ESSAYS IN LABOR ECONOMICS

DIVYA BHAGIA

A dissertation

submitted to the Faculty of

the department of Economics

in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Boston College
Morrissey College of Arts and Sciences
Graduate School

May 2022

ESSAYS IN LABOR ECONOMICS

DIVYA BHAGIA

Dissertation Committee:

THEODORE PAPAGEORGIOU (CHAIR)

SHAKEEB KHAN

KEVIN LANG

ARTHUR LEWBEL

Abstract

This dissertation contains three independent chapters on topics in labor economics. In the first chapter, I examine the sources of decline in the job-finding rate over the spell of unemployment. In particular, I distinguish between dynamic selection and the adverse effect of longer unemployment durations. In the second chapter, Carter Bryson and I explore the role of broad sectoral shifts in labor demand in explaining the divergence of employment outcomes of Black and White men during the second half of the 20th century. Finally, in the third chapter, I propose a nonparametric estimator for the discrete-time version of the Mixed Proportional Hazard (MPH).

In the first chapter, “Duration Dependence and Heterogeneity: Learning from Early Notice of Layoff,” I disentangle the sources of decline in the observed job-finding rate over the spell of unemployment. It is possible that an individual worker’s likelihood of exiting unemployment declines with the duration of unemployment (e.g., due to

employers discriminating against long-term unemployed). However, workers also differ in their employability, their desperation to find a job, or perhaps their ability to look for jobs. Such heterogeneity across workers, which is often unobserved, implies that the observed job-finding rate declines even when an individual worker's exit probability does not. As the more employable workers exit early, the still unemployed workers are increasingly the less employable ones. A long literature tries to disentangle the role of structural duration dependence from heterogeneity, but it has proved to be quite challenging. I develop and implement a novel approach to answer this question by leveraging variation in the length of notice an individual receives from their previous employer before being laid off.

The key idea behind my approach is that workers with longer notice start searching earlier and are more likely to exit unemployment early. If workers are heterogeneous, then the composition of long-notice workers should be worse at later durations as the more employable workers from this group have already left. This observation enables me to pin down the extent of heterogeneity and estimate structural duration dependence. My estimates imply that there is substantial heterogeneity in the likelihood of exiting unemployment across workers. I find that roughly 40% of the decline in the observed job-finding rate over the first five months is due to dynamic selection. Further, an individual's exit probability increases up to unemployment insurance (UI) exhaustion and remains constant after that. This is in contrast to the observed exit rate, which continues to decline even after benefit exhaustion. Recently, researchers have tried to explain this decline in the observed exit rate after UI exhaustion. For instance, it can be rationalized with storable offers (Boone and van Ours, 2012) or reference-dependent utility (DellaVigna et al., 2017). My estimates suggest that most of the decline in the observed exit rate after UI exhaustion is due to changes in the composition of workers.

In the second chapter which is joint work with Carter Bryson, “Understanding the Racial Employment Gap: The Role of Sectoral Shifts,” we quantify the extent to which sectoral reallocation can explain the divergence in employment outcomes of Black and White men during the last three decades of the 20th century. Using a shift-share strategy, we document that local employment-to-population ratios for Black men are relatively more responsive to local labor demand shocks. We also document substantial population responses for both groups of workers. Finally, we provide a framework incorporating frictional unemployment and imperfect mobility across locations to aggregate these local responses. We find that sectoral shifts can explain roughly half of the observed exacerbation in the employment-to-population ratio differential between Black and White workers over 1970–2000. Furthermore, our findings indicate that the increase in the differential due to sectoral shifts results from the greater responsiveness of Black workers to local labor demand shifts rather than a higher concentration of these shifts in areas or sectors with a higher share of Black workers.

In the final chapter, “The Discrete-Time Mixed Proportional Hazard Model”, I propose a nonparametric estimator for the discrete-time MPH model. Hazard models of event durations are widely employed in economics to analyze unemployment spells, retirement decisions, and an array of other topics. As the findings from the first chapter highlight, ignoring unobserved heterogeneity while analyzing duration data can lead to inaccurate inferences. The MPH model explicitly accounts for such heterogeneity but estimating this model can be challenging. I set up a discrete-time MPH model and propose an estimator for it that is based on the Generalized Method of Moments and is easy to implement.

ACKNOWLEDGMENTS

This work would not have been possible without the continued guidance and support of my advisors, Theodore Papageorgiou, Shakeeb Khan, Kevin Lang, and Arthur Lewbel. I am also grateful for India's affordable public education system, without which I would not have been able to undertake doctoral studies.

Finally, I thank my partner, Chloë, for her unconditional love and support, which kept me going during the most challenging times. I also thank our dog, Dobby, for reminding me to step away from my desk to take walks, stretch, and hydrate.

Contents

List of Tables

List of Figures

Chapter 1

Duration Dependence and Heterogeneity: Learning from Early Notice of Layoff

1.1 Introduction

A well-established empirical regularity is that the observed job-finding rate declines over the spell of unemployment, except for a spike around the time of unemployment insurance (UI) exhaustion. The observed decline in the job-finding rate could represent negative structural duration dependence — longer time out of work reduces a worker's likelihood of exiting unemployment. This would be true, for example, if employers discriminate against long-term unemployed workers (Kroft et al., 2013) or if workers search less aggressively once they get used to lower-income (DellaVigna et al., 2017). However, workers with different unemployment durations who may appear similar to researchers may actually be very different from each other. For instance, workers may differ in their employability, desperation to find a job, or perhaps their ability to look for jobs. Such heterogeneity across workers would imply that the observed job-finding rate declines even in the absence of structural duration dependence. As the more employable workers exit early, the still unemployed are increasingly the less employable ones.

Understanding how a worker's likelihood of exiting unemployment evolves over the spell and the extent of heterogeneity across workers are crucial for the design of unemployment policies.[1] Moreover, the magnitude and direction of structural duration dependence has consequences for the incidence of long-term unemployment and also affects the speed of recovery from economic downturns (Pissarides, 1992). Given its importance, a long literature has attempted to disentangle the sources of decline in the observed exit rate, though it has proved to be quite challenging to do so.

In this paper, I develop and implement a novel approach to empirically disentangle the role of structural duration dependence and unobserved heterogeneity in determining the evolution of the observed job-finding rate. I accomplish this by leveraging variation in the length of notice an individual receives from their previous employer before being laid off. The key idea behind my approach is that how the observed job-finding rate of long-term unemployed varies with notice length is informative about the underlying heterogeneity. To see this, consider two identical groups of workers. Workers in one of these groups get a notice, while workers in the other do not. Some workers who get a notice find a job during the notice period and exit unemployment right at the beginning of the spell. If workers are heterogeneous, then it is the more employable workers who find a job during the notice period. In which case, amongst

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

workers who are still unemployed at some later duration, those who got a notice should be compositionally worse than those who did not. On the other hand, if all workers are identical, there are no compositional differences between long-term unemployed workers who did or did not get a notice. Such compositional differences, if any, can be inferred from how the observed job-finding rate of long-term unemployed varies with notice length. Hence, I can pin down the extent of heterogeneity and estimate structural duration dependence.

I operationalize this intuition by formulating a Mixed Hazard (MH) model (Lancaster, 1979) in discrete time with multiple notice periods. [2] In this framework, an individual's hazard of exiting unemployment at any duration is specified as a product of their unobservable type and a structural hazard that varies with the duration of unemployment and the length of the notice period.[3] I show that it is possible to identify structural duration dependence while allowing for arbitrary heterogeneity across workers as long as the following two conditions hold: (a) the initial distribution of heterogeneity is identical across workers with different notice lengths and (b) the length of notice affects an individual's hazard at the onset of the unemployment spell but not at subsequent durations. Based on my identification result, I develop a robust method for estimating the model using the Generalized Method of Moments (GMM).

Relative to the existing literature on identification and estimation of the Mixed Hazard

[2] Hausman and Woutersen (2014) provides a succinct review of the literature.

[3] In the literature, it is more common to see the Mixed Proportional Hazard (MPH) model (Lancaster, 1979). In the MPH model, the structural hazard is restricted to be multiplicatively separable in the duration and the observable characteristic(s) as well.

(MH) model, my approach goes further in three dimensions. First, I do not impose any functional form restrictions on the distribution of heterogeneity. In a seminal paper, Heckman and Singer (1984) demonstrate that estimates of structural duration dependence are susceptible to misspecification of unobserved heterogeneity. Secondly, I impose restrictions on the structural hazard that are not arbitrary and can be derived from economic theory. [4] I assume that notice periods do not impact an individuals' hazard at later durations; this is consistent with a large class of job-search models. Finally, I provide a root-n consistent estimator for structural duration dependence and the moments of the distribution of unobserved heterogeneity. The estimator also easily incorporates right-censored data. The identification result here is akin to Brinch (2007) but in discrete-time; nonetheless, my proof is constructive.

I use data from the Displaced Worker Supplement (DWS), a biennial supplement of the Current Population Survey (CPS), for estimating the relevant moments. I compare workers who receive a notice of one to two months to workers who receive a notice of more than two months. I document that the observed job-finding rate in the first 12 weeks is 9-10 percentage points higher for workers with the longer notice. This is due to a greater proportion of long-notice workers transitioning directly to their next job without even entering unemployment. However, beyond the first 12 weeks, the

[4] Existing non-parametric identification results for the Mixed Hazard model rely on variation in some exogenous variable that enters the structural hazard multiplicatively (Elbers and Ridder, 1982; Heckman and Singer, 1984). The applicability of these results has been limited in practice due to difficulty in finding a variable that satisfies this criteria and because of lack of a convenient estimator. Another approach to identification is using multiple spell data (Honoré, 1993); however, identification there relies on assuming that unobserved characteristics of the job seeker remain constant across repeated spells.

observed job-finding rate is lower for workers with the longer notice. This is consistent with the composition amongst long-notice workers being worse at later durations as a greater proportion of them have already exited unemployment.

Consequently, estimates from the Mixed Hazard model uncover substantial heterogeneity in individual job-finding probabilities. The observed job-finding rate declines by 40% over the first five months; I find that almost half of this decline is due to the changing composition of workers over the spell of unemployment. Further, I find that after the first five months, an individual worker's job-finding probability increases until the time of their unemployment benefit exhaustion and remains constant after. This is in contrast to the observed-job finding rate, which keeps declining 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, 2020). I provide a much simpler explanation: as a large number of individuals exit right at benefit exhaustion, the composition of workers who are still unemployed after is significantly worse, which leads to a decline in the observed job-finding 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.

The estimation attributes the lower observed job-finding rate of the long-notice workers after the first 12 weeks to the presence of heterogeneity. However, this could also result from long-notice workers being negatively selected or a longer notice reducing a worker's exit probability after 12 weeks. In which case, my estimates would overestimate the extent of underlying heterogeneity. To assess the robustness of my results, I extend the analysis in several directions. I repeat the analysis conditional on observable characteristics of job-seekers and also allow for unobservable differences across the two groups. In both cases, my results are similar to the baseline estimates. I also allow for the notice length to affect an individual's job-finding rate even beyond the first 12 weeks. From this exercise, I can reject that the structural hazard for long-notice workers is lower than that of short-notice workers. Overall, the evidence that heterogeneity across workers plays a predominant role in determining the evolution of the observed job-finding rate is robust to alternative identifying assumptions.

My paper contributes to the large literature studying the dynamics of job-finding over the spell of unemployment. Given challenges with estimation, empirical studies employing the Mixed Hazard model often make strong functional form assumptions. As a consequence, in their review, Machin and Manning (1999) find mixed evidence on structural duration dependence from these studies. I exploit a new source of variation that enables me to estimate structural duration dependence with relatively minimal assumptions. Recently, Alvarez et al. (2016) revived this strand of work; they estimate a Mixed Hitting-Time (MHT) model (Abbring, 2012) using Austrian social security data on a selected sample of workers with multiple unemployment spells. They 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.[5] Mueller et al. (2021) utilize variation in expectations about job-finding from survey data to pin-down variation in actual job-finding rates. Both of these papers also document substantial heterogeneity across job-seekers; however, my estimator of structural duration dependence is flexible enough to capture changes around UI exhaustion.[6]

Given difficulties with estimating structural duration dependence, researchers have instead focussed on it's 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. [7] Several papers have documented how search effort or reservation wages evolve over the spell of unemployment (Krueger and Mueller, 2011; Marinescu and Skandalis, 2021; DellaVigna et al., 2020). Evidence provided in this paper suggests that call-back rates or other factors affecting returns to search matter initially; however, it is possible that changes in worker's optimizing behavior determines the likelihood of exiting unemployment at later durations.

Finally, a large number of papers document a spike in exit rates at UI exhaustion;

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

[6]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.

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

exit rate increases up to benefit exhaustion and declines thereafter.[8] 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, however; I do not find evidence of the decline after. Boone and van Ours (2012) suggest storable job offers to rationalize this spike. While, DellaVigna et al. (2017) argue that with reference dependence, search models predict that search effort decreases after benefits exhaustion, instead of staying constant. My estimates suggest that the decline in the exit rate after UI exhaustion can be attributed to a shift in composition of surviving workers as a huge proportion of workers exit unemployment right at benefit-exhaustion. However, the individual exit probability remains constant, consistent with the prediction of standard search models.

1.2 Econometric Framework

This section presents the main identification result of this paper. In Section 1.2.1, I formulate the Mixed Hazard model in discrete time with multiple notice periods and discuss the plausibility of underlying identifying assumptions. In Section 1.2.2, I show that the key components of this model are non-parametrically identified. Moreover, the analysis can be easily extended to deal with right-censored data and to include additional observed explanatory variables as shown in Section 1.2.3 and Section 1.2.4,

[8] 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).

respectively.

1.2.1 Mixed Hazard Model

The realized unemployment duration D is a random variable that takes values in $\{1,2,...\}$. Denote the cumulative and probability distribution function of unemployment duration by $G(.)$ and $g(.)$, respectively. The job-finding rate or hazard rate $h(d)$ is defined as follows:

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

In addition, the survival function $S(d) = 1 - G(d)$ is defined as the probability of unemployment duration D being greater than d. Then according to the preceding definitions, $S(d) = \prod_{s=1}^{d}(1-h(s))$.

Workers are heterogeneous and have an unobservable fixed type ν with the cumulative distribution $F(.)$. Before being laid off, workers receive a notice period of length L. I assume that the conditional hazard $h(d|\nu, l)$ for an individual with a notice of length l can be specified as a product of their type ν and a structural component $\psi_l(d)$ that is common to all individuals but varies with the duration of unemployment and the length of the notice period.

Assumption 1. *(Mixed Hazard) An individual's hazard rate conditional on receiving a notice of length l is given by*

$$h(d|\nu, l) = \psi_l(d)\nu$$

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

$\bar{\nu} = 1/\max_{d,l}\{\psi_l(d)\}$.

Here, the support of ν is chosen to ensure that individual job-finding rates lie between 0 and 1.

Lancaster (1979) was the first to introduce unobserved heterogeneity in models of duration data. He generalized the proportional hazard model (Cox, 1972) and specified the hazard rate as a product of a regression function that captures the effect of observed explanatory variables, a structural hazard that captures variation in the hazard over the spell, and a random variable that accounts for unobserved heterogeneity. The Mixed Hazard model formulated here is identical to Lancaster's Mixed Proportional Hazard (MPH) model except that the effect of the observable characteristic (i.e. the length of notice) is not restricted to enter the hazard proportionally. The assumption that workers have a fixed type, while standard in the literature, is not without loss of generality. It is possible that some workers can quickly find work but if they fail to do so, they suffer much larger adverse effects of longer unemployment. [9] However, the Mixed Hazard model is popular because it tractably incorporates two sources of variation in the average hazard rate over the duration of unemployment: structural duration dependence and heterogeneity across individuals. The following proposition clarifies this.

[9]On the other hand, it is also possible that some workers do not search for a job at all in the first couple of weeks but once they do start searching, they can instantaneously find a job.

Proposition 1. *Under Assumption 1, the observed exit rate at any duration is given by*

$$\tilde{h}(d|l) = \frac{g(d|l)}{1 - G(d-1|l)} = \psi_l(d)\mathbb{E}(\nu|D \geq d, l)$$

where $\mathbb{E}(\nu|D \geq d, l)$ *is the average type of workers who have survived until* d*. Additionally, for any* d*,* $\mathbb{E}(\nu|D \geq d, l) \geq \mathbb{E}(\nu|D \geq d+1, l)$*.*

Proof. See A.1. □

The observed hazard $\tilde{h}(d|l)$ is influenced by structural duration dependence $\psi_l(d)$ as well as the changing composition of workers over the spell of unemployment captured by $\mathbb{E}[\nu|D \geq d, l]$. Since Assumption 1 implies that workers with higher values of ν are more likely to exit unemployment at any duration, the surviving workers at later durations are disproportionately those with smaller values of ν. Hence, the average type of workers at later durations is lower than at earlier durations.

If there were no variation in the length of notice, the model specified so far would not be identified. Assume for instance, we observe that $\tilde{h}(d|l)$ declines over the spell of unemployment. Then we cannot distinguish between the following two scenarios which would both be consistent with the observed decline. One, there is no structural duration dependence, that is, $\psi_l(d)$ is constant over the spell of unemployment. However, there is significant heterogeneity across workers which causes the average type of workers and hence the average hazard to decline. Two, there is no heterogeneity across workers so the average type is constant but the structural hazard declines over the spell of unemployment.

With minimal variation in some observable characteristic, it is possible to identify key components of the model specified so far from the unemployment duration distribution. However, it is only possible to utilize this variation if we know how the distribution of heterogeneity and structural duration dependence vary by this characteristic. I formalize this notion in Theorem 2 in A.4. Elbers and Ridder (1982) and Heckman and Singer (1984) provide identification results for the MPH model where the effect of observable characteristics affects the hazard in the same way at any duration.[10] They also assume, as is typical in this literature, that the observable characteristic is independent of worker's unobservable type. So in other words, they show that identification can be achieved if the distribution of heterogeneity and structural duration dependence do not vary by the observable characteristic. I follow the previous literature and assume that notice periods are independent of worker's unobservable type. However, with respect to the structural hazards, I assume that notice length only affects the hazard initially and hence structural duration dependence does not vary with notice length after the initial period. Both of these are strong assumptions that I relax later.

The setting here is analogous to a duration model with time-varying observable variables. Brinch (2007) provides a non-constructive proof for this model in continuous time. A key difference here is that the exposition is in discrete time which allows me

[10] Ridder (1990) shows that MPH model is a special case of the Generalized Accelerated Failure-Time (GAFT) model and provides a non-parametric identification result for the GAFT model.

to consistently estimate the parameters of the model using GMM. [11] To the best of my knowledge, Alvarez et al. (2021) is the only other paper utilizing the resulting moment conditions from a discrete version of the Mixed Hazard model and constructing a GMM estimator. However, their identification result and estimator pertains to multiple spell data. [12]

Assumption 2. *(Independence) Notice length is independent of worker type, i.e. $L \perp \nu$.*

Independence ensures that the distribution of unobserved heterogeneity amongst workers with different lengths of notice is identical. Such that for any l, l', $F(\nu|l) = F(\nu|l')$. Note that, this does not imply that the type-distribution of surviving workers at each duration is identical for different lengths of notice, but only that the initial composition of workers is identical. In fact, if the length of the notice period affects an individual's job-finding rate at some point during the spell of unemployment, then at subsequent durations the composition of workers with different notice lengths will vary even if their initial composition was identical. This is precisely the variation that enables me to pin down heterogeneity.

In practice, it is possible that workers with certain characteristics are more likely to receive a longer notice than others. If these characteristics are also correlated with an individual's likelihood of finding a job, then Assumption 2 will be violated. To account

[11] Hausman and Woutersen (2014) provide an estimator that converges at the regular $\sqrt{N}$ rate for the semi-parametric MPH model with time-varying variables. Their estimator incorporates discrete measurement of duration.

