similar to the main analytical sample, but it excludes individuals who had not yet found employment at the time of the survey, had multiple jobs between their previous and current jobs, or had incomplete earnings information for other reasons. Robust standard errors are reported in the parenthesis.

C.3 Additional Descriptives

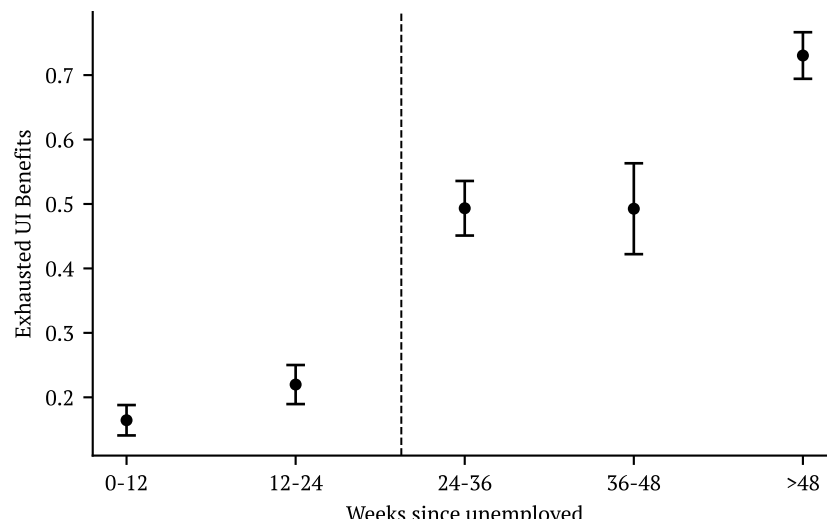
This section provides additional descriptive statistics. Table C3 presents the relationship between longer notice and earnings at the subsequent job. The table indicates that workers with longer notice tend to have higher earnings in their subsequent jobs. However, we cannot interpret this as a direct impact of longer notice because extended periods of unemployment can have a negative impact on wages (Schmieder et al., 2016), and as shown in this paper, a longer notice leads to shorter unemployment spells. Table C4 describes the incidence of UI take-up in the sample, and Figure C4 describes the timing of benefit exhaustion amongst UI takers. Figure C5 presents the data with unemployment duration binned in 4 and 9-week intervals.

TABLE C4: UNEMPLOYMENT INSURANCE TAKE-UP

Duration	Observations	Received UI Benefits
0 weeks	820	0.06
0-4 weeks	959	0.30
4-8 weeks	457	0.61
8-12 weeks	363	0.70
> 12 weeks	1562	0.82

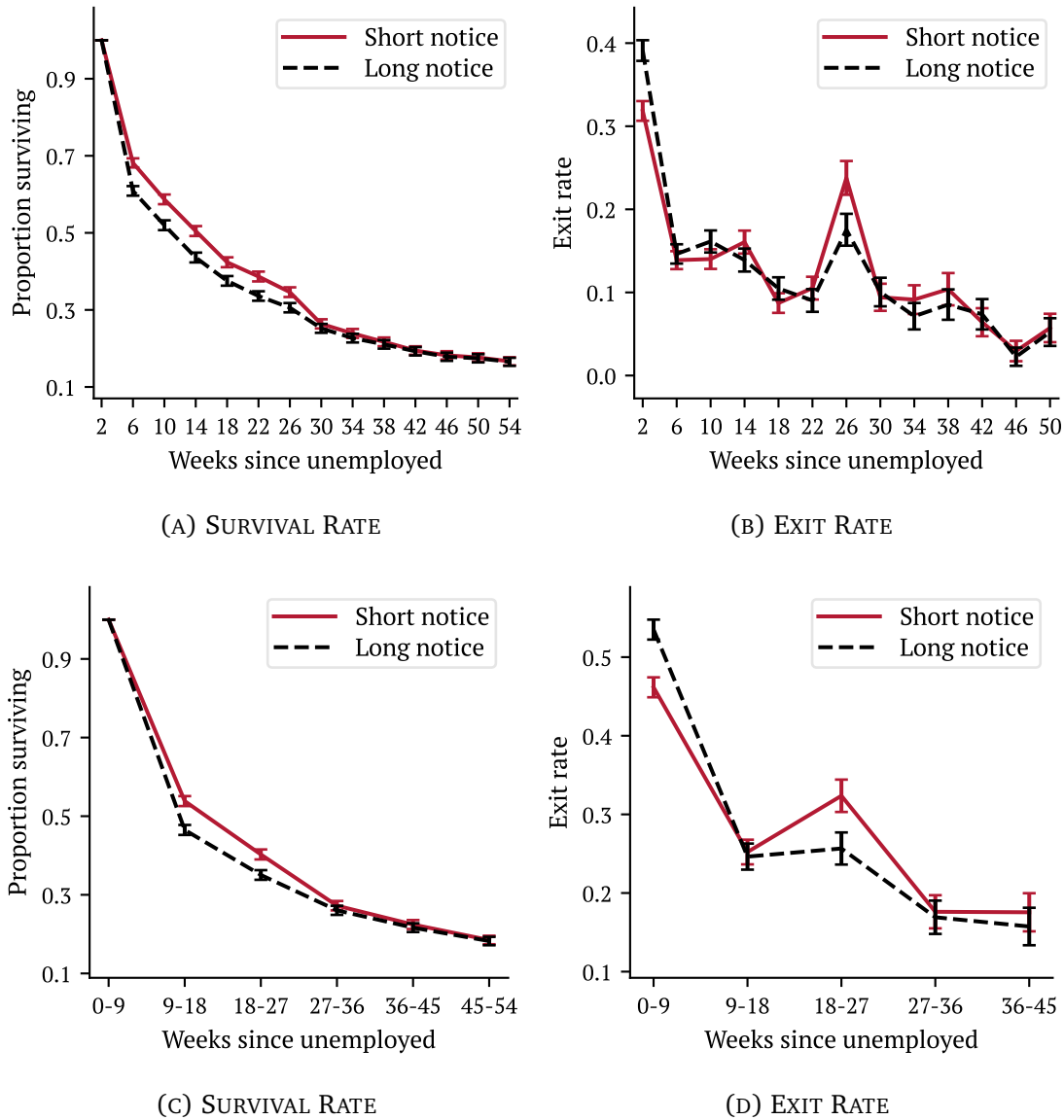
Notes: This table reports the percentage of individuals in the baseline sample who reported receiving UI benefits by the duration of unemployment.

FIGURE C4: TIMING OF BENEFIT EXHAUSTION



Note: The figure presents the proportion of individuals who report having exhausted their UI benefits by the duration of unemployment. The sample is restricted to individuals in the main analytical sample who reported receiving UI benefits, and duration is binned in 12-week intervals.

FIGURE C5: SURVIVAL AND EXIT RATES WITH ALTERNATIVE BINS



Note: Unemployment duration is binned in 4-week intervals for panels A and B, while it is binned in 9-week intervals for panels C and D. Panel A and C present the proportion of individuals who are unemployed at the beginning of each interval. Panel B and D present the proportion of individuals exiting unemployment in each interval amongst those who were still unemployed at the beginning of the interval. Error bars represent 90% confidence intervals.

APPENDIX D ROBUSTNESS CHECKS

In this subsection, I present a series of robustness checks.

D.1 Non-Parametric Estimates

As shown in the paper, the specified MH model is non-parametrically identified. Figure D1 presents the non-parametric estimate for the structural hazard alongside the baseline estimate, which assumes a log-logistic functional form for the hazard. This figure shows that the non-parametric point estimates are roughly equivalent to the baseline estimates. However, the standard errors are larger, specifically the standard error associated with the last data point, which blows up dramatically.