[12] van den Berg and van Ours (1996) also set up a discrete time MPH, however, they do not derive the distribution of their estimator.

for this possibility, in Section 1.2.4, I relax Assumption 2 to require the notice length to be independent of the worker-type conditional on observable characteristics. Additionally in A.4, I consider alternative assumptions on how type-distribution varies with notice period.

Finally, I assume that notice periods can affect an individual's job-finding probability only in the initial period but not beyond that.

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

$$\psi_l(d) = \psi(d)$$

The rationale behind this assumption is that workers with longer notice periods have more time to look for a new job before separating from their previous employer. Hence, it is possible that they are more likely to exit unemployment right at the beginning of their unemployment spell. However, after the initial period, assuming that notice length no longer matters for individual job-finding rates is equivalent to assuming that individual job-finding rates only vary with time elapsed since unemployed but not with time elapsed since start of job search. This later conjecture is reasonable if we think about human capital depreciation or employers discriminating against workers with long unemployment spells. This is because, we expect human capital to depreciate as a consequence of spending long periods of time out of work. Similarly, potential employers observe an individual's time of separation from their last job but not the length of their notice period. The assumption is also consistent with workers running down their savings and searching harder over time, as in Lentz and Tranæs (2005), as savings only start depleting once unemployed. [13]

Assumption 3 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). In my analysis, I group unemployment duration in 12-week intervals and require the job-finding rate after the first 12 weeks to not vary with notice periods. The mean vacancy duration ranges from 14 to 25 days (Davis et al., 2013) and so it is likely that workers have gone through all the jobs in stock in the first 12 weeks. In which case, my assumption is consistent with a stock-flow model of search. In A.4, I consider alternative assumptions on how the structural hazard varies with the length of the notice after the initial period.

1.2.2 Identification

Following Assumption 1, since ν has bounded support all the moments of ν exist and are finite. Denote the k^{th} moment of ν and the k^{th} moment of ν conditional on l by μ_k and $\mu_{l,k}$, respectively. Under Assumption 1, for any length of notice l and some

[13] It is also possible that workers get discouraged over time and stop trying. However, individuals who eventually drop out of the labor force are excluded from the analysis.

integer $\bar{D}$, the conditional distribution of unemployment duration $g(d|l)$ for $d = 1, .., \bar{D}$ is fully determined by the structural hazards $\{\psi_l(d)\}_{d=1}^{\bar{D}}$ and the conditional moments of type-distribution $\{\mu_{l,k}\}_{k=1}^{\bar{D}}$. To see this note,

$$g(d|l) = G(d|l) - G(d-1|l) = \psi_l(d)\mathbb{E}[\nu S(d-1|\nu, l)|l]$$

Expanding the above term for $d = 1, 2, 3, ...$, we can write:

$$\begin{aligned}
g(1|l) &= \psi_l(1)\mu_{l,1} \\
g(2|l) &= \psi_l(2)[\mu_{l,1} - \psi_l(1)\mu_{l,2}] \\
g(3|l) &= \psi_l(3)[\mu_{l,1} - [\psi_l(1) + \psi_l(2)]\mu_{l,2} + \psi_l(1)\psi_l(2)\mu_{l,3}] \\
&\vdots
\end{aligned}$$

Or more compactly,

$$g(d|l) = \psi_l(d)\sum_{k=1}^{\bar{D}} c_k(d, \boldsymbol{\psi}_l)\mu_{l,k} \tag{1.1}$$

where $\boldsymbol{\psi}_l = \{\psi_l(d)\}_{d=1}^{\bar{D}}$ and

$$c_k(d, \boldsymbol{\psi}_l) = \begin{cases} 1 & \text{for } k = 1 \\ c_k(d-1, \boldsymbol{\psi}_l) - \psi_l(d-1)c_{k-1}(d-1, \boldsymbol{\psi}_l) & \text{for } 1 \le k \le d \\ 0 & \text{for } k > d \end{cases}$$

If we observed only one length of notice, we would have $\bar{D}$ population moments but $2\bar{D}$ unknown parameters. In this case, the structural hazards and the moments of the heterogeneity distribution would not be identified from the distribution of unemployment duration.[14] Now suppose we observe two lengths of notice, l and l' and Assumptions 2 and 3 hold as well. Then we have $\mu_{l,k} = \mu_{l',k} = \mu_k$ for $k = 1, ..., \bar{D}$ and $\psi_l(d) = \psi_{l'}(d) = \psi(d)$ for $d = 2, ..., \bar{D}$. So now we have $2\bar{D}$ population moments and $2\bar{D}+1$ unknown parameters. The unknown parameters are the first $\bar{D}$ moments of ν and $\bar{D}+1$ structural hazard $\{\psi_l(1), \psi_{l'}(1), \psi(2), ..., \psi(\bar{D})\}$. With a scale normalization such as $\mu_1 = 1$, the necessary condition for identification would be satisfied. The scale normalization is unavoidable in Mixed Hazard models as rescaling of $\psi_l(d)$ and ν leads to the same distribution of D. The following theorem establishes that the sufficient condition for identification is satisfied as well.

Theorem 1. *Under Assumptions 1 to 3, for any l, l' with $\psi_l(1) \neq \psi_{l'}(1)$ and some integer $\bar{D}$. The structural hazards $\{\psi_l(1), \psi_{l'}(1), \{\psi(d)\}_{d=2}^{\bar{D}}\}$ and the moments of the type distribution $\{\mu_k\}_{k=1}^{\bar{D}}$ can be identified up to scale from $\{G(d|l), G(d|l')\}_{d=1}^{\bar{D}}$.*

Proof. See A.1. □

To see why the identification result holds, note that the observed hazard $\tilde{h}(1|l)$ de-

[14] Note that, even if we were not interested in estimating the moments of the heterogeneity distribution, the structural hazards would still not be identified in this case. See the discussion under Proposition 1.

pends on the structural hazard and the average type of workers as follows:

$$\tilde{h}(1|l) = \psi_l(1)\mu_1$$

Then since $\tilde{h}(2|l) = g(2|l)/S(1|l)$ where $S(1|l) = 1 - \tilde{h}(1|l)$, we can write the observed hazard at $d = 2$ as

$$\tilde{h}(2|l) = \psi(2)\left(\frac{\mu_1 - \psi_l(1)\mu_2}{1 - \psi_l(1)\mu_1}\right) = \psi(2)\mu_1\left(\frac{1 - \tilde{h}(1|l)(\mu_2/\mu_1^2)}{1 - \tilde{h}(1|l)}\right)$$

The second equality in the above expression follows from $\psi_l(1) = \tilde{h}(1|l)/\mu_1$. From here, we can see that $\tilde{h}(2|l)/\tilde{h}(1|l)$ will always be smaller than $\psi(2)/\psi_l(1)$ in the presence of heterogeneity. Further, higher the variance of ν, larger μ_2/μ_1^2 will be and hence further away $\tilde{h}(2|l)/\tilde{h}(1|l)$ will be from $\psi(2)/\psi_l(1)$. This is because with greater heterogeneity across workers, the composition of workers from the the first to second period changes more drastically. For instance, with no heterogeneity across workers $\mu_2/\mu_1^2 = 1$, the composition across both periods is unchanged and hence $\tilde{h}(2|l)/\tilde{h}(1|l) = \psi(2)/\psi_l(1)$. If we knew the extent of heterogeneity across workers as captured by μ_2/μ_1^2, we would know how composition changes from the first to the second period and be able to infer structural duration dependence $\psi(2)/\psi_l(1)$ from observed duration dependence $\tilde{h}(2|l)/\tilde{h}(1|l)$. The variation in notice lengths enables us to learn about the underlying heterogeneity and hence estimate structural duration dependence.

For two lengths of notice l and l',

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

WLOG, say $\tilde{h}(1|l') > \tilde{h}(1|l)$, then $\tilde{h}(2|l)/\tilde{h}(2/l') > 1$. Since more individuals with l' notice leave in the first period, the composition for that group worsens more going from the first to the second period. Moreover, $\tilde{h}(2|l)$ is going to be further above $\tilde{h}(2|l')$ when the variance across workers is higher. Using the above expression we can back out μ_2/μ_1^2 from $\tilde{h}(2|l)/\tilde{h}(2|l')$. A similar argument applies for the intuition for identification of structural hazards beyond the second period. [15]

1.2.3 Censoring

The identification result presented in the previous section pertains to the distribution of completed unemployment durations. However, in most datasets at least some individuals are still unemployed at the time of the survey. For these unemployed individuals, we observe how long they have been unemployed but we do not know if and when they will find a job. Let D_C denote the censoring time, that is, the time elapsed since an individual becomes unemployed to the time of the survey. In the data, we observe completed unemployment duration D for individuals who have already exited

[15]To see why higher moments determine the hazard at later duration, say, we are thinking about how the composition of workers changes going from $d = 2$ to $d = 3$. This will depend on the extent of heterogeneity across workers at the beginning of $d = 2$. If the distribution of heterogeneity had a positive skew, then the variance amongst surviving individuals at $d = 2$ would be lower than the variance amongst individuals at the beginning of $d = 1$. This is because the few individuals with really high likelihoods of exiting unemployment would have already exited reducing the variance amongst surviving workers.

unemployment at the time of the survey. On the other hand, for unemployed individuals, we observe the censoring time D_C. Specifically, for each individual, we observe $\Delta = \min\{D, D_C\}$ as well as an indicator for whether the individual was censored or not. Let $G^{\Delta}(.)$ denote the cumulative distribution of observed durations Δ.

It's useful to note that, we know the unemployment duration for individuals who are censored after some $\bar{D}$ and report a duration of less than $\bar{D}$. In particular, for $d < \bar{D}$, we have $G^{\Delta}(d|l, D_C > \bar{D}) = G(d|l, D_C > \bar{D})$. Now if we assume, as is common in the literature, that D_C is independent of the worker type, we would have $G(d|l, D_C > \bar{D}) = G(d|l)$ and the identification result in Theorem 1 would hold replacing $G^{\Delta}(d|l, D_C > \bar{D})$ for $G(d|l)$. [16] However, here we do not require this assumption. Instead, we just need to assume that the likelihood of being censored after $\bar{D}$ is identical across workers with different notice lengths.

Assumption 4. *Notice periods do not affect the likelihood of the censoring time being greater than $\bar{D}$,*

$$Pr(D_C > \bar{D}|l) = Pr(D_C > \bar{D})$$

The above assumption is more general than assuming $\nu \perp D_C$ or $Pr(D_C > \bar{D}|\nu) = Pr(D_C > \bar{D})$ as that along with Assumption 2 would imply Assumption 4. But more importantly, the assumption that the likelihood of being censored after some $\bar{D}$ is identical across workers with different notice lengths can be verified from the data.

[16] We are implicitly assuming that the structural hazard $\psi_l(d)$ is unaffected by the censoring time. In theory, it may be possible and would be more efficient to condition on $D_C > d$ at every duration d. However, in my dataset I only observe D_C lumpily, at intervals of one year.

Corollary 1. *Under Assumptions 1 to 4, for any* l, l' *and some integer* $\bar{D}$*. The structural hazards* $\{\psi_l(1), \psi_{l'}(1), \{\psi(d)\}_{d=2}^{\bar{D}}\}$ *and the conditional moments of the type distribution* $\{\mathbb{E}(\nu^d|D_C > \bar{D})_k\}_{d=1}^{\bar{D}}$ *can be identified up to scale from* $\{G^{\Delta}(d|D_C > \bar{D}, l), G^{\Delta}(d|D_C > \bar{D}, l)\}_{d=1}^{\bar{D}}$.

Proof. See A.1. □

1.2.4 Observable Characteristics

So far we have assumed that the length of notice is independent of worker's type. However, it is possible that workers who have higher job-finding rates are also more likely to receive a longer notice, or vice versa. To account for this possibility, I extend the analysis to control for observable characteristics. Let X denote a vector of pre-notice variables.

Assumption 1a. *Job-finding rate for an individual with some vector of observable characteristics* X *and notice* l *is given by:*

$$h(d|\nu, l, X) = \psi_l(d, X)\nu$$

where $\nu \in (0, \bar{\nu}]$ *with* $\bar{\nu} = 1/\max_{d,l}\{\psi_l(d, X)\}$.

Now instead of assuming that the notice length is independent of worker type, I follow the treatment effects literature and assume unconfoundedness (Rosenbaum and Rubin, 1983).

Assumption 2a. *Notice length is independent of the worker type conditional on a vector of observed characteristics i.e.* $L \perp \nu | X$.

Finally, we also need to assume that the stationarity assumption holds for each structural hazard $\psi(d, X)$.

Assumption 3a. *For all* l *and* $d > 1$, $\psi_l(d, X) = \psi(d, X)$.

Corollary 2. *Under assumptions* $1a - 3a$*, for any* l, l' *and some integer* $\bar{D}$*. The structural hazards* $\{\psi_l(1, X), \psi_{l'}(1, X), \{\psi(d, X)\}_{d=2}^{\bar{D}}\}$ *and the conditional moments of the type distribution* $\{\mathbb{E}(\nu^d | D_C > \bar{D}, X)\}_{d=1}^{\bar{D}}$ *can be identified up to scale from* $\{G(d|l, X), G(d|l', X)\}_{d=1}^{\bar{D}}$.

Proof. Follows directly from Theorem 1. □

Note that, we can only incorporate discrete observable characteristics as the object $G(d|l, X)$ is not well defined for continuous X's.[17]

1.3 Context and Data

In this section, I first describe the institutional setting and the data. I then document how the observed job-finding rate varies with the length of notice. In the data, the job-finding rate for workers with longer notice periods is higher early in the spell but lower later. As discussed in the previous section, the observed difference at later durations in

[17] To incorporate continuous X's we would not only need to specify how the structural hazard varies with observable characteristics but also specify how the distribution of heterogeneity varies with these characteristics. Even the propensity score matching approach doesn't work here due to non-linearity.

the job-finding rates of workers with longer vs. shorter notices is informative about the underlying heterogeneity across workers.

1.3.1 Institutional Details

In the US, employers are only required to give a notice under certain circumstances. In particular, the Worker Adjustment and Retraining Notification (WARN) act is a federal law that requires employers with 100 or more full-time employees to provide a 60 calendar-day advance notification of plant closings and mass layoffs of employees. [18] While the federal law applies to employers in all states, some states have implemented their own WARN laws that broaden the scope of who is covered under the law. States with significant differences from the federal WARN are California, New York, and Illinois. However, it 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 the UI side, workers who are terminated without cause, in most cases, are eligible to receive UI benefits for a limited duration. The UI program in the US is a federal program, but benefit levels and durations are set by individual states. Eligibility and benefits are determined based on earnings or hours/weeks of work during a base period. Typically this base period is the first four out of five completed calendar quarters that

[18] A plant closing is defined as the shutdown of an employment site (or units within a site) resulting in an employment loss for 50 or more employees during a 30-day period. While, a mass layoff is defined as employment loss during a 30-day period for 500 or more employees, or for 50-499 employees if they make up one-thirds of the employer's active workforce. The law only applies to terminations resulting in a layoff exceeding 6 months, other than discharge for cause, voluntary departure, or retirement.

precede the filing of a claim. For most states, the maximum period for receiving benefits is 26 weeks. Since a 1970 amendment to Federal Unemployment Tax Act (FUTA), there has been an extended benefit program that may be triggered by the state unemployment rate. There are also temporary programs to extend benefits during recessions. Nine states have a uniform benefit duration of 26 weeks. The remaining states have benefit durations that can vary in length based on the applicant's earnings history during the base period.

1.3.2 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 3 years due to a plant closure, their position being abolished, or having insufficient work at their previous employment. Apart from details on workers' lost and current job, 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 my analysis, I consider individuals aged 20 to 60 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 notice of either 1-2 months (*short notice*) or greater than 2 months (*long notice*). I exclude individuals who did not receive any notice at all because it is uncertain whether these workers were displaced or quit their jobs voluntarily. [19] I also exclude a small number of individuals who report receiving a notice period of less than 1 month. The choices are made to minimize observable differences between individuals with long and short notice. For additional details on data construction see B.2.

For some individuals in the data, we do not observe completed unemployment spells. For this reason, I only include individuals who had lost their previous job at least one year before the survey. For individuals who are followed for at least 52 weeks, we can calculate the proportion exiting unemployment at each duration before 52 weeks. Corollary 1 formally states that it is possible to identify structural hazards at durations less than some arbitrarily long duration by restricting our analysis to individuals who are not censored until then. For Corollary 1 to hold, we need that the likelihood of being observed more than a year after separation is identical regardless of notice length. Panel A of Table 1.1 shows that this assumption is not rejected by the data.

Panel B of Table 1.1 presents the summary statistics for the sample used in estimation. Relative to Panel A, this sample excludes workers who had lost their previous job in the last year. In my sample, 1333 individuals received a notice of one months which I refer to as the *short notice*. 1554 individuals received a notice longer than two months which I refer to as the *long notice*. Individuals with the longer notice period are older, have

[19]Even though the DWS is meant to capture separations that are due to firms facing economic challenges, the distinction between quits and voluntary layoffs is blurred. Firms willing to cut costs may reduce or fail to raise wages or cut hours instead of laying off workers. This can encourage workers, specially those who have better alternatives to quit (Farber, 2017).

TABLE 1.1: DESCRIPTIVES BY NOTICE LENGTH

	1-2 months (1)	>2 months (2)	(2)-(1)
Panel A: Time elapsed since job-loss			
Surveyed >52 weeks after	0.69	0.70	0.02
	(0.01)	(0.01)	(0.01)
Observations	1944	2205	
Panel B: Sample used for analysis			
Age	42.02	43.20	1.18***
	(0.27)	(0.25)	(0.37)
Female	0.44	0.46	0.02
	(0.01)	(0.01)	(0.02)
Married	0.61	0.65	0.04**
	(0.01)	(0.01)	(0.02)
Black	0.10	0.09	-0.01
	(0.01)	(0.01)	(0.01)
High School Graduate	0.56	0.57	0.02
	(0.01)	(0.01)	(0.02)
College Graduate	0.42	0.39	-0.03
	(0.01)	(0.01)	(0.02)
Plant Closure	0.45	0.63	0.18***
	(0.01)	(0.01)	(0.02)
Union Membership	0.14	0.15	0.01
	(0.01)	(0.01)	(0.01)
Years of Tenure	6.94	8.99	2.06***
	(0.18)	(0.19)	(0.26)
Log Earnings	6.52	6.56	0.03
	(0.02)	(0.02)	(0.02)
Observations	1333	1554	

Notes: Sample for Panel A consists of individuals between ages 20-60 in the DWS who had worked full-time for at least six months at their previous employment, were provided health insurance at their lost job, and did not expect to be recalled. Panel B further restricts the sample and excludes individuals surveyed within one year of job-loss. Individuals with missing or zero earnings (390 observations) or missing tenure (26 observations) are excluded when calculating statistics pertaining to the missing variable but included otherwise.

had longer attachment to their previous jobs, and are more likely to be laid off due to a plant closure. Individuals with different lengths of notice do appear to have similar levels of education and be in jobs that do not pay too differently. Figure A.1 presents the probability of receiving a longer notice over the business cycle. Workers who get laid off during a recession are less likely to receive a longer notice. This implies that individuals with shorter notice periods tend to be disproportionately laid off during economic downturns. In my estimation, I explicitly account for these differences in observable characteristics across the two groups.

1.3.3 Distribution of Unemployment Duration

If workers start searching for a job once they get a notice, then those with longer notice are more likely to find a job during the notice period. In the data, around 13 percent of workers who get one to two months of notice do not even enter unemployment. These workers may find a job while previously employed and delay their start date or negotiate with their past employers to leave before the end of the notice period. Panel A of Table 2 shows that the likelihood of not entering unemployment is around ten percentage points higher for workers who get a notice of more than two months. This suggests that workers with a longer notice start their search earlier. Moreover, the effect of longer notice on the likelihood of making such job-to-job transitions is robust to controlling for a rich set of worker characteristics and labor market conditions.

To examine how the observed job-finding rate varies with the length of notice over

TABLE 1.2: OBSERVED EXIT RATE – EARLY IN THE SPELL

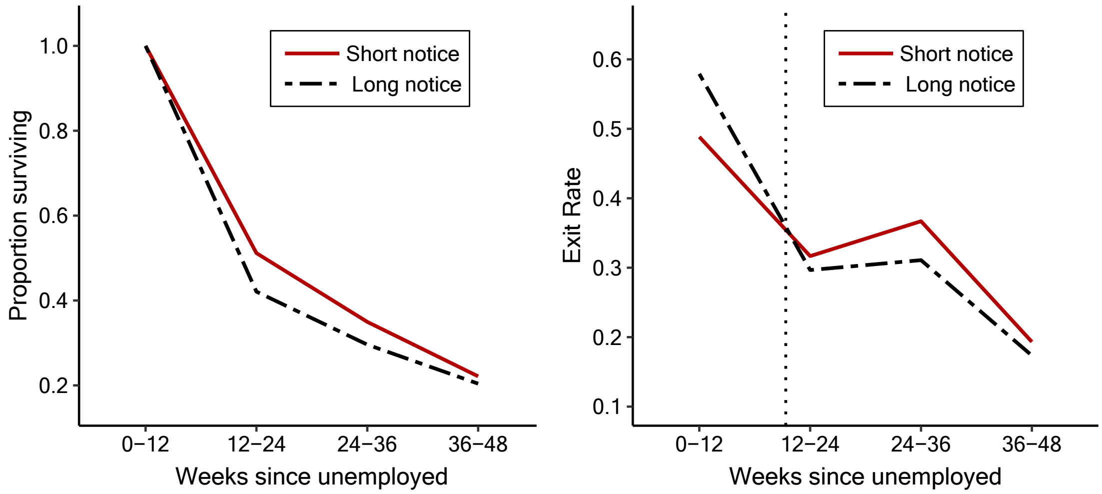
	(1)	(2)	(3)	(4)
PANEL A. $\mathbb{I}\{$UNEMPLOYMENT DURATION = 0 WEEKS$\}$				
>2 month notice	0.105***	0.099***	0.099***	0.090***
	(0.014)	(0.016)	(0.016)	(0.018)
PANEL B. $\mathbb{I}\{$UNEMPLOYMENT DURATION $\leq$ 12 WEEKS$\}$				
>2 month notice	0.091***	0.106***	0.105***	0.091***
	(0.019)	(0.020)	(0.021)	(0.022)
Additional Controls		X	X	X
State FEs			X	X
Industry × Year FEs				X
Observations	2887	2486	2486	2472

Notes: The table shows results from linear regressions of a variable that indicates whether an individual reported an unemployment duration of 0 weeks (Panel A) or less than 12 weeks (Panel B) on an indicator variable that takes value 1 if the individual received a notice of more than 2 months and 0 if they received a notice of 1-2 months. Column (1) includes all individuals in the sample. Columns (2)-(4) exclude individuals with missing values on one or more included controls. Additional controls include age, gender, race, marital status, education, union status, reason for displacement, and tenure and earnings at lost job. Industries are categorized into 14 categories. Year refers to year of displacement.

the spell of unemployment, I bin the data into 12-week intervals.[20] In which case, the first 12 weeks correspond to the first period wherein we let the structural hazard vary with the length of the notice period. Panel B in Table 1.2 shows that workers with more than two months of notice are 9 percentage points less likely to be unemployed for more than 12 weeks. Figure 1.1 presents the survival and job-finding rate for both lengths of notice. Around 57% of individuals with the long-notice exit within the first 12 weeks. On the other hand, only 48% of those with shorter notice periods do so. However, over

[20]See Figure A.2 for presentation of data with alternative binning definitions.

FIGURE 1.1: SURVIVAL AND EXIT RATE – LATER IN THE SPELL



(A) SURVIVAL RATE

(B) EXIT RATE

Notes: 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 proportion of individuals who are unemployed at the beginning of each interval. Panel B presents the proportion of individuals exiting unemployment in each interval amongst those who were still unemployed at the beginning of the interval.

the spell of unemployment, individuals with shorter notice catch up, and the survival rate for both groups is almost identical by 48 weeks of being unemployed. As we can see from Panel B, at all durations after 12 weeks, the observed job-finding rate is greater for individuals with the shorter notice.

If the composition of both groups is identical at the time of layoff and a longer notice does not reduce an individual worker's exit probability at later durations, the higher observed exit rate after 12 weeks for the short-notice group points towards heterogeneity. Suppose all workers are identical in their odds of finding a job, then the observed exit rate for short-notice workers should not be higher than that of long-notice workers. On the other hand, if workers differ significantly in their employability or ability to find a job, the more employable workers find a job faster. Now, if we look at workers who have been unemployed for three months, workers in the long-notice group have been searching for about five months. In contrast, workers in the short-notice group have been searching for about four. Then it follows that amongst workers unemployed for three months, those who got a longer notice are increasingly the less employable workers.

However, assuming that the composition of long- and short-notice workers is identical at the time of layoff and that a longer notice does not reduce an individual worker's exit probability at later durations are both strong assumptions. We already saw that workers with two lengths of notice differ on several observable characteristics. In the empirical implementation, I will not only account for these observable differences but also allow for unobservable differences across the two groups. Moreover, I will also allow the length of notice to impact an individual's exit probability even beyond the first 12 weeks.

1.4 Estimation

Generalized Method of Moments (GMM). We can now use the identification result given in Theorem 1 to construct a consistent estimator for the structural hazards and the moments of the distribution of unobserved heterogeneity using Generalized Method of Moments (GMM). The model is identified up to a scale and so I normalize the mean of the heterogeneity distribution and set $\mu_1 = 1$. For each individual i define the following moment condition:

$$m_i(l,d,\Theta) = \mathbb{I}\{L_i = l\}\left[\mathbb{I}\{D_i = d\} - g(d|l;\Theta)\right]$$

where $\Theta = \{\{\psi_l(1)\}_{l=1}^{J}, \{\psi(d)\}_{d=2}^{\bar{D}}, \{\mu_k\}_{k=2}^{\bar{D}}\}$ and $g(d|l;\Theta)$ is given by the right-hand side of eq. (1.1) under Assumptions 1 to 3 and $\mu_1 = 1$. Now stack moment conditions pertaining to different lengths of notice and durations in one vector, $m_i(\Theta) = \{\{m_i(l,d,\Theta)\}_{d=1}^{\bar{D}}\}_{l=1}^{J}$. Since eq. (1.1) holds, $\mathbb{E}[m_i(\Theta)] = 0$. Note that, $m_i(\Theta)$ contains $J \times \bar{D}$ moment conditions and Θ contains $2 \times \bar{D}$ parameters. As shown by Theorem 1, as long as $J > 1$, our parameters of interest are identified from the moment conditions.

To construct the GMM estimator, note that, the corresponding sample average for $\mathbb{E}[m_i(\Theta)]$ can be written as:

$$\hat{m}(\Theta) = \frac{1}{N}\sum_{i=1}^{N} m_i(\Theta) = \{\pi_l[\hat{g}(d|l) - g(d|l;\Theta)]\}_{d=1}^{\bar{D}}\}_{l=1}^{J}$$

where $\pi_l = N_l/N$, N_l is the number of individuals with length l notice in the sample, N is total number of individuals in the sample, and $\hat{g}(d|l)$ is the sample counterpart of the duration distribution. Then the GMM estimator $\hat{\Theta}$ is given by,

$$\hat{\Theta} = \arg\max_{\Theta} \hat{m}(\Theta)' \hat{W} \hat{m}(\Theta)$$

When the model is just-identified, $\hat{W}$ is given by the identity matrix. When the model is over-identified, the efficient weighting matrix is given by $\hat{W} = \left[\frac{1}{N}\sum_{i=1}^{N} m_i(\hat{\Theta}) m_i(\hat{\Theta})'\right]^{-1}$

and we can compute $\hat{\Theta}$ using two-step estimation. Asymptotic distribution of this estimator is given $\sqrt{N}(\hat{\Theta}-\Theta) \to N(0,(\hat{M}'\hat{\Omega}^{-1}\hat{M})^{-1})$ where $\hat{M} = \partial\hat{m}(\hat{\Theta})/\partial\Theta$.

We can construct a GMM estimator from data on right-censored spells in a similar fashion. In particular, the sample moments for estimation will now pertain to the distribution of observed durations conditional on the censoring time being greater than some $\bar{D}$. And the identified parameters are $\Theta^{\Delta} = \{\{\psi_l(1)\}_{l=1}^{J}, \{\psi(d)\}_{d=2}^{\bar{D}}, \{\mathbb{E}(\nu^k|D_C > \bar{D})\}_{k=1}^{\bar{D}}\}$.[21] Similarly, we can also construct a GMM estimator for all the parameters conditional on some observable characteristics X by further conditioning the distribution of observed durations. However, to be able to calculate this conditional distribution, $G^{\Delta}(d|D_C > \bar{D}, l, X)$, for any X and l, we need a considerable number of individuals in each cell. For this reason, I only use discrete-variables or sub-classifications of continuous variables for estimation. In addition, when appropriate, I report the average estimate for the parameters over different values of X.

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

$$\psi(d) = \frac{(\alpha_2/\alpha_1)(d/\alpha_1)^{\alpha_2-1}}{1+(d/\alpha_1)^{\alpha_2}} \tag{1.2}$$

where $\alpha_1 > 0, \alpha_2 > 0$. The hazard function in eq. (1.2) is monotonically decreasing when $\alpha_2 \leq 1$ and is unimodal, initially increasing and subsequently decreasing, when $\alpha_2 > 1$.

[21] See Section 1.2.3 for details.

The mode or the turning point is at $\alpha_1(\alpha_2-1)^{1/\alpha_2}$. This provides a flexible parametrization for the structural hazard relative to other commonly used parametrization, such as Weibull or Gompertz, as it allows the structural hazard to be non-monotonic. However, I will also present estimates with alternative functional form restrictions and non-parametric estimates.

1.5 Duration Dependence and Heterogeneity

1.5.1 Baseline Estimates

Table 1.3 presents the baseline estimates. Since we normalized the mean of the heterogeneity distribution to be equal to one in the first period, the estimated individual job-finding rate during the first 12 weeks for short- and long- notice individuals exactly coincide with their corresponding observed job-finding rate in the data. The estimation corresponds to four periods so we can back out up to the fourth moment of the heterogeneity distribution. Estimate for the second moment points towards substantial heterogeneity in job-finding rates across individuals with a variance of 0.16 for the initial period. Given that the third and fourth moment are imprecisely estimated, we cannot make additional inferences about the shape of the heterogeneity distribution.

The last two rows in panel A of Table 1.3 present estimates for parameters of structural duration dependence as specified in eq. (1.2). Structural hazard implied by these parameter values is presented in panel A of Figure 1.2 along with the observed hazard from the data. The estimated hazard is above the observed throughout the spell of

TABLE 1.3: ESTIMATION RESULTS

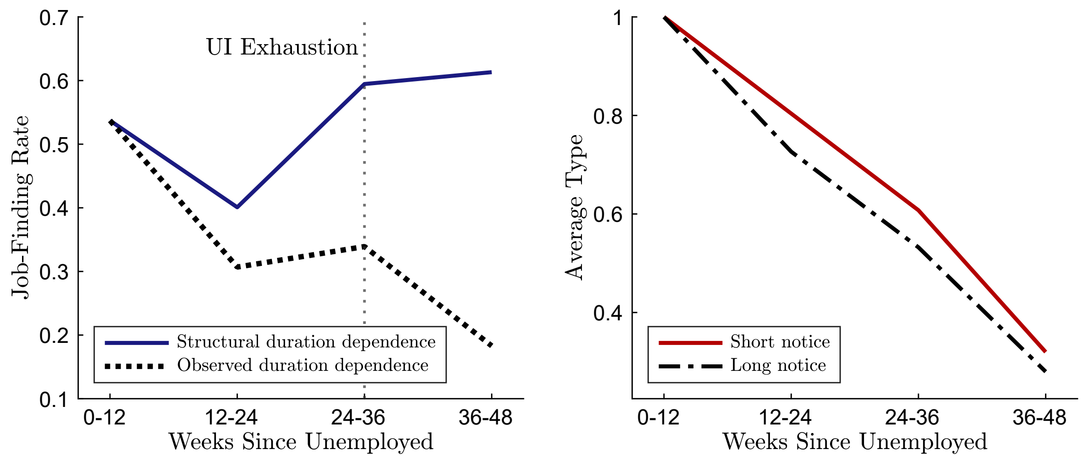
Parameter	Explanation	Estimate	SE
Panel A: Estimated Parameters			
$\psi_S(1)$	Job-finding rate 0-12 weeks: Short notice	0.49	0.01
$\psi_L(1)$	Job-finding rate 0-12 weeks: Long notice	0.58	0.01
μ_2	Second moment of type distribution	1.20	0.10
μ_3	Third moment of type distribution	1.44	0.51
μ_4	Fourth moment of type distribution	1.63	1.84
α	Scale parameter for $\psi(d)$	2.82	0.19
β	Shape parameter for $\psi(d)$	3.24	0.51
Panel B: Duration Dependence			
$\psi(1)$	Structural hazard: 0-12 weeks	0.54	0.01
$\psi(2)$	Structural hazard: 12-24 weeks	0.40	0.06
$\psi(3)$	Structural hazard: 24-36 weeks	0.59	0.11
$\psi(4)$	Structural hazard: 36-48 weeks	0.61	0.12
Hansen-Sargan Test			
	Test statistic: 0.37	Critical value, $df = 1, \chi^2_{0.05}$: 3.84	

Notes: The table shows estimation results from the Mixed Hazard model specified in Section 1.2. Estimation assumes that structural duration dependence is specified by eq. (1.2).

unemployment, reflecting the role of dynamic sorting. The estimates imply that the structural hazard declines from 0-12 to 12-24 weeks by over 25%. However, going from 12-24 to 24-36 weeks structural hazard increases and remains constant thereafter. This is in contrast to the observed hazard which decreases by 40% from 12-24 to 36-48 weeks. The results suggest that most of the depreciation in the observed job-finding rate from 12-48 weeks is due to dynamic sorting instead of having a structural explanation.

Panel (b) of Figure 1.2 shows the implied average type for individuals with short- and long-notice over the spell of unemployment. For individuals with both lengths of notice, the average-type declines over the spell of unemployment. However, for individuals

FIGURE 1.2: BASELINE ESTIMATES



(A) STRUCTURAL HAZARD

(B) AVERAGE TYPE

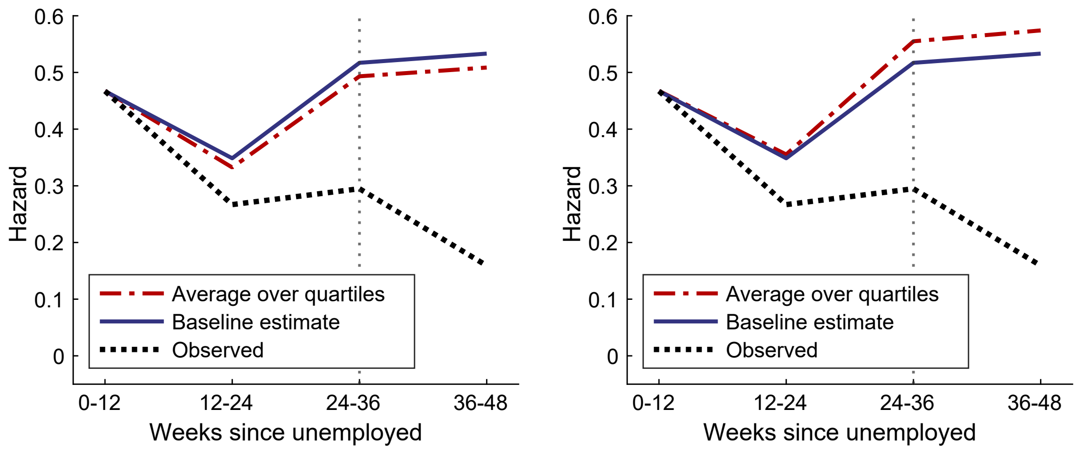
Notes: Solid line in panel A presents the estimates for the structural hazard as implied by the estimated parameters in Table 1.3. Dotted line in panel A presents the observed job-finding rate from the data. Panel B presents the implied average-type at each duration for those with short and long notice.

with longer notice length, the composition worsens more from 0-12 to 12-24 weeks, reflecting that the job-finding rate in the initial period is greater for those with longer notice. By the end of 36 weeks, since a large number of individuals have already left the unemployment, there is little difference in the average-type across both groups.

1.5.2 Robustness

In this section, I assess the robustness of the main results to alternative functional form restrictions on the baseline hazard, using alternative moments for estimation, accounting for observables in estimation, and relaxing the assumptions of the baseline model.

Figure 1.3: Controlling for Observables



(a) Within Age Quartiles

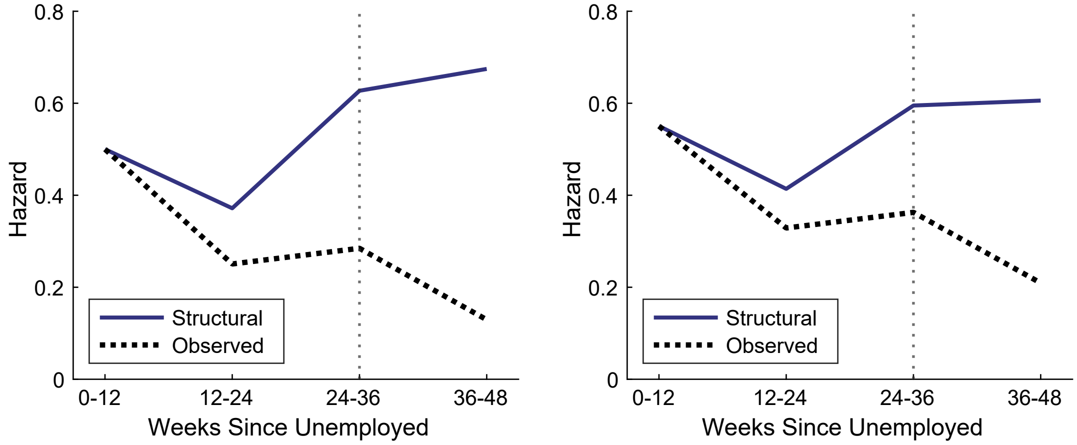
(b) Within Tenure Quartiles

Notes: Panel A presents the average of estimates for the structural hazard from the Mixed Hazard model estimated separately for each quartile of the tenure distribution. Panel B presents the corresponding average for age quartiles. Standard errors are presented in Table A.6.

As we saw in Section 1.2, the structural hazard is non-parametrically identified. However, I impose a log-logistic functional form on the hazard to minimize the number of estimated parameters. Figure A.5 shows that non-parametric estimates for the structural hazard move in a similar fashion as the baseline estimates over the spell of unemployment. Figure A.2 presents estimates from data binned in 9 weeks instead of 12 weeks. Once again, the structural hazard increases more than the observed hazard up until UI exhaustion and remains constant thereafter.

From Table 1.1 we saw that individuals with longer notice are older and have had longer tenure at their previous jobs. If seniority impacts the direction of duration dependence, then our estimates could be biased. For this reason, I re-estimate the model separately within each quartile of the tenure and age distribution. Panel A and

FIGURE 1.4: DURATION DEPENDENCE OVER THE BUSINESS CYCLE



(A) RECESSION YEARS (B) NORMAL YEARS

Notes: Solid line in panel A presents the estimates for the structural hazard from the Mixed Hazard model for the years 2000-2001 and 2007-2009. Dotted line in panel A presents the observed job-finding rate observed in the data for these years. Panel B presents the corresponding estimates and data for the remaining years in the sample. Standard errors are presented in Table A.5.