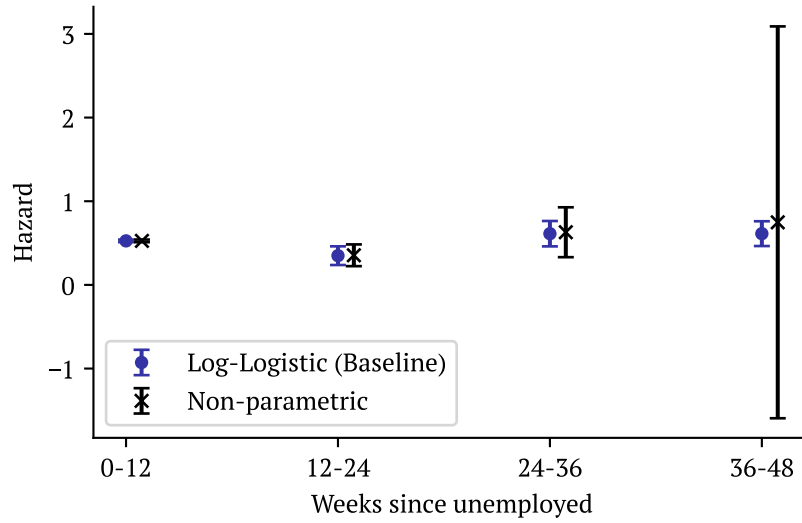
D.2 Unweighted Sample

Figure D2 displays the data and estimates using the unweighted sample. Panel A shows that the pattern of exit rates and how they vary by notice length is similar to that observed in the weighted sample in Figure 1. Similarly, the estimates based on the unweighted sample, presented in panel B, closely resemble the baseline estimates shown in Figure 2. This suggests that observable characteristics have a limited role in explaining the raw differences in exit rates between the two notice groups.

D.3 Alternative Notice Categories

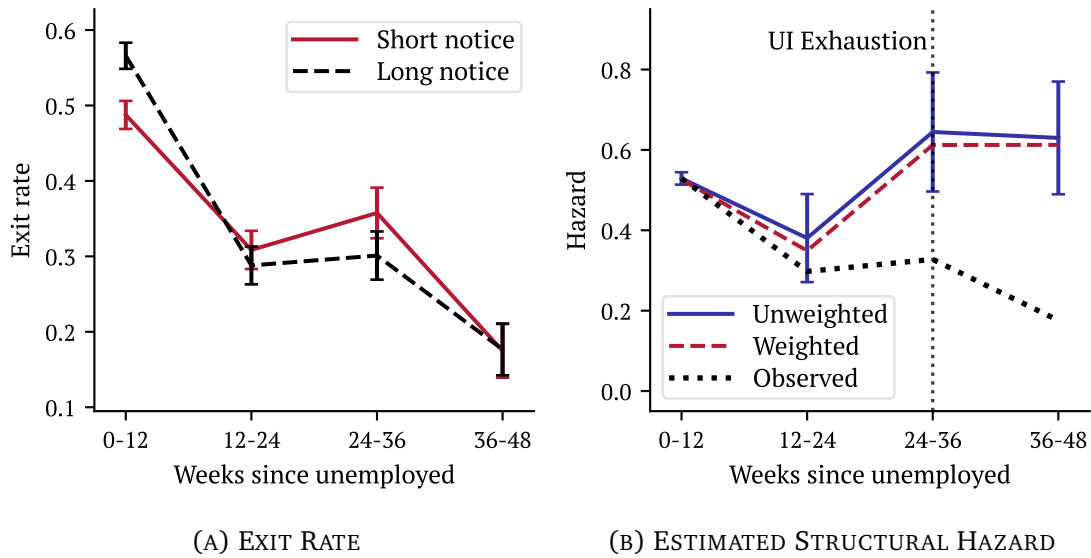
As noted in Section II, the question on notice length in the DWS is categorical and includes categories of no notice, <1 month, 1-2 months, and >2 months. So far, I have only focused on the latter two categories, referred to as short and long notice. I now test the robustness of my findings with respect to this restriction. First, I include individuals who received no notice as a third group in the original estimation sample. Then, I

FIGURE D1: NON-PARAMETRIC ESTIMATES



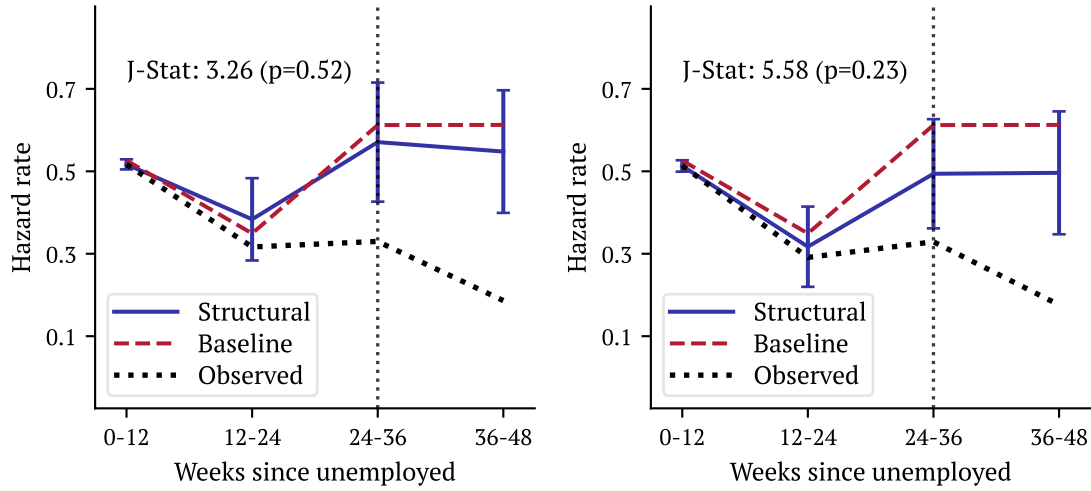
Note: The figure compares non-parametric estimates of the structural hazard from the Mixed Hazard model with baseline estimates that assume a log-logistic functional form for the structural hazard. Error bars represent 90% confidence intervals.

FIGURE D2: DATA AND ESTIMATES USING THE UNWEIGHTED SAMPLE



Note: The figure presents data and estimates for the unweighted analytical sample. Panel A displays the exit rate from the data separately for long and short-notice workers. In panel B, the solid blue line shows the estimated structural hazard from the Mixed Hazard model using unweighted moments, the dashed red line shows the estimate using weighted moments, and the dotted black line represents the average exit rate for short and long-notice workers in the data.

FIGURE D3: ESTIMATES WITH ALTERNATIVE NOTICE CATEGORIES



(A) NO NOTICE, 1-2 MONTHS, >2 MONTHS (B) <1 MONTH, 1-2 MONTHS, >2 MONTHS

Note: Panel A presents the estimated hazard for the sample consisting of individuals who received no notice, along with those who received notice periods of 1-2 months and >2 months, as in the main analytical sample. Panel B estimates are for individuals who received notice of <1 month, in addition to the groups from the main analytical sample. J-Stat refers to the Sargan-Hansen statistic for overidentifying restrictions.

repeat the process, adding individuals with <1 month's notice as the third group.³⁰ I use the same covariates as before to fit the propensity score models using multinomial logit and weight the data accordingly. Observations where the propensity scores fall outside the 0.1 to 0.9 range are excluded. I then re-estimate the Mixed Hazard model with three notice categories for both the augmented samples.

The estimates from the above exercise are presented in Figure D3. Panel A corresponds to the sample where the third category is no notice, while panel B corresponds to the sample where the third category is <1 month notice. After excluding extreme

³⁰I do not include all four categories simultaneously for two reasons. The first, and less significant, reason is that the overlap in propensity scores is already suboptimal with three categories and worsens with four. The primary reason is that the exit rate in the first period for the <1 month group is lower than for the no notice group, possibly due to selection, since separations with no notice may also include voluntary quits. However, the first-period exit rates for both are lower than those for the 1-2 month and >2 month notice groups.