B of Figure 1.3 presents the average of estimated structural hazards in each quartile of tenure and age distribution, respectively. We can see that even after accounting for seniority, the estimates for structural duration dependence remain unchanged. In Figure 1.4, I present estimates separately for individuals laid off during recessions and those laid off otherwise. We can see that qualitatively both the estimates for structural duration dependence are similar. However, for individuals laid-off during recessions, the individual job-finding rate increases slightly even after 24-36 weeks. This probably reflects the fact that extended benefits during downturns push the UI exhaustion further in the spell.[22] I also reestimate the model excluding individuals who were laid off as a

[22]Figure A.3, shows that this is true in the data.

result of a plant closure. Estimates are once again qualitatively similar and are presented in Figure A.6.

In A.4, I relax the assumptions of my model in two dimensions. First, I allow the mean of the heterogeneity distribution to be different for workers with different notice lengths. Second, I allow the structural hazards beyond the initial period to vary for workers with different lengths of notice up to some constant. The results, presented in Figure A.9, point towards no mean differences and the structural hazard being slightly greater for individuals with a longer notice even beyond the first 12 weeks.

1.6 A Model of Job-Search

Estimates from the Mixed Hazard model imply that an individual worker's job finding rate declines over the first 5 months. This is consistent with evidence from audit studies (Kroft et al., 2013) that document decline in callback rates over the spell of unemployment. I also document that an individual's exit probability increases up to UI exhaustion. Both of these findings together, imply that an individual's job-finding rate is determined by the actions of optimizing agents reacting to their changing payoffs over the spell of unemployment, as well as, external factors that may directly influence a worker's employment prospects. Moreover, I do not find evidence for an individual's job-finding rate to decline after UI exhaustion. This is in contrast to the observed job-finding rate which continues to decline even after exhaustion. Researchers have tried to explain the decline in the observed rate after exhaustion using behavioral explanations (Boone and van Ours, 2012; DellaVigna et al., 2017, 2020). In this section, I set up and calibrate a search model without any behavioral modifications but with heterogeneous workers. I show that my empirical findings can be rationalized in this framework and are broadly consistent with evidence on callback rates from Kroft et al. (2013).

1.6.1 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.[23] 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:

$$\lambda(d,\nu,s)=\delta(d)\nu s$$

The offer arrival rate $\delta(d)\nu$ varies over the duration of unemployment as well as across workers of different types. Once a worker finds a job, they become employed and remain employed until the end. A worker receives unemployment insurance (UI) benefits $b(d)$ when unemployed and wages w when employed. The flow utility from consumption is

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

given by the function u. Then the value function for a worker of type ν unemployed at duration d is given by:

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

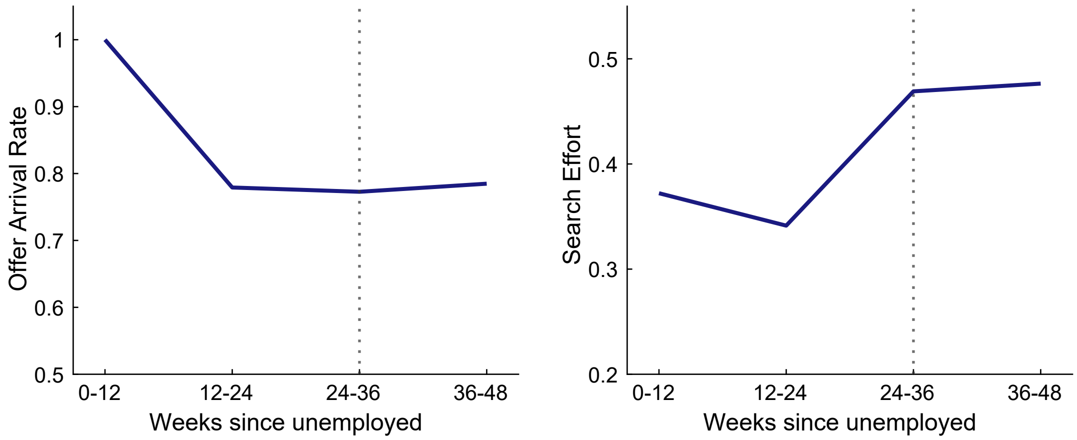
Here, β is the discount rate and V_e is the value of employment given by $V_e = u(w) + \beta V_e$. The UI benefits $b(d)$ are equal to b for $d \leq D_B$ and equal to 0 otherwise. I also assume that after some time $D_T \geq D_B$ the job-finding function $\lambda(d,s,\nu)$ does not depend on the duration of unemployment d, such that for $d > D_T$, $\delta(d) = \delta_T$. This ensures that after D_T, job seekers 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 a high arrival rate.

1.6.2 Numerical Analysis

I now calibrate the model to separate out the role of worker's optimizing behavior and external forces that change dynamically over the spell of unemployment. 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 $\delta(d)\nu s(d,\nu)$. So a worker's job-finding 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 $\tilde{h}(d) = \mathbb{E}[h(d|\nu)|D \geq d]$ also changes due to changes in composition over the spell of unemployment. Since

FIGURE 1.5: CALIBRATION



(A) OFFER ARRIVAL RATE

(B) SEARCH EFFORT

Notes: The figure presents search effort and offer arrival rate from the calibration of the search model. Search effort is averaged over two types of workers.

we already have an estimate for how an individual's job-finding rate evolves over the spell of unemployment that is unaffected by compositional changes, I use this estimate to target structural duration dependence $h(d) = \mathbb{E}[h(d|\nu)]$ from the model.[24] Further details for the calibration are provided in Appendix A.5.1.

Figure 1.5 presents the search effort and offer arrival rate as implied by the calibration. The offer arrival rate declines during the first 5 months and is constant thereafter. An individual's search effort declines slightly during the first 5 months but then increases up to UI exhaustion and remains constant thereafter. The initial decline in the offer arrival rate potentially reflects the decline in callback rates over the spell of unemployment as

[24]Note that the search model does not correspond exactly 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 A.5.2, I simulate data from the search model with notice periods and show that my estimator captures movements in $\mathbb{E}[h(d|\nu)]$ well.

documented by Kroft et al. (2013).

1.7 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 job-finding 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, at later durations, the observed exit rate is lower for these workers. 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.

Estimates from the hazard model uncover substantial heterogeneity in individual job-finding probabilities. The observed job-finding rate declines by over 40% over the first 5 months; in contrast, the estimated individual hazard only declines by 25% over this period. Moreover, I find that after the first 5 months, none of the depreciation in the observed job-finding rate is due to structural duration dependence. Instead, an individual's job-finding probability increases up to UI exhaustion and remains constant after. The observed job-finding rate continues to decline after exhaustion as well, which has led researchers to suggest behavioral explanations for this pattern. I provide a much simpler explanation for this pattern which is the presence of heterogeneity. I show that my estimates can be rationalized within a standard search model with heterogeneous workers. My findings point towards importance of accounting for heterogeneity while estimating and calibrating search models.

CHAPTER 2

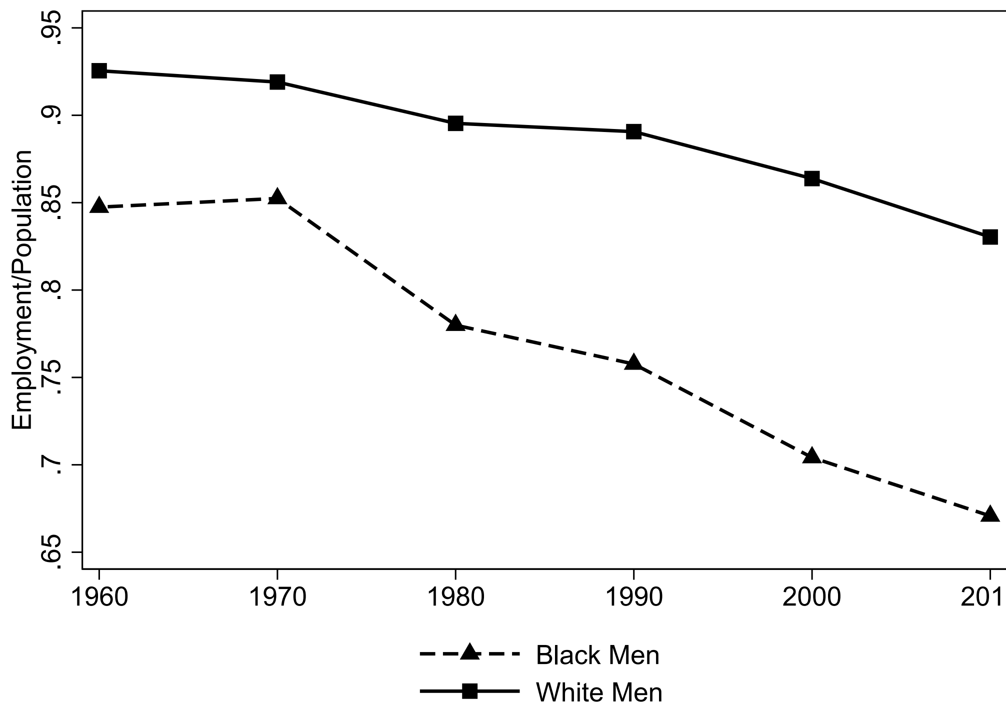
UNDERSTANDING THE RACIAL EMPLOYMENT GAP: THE ROLE OF SECTORAL SHIFTS

2.1 Introduction

Despite apparent progress in civil rights for African American individuals in the United States, the labor market outcomes of Black men deteriorated significantly compared to White men in the last three decades of the 20th century. As Figure 2.1 shows, employment rates for both groups of workers declined during this period; however, the decline for Black men was more pronounced. One proposed explanation for this phenomenon is that sectoral reallocation of economic activity over this period was especially detrimental for Black men.[1] This disparate impact could stem from Black workers being more exposed to sectoral shifts due to being overrepresented in declining sectors or being excessively located in areas with a concentration of declining sectors. However, it is also possible that adjustment to changes in labor demand differs by race,

[1]This explanation was first proposed by Wilson (1987). Bound and Freeman (1992) and Bound and Holzer (1993) link the decline in the manufacturing sector to lower wages and employment for Black relative to White men in the 1970s and 1980s. Also, see Boustan (2017) for a recent discussion.

FIGURE 2.1: EMPLOYMENT-TO-POPULATION RATIOS FOR BLACK AND WHITE MEN



Notes: Data are from the United States Census, accessed through Integrated Public Use Micro Samples (IPUMS) (Ruggles et al., 2021). Employment and population is calculated for Black and White men between the ages of 25-55 who are not in the armed forces and do not reside in institutionalized group quarters.

say, due to differences in ability to move away from affected sectors or locations.

In this paper, we quantify the extent to which sectoral reallocation can explain the divergence in employment outcomes of Black and White men over 1970-2000.[2] In addition, we assess the degree to which this effect is driven by differences in exposure vs. differences in adjustment to sectoral shifts across two groups of workers. In order to accomplish these goals, we first exploit regional variation in exposure to sectoral shifts

[2]We focus on employment as a share of the total population for each group to avoid selection problems involved with analyzing other outcomes such as wages due to differences in non-participation across groups and over time.

for each group to uncover differential patterns of labor market adjustment. In particular, we estimate the race-specific elasticity of the local employment-to-population ratio and population with respect to local labor demand shifts. We then provide a framework that enables us to aggregate the local employment responses while accounting for population movements across locations and investigate the contribution of various margins of adjustment.

We use data from the Census Integrated Public Use Micro Samples (IPUMS) and measure changes in our outcome variables over 1970-2000 at the level of commuting zones (CZs). To measure changes in local labor demand, we create a *shift-share* measure by combining local employment shares by sectors in 1960 and national changes in sectoral employment over 1970-2000. We document that the local employment-to-population ratio as well as population vary positively with local labor demand shifts for both groups of workers. We find an elasticity of 0.2 for Black men relative to 0.05-0.1 for White men for the employment-to-population ratio. We find substantial population responses; the estimated elasticity of population for both Black and White workers is around 1. This is in line with the previous literature that finds sizable population responses to shifts in local labor demand (Blanchard and Katz, 1992; Glaeser et al., 1995; Amior and Manning, 2018).

The large population responses highlight the importance of accounting for such movements when considering the aggregate impact of sectoral shifts on employment. In order to do so, one needs to impose some structure on how shifts in labor demand in one region affect the population in other regions. To this end, we set up a stylized model of local labor markets that incorporates matching frictions within markets and mobility frictions across markets. The model implies that changes in the local employment-to-population ratio and population in response to local demand shifts depend on parameters that govern these frictions, which can vary across different groups of workers. Hence, the parameters of the model can be inferred from our estimated elasticities. Further, we show that aggregating these local responses implies that the change in aggregate employment depends on how employment changes *within* local markets and how population adjusts *across* markets for each group. The contribution of both of these margins of adjustment to the aggregate employment gap depends on the initial composition of workers across locations and sectors, which determines how exposed workers of a particular group are to demand shifts.

Combining the estimated elasticities with our framework, we find that sectoral reallocation can explain up to half of the increase in the Black-White employment gap on aggregate. We find that sectoral shifts cannot explain any of the decline in employment for White men over 1970-2000. While labor demand shifts do decrease employment for White men within labor markets, individuals are able to mitigate the impact of these shifts by moving out of harder hit locations into less hard hit locations. We find that a 2 percentage point decrease in the employment-population ratio for Black workers can be attributed to sectoral shifts. This is because, while the population responses across labor markets play out similarly for both groups of workers, the decrease in employment within labor markets for Black men is three-fold that of White men. Furthermore, we provide an intuitive and comprehensive decomposition for understanding these effects and apply it to our sample data. Our findings indicate that the increase in the Black-White employment gap due to sectoral shifts results from the greater responsiveness of Black workers to local labor demand shifts rather than a higher concentration of these shifts in areas or sectors with a higher share of Black workers.

This paper is the first to our knowledge to provide a framework to formally aggregate local effects and decompose the overall response into differential exposure vs. differential response by group in the context of the Black-White employment gap. Existing empirical studies exploit variation across regions in exposure to sectoral shifts and document that employment for Black men changes by more relative to White men in response to changes in local labor demand (Bound and Holzer, 2000; Batistich and Bond, 2019; Gould, 2020). While these studies provide robust evidence at the local level on how different racial groups are affected by changes in labor demand, a growing literature spanning the macro, labor, trade, and urban fields has shown that cross-regional estimates are only informative about aggregate responses under specific assumptions.[3] Our work builds on this growing literature and we argue that fully evaluating the effects of local labor demand shocks in the aggregate requires accounting explicitly for migration responses as well as regional employment composition, which we take into account in our framework.

[3]See Chodorow-Reich (2020) for a recent review in the context of the macroeconomics literature. See Adão et al. (2019) for a detailed discussion on recovering the general equilibrium effects of economic shocks from cross-regional estimates.

Our analysis indicates that sectoral shifts that played out differently across different regions explain a large portion of the widening in the Black-White employment gap since 1970. An extensive literature on the wage gap between Black and White workers has traditionally attributed a portion of this differential to observable characteristics of workers and explained the remainder through theories of labor market discrimination.[4] However, given that over the 20th century, measures of racial prejudice declined steadily (Lang and Lehmann, 2012) and the skills gap converged (Card and Krueger, 1992; Neal, 2005), this approach has trouble capturing the persistence of economic disparities by race. Two recent studies by Bayer and Charles (2018) and Hurst et al. (2021) focus on slowed convergence in the earnings and wage gaps, respectively, since the 1970s-1980s and propose that the effects of decreased discrimination and increased educational attainment among Blacks were offset by increasing returns to certain types of skills that disproportionately benefited White relative to Black individuals. Our paper is complementary to these studies in the sense that we consider a different dimension of changes in labor demand, sectoral, rather than skill specific, but share their emphasis on explaining stalled racial progress over this time horizon. Furthermore, while these authors consider trends at the level of aggregation of skill or occupation, we leverage the spatial dimension of the sectoral reallocation patterns that occurred during the 20th century by using local labor markets as our unit of analysis.

The rest of the paper is structured as follows. Section 2.2 presents our empirical

[4] Early studies include Smith and Welch (1977) and Brown (Brown). See Lang and Lehmann (2012) for a recent review.

analysis. In this section, we provide details on the methodology we use to estimate the relationship between local labor demand shifts and employment and population outcomes as well as the associated data sources we use to do so. In Section 2.3, we outline a model of labor market frictions and regional mobility that delivers key predictions about how employment and population respond to changes in labor demand for different groups. We show how the parameters in the model can be recovered from our estimates in the previous section. Section 2.4 derives a framework to analyze the aggregate effect of local labor demand shifts on the employment-to-population ratio gap between Black and White workers and decomposes the relevant margins of adjustment. Section 2.5 provides a discussion of the importance of such a framework as well as suggestions to researchers interested in aggregating local responses to labor demand shifts. Section 2.6 concludes.

2.2 Data and Empirical Analysis

In this section we first describe the data used in our analysis. We then outline how we measure local labor demand shifts and present their geographical distribution. Finally, we document employment and population responses to local labor demand shifts separately for Black and White men over the period of our analysis.

2.2.1 Data Description

We study changes in local employment and population in response to changes in local labor demand over 1970-2000 separately for Black and White men. We select this period for our analysis as it is over this period that the employment outcomes of Black and White men diverged considerably.[5] For the purpose of our analysis, we need a measure for changes in labor demand over this period at the level of local labor markets. To construct this measure, we use local employment shares by sectors in 1960 and national changes in sectoral employment over 1970-2000.[6] We use long differences from 1970-2000 in our analysis instead of separately analyzing changes over each decade to avoid conflating short- and long-run responses (Jaeger et al., 2018).

All the variables used in our analysis are constructed using Census Integrated Public Use Micro Samples (IPUMS) (Ruggles et al., 2021) for the years 1960, 1970, and 2000.[7] Census data is particularly suited to our analysis as its large sample size enables us to conduct detailed regional analysis separately by race. We use commuting zones (CZs) as our measure of local labor markets. Commuting zones were developed by Tolbert and Sizer (1996) who used county-level commuting data from the 1990 Census data to create 741 clusters of counties based on the strength of commuting ties across counties. We use measures of geography provided in the Census data, in particular, public use microdata areas (PUMAs) for years 1960 and 2000 and county groups for the year 1970, to match Census data to commuting zones. We are able to obtain a consistent sample

[5]See fig. 2.1. We only focus on men due to massive changes in labor force participation of women over this period which makes analysis for women more complicated.

[6]For seminal papers using this approach to measure local labor demand shocks, see Bartik (1991) and Blanchard and Katz (1992). Large number of recent studies in international trade have used shift-share measures to document the impact of trade shocks, for instance, see Autor et al. (2013); Topalova (2010); Dix-Carneiro and Kovak (2017) among others.

[7]In particular, we use the following samples: 1960 5%, 1970 1% Form 1 Metro, 1970 1% Form 2 Metro, and 2000 5% sample.

of 728 commuting zones for each year used in our analysis.[8]

We measure all outcomes based on men between ages 25-55 who are not in the armed forces and do not reside in institutional group quarters. We restrict our sample to commuting zones that had at least 250 Black men who satisfied this criteria and were employed in 1960. This is to ensure a large enough sample size for calculating race-specific outcomes at the level of commuting zones. Our final sample includes 328 commuting zones. The restriction mainly excludes commuting zones in the Mid-Western plains and Rocky mountains. Commuting zones in our sample accounted for 87% of the total employment in 2000. Table 2.1 presents summary statistics for our sample as well as for all commuting zones in 1970 and 2000. Since we restrict our sample on size of employment for Black men, the commuting zones in our sample are on an average larger and have a higher share of employment in manufacturing. However, the average commuting zones in our sample is comparable to the average commuting zone in the United States in terms of employment and wages.