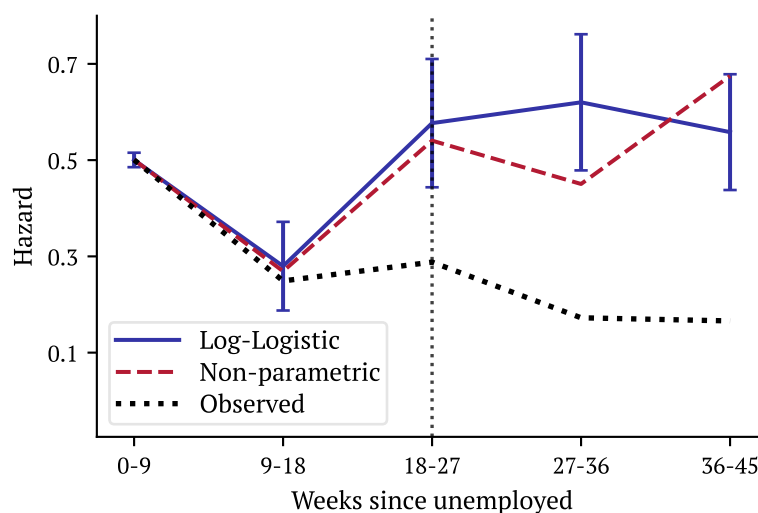
propensity score values, 7,122 of 12,061 observations remain for the data used in panel A, and 5,667 of 5,898 observations remain for the data used in panel B. In both cases, the estimates of structural duration dependence are qualitatively similar to the baseline results—the structural hazard consistently exceeds the exit rate from the data, initially decreasing, then increasing up to the third interval, and remaining constant thereafter. However, in panel B, the estimated structural hazard increases less leading up to benefit exhaustion compared to the baseline estimate. This could reflect differences in patterns of heterogeneity or duration dependence across the samples. More importantly, because we have more than two notice categories, the model is over-identified. With $3 \times 4 = 12$ moment conditions and 8 parameters, there are 4 degrees of freedom. The Sargan-Hansen J-statistic for testing overidentifying restrictions is reported in both panels, along with the respective p -values.³¹ In both cases, the large p -values indicate that there is no strong evidence against the null hypothesis.

D.4 Binning Unemployment Duration

Finally, a limitation of the application in the paper is the need to bin unemployment durations into 12-week intervals due to the small sample size of the DWS and the potentially noisy measurement of duration caused by retrospective bias. While Figure C5 shows that the patterns of how exit rates vary with duration and notice length are similar when binning data into 4- or 9-week intervals, I now present estimates from the MH model using unemployment duration data binned into 9-week intervals in Figure D4. Once again, we uncover the same pattern in estimates of structural duration dependence, with a decline from the first to the second interval, followed by a rise in the third interval corresponding to 18-27 weeks, which aligns with UI exhaustion at 26 weeks. I also present non-parametric estimates, which closely follow the log-logistic estimates.

³¹The estimated model in the main analysis is also over-identified due to the imposition of a log-logistic functional form on the structural hazard, resulting in one free parameter. However, here we are testing a more extensive set of overidentifying restrictions, with the caveat that the power of this test is lower.

FIGURE D4: ESTIMATES WITH UNEMPLOYMENT DURATION BINNED IN 9-WEEK INTERVALS



Note: The figure presents estimates from the Mixed Hazard model using data with unemployment duration binned in 9-week intervals. The solid blue line represents estimates for the structural hazard assuming the Log-Logistic functional form. The dashed red line represents non-parametric estimates, while the dotted black line represents the observed exit rate from the data.

I conduct another exercise to evaluate the effect of interval sizes directly on estimates from the MH model. Specifically, I take exit rates implied by the model for a given data-generating process (DGP) and compute the exit rates that would be observed if unemployment durations were binned into different-sized intervals. I then compare how the estimates change with the size of the bins.

The data generating process (DGP) for this exercise is defined as follows: There are two groups of notice, each with a probability of 0.5, and the first-period structural hazards are set at 0.1 and 0.2 for the two groups, respectively. For periods after the first ($d > 1$), the structural hazards, which are common to both groups, are generated in four different cases using the Weibull function bkd^{d-1} to ensure our conclusions are not an artifact of a specific DGP. These four cases are: (1) increasing ($b = 0.2, k = 1.2$), (2) decreasing ($b = 0.2, k = 0.75$), (3) constant ($b = 0.15, k = 1$), and (4) non-monotonic, where the Weibull function is increasing as in (1) until $d = 7$ and then follows (2) multiplied by 1.75. The unobserved type ν is a mixture of three Beta distributions: $\nu_1 \sim \text{Beta}(0.1, 0.1)$ with probability 0.5, $\nu_2 \sim \text{Beta}(0.3, 0.5)$ with probability 0.1, and

$\nu_3 \sim \text{Beta}(0.25, 0.5)$ with probability 0.4. I generate the exit rates for 12 time periods, bin them into intervals of four different sizes—1, 2, 3, and 4—and estimate the model in each case.

The dashed black line in Figure D5 represents the profile of exit rates corresponding to different bin sizes used in the estimation for the four DGPs from panels A-D. The solid dark-grey line shows the estimated structural hazard in each case. Comparing the estimates across different bin sizes in each panel, one can see that increasing the bin size leads to potentially missing the full extent of changes in the structural hazard over time. Specifically, when the structural hazard is increasing, the binned estimate increases by less (compare column 4 to column 1 in panel A). Similarly, when the structural hazard is decreasing, the binned estimate decreases by less (panel B). This is a direct consequence of the binned exit rates not fully capturing the extent of changes in the exit rates over time, making it difficult for the estimates using those exit rates to capture the full dynamics.

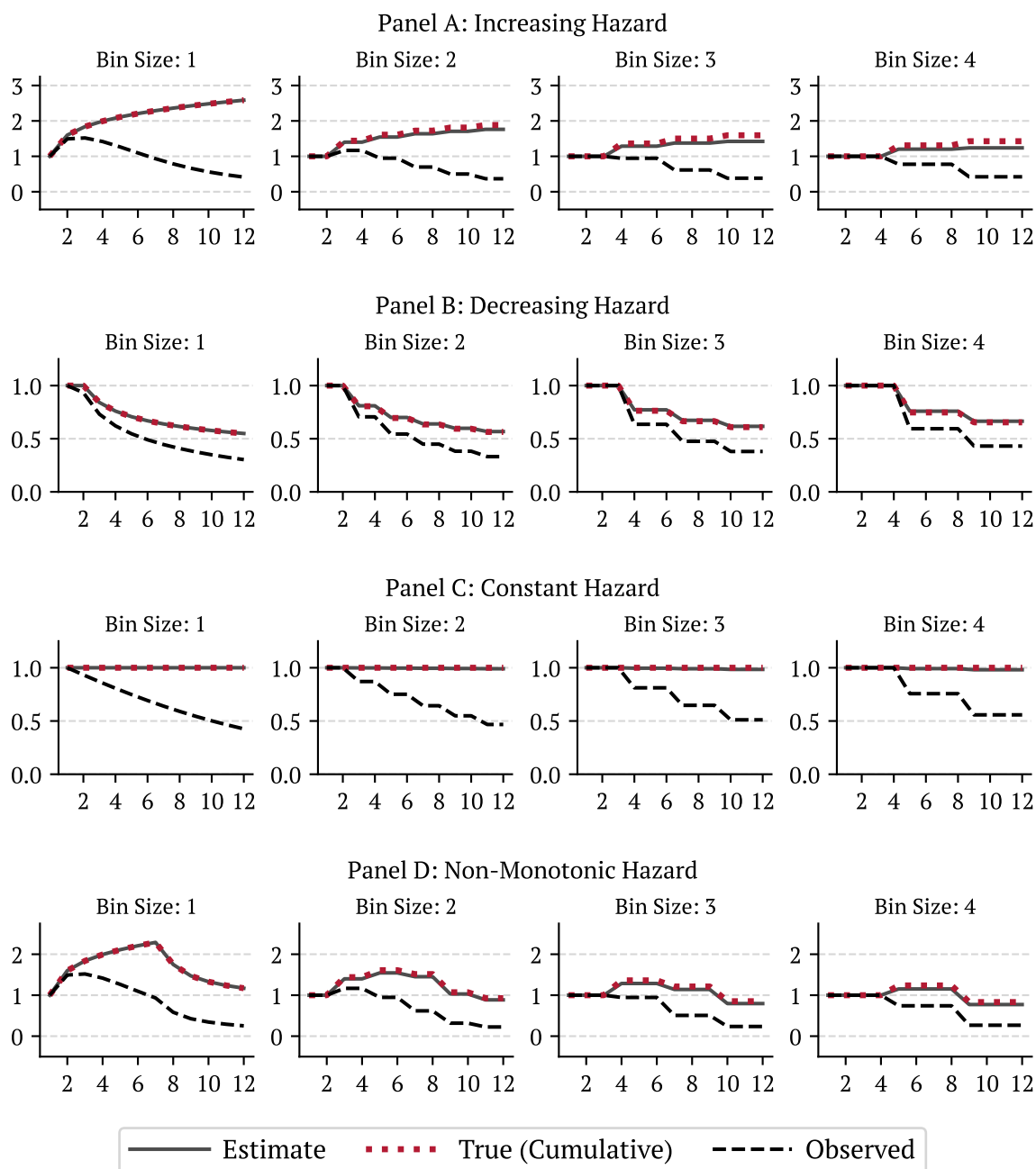
However, note that regardless of bin size, the estimates are able to uncover the broader trends in the structural hazard. This is because the estimates correspond to the cumulative structural hazard within a given interval for the average type, represented by the dotted red line on the plots. Specifically, the cumulative structural hazard is defined as the probability of an individual exiting unemployment in a given interval, provided that the individual hasn't exited by the start of the interval.³² So, while the estimates may miss some nuances within the intervals, the extent of what is missed is limited to what would be missed if we were analyzing the true structural exit probabilities in broader intervals. This is also evident from the implied average type for each estimate presented in Figure D6. From this figure, it can be seen that the average type in each panel falls similarly over time for different bin sizes, implying that we are able