2.2.2 Measurement of Local Labor Demand Shifts

We are interested in quantifying how local employment and population respond to changes in local labor demand. However, shifts in labor demand are not directly observable. Therefore we use employment shares by sectors in 1960 and national changes

[8]Census data is matched to commuting zones using crosswalk files provided on David Dorn's website https://www.ddorn.net/data.htm for years 1970 and 2000. Crosswalk for 1960 was obtained from Evan K. Rose's website https://ekrose.github.io/resources/.

TABLE 2.1: SUMMARY STATISTICS FOR COMMUTING ZONES

	All		Sample	
	1970	2000	1970	2000
	(1)	(2)	(3)	(4)
Log of Population	9.44	9.93	10.52	11.06
Share of Manufacturing	0.23	0.21	0.28	0.24
Employment Rate: Black Men	0.82	0.71	0.84	0.70
Employment Rate: White Men	0.92	0.86	0.92	0.86
Log of Wages: Black Men	9.87	9.91	9.78	9.90
Log of Wages: White Men	10.30	10.24	10.33	10.31
Observations	741	741	328	328

Notes: All statistics are calculated based on non-institutionalized civilian men between the ages of 25-55. Columns (1) and (2) present averages of variables across all commuting zones. While, columns (3) and (4) present averages for commuting zones that had at least 250 non-institutionalized civilian Black men between the ages of 25-55 who were employed in 1960. Employment Rate is calculated by diving total race-specific employment by population in each commuting zone. Wages refers to total wage and salary income of employed workers.

in sectoral employment over 1970-2000 to create a proxy for these shifts. In particular, we construct a proxy for changes in local labor demand, denoted by ΔA_{lg}, as follows:

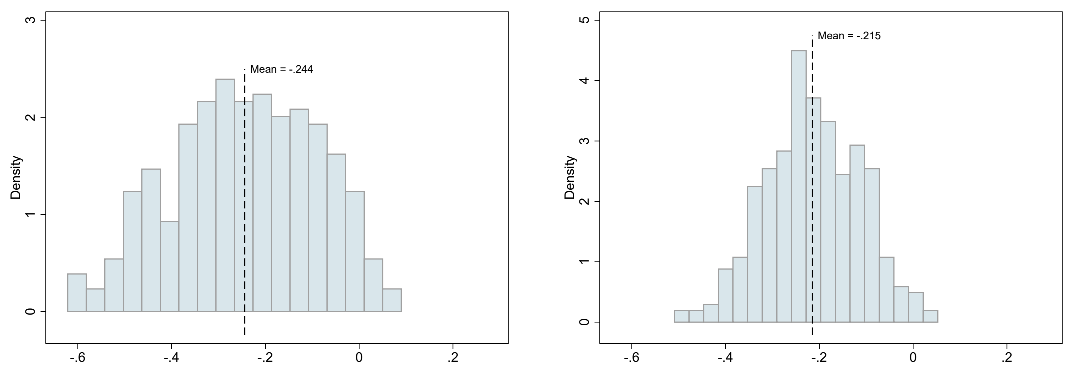
$$\Delta A_{lg} = \sum_{s} \frac{L_{lsg,1960}}{L_{lg,1960}} \cdot \left(\ln(L_{s,2000}) - \ln(L_{s,1970})\right) \tag{2.1}$$

Here, $L_{s,y}$ represents total employment in industry s in year y and $L_{lsg,y}$ represents employment in commuting zone l for group g in sector s in year y. Figure B.1 presents changes in sectoral employment over 1970-2000 for all 33 industries in our analysis.[9]

Our measured demand shifts will be more negative in locations with a larger initial

[9]We use time consistent industry codes at the three digit level from IPUMS contained in the variable `ind1990`. IPUMS constructs these codes from underlying Census Bureau industrial classifications to reflect the same broad set of industries across different samples. Based on these, we construct a broad classify industries into 33 broad categories.

FIGURE 2.2: DISTRIBUTION OF LOCAL LABOR DEMAND SHIFTS



(A) BLACK MEN (B) WHITE MEN

Notes: Figure shows the histogram for our proxies for labor demand shifts for Black and White workers. The proxy is constructed using employment shares by sectors in 1960 and national changes in sectoral employment over 1970-2000 according to eq. (2.1).

share of employment in declining industries. Also, note that we allow the employment shares that determine the weight each sector gets in the overall measure to be race-specific. Suppose for instance, that the demand for durable goods goes down. If 30% of Black (B) men in Detroit-Flint, MI are employed in durable goods manufacturing vs. 20% of White (W) men, then our measured labor demand shift will be more negative for Black workers in that location. In other words, our proxy implicitly takes into account differential initial sectoral composition of Black and White workers. Figure 2.2 shows the distribution of local labor demand shifts ΔA_{lg} for Black and White workers. Given similar means of $\Delta A_{l,B}$ and $\Delta A_{l,W}$, it appears that both groups of worker's average exposure to broad shifts in labor demand was of similar magnitude. However, Black workers faced a larger dispersion of demand shifts across regions, as the standard deviation of $\Delta A_{l,B}$ is roughly 3 times that of $\Delta A_{l,W}$.

We also map the geographic dispersion of our proxy for local labor demand shifts for

Black and for White men in fig. 2.3. It is apparent that there is a higher concentration of large, negative labor demand shifts in regions of the United States that were hardest hit by the decline of the manufacturing sector during the end of the 20th century. In the maps, areas such as the Rust Belt, Appalachia, and the North East display larger negative labor demand shifts. Though our proxies are designed to capture shifts in labor demand across *all* industries, the decline of the manufacturing sector in the U.S. represents a large change in employment demand during the time period under consideration and therefore appears prominently in these measures.

2.2.3 Employment and Population Responses

We estimate the following equations to measure the elasticity of local employment and population with respect to local labor demand shifts:

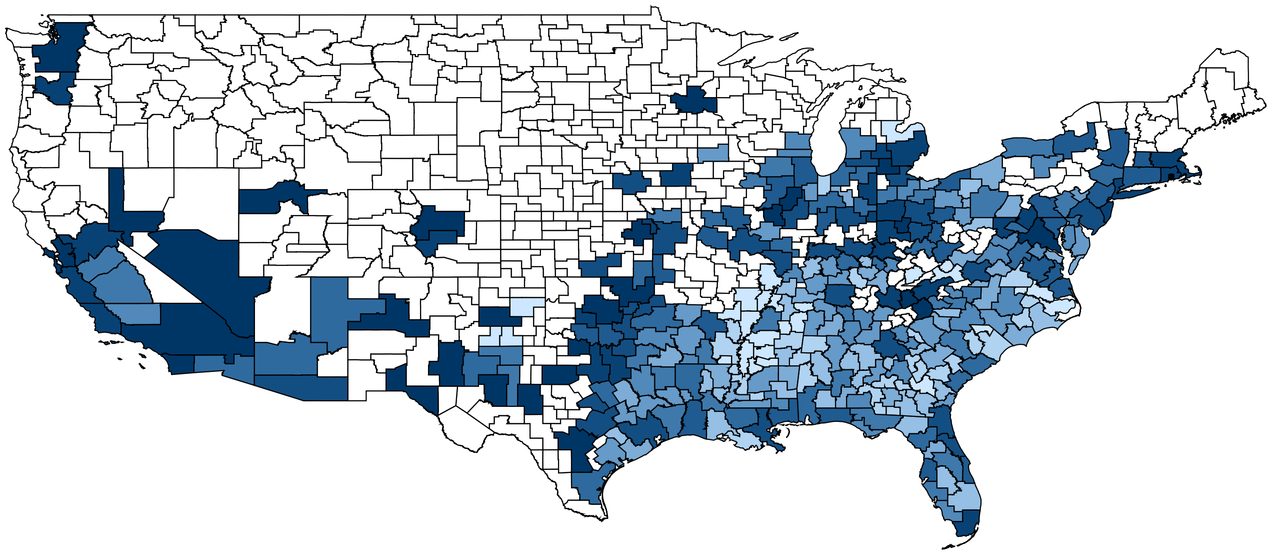
$$\Delta \ln l_{lg} = \delta_{1g} + \beta_{1g}\Delta A_{lg} + \zeta_{1g} X_{lg} + \varepsilon_{lg} \tag{2.2}$$

$$\Delta \ln P_{lg} = \delta_{2g} + \beta_{2g}\Delta A_{lg} + \zeta_{2g} X_{lg} + \nu_{lg} \tag{2.3}$$

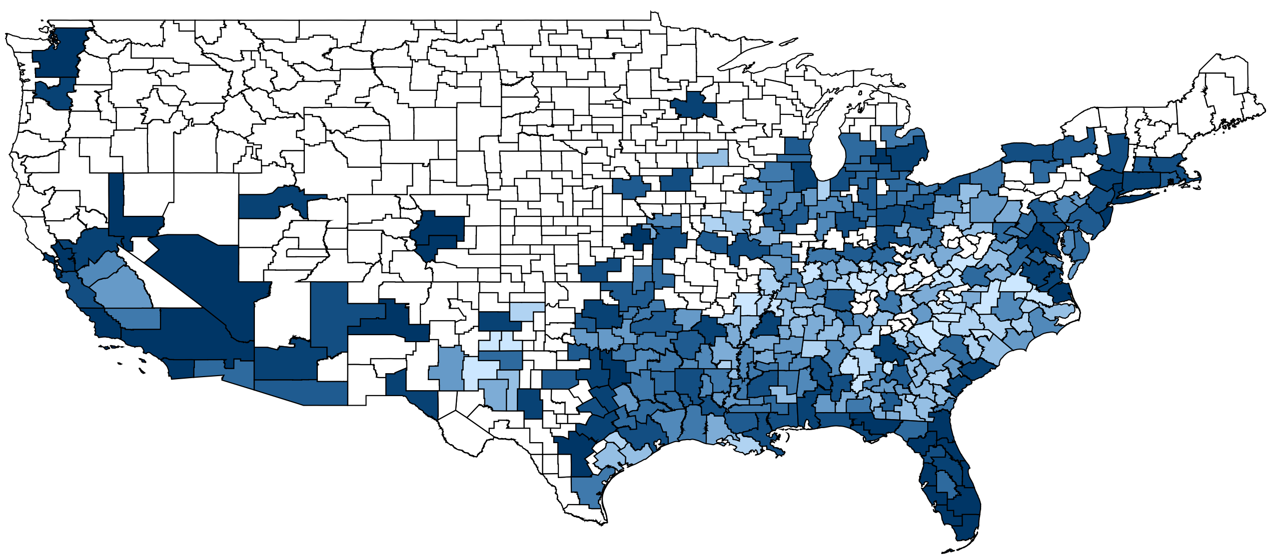
In these equations, $\Delta \ln l_{lg}$ represents the change in log of employment-to-population ratio for group g in commuting zone l from 1970 to 2000. While, $\Delta \ln P_{lg}$ represents the corresponding change in log of population. X_{lg} contains controls that capture the potential effects of demographic composition across commuting zones and may also vary across groups. δ_{1g} and δ_{2g} are intercept terms that we estimate separately for each racial group. Under the identifying assumption that the sectoral composition of

FIGURE 2.3: GEOGRAPHICAL EXPOSURE TO SECTORAL SHIFTS

(A) BLACK MEN



(B) WHITE MEN



Notes: Maps show the geographic distribution of our proxy for local labor demand shifts across the commuting zones in our sample for Black and White workers. The proxy is constructed using employment shares by sectors in 1960 and national changes in sectoral employment over 1970-2000 according to eq. (2.1). Darker shaded areas represent lower values and lighter shaded areas represent higher values. Areas that are not shaded are not included in our sample, as we only include commuting zones which had at least 250 non-institutionalized civilian Black men between the ages of 25-55 who were employed in 1960. Our sample accounted for 87% of total employment in 2000.

TABLE 2.2: EMPLOYMENT AND POPULATION RESPONSES TO LABOR DEMAND SHIFTS OVER 1970-2000

	(1)	(2)	(3)
Panel A: Black Workers			
Outcome: log change in employment rate			
Labor Demand Shifts, $\hat{A}_{l,B}$	0.20***	0.22***	0.21***
	(0.04)	(0.05)	(0.05)
Outcome: log change in population			
Labor Demand Shifts, $\hat{A}_{l,B}$	0.97***	0.95***	0.81**
	(0.20)	(0.27)	(0.30)
Panel B: White Workers			
Outcome: log change in employment rate			
Labor Demand Shifts, $\hat{A}_{l,W}$	0.05**	0.08***	0.10***
	(0.02)	(0.02)	(0.03)
Outcome: log change in population			
Labor Demand Shifts, $\hat{A}_{l,W}$	1.01**	1.48**	1.18***
	(0.47)	(0.58)	(0.38)
Additional Controls		X	X
State Fixed Effects			X
Number of Observations	328	328	328

Notes: Sample is restricted to commuting zones that had at least 250 noninstitutionalized civilian Black men between the ages of 25-55 who were employed in 1960. Table reports results from linear regressions of log differences in employment-to-population ratio and population over 1970-2000 on our measure of local labor demand shifts. Panel A shows the results for Black workers, while Panel B shows the results for White workers. Labor demand shifts for each group are constructed using employment shares of the respective group by sectors in 1960 and national changes in sectoral employment over 1970-2000 according to eq. (2.1). Additional controls include the share of incarcerated population and educational composition for each group as well as the share of foreign-born population in 1970. Robust standard errors, clustered at state-level are in the parentheses. $^{*}p \leq 0.10$;$^{**}p \leq 0.05$;$^{***}p \leq 0.01$

employment for each group in 1960 is uncorrelated with other changes at the local level over 1970-2000 that might impact employment or induce migration, we will obtain unbiased and consistent estimates for β_{1g} and β_{2g} (Goldsmith-Pinkham et al., 2020).

Table 2.2 shows the results of estimating eqs. (2.2) and (2.3) separately for Black and White workers. We obtain statistically significant and economically meaningful estimates of the degree to which changes in employment rates and population shares over 1970-2000 vary with local demand shifts. In general, positive changes in local demand induce positive responses in both employment rates and population shares. Moreover, our results are robust to various additional controls for initial demographic composition as well as fixed effects at the state level.

Column (1) of the table shows our results when we include no controls. Column (2) shows our results when we control for the respective share of incarcerated population, the respective share of the population with a college degree, the respective share of the population who dropped out of high school, and the share of foreign-born population in 1970. Column (3) shows our results when we include these additional controls as well as state fixed effects. The estimates are fairly stable across specifications, which gives us confidence that our measures of local labor demand shifts capture sectoral reallocation patterns over this time period that are unrelated to the initial demographic composition of different regions.

As we can see from the coefficients in the first and third rows, local employment of Black men during this time period was more sensitive to changes in local labor demand than that of White men. Estimates imply that in response to the measured labor demand shifts, for an average commuting zone in our sample, the employment-to-population ratio for Black men decreased by 4 percentage points relative to a 1 percentage point decrease for White men. This reflects previous findings in the literature that have argued that although de-industrialization in the 1970s and 1980s produced negative consequences for the average worker, it was relatively more damaging to labor market outcomes of Black workers.

From the second and fourth rows, we can see that both Black and White men out-migrate in response to declining labor demand in local markets. In addition, the coefficients on the changes in population are economically as well as statistically similar for both groups of workers. Reduced-form population responses can inform us about how population in a local area changes due to changes in labor demand in that area. However, local populations are also affected by changes in labor demand in other areas. In order to fully capture population responses due to simultaneous changes in labor demand in several areas, we need to impose some structure on how individuals move across locations. We do so in the next section.

2.3 A Model of Labor Market Frictions

In this section, we present a simple model of local labor markets that incorporates matching frictions within markets and mobility frictions across markets. The model is based on Kim and Vogel (2021) and allows us to characterize the relationship between local labor demand and employment outcomes for different racial groups. We show that structural parameters of our model can be inferred from the elasticities estimated in the previous section.

2.3.1 Model Setup

The economy consists of K local labor markets indexed by l. Workers belong to different, non-overlapping groups indexed by g. The total number of workers in the population for each group, denoted by P_g, is fixed. There are P_{lg} workers of group g and V_{lg} vacancies for workers of group g in local labor market l. If employed, a worker belonging to group g in location l produces flow output A_{lg}.[10] Workers choose a location based on their expected utility and search for employment opportunities in that location. Firms decide how many vacancies to post for each group in each location to maximize their profits. Employment for group g in location l, L_{lg}, is determined by the labor market tightness, $\theta_{lg} \equiv V_{lg}/L_{lg}$. Hence, markets are effectively indexed by location-group pairs lg. Both firms and workers are risk neutral.

Worker's location choice

Workers search for employment in the location that provides them the highest expected utility. We will assume that the expected utility for a worker belonging to group

[10]We assume that workers of different groups have different productivity within each local labor market to account for differences in composition across sectors within each location.

g from searching in location l is given by:

$$u_{lgi} = w_{lg} l_{lg} \varepsilon_{lgi}$$

where w_{lg} and l_{lg}, respectively, represent the wage and job-finding probability in location l for workers belonging to group g. ε_{lgi} represents an idiosyncratic utility component which captures individual-specific preferences for living in l. The cumulative density function for $\{\varepsilon_{lgi}\}_{l=1}^{K}$ is given by:

$$F_g(\varepsilon_1, ..., \varepsilon_K) = \exp\left(-\sum_{l=1}^{K} \varepsilon_l^{-1/\kappa_g}\right)$$

As will become clear in Section 2.3.4, the parameter κ_g governs the elasticity of labor supply across labor markets and can be interpreted as capturing costs of moving across locations.

Vacancy Posting

Firms incur a vacancy posting cost in each market F_{lg}. We assume free entry such that firms post vacancies until their profits from a new vacancy are zero. Therefore, total vacancies in each market are determined by the following condition which equates marginal costs to marginal benefits of vacancy posting:

$$(A_{lg} - w_{lg}) q_{lg} = F_{lg} \tag{2.4}$$

where q_{lg} denotes the probability of filling a vacancy for a firm. Note that, $A_{lg} - w_{lg}$ is the net gain for the firm from filling a vacancy.

Matching and Wage Setting

We assume that labor markets are frictional and the total number of matches in any market is determined by a Cobb-Douglas matching function as follows:

$$L_{lg} = m(V_{lg}, P_{lg}) = \gamma_{lg} V_{lg}^{\alpha_g} P_{lg}^{1-\alpha_g} \tag{2.5}$$

where $\gamma_{lg} > 0$ represents the efficiency of the matching technology and $\alpha_g \in (0,1)$ is the elasticity of matching with respect to vacancies. Given the above matching technology, the job-finding rate for a worker of group g in location l is given by:

$$l_{lg} = \frac{L_{lg}}{P_{lg}} = \gamma_{lg} \theta_{lg}^{\alpha_g}$$

The employment rate for workers of the group with a higher α_g is relatively more sensitive to market tightness. The match technology also determines the probability of filling a vacancy, $q_{lg} = L_{lg}/V_{lg} = \gamma_{lg} \theta_{lg}^{-(1-\alpha_g)}$.

Finally, we will assume that wages are determined by Nash Bargaining such that:

$$w_{lg} = \omega_{lg} A_{lg} \tag{2.6}$$

where ω_{lg} represents the bargaining power of workers.

2.3.2 Equilibrium Outcomes

We now discuss the derivation of the equilibrium. From eqs. (2.4) to (2.6) we can solve for equilibrium labor market tightness to obtain:

$$\theta^*_{lg} = \left[\frac{(1-\omega_{lg})\gamma_{lg}}{F_{lg}}\right]^{\frac{1}{1-\alpha_g}} A_{lg}^{\frac{1}{1-\alpha_g}}$$

This implies that the equilibrium job finding probability is given by,

$$l^*_{lg} = \gamma_{lg}\theta^{*\alpha_g}_{lg} = \gamma_{lg}\left[\frac{(1-\omega_{lg})\gamma_{lg}}{F_{lg}}\right]^{\frac{\alpha_g}{1-\alpha_g}} A_{lg}^{\frac{\alpha_g}{1-\alpha_g}} \tag{2.7}$$

Note that this expression depends only on parameter values and flow output in a market A_{lg}, which we take to be exogenously given. Then, combining these results with workers' location choices allows us to derive a simple expression for population shares in each location.