³²For instance, for a bin size of 2, the cumulative structural hazard corresponding to the second interval for the average worker is given by: $\Pr(3 \leq D \leq 4 | D \geq 2, L, \nu = \mu) = \psi(3)\mu + [1 - \psi(3)\mu]\psi(4)\mu$.

to capture the contribution of underlying heterogeneity in the observed exit rates very closely, regardless of the bin size.

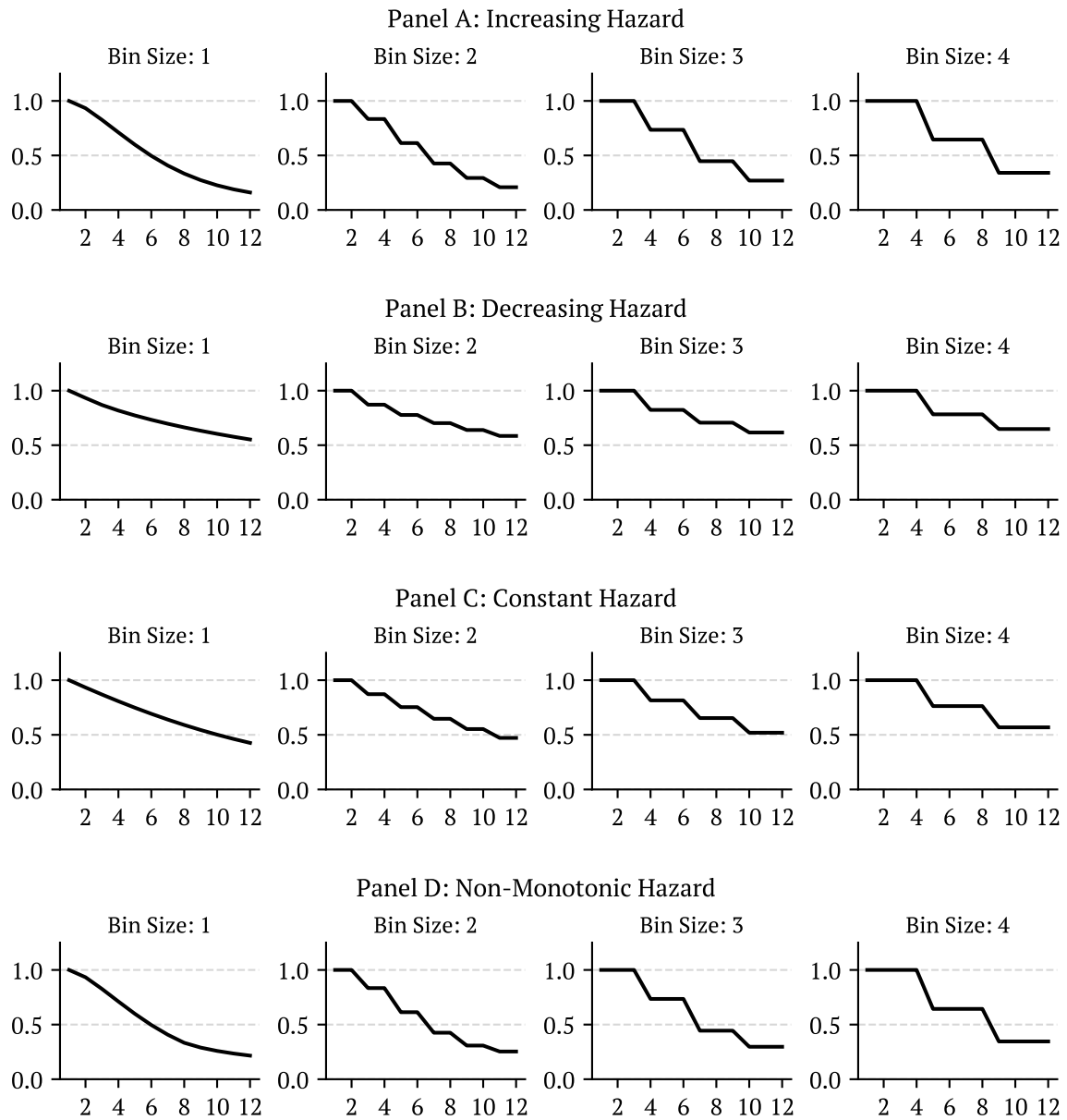
Overall, this shows that the assumptions of the model do not interact with bin sizes in a way that leads to systematically biased estimates, ensuring that the estimates reported in this paper correspond to meaningful underlying patterns.

FIGURE D5: ESTIMATION ON BINNED SIMULATED DATA: HAZARD



Note: The figure presents results from a robustness exercise evaluating the impact of binning data into intervals. I calculate the moments used for estimation from the Mixed Hazard model with 12 durations. These moments are then binned into intervals of lengths ranging from 1 to 4, and the model is estimated using the binned moments. Along with the estimates of the structural hazard (dark grey solid line), I present two additional quantities binned similarly for comparison. The dotted red line represents the true structural cumulative hazard of exiting in the specific interval, while the black dashed line represents the average observed hazard.

FIGURE D6: ESTIMATION ON BINNED SIMULATED DATA: AVERAGE TYPE



Note: The figure presents results from a robustness exercise evaluating the impact of binning data into intervals. I calculate the moments used for estimation from the Mixed Hazard model with 12 durations. These moments are then binned into intervals of lengths ranging from 1 to 4, and the model is estimated using the binned moments. The plots represent the average type at each duration, as implied by the estimated structural hazard.

APPENDIX E GENERALIZATION

The main identification result in the paper relies on two crucial assumptions: (i) the notice length is independent of the worker type (conditional on observables), and (ii) the structural hazard after the initial period does not vary with notice length. In this section, I generalize the identification result showing that we can identify structural duration dependence and the moments of heterogeneity distribution, provided we know how the structural hazard following the initial period and the distribution of heterogeneity vary with notice length. Previously, the model operated under the assumption that the distribution of heterogeneity and the structural hazards at later durations did not vary with notice length. Now, we are considering alternative assumptions about how these factors might vary. This is useful as it allows us to assess estimates under different assumptions and compare them to the original results.