Proposition 2. *For each group g, the share of workers who work in location l is given by:*

$$\frac{P^*_{lg}}{P_g} = \frac{\tilde{c}_{lg} A_{lg}^{\frac{1}{\kappa_g(1-\alpha_g)}}}{\sum_l \tilde{c}_{lg} A_{lg}^{\frac{1}{\kappa_g(1-\alpha_g)}}} \tag{2.8}$$

where $\tilde{c}_{lg} = \left(\omega_{lg}\gamma_{lg}\left[\frac{(1-\omega_{lg})\gamma_{lg}}{F_{lg}}\right]^{\frac{\alpha_g}{1-\alpha_g}}\right)^{1/\kappa_g}$.

(Proofs of all propositions are provided in the Appendix.)

This result allows us to also write simple, closed form expressions for population shares that depend solely on parameter values and flow output A_{lg}.

2.3.3 Comparative Statics

We next derive comparative statics to illustrate how changes in local labor demand induce changes in employment rates and population shares depending on the strength of labor market frictions and mobility costs. Our model predicts that equilibrium responses of employment and population shares for each group depend in a simple way on the parameters that govern the degree to which each group faces locational and matching frictions.

Proposition 3. *For some variable x_{lg}, denote $\hat{x}_{lg} = \partial \ln x_{lg}$. The equilibrium responses of employment rates $\hat{l}^*_{lg}$ and population shares $\hat{P}^*_{lg}$ to changes in local labor demand $\hat{A}_{lg}$ are given by*

$$\hat{l}^*_{lg} = \frac{\alpha_g}{1-\alpha_g}\hat{A}_{lg} \tag{2.9}$$

$$\hat{P}^*_{lg} = \frac{1}{\kappa_g(1-\alpha_g)}\left(\hat{A}_{lg} - \sum_l \pi^p_{lg}\hat{A}_{lg}\right) \tag{2.10}$$

where $\pi^p_{lg} = P^*_{lg}/P_g$.

Equation 2.9 shows that if a market lg experiences a negative labor demand shock, employment will decline in a manner proportional to the size of the match elasticity for workers in group g. This is because employment prospects worsen for workers in that market, as firms earn less profit per vacancy and post fewer vacancies as a result of the decline in A_{lg}. The degree to which employment declines depends on the parameter α_g, such that groups with higher α_g experience larger drops in employment for a given shock to labor demand. This is due to the fact that α_g governs the elasticity of the matching function – in other words, how the total number of matches vary with the total number of vacancies posted by firms in a market.

Equation 2.10 shows that in response to a negative shock to labor demand, we expect to see out-migration of workers as they decide to search for better employment opportunities in other areas. This happens so long as the shock in market lg is more negative than the shock in the average market (the second term in the parentheses). The magnitude of out-migration depends on the parameter κ_g; hence, we interpret this parameter as representing costs of migration, since a larger κ_g implies a lower population response, ceteris paribus. We also see that α_g helps determine the response of population shares, since workers who choose to migrate must look for new jobs and will encounter matching frictions in other areas. In this equation, a larger $1-\alpha_g$ corresponds to a higher propensity of being crowded out by other job seekers, since it is the elasticity of matches with respect to population size in the matching function. Hence, groups with a higher $1-\alpha_g$ will be less likely to migrate since they internalize the fact that other workers moving to other markets may crowd them out as well.

TABLE 2.3: STRUCTURAL PARAMETERS

Parameter	Explanation	Estimate	SE
α_B	Match Elasticity: Black men	0.17	(0.025)
α_W	Match Elasticity: White men	0.05	(0.022)
κ_B	Mobility Costs: Black men	1.24	(0.025)
κ_W	Mobility Costs: White men	1.04	(0.022)

Notes: Estimates are for parameters of the model as implied by estimated elasticities reported in Table 2.2. Standard errors (SE) are computed using the delta method.

2.3.4 Structural Parameters

The equations we estimated in section 2.2 are the empirical counterparts of eqs. (2.9) and (2.10) and allow us to recover the structural parameters of the model from the estimated elasticities.[11] Table 2.3 shows the estimates for the structural parameters implied by our results. We can see that α_B is more than double α_W, implying that the match elasticity with respect to vacancy posting is much larger for Black than for White workers. In other words, the employment prospects of Black men over this time horizon were much more sensitive to labor demand shifts to local areas. We can also see that since $\kappa_B > \kappa_W$, mobility costs to re-location are higher for Black workers than for White workers.[12]

[11] Recall that we constructed proxies for labor demand shifts $\hat{A}_{lg}$ using the following formula $\Delta A_{lg} = \sum_s (L_{lsg,1960}/L_{lg,1960}) \cdot (\ln(L_{s,2000}) - \ln(L_{s,1970}))$. One concern with this measure could be that the change in sectoral employment differs in magnitude from the change in sectoral labor demand. However, our proxies will still capture the variation in local labor demand due to sectoral shifts under the condition that the degree to which employment varies with labor demand is the same across sectors.

[12] Suppose instead that workers' decision choices were based only on wages in a local area, rather than on wages and employment probabilities. Then, the mobility cost parameter κ_g would be $1/\beta_{2g}$.

2.4 Aggregation and Decomposition

As we mentioned before, it is not possible to directly infer what happens to aggregate statistics from the local responses documented in the previous section. Our estimates suggest that there are large movements in local employment in response to labor demand shifts, however, workers also out-migrate from harder hit areas. The impact on the aggregate depends on the extent to which population responses are able to mitigate the employment responses. Additionally, the size of harder hit areas relative to the rest of the economy also determines the contribution of local demand shifts to aggregate employment. The next proposition formally states this idea.

Proposition 4. *The change in aggregate employment for group g, L_g, in response to shifts in local labor demand can be written as follows:*

$$\hat{L}_g = \frac{\alpha_g}{1-\alpha_g}\sum_l \pi_{lg}\hat{A}_{lg} - \frac{1}{\kappa_g(1-\alpha_g)}\sum_l \left(\frac{l_g - l_{lg}}{l_{lg}}\right)\pi_{lg}\hat{A}_{lg} \tag{2.11}$$

From this proposition, we can see that the change in aggregate employment for each group can be written as the sum of two terms — the first term captures the *employment* response while the second captures the *population* response. The magnitude of the employment response for group g depends on the overall intensity of the shock for that group, the weighted average of $\hat{A}_{lg}$, and the responsiveness of group g's local employment to changes in local labor demand captured by α_g. Note that, the intensity of a negative shock can be higher for one group if local shifts are more negative for this group in all locations or most affected locations have a higher share of employment for that group.[13]

Workers can mitigate the impact of negative shifts by moving away from affected areas. The second term in eq. (2.11) illustrates that population responses can mitigate employment responses only to the extent that there are locations with very different employment opportunities. For instance, if all locations were identical in terms of probability of finding employment, such that $l_{lg} = l_g$, workers moving from one location to another won't gain much and we would see that the contribution of the population response to aggregate employment would be zero. The population response, in addition, also depends on the elasticities governing the extent of labor market frictions, α_g and κ_g.

2.4.1 Aggregation Exercise

We now combine estimates of the structural parameters presented in table 2.3 with data on regional employment shares in 1960 and our measure of $\hat{A}_{lg}$ over 1970-2000 to document the change in aggregate employment for Black and White men over this period that can be explained by our measured labor demand shifts. Table 2.4 presents the results from this exercise. We report the actual as well as the predicted changes in employment-to-population ratio for Black and White men over 1970-2000.

[13] Given the way we measure $\hat{A}_{lg}$, the reason shifts are different across locations is due to initial sectoral composition. So differences in intensity of shifts in this framework result from differences in initial composition of employment across sectors and locations.

Table 2.4: Actual and Predicted Changes in Employment-to-Population Ratios

	Employment-to-Population Ratio			Log change (1970-2000)	
	1970	2000	Predicted	Actual	Predicted
	(1)	(2)	(3)	(4)	(5)
Black Men	0.82	0.71	0.80	-0.137	-0.024
White Men	0.88	0.81	0.88	-0.080	0.000
Gap: Black-White	-0.06	-0.10	-0.08	-0.058	-0.024

Notes: Sample is restricted to commuting zones that had at least 250 non-institutionalized civilian Black men between the ages of 25-55 who were employed in 1960. Columns (1) and (2) show actual values in our sample for 1970 and 2000, respectively. Column (3) shows the predicted value for 2000 under the counterfactual scenario where only sectoral reallocation affects the employment-to-population ratio. Column (4) shows the log difference between the values in Columns (2) and (1), while Column (5) shows the log difference between the values in Columns (3) and (1). The last row shows the difference between the values in the preceding two rows.

From table 2.4, we can see that our measured labor demand shifts capture little of the decline in the employment-to-population ratio within groups, but does a reasonable job explaining the increase in the gap between Black and White workers. We can see that over 1970-2000, the employment-to-population ratio fell from 0.88 to 0.81 for White workers and from 0.82 to 0.71 for Black workers. Our measured shifts cannot account for any of the change in the employment-to-population ratio for White workers, however, they do account for a decrease of 2 percentage points for Black workers. Consequently, sectoral reallocation helps us explain a significant proportion of the divergence in the employment-to-population ratio between these groups of workers. Over 1970 to 2000, the actual gap between Black and White workers widened from -0.06 to -0.10, whereas our results imply that labor demand shifts caused it to widen to -0.08, 50% of the observed increase.

TABLE 2.5: EMPLOYMENT RESPONSE VS. POPULATION RESPONSE

	Employment Response (1)	Population Response (2)	Total Response (3)
Black Men	-0.031	0.006	-0.024
White Men	-0.006	0.006	0.000

2.4.2 Decomposition Exercise

In order to further understand what forces are driving changes in the employment-to-population ratio, we next decompose the aggregate change in employment separately into the population and employment response. This decomposition is presented in table 2.5.

We can see from the table that the reason why our model predicts no change in the employment-to-population ratio for White workers is that the employment response and population response exactly offset. White workers were negatively affected by labor demand shifts in terms of employment, but in aggregate, were able to mitigate the negative impact of these shifts by moving to less affected areas. Hence, our framework implies a total response of zero for White men over 1970-2000. This decomposition also allows us to understand where the negative effect on the employment-to-population ratio for Black workers is coming from. Black workers experienced an employment response that was much larger in magnitude than White workers and a population response of similar magnitude. However, the population response was not large enough to offset the large employment response for Black workers, leading to an overall negative impact on aggregate.

Next, we investigate what accounts for the larger employment response of Black workers. Conceptually, it may stem from two sources: greater exposure to labor demand shifts or larger response to a similar set of shifts. In order to formalize the contributions of these margins, we present results from the following two decompositions:

$$\text{(i)}.\ \ \hat{L}_B - \hat{L}_W = \sum_l \pi_{lB}\hat{A}_{lB} \underbrace{\left(\frac{\alpha_B}{1-\alpha_B} - \frac{\alpha_W}{1-\alpha_W}\right)}_{\text{Differential Response}} + \tfrac{\alpha_W}{1-\alpha_W} \underbrace{\sum_l \left(\pi_{lB}\hat{A}_{lB} - \pi_{lW}\hat{A}_{lW}\right)}_{\text{Differential Exposure}}$$

$$\text{(ii)}.\ \ \hat{L}_B - \hat{L}_W = \sum_l \pi_{lW}\hat{A}_{lW} \underbrace{\left(\frac{\alpha_B}{1-\alpha_B} - \frac{\alpha_W}{1-\alpha_W}\right)}_{\text{Differential Response}} + \tfrac{\alpha_B}{1-\alpha_B} \underbrace{\sum_l \left(\pi_{lB}\hat{A}_{lB} - \pi_{lW}\hat{A}_{lW}\right)}_{\text{Differential Exposure}}$$

These equations show two similar ways of breaking down the evolution of the employment gap.[14] In each, we can see that changes in the employment gap between Black and White workers may be attributed to either a higher response of a certain group to a given set of labor demand shifts *(Differential Response)* or a higher incidence of labor demand shifts for a certain group of workers *(Differential Exposure)*. Notice that the Differential Response term depends on the structural parameters that govern the responsiveness of employment for each group to changes in labor demand. The Differential Exposure term depends on the average size of labor demand shifts experienced by each group as well as the composition of employment across regions for each group.

[14]These equations can be derived by writing out eq. (2.11) separately for Black and White workers and taking the difference.

TABLE 2.6: DIFFERENTIAL RESPONSE VS. DIFFERENTIAL EXPOSURE

	Differential Response (1)	Differential Exposure (2)
(i)	-0.023	-0.002
(ii)	-0.017	-0.007

table 2.6 shows the results of this exercise. In each case, we can see that the Differential Response term dominates the Differential Exposure term in magnitude (note that the only difference between (i) and (ii) are the weights used). Hence, what explains the employment response across groups is the fact that for Black workers, employment drops by much more after a local labor demand shock of a given magnitude. We do not find much evidence that local labor demand shifts were simply more heavily concentrated in areas that had a higher composition of Black employment.

2.5 Discussion

In this section, we highlight the importance of accounting for population responses when aggregating local employment responses. Not doing so would lead to overestimating the impact of a set of economic shocks on aggregate employment, depending on the extent to which population responses mitigate the impact of these shocks. The extent of overestimation will depend on whether changes in local employment-to-population ratios or local employment is the dependent variable. To illustrate this, consider the following. Let $\hat{x}$ denote the log change in outcome x following some shock(s) to the fundamentals of the economy. Then, for group g in location l, the log change in employment, $\hat{L}_{lg}$, is given by the sum of the log change in employment-to-population ratio, $\hat{l}_{lg}$, and population, $\hat{P}_{lg}$:

$$\hat{L}_{lg} = \hat{l}_{lg} + \hat{P}_{lg}$$

Next, the log change in aggregate employment for group g, $\hat{L}_g$, can be written as:

$$\hat{L}_g = \sum_l \pi_{lg} \hat{L}_{lg}$$

where $\pi_{lg} = L_{lg}/L_g$ is the employment share in location l for group g. Furthermore, changes in population for group g in location l can be written as a sum of two components:

$$\hat{P}_{lg} = \hat{P}_{lg}^{PE} + \hat{P}_{lg}^{GE}$$

Here $\hat{P}_{lg}^{PE}$ is the Partial Equilibrium (PE) term that represents changes in population that are in response to shocks affecting location l. On the other hand, $\hat{P}_{lg}^{GE}$ is the General Equilibrium (GE) term capturing changes in population in location l that are due to shocks in other locations. This term arises because shocks to other location induce migration responses into or out of location l. Separately regressing log changes in employment-to-population ratio and population on shocks at the local level recovers $\hat{l}_{lg}$ and $\hat{P}_{lg}^{PE}$, respectively. Regressing log changes in local employment on shocks at the local level recovers $\hat{L}_{lg} = \hat{l}_{lg} + \hat{P}_{lg}^{PE}$.[15] Note that neither approach estimates the $\hat{P}_{lg}^{GE}$

[15] This is a version of the so-called "missing intercept" problem — local regressions are unable to recover aggregate general equilibrium impact of a shock when regions are spatially correlated. See Adão et al. (2019) for a brief discussion of different approaches to deal with this problem.

term, which is a crucial component in the aggregate group response.

In our approach, we impose additional structure on population movements in order to express $\hat{P}_{lg}^{GE}$ as a function of the estimated elasticities from our local regressions and the observed shocks.[16] This enables us to back out the aggregate impact of sectoral shifts. Without this additional structure, one has two options for aggregating local responses: (1) estimate and sum up $\hat{l}_{lg}$ across locations or (2) estimate and sum up $\hat{L}_{lg}$ across locations. We emphasize that while both of these will lead us to inexact conclusions about the aggregate impact, the former is better than latter. This is because changes in local employment-to-population ratios capture changes in employment net of population movements. Although, this approach still ignores the indirect adjustment in the likelihood of finding employment in a region due to shocks in other regions.

However, estimating and aggregating changes in local employment (instead of the local employment *rate*) results in erroneously finding a very large aggregate elasticity of employment with respect to shocks. This is because that procedure incorporates the partial equilibrium population term, $\hat{P}_{lg}^{PE}$, while ignoring the general equilibrium term, $\hat{P}_{lg}^{GE}$. In other words, population (and hence employment) in more negatively affected regions is allowed to decline, while ignoring that this would lead to an increase in population (and hence employment) in less affected regions.[17] Therefore our suggestion for other researchers interested in aggregating employment from the local responses

[16]In particular, according to our model $\hat{P}_{lg}^{PE} = \beta_{2g}\hat{A}_{lg}$ and $\hat{P}_{lg}^{GE} = -\beta_{2g}\sum_{l'}\pi_{lg}^{p}\hat{A}_{l'g}$, see eq. (2.10).

[17]One exception to this would be if individuals moving to other locations do not find jobs upon moving. However, this is ruled out empirically by our estimated positive elasticity of employment-to-population ratios with respect to local labor demand shocks.

is to use employment-to-population ratios and acknowledge the estimates as upper bounds on true employment effects.

2.6 Conclusion

We provide a simple framework to assess the degree to which sectoral shifts in labor demand played a role in the widening of the employment gap between Black and White workers over 1970-2000. We find that sectoral reallocation can explain about half of the increase in the gap between the employment-to-population ratio for Black men and White men, but little of the evolution of the employment-to-population ratios for these groups individually over this period. Mirroring other results in the literature, we find that employment for Black workers is more responsive to changes in labor demand. Furthermore, most of the increase in the employment gap can be attributed to the differential response of Black workers to local labor demand shifts, rather than a higher incidence of shifts to sectors or regions in which Black workers are overrepresented. In future work, we plan to further investigate what gives rise to these differential responses across groups.

Our results show that sectoral shifts were an important factor behind the decrease in Black relative to White employment over the second half of the 20th century, after several decades of convergence. The decline in the manufacturing sector, which some authors have argued hit Black workers harder than White workers, is an important example of these structural changes. However, they explain only about half of the evolution of the Black-White employment gap and fail to account for why both Black and White workers experienced a decline in their employment-to-population over this time period. In decomposing these effects, we highlight the importance of different margins of labor market adjustment arising from both imperfect regional mobility and frictional labor markets that may have differential effects across groups of workers. Our findings suggest that future studies should pay close attention to the aggregation of local effects, which may be offset by migration patterns, and look for other sources of divergence in employment opportunities for Black and White individuals in recent years.

Chapter 3

The Discrete-Time Mixed Proportional Hazard Model

3.1 Introduction

Hazard models of event durations are widely employed in the field of economics to analyze unemployment spells, retirement decisions, and an array of other topics. Consequently, the estimation of duration models has been the subject of significant research in econometrics. It is now well-known that ignoring unobserved heterogeneity in duration data models can lead to erroneous inferences. To account for unobserved heterogeneity, Lancaster (1979) introduced the Mixed Proportional Hazard (MPH) model, a generalization of Cox (1972)'s proportional hazard model. In the MPH model, the hazard rate is specified as the product of a regression function that captures the effect of observed explanatory variables, a baseline hazard that captures variation in the hazard over the spell, and a random variable that accounts for the unobserved heterogeneity.

The shape of the baseline hazard and the distribution of the unobserved heterogeneity are nonparametrically identified in the MPH model (Elbers and Ridder, 1982;

Heckman and Singer, 1984). However, applied researchers often impose parametric assumptions on one or both of these components of the model due to the lack of a convenient nonparametric estimator. Unfortunately, such parametric assumptions are not innocuous in this setting, and the estimates of the baseline hazard can be quite sensitive to the specification of the heterogeneity distribution. In this paper, I propose an easily implementable nonparametric estimator for the discrete-time version of the MPH model. Similar to the literature on the continuous-time MPH model, the identification of the baseline hazard in this paper relies on exogenous variation in some observable characteristic. However, I assume that this variable is discrete. Discrete exogenous variables are frequently employed in empirical work in economics. Moreover, we can relax the exogeneity assumption to selection on observables with a discrete variable as in the treatment effects literature.

Related Literature

Several papers have proposed semi-parametric estimators for the continuous-time MPH model. In seminal work, Heckman and Singer (1984) consider the nonparametric maximum likelihood estimator (NPMLE) of the MPH model with a parametric baseline hazard and regression function. However, the speed of convergence and the asymptotic distribution of this estimator is not known. Baker and Melino (2000) provide Monte Carlo evidence on the behavior of the NPMLE. They find that the NPMLE gives biased estimates of all the parameters in the MPH model if the baseline hazard is left relatively free.

Han and Hausman (1990) and Meyer (1990) propose an estimator for the MPH model that allows a flexible baseline hazard but assumes gamma heterogeneity. Hausman and Woutersen (2014) present simulations and a theoretical result to show that using a nonparametric estimator of the baseline hazard with gamma heterogeneity yields inconsistent estimates for all parameters and functions if the true mixing distribution is not a gamma.

Horowitz (1999) was the first to propose a nonparametric estimator for both the baseline hazard and the distribution of the unobserved heterogeneity. However, Hahn (1994) shows that his model cannot be estimated at the rate $\sqrt{N}$, where N is the sample size. Ishwaran (1996) shows the fastest possible rate at which this model can be estimated is $N^{2/5}$, and the rate of convergence of Horowitz (1999)'s estimator is arbitrarily close to this rate. Hausman and Wouterson (2014) propose an estimator for the MPH model that allows for a nonparametric baseline hazard and discrete measurement of durations. In their model, the estimator of the integrated baseline hazard rate converges at the regular rate, $\sqrt{N}$. They avoid estimating the unobserved heterogeneity distribution. Their estimator relies on a modified version of Kendall's rank correlation and can be computationally burdensome. In contrast, my estimator for the discrete-time MPH model is easy to implement and uses the Generalized Method of Moments (GMM) Hansen (1982). Modeling durations as discrete allows the model to be expressed as a set of moment conditions which can then be used to construct an estimator using GMM.

There are two other papers that setup the discrete-time MPH model. van den Berg and van Ours (1996) set up and estimate such a model using seemingly unrelated nonlinear regression (SUNR) on aggregate data. However, they do not study the properties of their estimator. Finally, Alvarez et al. (2021) develop an estimator for the discrete-time MPH model with multiple spells using GMM.

The rest of this paper is organized as follows. In Section 3.2, I formulate the discrete-time MPH model and discuss it's identification. Monte Carlo evidence on the behavior of my proposed estimator is presented in Section 3.3. Finally, Section 3.4 concludes.

3.2 Model Setup

In the MPH model the hazard is a function of an exogenous regressor X, unobserved heterogeneity ν, and the baseline hazard $\psi(t)$:[1]

$$h(t|X,\nu) = \psi(t)\phi(X)\nu$$

In discrete-time, this implies the following survival probabilities:

$$P(T \geq t|X,\nu) = S(t|X,\nu) = \prod_{s=1}^{t-1} \big(1-\psi(s)\phi(X)\nu\big)$$

$$P(T \geq t|X) = \mathbb{E}[S(t|X,\nu)] = \mathbb{E}\left[\prod_{s=1}^{t-1} \big(1-\psi(s)\phi(X)\nu\big)\right]$$

[1] $\phi(X)$ is often specified as $e^{X\beta}$. This is not necessary here and we will consider a fully non-parametric model.

where $\mathbb{E}[.]$ denotes expectation with respect to ν. Then the probability distribution function is given by:

$$g(t|X) = \mathbb{E}[S(t-1|X,\nu)] - \mathbb{E}[S(t|X,\nu)] = \psi(t)\phi(X)\mathbb{E}[\nu S(t|X,\nu)] \tag{3.1}$$

Let $g_l(t) = g(t|X = l)$ and $\mu_k = \mathbb{E}[\nu^k]$. Then expanding eq. (3.1) for $t = 1, 2, 3, \ldots$

$$\begin{aligned}
g_l(1) &= \psi(1)\phi(l)\mu_1 \\
g_l(2) &= \psi(2)\phi(l)[\mu_1 - \psi(1)\phi(l)\mu_2] \\
g_l(3) &= \psi(3)\phi(l)[\mu_1 - \psi(1)\phi(l)\mu_2 - \psi(2)\phi(l)\mu_2 + \psi(1)\psi(2)\phi(l)^2\mu_3] \\
&\vdots
\end{aligned}$$

More generally we can write,

$$g_l(t) = \psi(t)\sum_{k=1}^{d}\phi(l)^k c_k(t)\mu_k$$

where $c_1(1) = 1$ and for $1 < k \leq t$,

$$c_k(t) = c_k(t-1) - \psi(t-1)c_{k-1}(t-1)$$

Note that $c_k(t)$ is function of $\{\psi(1), \psi(2), \ldots, \psi(t-1)\}$.

Now let's say X is a binary variable. Then if we normalize $\phi(0)\mu_1 = 1$ and denote

$\phi(1) = \gamma$, for $t = 1,2$ and $l = 1,0$ we have:

$$g_1(1) = \psi(1)\gamma$$

$$g_0(1) = \psi(1)$$

$$g_1(2) = \psi(2)[\gamma - \gamma^2\psi(1)\lambda_2]$$

$$g_0(2) = \psi(2)[1 - \psi(1)\lambda_2]$$

where $\lambda_k = \mu_k/\mu_1^k$. From the first two equations we can solve for $\psi(1) = g_0(1)$ and $\gamma = g_1(1)/g_0(1)$. We can then plug that in the next two equations to solve for $\psi(2)$ and λ_2. In general, we can extend this logic to solve this system of equations recursively. To see this note that we have two sets of equation for each time period:

$$g_1(t) = \psi(t)\sum_{k=1}^{t} c_k(t)\gamma^k\lambda_k \tag{3.2}$$

$$g_0(t) = \psi(t)\sum_{k=1}^{t} c_k(t)\lambda_k \tag{3.3}$$

Now say, we have already solved for $\{\lambda_1, \lambda_2, ..., \lambda_{t-1}, \psi(1), \psi(2), ..., \psi(t-1)\}$. Then eqs. (3.2) and (3.3) represent two non-linear equations with two unknowns, $\psi(t)$ and λ_t. These unknowns would be locally identified if the gradient is full rank. So the condition for local identification is:

$$\psi(t)c_t(t)\left[\sum_{k=1}^{t} c_k(t)\gamma^k\lambda_k\right] \neq \psi(t)\gamma^t c_t(t)\left[\sum_{k=1}^{t} c_k(t)\lambda_k\right]$$

or more compactly $\mathbb{E}[\nu S_1(t,\nu)] \neq \gamma^t \mathbb{E}[\nu S_0(t,\nu)]$.

Another way to see this is by noting that,

$$g_1(t) - \gamma^t g_0(t) = \psi(t)\left[\sum_{k=1}^{t} c_k(t)\gamma^k \lambda_k - \gamma^t \sum_{k=1}^{t} c_k(t)\lambda_k\right]$$

And so we can solve for $\psi(t)$ as follows:

$$\psi(t) = \frac{g_1(t) - \gamma^t g_0(t)}{\sum_{k=1}^{t-1} c_k(t)\gamma^k \lambda_k - \gamma^t \sum_{k=1}^{t-1} c_k(t)\lambda_k}$$

Note that the above expression only depends on parameters from before period t. Now if $\mathbb{E}[\nu S_1(t,\nu)] \neq \gamma^t \mathbb{E}[\nu S_0(t,\nu)]$ doesn't hold, both the denominator and numerator in this equation will be equal to 0. Similarly, we can also solve for λ_t by taking the ratio of eqs. (3.2) and (3.3) as follows:

$$\lambda_t = \frac{g_0(t)\sum_{k=1}^{t-1} c_k(t)\gamma^k \lambda_k - g_1(t)\sum_{k=1}^{t-1} c_k(t)\lambda_k}{\left(\prod_{s=1}^{t-1}\psi(s)\right)(g_1(t) - \gamma^t g_0(t))}$$

In general it is hard to further simply these expressions for the general t period. However, in the appendix I provide expressions for $t = 2, 3$.

Even though we can solve this system recursively, using an estimator based on recursive substitution will not be efficient. So instead, I propose to use GMM to estimate this system of equations. If we have $\bar{T}$ time periods then with a binary X, we will have $2 \times \bar{T}$ moments and $2 \times \bar{T}$ parameters. So the model will be just identified. In general with a categorical variable with J values, we will have $J \times \bar{T}$ moments and $2(\bar{T}-1)+J$ parameters, and the model will be overidentified. In the next section, I study properties of these estimators using simulations.

3.3 Monte Carlo Experiments

I investigate the performance of these estimators in a variety of settings. These settings vary by the true data generating process (DGP) for the data and the number of sample observations used in estimation. With respect to the DGP, I assume that $\nu \sim \text{Beta}(\alpha, \beta)$ with parameters $\alpha = 2$ and $\beta = 3$. I generate 100 random samples of size 100,000 from this distribution. I consider two cases for the baseline hazard, one where the baseline hazard decreases and the other where it increases. In particular, I assume that the baseline hazard is Weibull:

$$\psi(t) = (b/a)(t/a)^{b-1}$$

I use two sets of values for the hazard, $a = 3, b = 1.25$ and $a = 2, b = 0.75$ and estimate the model for 8 time periods. The beta distribution and the baseline hazards implied by these set of parameter values are presented in Panel (A) of Figure 3.1. For the categorical variable, I consider estimation with two ($J = 2$) and three ($J = 3$) categories where $\phi(1) = 1, \phi(2) = 2, \phi(3) = 1.5$. Panel (B) of Figure 3.1 shows the duration distribution function implied by this DGP. Finally, I also analyze how the sample size affects the estimators by estimating the model on 100,000 observations and the first 10,000 observations.

Results from this exercise are presented in Figure 3.2 and Figure 3.3 for the case where the true hazard is decreasing and increasing, respectively. From Figure 3.2 and Figure 3.3, we can see that the average prediction from simulations follows the true baseline hazard at all durations. The confidence bands are narrower for the cases with larger samples. Surprisingly, going from $J = 2$ to $J = 3$ does not significantly alter the precision of the estimates. Bias and standard error for the other parameters of the model are presented in Table 3.1. These tables show that the γ parameter and the first three moments of the unobserved heterogeneity distribution are precisely estimated. However, the standard errors for the higher moments are too large and increase with the duration.

3.4 Conclusion

Failure to account for unobserved heterogeneity in models for duration data makes the estimated hazard rate decrease more with the duration than the hazard rate of a randomly selected member of the population. In this paper, I study the properties of a new estimator for the discrete-time MPH model that is fully non-parametric. I show that modeling durations as discrete allows the model to be expressed as a set of moment conditions which can then be used to construct an estimator using GMM. In particular, the density at any duration t can be expressed in terms of baseline hazards up to that duration and the first $t-1$ moments of the unobserved heterogeneity distribution. The estimators for the baseline hazards perform well in both small and large samples. However, the higher moments of the unobserved heterogeneity cannot be estimated precisely, even with large samples.

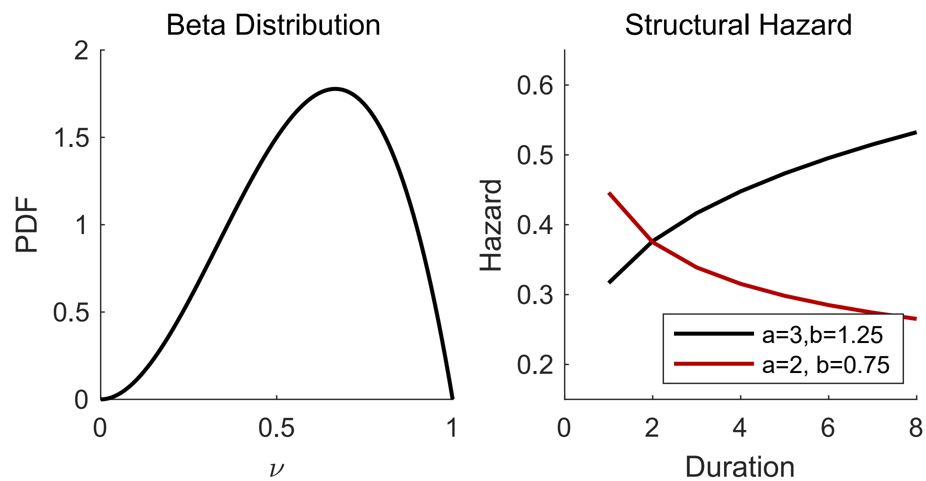
Figures and Tables

TABLE 3.1: OTHER PARAMETERS

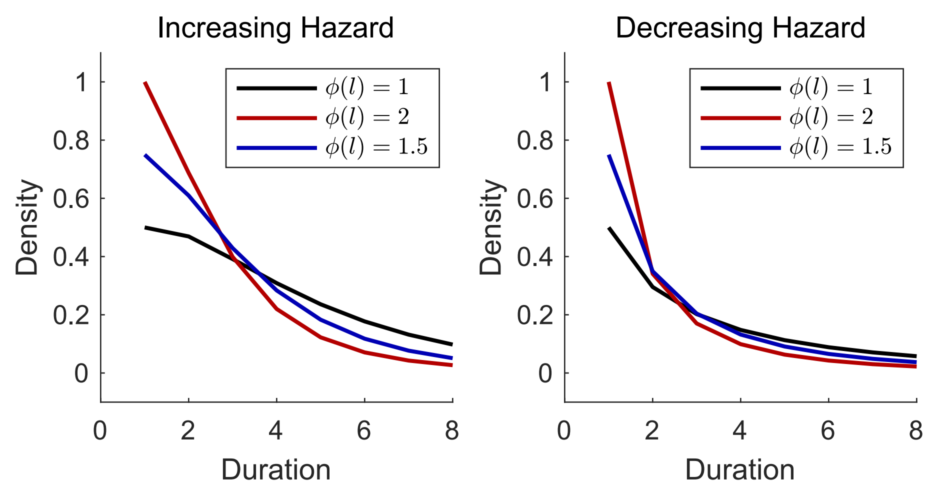
Panel A: True Hazard is Decreasing								
	$J=2$				$J=3$			
	$N=10000$		$N=100000$		$N=10000$		$N=100000$	
Parameter	Bias	SE	Bias	SE	Bias	SE	Bias	SE
γ_1	-0.00	0.05	0.00	0.02	0.01	0.06	0.00	0.02
μ_2	-0.00	0.02	-0.00	0.01	0.00	0.02	0.00	0.01
μ_3	-0.01	0.05	-0.00	0.01	-0.00	0.05	0.00	0.02
μ_4	-0.01	0.11	-0.00	0.03	-0.00	0.10	0.00	0.04
μ_5	-0.03	0.23	-0.01	0.07	-0.01	0.22	0.00	0.08
μ_6	-0.04	0.52	-0.01	0.15	-0.01	0.57	0.00	0.20
μ_7	-0.03	1.38	-0.01	0.39	0.03	1.58	-0.00	0.50
μ_8	0.12	3.90	0.04	1.07	0.20	4.35	-0.02	1.27
Panel B: True Hazard is Increasing								
	$J=2$				$J=3$			
	$N=10000$		$N=100000$		$N=10000$		$N=100000$	
Parameter	Bias	SE	Bias	SE	Bias	SE	Bias	SE
γ_1	0.00	0.06	-0.01	0.02	0.01	0.08	0.00	0.02
μ_2	-0.01	0.04	-0.00	0.01	-0.00	0.04	-0.00	0.01
μ_3	-0.02	0.11	-0.01	0.04	-0.01	0.12	-0.00	0.03
μ_4	-0.04	0.25	-0.02	0.08	-0.03	0.27	-0.01	0.07
μ_5	-0.08	0.52	-0.03	0.14	-0.08	0.56	-0.02	0.14
μ_6	-0.12	1.07	-0.04	0.26	-0.16	1.13	-0.04	0.28
μ_7	-0.14	2.16	-0.06	0.50	-0.29	2.30	-0.08	0.55
μ_8	-0.09	4.38	-0.08	1.00	-0.45	4.69	-0.16	1.10

Notes: This table reports bias and standard error for estimates across 100 samples. N is the sample size, and J denotes the number of categories of the discrete observable variable, X. For $J=2$, $X=\{0,1\}$ with $\phi(0)=1$ and $\phi(1)=2$. While for $J=3$, $X=\{0,1,2\}$ with $\phi(0)=1, \phi(1)=2$, and $\phi(2)=1.5$. The true hazard is assumed to be Weibull, with parameters $a=2, b=0.75$ and $a=3, b=1.25$ in Panel A and B, respectively. $\gamma_k=\phi(1)/\phi(0)$ and λ_k denotes the k^{th} normalized moment of the unobserved heterogeneity distribution.

FIGURE 3.1: DATA GENERATING PROCESS



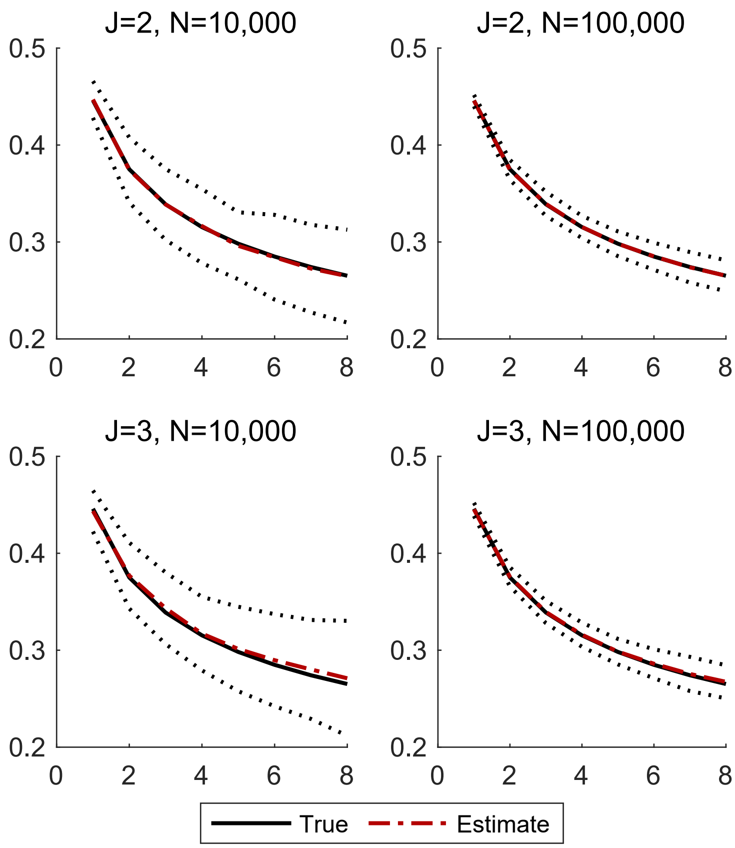
(A) UNOBSERVED HETEROGENEITY AND STRUCTURAL HAZARD



(B) DATA IMPLIED BY THE DGP

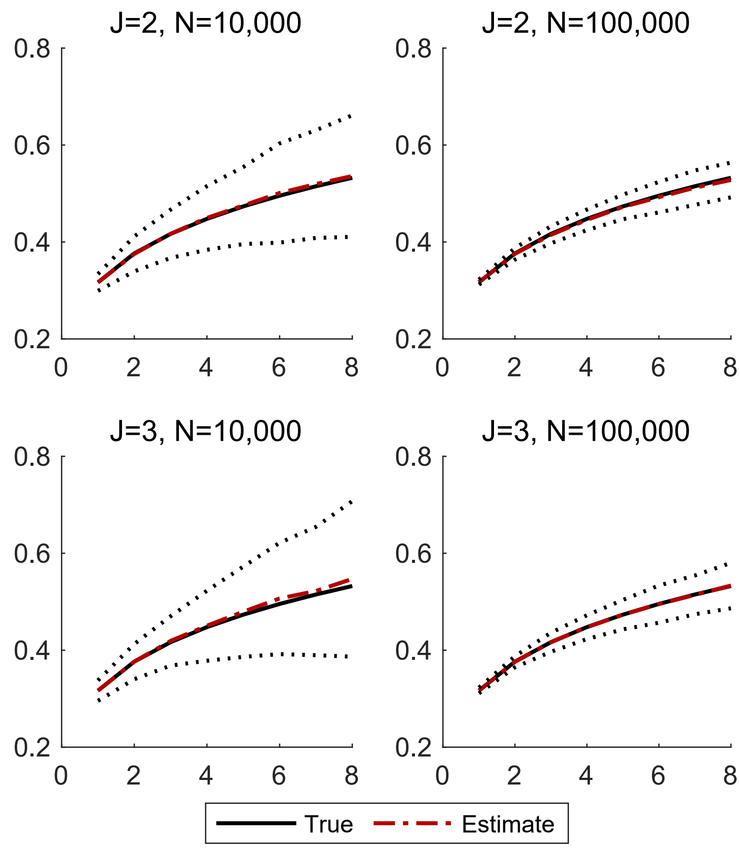
Notes: Panel A presents the underlying Data Generation Process (DGP) for the simulations. The distribution of unobserved heterogeneity is assumed to be beta with parameters $\alpha = 2$ and $\beta = 3$. Structural hazard is assumed to be Weibull: $\psi(t) = (b/a)(t/a)^{b-1}$. Panel B presents the duration density function implied by this DGP for three different values of $\phi(l)$.