For brevity, I ignore observable characteristics while presenting the proof in this section and work with a version of the model where an individual worker's exit probability is given by $h(d|\nu, L) = \psi_L(d)\nu$, with ν being exogenous. However, the result can be straightforwardly extended to incorporate observable characteristics as in the model with $h(d|\nu, L, X) = \psi_L(d)\phi(X)\nu$, under conditional independence. In fact, in the implementation of this result in Section E.2, I do use exit rates weighted by the inverse of propensity scores.

Unlike before, now we will not assume stationarity or independence; instead, we will allow the distribution of heterogeneity and structural hazards after the initial period to vary based on certain known parameters. In particular, for two lengths of notice ℓ and ℓ' , define κ_d as the difference between the d^{th} moment of ν conditional on ℓ' and ℓ as follows:

$$\kappa_d = \mu_{\ell',d} - \mu_{\ell,d}$$

where $\mu_{\ell,d} = E(\nu^d|\ell)$. So κ_1 is the difference between the average type of workers

with ℓ' and ℓ notice lengths. Additionally, define γ_d as the ratio of structural hazards at duration d for two lengths of notice,

$$\gamma_d = \frac{\psi_{\ell'}(d)}{\psi_{\ell}(d)}$$

The identification result presented below states that if for some $\bar{D}$ we know κ_d for $d = 1, \dots, \bar{D}$ and γ_d for $d = 2, \dots, \bar{D}$, we can identify the first $\bar{D}$ structural hazards and moments of type distribution for each notice length up to scale.³³

E.1 General Identification Result

Theorem 2. For some ℓ, ℓ' , define $\kappa_d = \mu_{\ell',d} - \mu_{\ell,d}$ and $\gamma_d = \psi_{\ell'}(d)/\psi_{\ell}(d)$. Then for some $\bar{D}$, if $\{\kappa_d\}_{d=1}^{\bar{D}}$ and $\{\gamma_d\}_{d=2}^{\bar{D}}$ are known, then the baseline hazards $\{\psi_{\ell}(d), \psi_{\ell'}(d)\}_{d=1}^{\bar{D}}$ and the conditional moments of the type distribution $\{\mu_{\ell,d}, \mu_{\ell',d}\}_{d=1}^{\bar{D}}$ are identified up to a scale from $\{h(d|\ell), h(d|\ell')\}_{d=1}^{\bar{D}}$.

Proof. First note that we can write,

$$g(d|\ell) = \psi_{\ell}(d) \sum_{k=1}^d c_k(\psi_{\ell,d-1}) \mu_{\ell,k} \quad (13)$$

where $\psi_{\ell,d-1} = \{\psi_{\ell}(s)\}_{s=1}^{d-1}$, $c_k(\psi_{\ell,0}) = 1$, and

$$c_k(\psi_{\ell,d-1}) = \begin{cases} c_k(\psi_{\ell,d-2}) & \text{for } k = 1 \\ c_k(\psi_{\ell,d-2}) - \psi_{\ell}(d-1)c_{k-1}(\psi_{\ell,d-2}) & \text{for } 1 < k \leq d \\ 0 & \text{for } k > d \end{cases}$$

Now, we can prove the statement of the theorem by induction. First, note that the

³³Alternatively, we could know κ_d for $d = 2, \dots, \bar{D}$ and γ_d for $d = 1, \dots, \bar{D}$. Also, in theory, the choice of defining γ_d and κ_d as a ratio or a difference does not impact the proof of identification. In this case, I define κ_d as a difference and γ_d as a ratio for the convenience of varying these parameters when examining the changes in estimates.

statement is true for $\bar{D} = 1$. To see this, note that

$$g(1|\ell) = \psi_l(1)\mu_{\ell,1} \quad g(1|\ell') = \psi_{\ell'}(1)(\mu_{\ell,1} + \kappa_1)$$

Normalizing $\mu_{\ell,1} = 1$, we can solve for $\psi_l(1) = g(1|\ell)$ and $\psi_{\ell'}(1) = \frac{g(1|\ell')}{1+\kappa_1}$.

Now, let us assume that the statement is true for $\bar{D} = d - 1$. Then we can identify $\{\psi_l(s), \psi_{\ell'}(s)\}_{s=1}^{d-1}$ and $\{\mu_{\ell,s}, \mu_{\ell',s}\}_{s=1}^{d-1}$ from $\{g(s|\ell), g(s|\ell')\}_{s=1}^{d-1}$. To complete the proof, we need to prove that the statement is true for $\bar{D} = d$ as well.

Denote $\Gamma_d = \prod_{s=1}^d \gamma_s$ and $\Psi_\ell(d) = \prod_{s=1}^d \psi_\ell(s)$. Now note that,

$$\begin{aligned} g(d|\ell) &= \psi_l(d) \sum_{k=1}^d c_k(\psi_{\ell,d-1})\mu_{\ell,k} \\ &= \psi_l(d) \left[\sum_{k=1}^{d-1} c_k(\psi_{\ell,d-1})\mu_{\ell,k} + c_d(\psi_{\ell,d-1})\mu_{\ell,d} \right] \\ &= \psi_l(d) \left[\sum_{k=1}^{d-1} c_k(\psi_{\ell,d-1})\mu_{\ell,k} + (-1)^{d-1}\Psi_\ell(d-1)\mu_{\ell,d} \right] \end{aligned}$$

From the above equation we can solve for $\mu_{\ell,d}$ as follows:

$$\mu_{\ell,d} = \frac{(-1)^d}{\Psi_\ell(d-1)} \left[\sum_{k=1}^{d-1} c_k(\psi_{\ell,d-1})\mu_{\ell,k} - \frac{g(d|\ell)}{\psi_l(d)} \right] \quad (14)$$

Using the fact that $\mu_{\ell',d} = \kappa_d + \mu_{\ell,d}$, we can write $g(d|\ell')$ as follows:

$$g(d|\ell') = \psi_{\ell'}(d) \left[\sum_{k=1}^{d-1} c_k(\psi_{\ell',d-1})\mu_{\ell',k} + (-1)^{d-1}\Psi_{\ell'}(d-1)(\kappa_d + \mu_{\ell,d}) \right]$$

By plugging in $\mu_{\ell,d}$ from equation (14) in the above expression, we can solve for $\psi_{\ell'}(d)$ as follows:

$$\psi_{\ell'}(d) = \frac{g(d|\ell') - \Gamma_d g(d|\ell)}{\sum_{k=1}^{d-1} c_k(\psi_{\ell',d-1})\mu_{\ell',k} - \Gamma_{d-1} \sum_{k=1}^{d-1} c_k(\psi_{\ell,d-1})\mu_{\ell,k} + (-1)^{d-1}\kappa_d \Psi_{\ell'}(d-1)}$$

Plugging this back in expression for $\mu_{\ell',d}$, we can solve for

$$\mu_{\ell',d} = \frac{(-1)^d}{\Psi_{\ell'}(d-1)} \left[\frac{g(d|\ell')\Gamma_{d-1} \sum_{k=1}^{d-1} c_k(\psi_{\ell,d-1})\mu_{\ell,k} - \Gamma_d g(d|\ell) \sum_{k=1}^{d-1} c_k(\psi_{\ell',d-1})\mu_{\ell',k} - (-1)^{d-1} g(d|\ell')\kappa_d \Psi_{\ell'}(d-1)}{g(d|\ell') - \Gamma_d g(d|\ell)} \right]$$

So as long as the denominators in the expressions for $\psi_{\ell'}(d)$ and $\mu_{\ell',d}$ are not zero, we would have identification. $\square$

We can see that with $\kappa_d = 0$ for $d = 1, \dots, \bar{D}$ and $\gamma_d = 1$ for $d = 2, \dots, \bar{D}$, the above theorem is equivalent to the result in the main text. Also, note that the theorem can more generally be applied to situations with other observable characteristics. For instance, with $\kappa_d = 0$ for $d = 1, \dots, \bar{D}$ and $\gamma_d = \gamma$ for $d = 1, \dots, \bar{D}$, the above is equivalent to the discrete MPH model. In the following subsection, I investigate how the estimates of structural hazard vary under different assumptions on κ_d and γ_d .

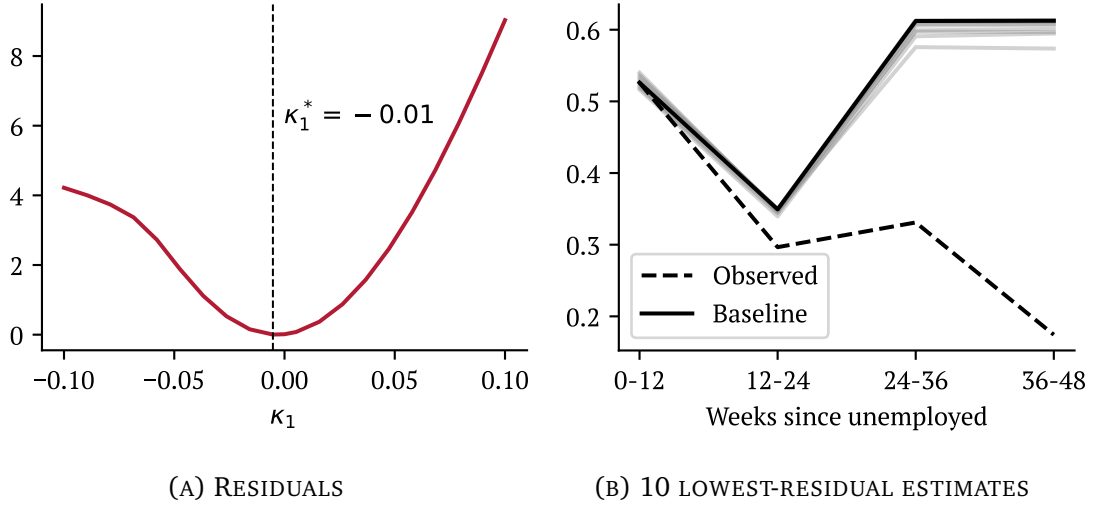
E.2 Implementation

In our estimation, we utilized two lengths of notice, 1-2 months (S) and >2 months (L). Let's define $\kappa_d = \mu_{L,d} - \mu_{S,d}$ and $\gamma_d = \psi_L(d)/\psi_S(d)$. For our baseline estimates, we assumed that the distribution of heterogeneity for individuals with these different notice lengths was identical, i.e., $\kappa_d = 0$ for all d . We also assumed that after the first period, the structural hazards for both the groups were the same, so $\gamma_d = 1$ for $d > 1$. I now study how our estimates change if the underlying distribution of heterogeneity and/or the structural hazards after the initial period are different for workers with different lengths of notice. In particular, I perform the following three exercises.

1. Allow average type to vary

I relax the assumption that notice length is independent of a worker's type and let the mean of the heterogeneity distribution vary across the two groups. I assume that apart from the mean, the rest of the shape of the distribution for the two groups is identical. Since we have $\bar{D} = 4$, this implies that in the 2nd, 3rd, and 4th central moment, the

FIGURE E1: ALLOW AVERAGE TYPE TO VARY



Note: The figure presents results from the estimation of a more generalized Mixed Hazard model, where the mean of the heterogeneity distribution for individuals with different lengths of notice is allowed to vary according to the parameter κ_1 . Panel A presents the residuals from GMM estimation for different values of κ_1 . Panel B presents the structural hazard estimates from the ten best models with the lowest-valued residuals (light grey lines), compared to the baseline estimate (solid line) and the observed hazard in the data (dashed line).

variance, skewness, and kurtosis for the two groups are identical. Scale changes would impact the non-central moments, so all four κ_d s will be non-zero. Denote central moments by $\tilde{\mu}$. Note that, $\tilde{\mu}_2 = \mu_2 - \mu_1^2$. Then since we need $\tilde{\mu}_{S,2} = \tilde{\mu}_{L,2}$,

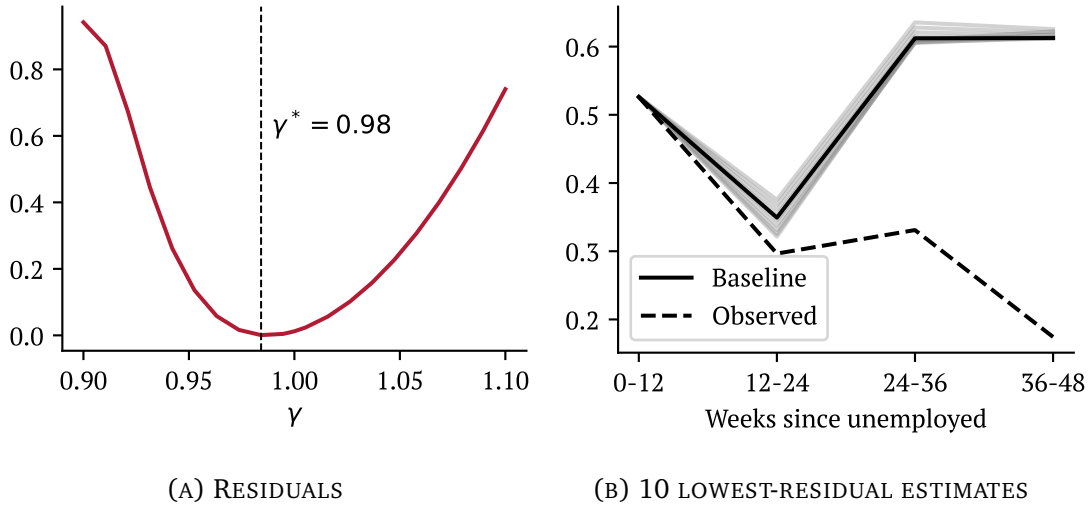
$$\mu_{S,2} - \mu_{S,1}^2 = \mu_{S,2} + \kappa_2 - (\mu_{S,1} + \kappa_1)^2 \rightarrow \kappa_2 = \kappa_1(\kappa_1 + 2\mu_{S,1})$$

Similarly, noting that $\tilde{\mu}_3 = \mu_3 - 3\mu_1\mu_2 + 2\mu_1^3$ and setting $\tilde{\mu}_{S,3} = \tilde{\mu}_{L,3}$, implies $\kappa_3 = \kappa_1(\kappa_1^2 + 3\kappa_1\mu_{S,1} + 3\mu_{S,2})$. And since, $\tilde{\mu}_4 = \mu_4 - 4\mu_3\mu_1 + 6\mu_2\mu_1^2 - 3\mu_1^4$, then setting $\tilde{\mu}_{S,4} = \tilde{\mu}_{L,4}$, we would have $\kappa_4 = \kappa_1(\kappa_1^3 + 4\kappa_1^2\mu_{S,1} + 6\kappa_1\mu_{S,2} + 4\mu_{S,3})$.

Now, assuming $\gamma_d = 1$ for $d > 1$ and normalizing $\mu_{S,1} = 1$, I reestimate the model with κ_1 values spaced at 0.01 intervals within the range $[-0.1, 0.1]$.³⁴ κ_2, κ_3 and κ_4

³⁴For values beyond this interval, the model fit deteriorates drastically, and the estimated moments of the heterogeneity distribution blow up in either direction.

FIGURE E2: ALLOW STRUCTURAL HAZARDS AFTER THE FIRST PERIOD TO VARY



Note: The figure presents results from the estimation of a more generalized Mixed Hazard model, where the structural hazard after the initial period for individuals with different lengths of notice is allowed to vary according to the parameter γ . Panel A presents the residuals from GMM estimation for different values of γ . Panel B presents the structural hazard estimates from the 10 best models with the lowest-valued residuals (light grey lines), compared to the baseline estimate (solid line) and the observed hazard in the data (dashed line).

are defined as above. Residuals from this exercise are presented in panel A of Figure E1, from which we can see that the residual-minimizing value of κ is close to zero. In panel B, I present the estimates for structural duration dependence for the values of κ from the ten models with the lowest residuals. Results from this exercise suggest that the assumption $\kappa = 0$ is not unreasonable and likely does not affect the conclusions.

2. Allow structural hazards after the first period to vary

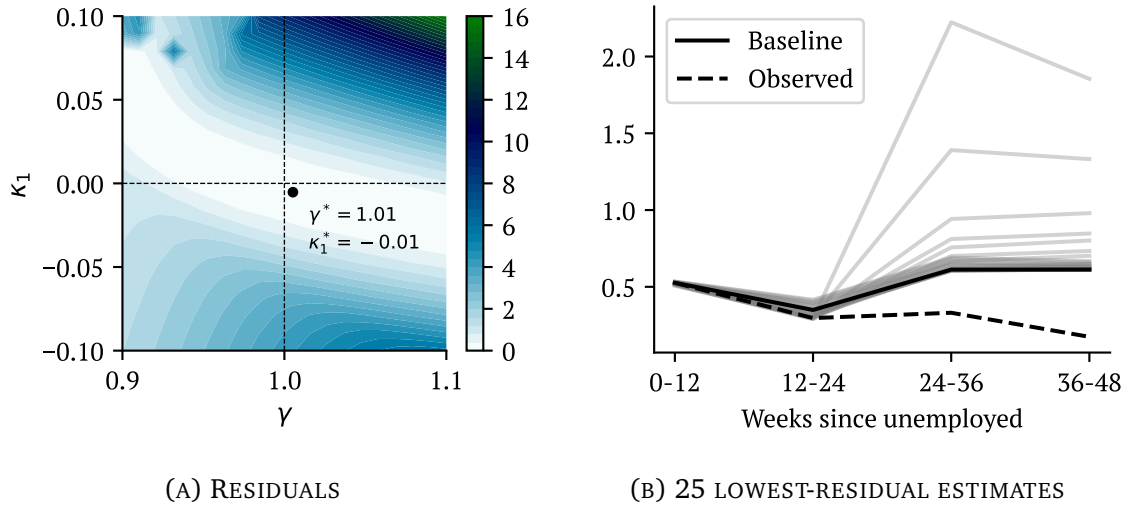
Now, I assume notice length to be independent of worker type but allow structural hazards beyond the initial period to vary for workers with different lengths of notice up to some constant γ . This corresponds to assuming $\kappa_d = 0$ for $d = 1, \dots, \bar{D}$ and $\gamma_d = \gamma$ for $d = 2, \dots, \bar{D}$. I estimate the model for values of γ spaced at intervals of 0.01 within the range $[0.9, 1.1]$. Results from this exercise are presented in Figure E2. As before, panel A shows model residuals for different γ values, while panel B presents the estimates for structural duration dependence for the values of γ from the ten models with the lowest residuals. As we can see, the residual-minimizing value of γ is close to one, giving

support to the assumption $\gamma = 1$.

3. Allow the average type and structural hazards after the first period to vary

Finally, I create a grid for values of κ and γ used in the above two exercises and reestimate the model for each point in the grid. Panel A of Figure E3 presents the residuals for different values in the grid, while panel B of Figure E3 shows the estimates for the set of κ and γ values that result in the 25 models with the lowest residuals. Mirroring the results from the above two exercises, the findings indicate no significant mean differences between short- and long-notice groups, nor a difference in structural hazards for long- vs. short-notice workers beyond the initial period.

FIGURE E3: ALTERNATIVE ASSUMPTIONS ON STRUCTURAL HAZARDS AND HETEROGENEITY DISTRIBUTION

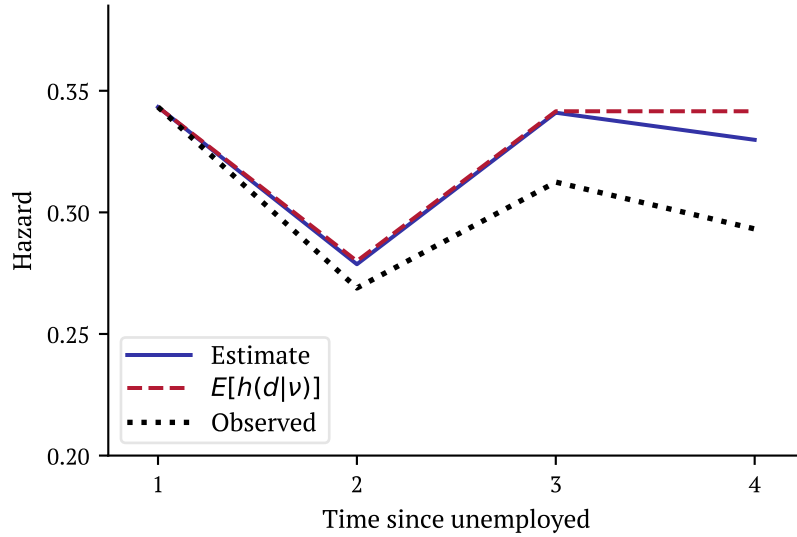


Note: The figure presents results from the estimation of a more generalized Mixed Hazard model. The mean of the heterogeneity distribution for individuals with different lengths of notice is allowed to vary according to the parameter κ_1 . The structural hazard after the initial period for individuals with different lengths of notice is allowed to vary according to the parameter γ . Panel A presents the residuals from GMM estimation for different values of κ_1 and γ . Panel B presents the structural hazard estimates from the 25 best models with the lowest-valued residuals (light grey lines), compared to the baseline estimate (solid line) and the observed hazard in the data (dashed line).

APPENDIX F SEARCH MODEL SIMULATION

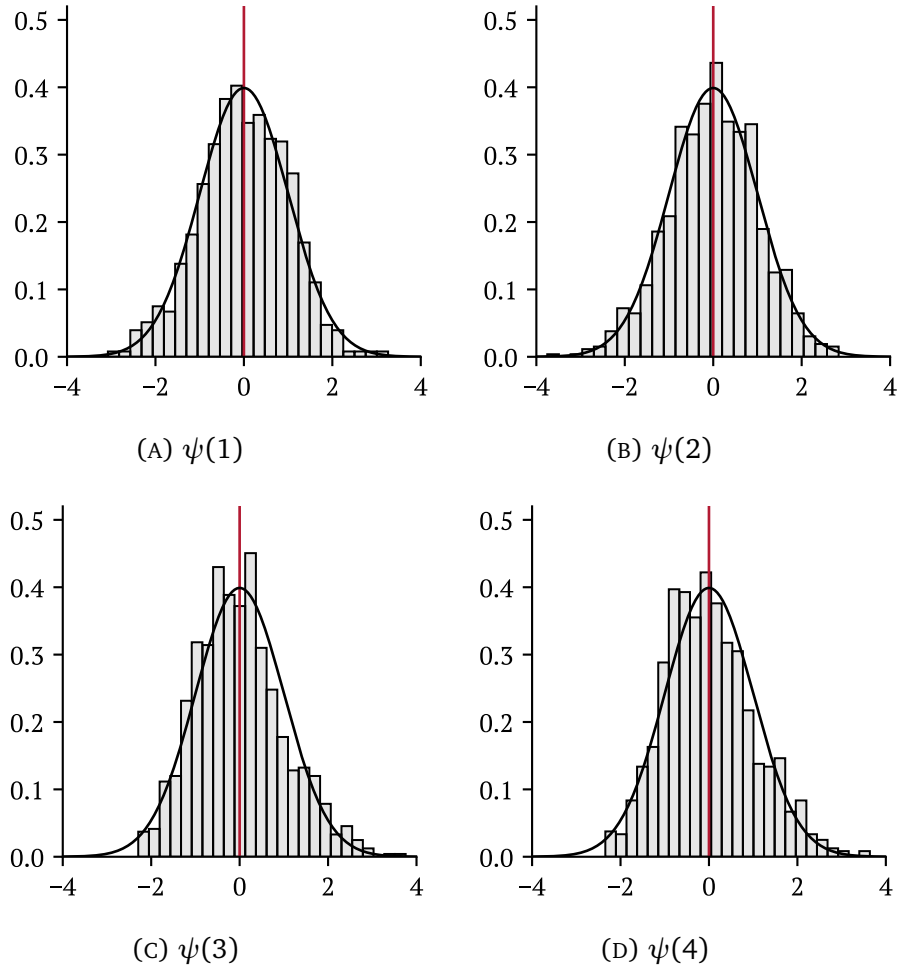
In this section, I simulate data from the search model presented in the main text. To incorporate multiple notice periods, I let the offer rate in the first period be different for long (L) and short (S) notice individuals. I set $\nu_H = 1$, $\nu_L = 0.5$ and $\pi = 0.5$, $\delta_L(1) = 1.25$, $\delta_S(1) = 1$, and $\delta(d) = 0.95$ for $d = 2, 3, 4$. The rest of the parameters are set as in the calibration of the model in the main text. I assume there are 3000 individuals, half of whom receive the L length notice. I simulate data on exit rates for this model 1000 times. The average of estimates for the structural hazard is presented in Figure F1, while the distribution of the estimates is presented in Figure F2.

FIGURE F1: SIMULATION: AVERAGE ESTIMATE



Note: The solid blue line presents the average estimate from 1000 simulations of the search model. The dashed red line presents the structural duration dependence $\mathbb{E}[h(d|\nu)]$ implied by the model. While the dotted black line presents the observed structural duration dependence $\mathbb{E}[h(d|\nu)|D \geq d]$ implied by the model.

FIGURE F2: ESTIMATES USING SIMULATED DATA FROM THE SEARCH MODEL



Note: The figure presents the normalized distribution of structural duration dependence estimated on simulated data from the search model. The vertical lines represent the mean of the distribution for each structural hazard. Standard normal density is overlaid for reference.

REFERENCES

- Abbring, J. H. (2012). Mixed Hitting-Time Models. *Econometrica* 80(2), 783–819.
- Alvarez, F. E., K. Borovičková, and R. Shimer (2016, April). Decomposing Duration Dependence in a Stopping Time Model. Working Paper w22188, National Bureau of Economic Research.
- Alvarez, F. E., K. Borovičková, and R. Shimer (2021, July). Consistent evidence on duration dependence of price changes. Working Paper 29112, National Bureau of Economic Research.
- Boone, J. and J. C. van Ours (2012, December). Why is There a Spike in the Job Finding Rate at Benefit Exhaustion? *De Economist* 160(4), 413–438.
- Brinch, C. N. (2007, April). Nonparametric Identification of the Mixed Hazards Model with Time-Varying Covariates. *Econometric Theory* 23(2), 349–354.
- Burdett, K. and T. Vishwanath (1988, October). Declining Reservation Wages and Learning. *The Review of Economic Studies* 55(4), 655–665.
- Coles, M. G. and E. Smith (1998). Marketplaces and Matching. *International Economic Review* 39(1), 239–254.
- Cox, D. (1972). Regression Models and Life-Tables.
- DellaVigna, S., J. Heining, J. F. Schmieder, and S. Trenkle (2021, 10). Evidence on Job Search Models from a Survey of Unemployed Workers in Germany*. *The Quarterly Journal of Economics* 137(2), 1181–1232.
- DellaVigna, S., A. Lindner, B. Reizer, and J. F. Schmieder (2017, November). Reference-Dependent Job Search: Evidence from Hungary*. *The Quarterly Journal of Economics* 132(4), 1969–2018.

- Elbers, C. and G. Ridder (1982, July). True and Spurious Duration Dependence: The Identifiability of the Proportional Hazard Model. *The Review of Economic Studies* 49(3), 403–409.
- Farber, H. S. (2017). Employment, Hours, and Earnings Consequences of Job Loss: US Evidence from the Displaced Workers Survey. *Journal of Labor Economics* 35(S1), 235–272.
- Ganong, P. and P. Noel (2019, July). Consumer Spending during Unemployment: Positive and Normative Implications. *American Economic Review* 109(7), 2383–2424.
- Gonzalez, F. M. and S. Shi (2010). An Equilibrium Theory of Learning, Search, and Wages. *Econometrica* 78(2), 509–537.
- Heckman, J. and B. Singer (1984). A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data. *Econometrica* 52(2), 271–320.
- Honoré, B. E. (1993). Identification Results for Duration Models with Multiple Spells. *The Review of Economic Studies* 60(1), 241–246.
- Jarosch, G. and L. Pilossoph (2019, July). Statistical Discrimination and Duration Dependence in the Job Finding Rate. *The Review of Economic Studies* 86(4), 1631–1665.
- Katz, L. F. and B. D. Meyer (1990, February). The impact of the potential duration of unemployment benefits on the duration of unemployment. *Journal of Public Economics* 41(1), 45–72.
- Kolsrud, J., C. Landais, P. Nilsson, and J. Spinnewijn (2018, April). The Optimal Timing of Unemployment Benefits: Theory and Evidence from Sweden. *American Economic Review* 108(4-5), 985–1033.

- Kroft, K., F. Lange, and M. J. Notowidigdo (2013, August). Duration Dependence and Labor Market Conditions: Evidence from a Field Experiment*. *The Quarterly Journal of Economics* 128(3), 1123–1167.
- Krueger, A. B. and A. Mueller (2011). Job Search, Emotional Well-Being, and Job Finding in a Period of Mass Unemployment: Evidence from High Frequency Longitudinal Data [with Comments and Discussion]. *Brookings Papers on Economic Activity*, 1–81.
- Lancaster, T. (1979). Econometric Methods for the Duration of Unemployment. *Econometrica* 47(4), 939–956.
- Lentz, R. and T. Tranæs (2005, July). Job Search and Savings: Wealth Effects and Duration Dependence. *Journal of Labor Economics* 23(3), 467–489.
- Machin, S. and A. Manning (1999, January). Chapter 47 The causes and consequences of longterm unemployment in Europe. In *Handbook of Labor Economics*, Volume 3, pp. 3085–3139. Elsevier.
- Marinescu, I. and D. Skandalis (2021, May). Unemployment Insurance and Job Search Behavior*. *The Quarterly Journal of Economics* 136(2), 887–931.
- Mortensen, D. T. (1986, January). Chapter 15 Job search and labor market analysis. In *Handbook of Labor Economics*, Volume 2, pp. 849–919. Elsevier.
- Mueller, A. I., J. Spinnewijn, and G. Topa (2021, January). Job Seekers’ Perceptions and Employment Prospects: Heterogeneity, Duration Dependence, and Bias. *American Economic Review* 111(1), 324–363.
- Pavoni, N. (2009). Optimal Unemployment Insurance, with Human Capital Depreciation, and Duration Dependence*. *International Economic Review* 50(2), 323–362.
- Pavoni, N. and G. L. Violante (2007, January). Optimal Welfare-to-Work Programs. *The Review of Economic Studies* 74(1), 283–318.

- Pissarides, C. A. (1992, November). Loss of Skill During Unemployment and the Persistence of Employment Shocks*. *The Quarterly Journal of Economics* 107(4), 1371–1391.
- Rosenbaum, P. R. (1987, 2024/07/01/). Model-based direct adjustment. *Journal of the American Statistical Association* 82(398), 387–394.
- Rosenbaum, P. R. and D. B. Rubin (1983, April). The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1), 41–55.
- Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and non-randomized studies. *Journal of educational Psychology* 66(5), 688.
- Schmieder, J. F., T. von Wachter, and S. Bender (2016, March). The effect of unemployment benefits and nonemployment durations on wages. *American Economic Review* 106(3), 739–77.
- Shimer, R. and I. Werning (2006, May). On the Optimal Timing of Benefits with Heterogeneous Workers and Human Capital Depreciation. Technical Report w12230, National Bureau of Economic Research.
- van den Berg, G. J. (1990). Nonstationarity in Job Search Theory. *The Review of Economic Studies* 57(2), 255–277.
- van den Berg, G. J. and J. C. van Ours (1996). Unemployment Dynamics and Duration Dependence. *Journal of Labor Economics* 14(1), 100–125.