FIGURE 3.2: ESTIMATES OF BASELINE HAZARD: TRUE HAZARD IS DECREASING



Notes: The solid lines represent the true baseline hazard that is assumed to be Weibull with parameters $a = 2$ and $b = 0.75$. The dash-dot lines present the average estimate of the baseline hazard across 100 samples. N is the sample size, and J denotes the number of categories of the discrete observable variable, X. For $J = 2$, $X = \{0, 1\}$ with $\phi(0) = 1$ and $\phi(1) = 2$. While for $J = 3$, $X = \{0, 1, 2\}$ with $\phi(0) = 1, \phi(1) = 2$, and $\phi(2) = 1.5$. The dotted lines represent 95% confidence intervals.

FIGURE 3.3: ESTIMATES FOR THE BASELINE HAZARD: TRUE HAZARD IS INCREASING



Notes: The solid lines represent the true baseline hazard that is assumed to be Weibull with parameters $a = 3$ and $b = 1.25$. The dash-dot lines present the average estimate of the baseline hazard across 100 samples. N is the sample size, and J denotes the number of categories of the discrete observable variable, X. For $J = 2$, $X = \{0, 1\}$ with $\phi(0) = 1$ and $\phi(1) = 2$. While for $J = 3$, $X = \{0, 1, 2\}$ with $\phi(0) = 1$, $\phi(1) = 2$, and $\phi(2) = 1.5$. The dotted lines represent 95% confidence intervals.

APPENDIX A

DURATION DEPENDENCE AND HETEROGENEITY: LEARNING FROM EARLY NOTICE OF LAYOFF

A.1 Proofs and Derivations

A.1.1 Proof of Proposition 1

Proof. Since $h(d|\nu, l) = \psi_l(d)\nu$, we have that

$$S(d-1|\nu, l) - S(d|\nu, l) = \psi_l(d)\nu S(d-1|\nu, l)$$

Note that, for any d, $S(d|l) = \mathbb{E}[S(d|\nu, l)|l]$. In which case we can write

$$\tilde{h}(d|l) = \frac{S(d-1|l) - S(d|l)}{S(d-1|l)} = \psi_l(d)\mathbb{E}\left(\nu \cdot \frac{S(d-1|\nu, l)}{S(d-1|l)}\bigg| l\right)$$

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

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

where the first inequality follows from the Bayes rule.

Now, for any d and $\nu_H > \nu_L$,

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

The above equation implies that,

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

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

A.1.2 Proof of Theorem 1

Proof. Denote, μ_k as the k^{th} moment of ν. Now define $\tilde{S}(d|\nu)$ as follows,

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

Then by Assumption 3, for all l and $d > 1$

$$S(d|\nu, l) = [1 - \psi_l(1)\nu]\tilde{S}(d|\nu) \tag{A.1}$$

Note that $1 - G(d|l) = \mathbb{E}[S(d|\nu, l)|l]$, then by Assumption 2,

$$1 - G(d|l) = \mathbb{E}[S(d|\nu, l)] \tag{A.2}$$

In which case,

$$g(d|l) = \psi_l(d)\mathbb{E}[\nu S(d-1|\nu, l)] \tag{A.3}$$

with $S(0|\nu, l) = 1$. Consider any $l, l' \in \Omega$. Plugging in eq. (A.1) in eq. (A.2) and eq. (A.3) and taking the difference between l' and l, we get

$$G(d-1|l') - G(d-1|l) = [\psi_{l'}(1) - \psi_l(1)]\mathbb{E}[\nu\tilde{S}(d-1|\nu)] \tag{A.4}$$

$$g(d|l') - g(d|l) = -\psi(d)[\psi_{l'}(1) - \psi_l(1)]\mathbb{E}[\nu^2\tilde{S}(d-1|\nu)] \tag{A.5}$$

We will normalize, $\mu_1 = 1$. In which case, $\psi_l(1) = g(1|l)$ and $\psi_{l'}(1) = g(1|l')$. Plugging this in eq. (A.4), we can solve for

$$\mathbb{E}[\nu\tilde{S}(d-1|\nu)] = \frac{G(d-1|l') - G(d-1|l)}{g(1|l') - g(1|l)} \tag{A.6}$$

Similarly from eq. (A.5) we can solve for,

$$\mathbb{E}[\nu^2\tilde{S}(d-1|\nu)] = -\frac{g(d|l') - g(d|l)}{\psi(d)[g(1|l') - g(1|l)]} \tag{A.7}$$

Note that plugging in eq. (A.1) in eq. (A.3) gives us,

$$g(d|l) = \psi(d)E[\nu S_l(d-1,\nu)] = \psi(d)\big[\mathbb{E}[\nu\tilde{S}(d-1|\nu)] - g(1|l)\mathbb{E}[\nu^2\tilde{S}(d-1|\nu)]\big]$$

Plugging the expressions from eq. (A.6) and eq. (A.7) in the above equation, for $d > 1$ we can find

$$\psi(d) = \frac{g(d|l)g(1|l') - g(d|l')g(1|l)}{G(d-1|l') - G(d-1|l)}$$

Here, the denominator is not equal to zero as we assumed $\psi_l'(1) \neq \psi_l(1)$. This proves the identification of $\{\psi_l(1), \psi_{l'}(1), \{\psi(d)\}_{d=2}^{\bar{D}}\}$. To see that the first $\bar{D}$ moments of F are identified, note that we can write from eq. (1.1),

$$g(d|l) = \psi_l(d)\sum_{k=1}^{\bar{D}} c_k(d, \boldsymbol{\psi_l})\mu_k$$

where $\boldsymbol{\psi_l} = \{\psi_l(d)\}_{d=1}^{\bar{D}}$ and

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

Denote $\boldsymbol{g}_l = \{g(d|l)\}_{d=1}^{\bar{D}}$ and $\boldsymbol{\mu} = \{\mu_k\}_{k=1}^{\bar{D}}$. Then we can write $\boldsymbol{g}_l = C(\boldsymbol{\psi}_l)\boldsymbol{\mu}$ where $C(\boldsymbol{\psi}_l)$ is an upper triangular matrix with $C_{s,k}(\boldsymbol{\psi}_l) = \psi_l(s)c_k(s,\boldsymbol{\psi}_l)$. In addition, the diagonal elements of $C(\boldsymbol{\psi}_l)$ are non-zero. To see this note that, $C_{d,d}(\boldsymbol{\psi}_l) = (-1)^{d-1}\prod_{s=1}^{d}\psi_l(s)$ and each $\psi_l(s) > 0$. Hence, $C(\boldsymbol{\psi}_l)$ is invertible and we can plug in $\boldsymbol{\psi}_l$ in $\boldsymbol{g}_l = C(\boldsymbol{\psi}_l)\boldsymbol{\mu}$ to solve for $\boldsymbol{\mu}$. □

A.1.3 Proof of Corollary 1

Proof. First note that for $d < \bar{D}$,

$$\begin{aligned} G^{\Delta}(d|\nu, l, D_C > \bar{D}) &= 1 - Pr(\Delta > d|\nu, l, D_C > \bar{D}) \\ &= 1 - Pr(D > d, D_C > d|\nu, l, D_C > \bar{D}) \\ &= G(d|\nu, l, D_C > \bar{D}) \end{aligned}$$

where the second equality follows from $\Delta = \min\{D, D_C\}$ and the third equality follows from $d < \bar{D} < D_C$. Next, note that by Bayes Rule

$$f(\nu|l, D_C > \bar{D}) = \frac{Pr(\nu, l, D_C > \bar{D})}{Pr(l, D_C > \bar{D})} = \frac{Pr(\nu, l|D_C > \bar{D})Pr(D_C > \bar{D})}{Pr(D_C > \bar{D}|l)Pr(l)}$$

By Assumption 4, $Pr(D_C > \bar{D}|l) = Pr(D_C > \bar{D})$. Then,

$$f(\nu|l, D_C > \bar{D}) = \frac{Pr(\nu, l|D_C > \bar{D})}{Pr(l)} = \frac{Pr(\nu|D_C > \bar{D})Pr(l|D_C > \bar{D})}{Pr(l)} = f(\nu|D_C > \bar{D})$$

where the second equality follows from $\nu \perp L$ as assumed in Assumption 2 and third equality follows from Assumption 4. In which case, we can write

$$1 - G^{\Delta}(d|l, D_C > \bar{D}) = \mathbb{E}[S(d|\nu, l)|D_C > \bar{D}]$$

Rest of the proof follows from Theorem 1 by exchanging moments conditional on l by moments conditional on l and $D_C > \bar{D}$. □

A.2 Data

A.2.1 Data Construction

While the DWS is available starting 1984, the variable on length of notice was not available for the first two samples. In addition, the definition of displaced workers for the purpose of the survey has changed over time. [1] 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 mostly missing for the year 1994. For these reasons, I start my analysis from 1996. Moreover, to maintain consistency across years, I exclude self-employed individuals or those who expected to be recalled from the 1996 sample.

[1] The recall window was 5 years instead of 3 before 1994.

Individuals with missing notice length or unemployment duration are excluded from the sample. I also exclude individuals reporting switching more than 2 jobs since the lost job to minimize retrospective bias. The sample is restricted to those who lost a job at least one year before the survey. Out of those who had not found a job when the survey was administered, I exclude individuals who report an unemployment duration of less than 52 weeks or those who are no longer in the labor force. Since 2012, tenure at 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 A.1 presents the descriptive statistics of my baseline sample compared to all individuals in the DWS as well as the CPS over my sample period. Relative to CPS and DWS, individuals in my sample are more highly educated, less likely to not be employed, and older.

A.2.2 Additional Descriptives

In this subsection, I present additional descriptives about the sample used in estimation. In particular, Figure A.1 presents the likelihood of receiving a long notice by displacement year. Table A.2 shows the effect of longer notice on earnings at subsequent job amongst job-finders. Figure A.2 presents the data with alternative bins. Finally, Table A.3 and Figure A.3 describes incidence of UI take-up and timing of benefit exhaustion amongst UI takers for the baseline sample.

TABLE A.1: COMPARISON OF THE BASELINE SAMPLE TO ALL INDIVIDUALS IN THE DISPLACED WORKER SUPPLEMENT (DWS) AND THE CURRENT POPULATION SURVEY (CPS)

Variables	Sample (1)	DWS (2)	CPS (3)
Age	42.65	39.29	40.15
Female	0.45	0.44	0.52
Black	0.09	0.11	0.10
Married	0.63	0.53	0.58
Educational Attainment			
HS Dropout	0.04	0.09	0.09
HS Graduate	0.56	0.65	0.61
College Degree	0.40	0.26	0.30
Employment Status			
Employed	0.90	0.68	0.76
Unemployed	0.08	0.21	0.04
NILF	0.02	0.12	0.19
Number of Observations	2887	43513	920207

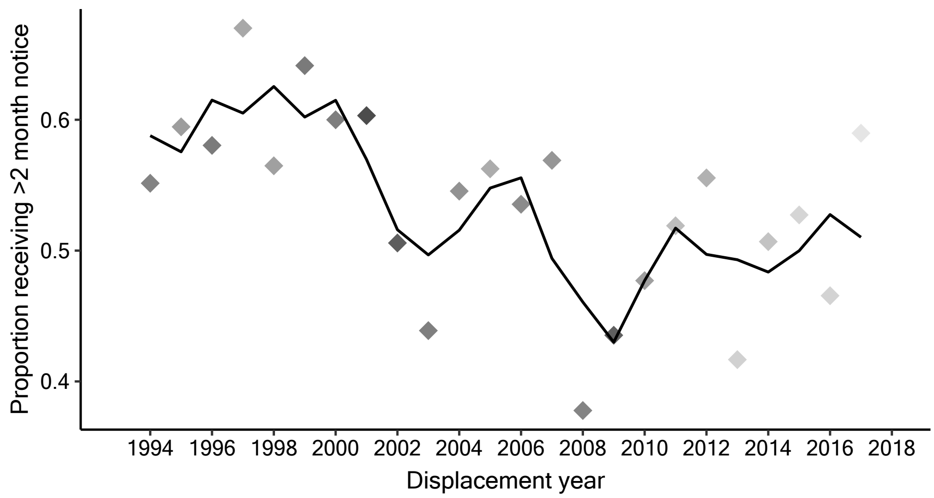
Notes: All samples are restricted to individuals between the ages of 20-60 and pertain to years 1996-2020. Column (1) includes individuals from the DWS who lost their job at least one year before the survey, worked full-time for at least six months and were provided health insurance at their lost job, did not expect to be recalled, and received a notice of either 1-2 months or greater than 2 months. Column (2) and (3) include all individuals in the DWS and the monthly CPS, respectively, over the sample period.

A.3 Robustness

Even though unemployment durations are reported in weeks in the DWS, I calculate the moments used in estimation at 12-week intervals. This is simply because the data at smaller intervals is very noisy due to small sample sizes. However, it is possible to conduct the analysis at 9-week intervals. Panels C and D of Figure A.2 present the data with the unemployment duration binned in 9 weeks. While, Figure A.4 presents estimate using this data. [2] Once again, the estimated structural hazard is above the observed

[2]Note that, the period corresponding to unemployment duration 0 and interval 1-9 weeks are binned together in estimation.

FIGURE A.1: LENGTH OF NOTICE OVER TIME



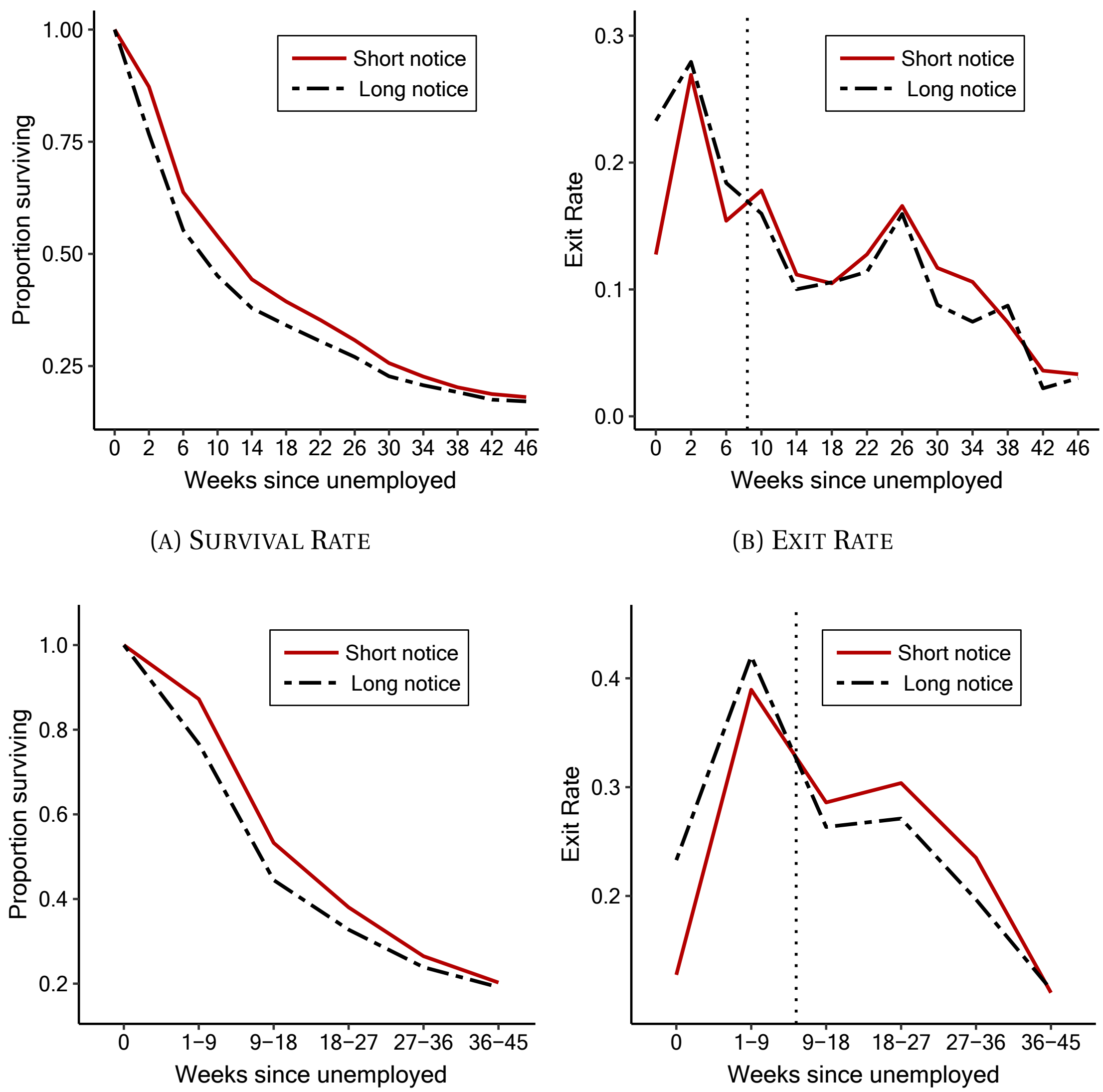
Notes: The diamonds present the proportion of individuals who received a notice of more than 2 months amongst individuals in the sample who were displaced in a given year. The solid line represents a 3-year moving average this proportion. The shading for the diamonds is darker if the sample size in a given year is larger.

TABLE A.2: EARNINGS AT NEXT JOB

	Weekly Log Earnings			
	(1)	(2)	(3)	(4)
>2 month notice	0.083*	0.116***	0.115***	0.127***
	(0.043)	(0.042)	(0.042)	(0.045)
Additional Controls		X	X	X
State FEs			X	X
Industry × Year FEs				X
Observations	1967	1664	1664	1654

Notes: The table shows results from linear regressions of log weekly wages on an indicator variable that takes value 1 if the individual received a notice of more than 2 months and 0 if they received a notice of 1-2 months. Sample includes all individuals in the baseline sample with non-missing earnings at their next job. Individuals who changed more than one job are excluded from the sample. Columns (2)-(4) also exclude individuals with missing values on one or more included controls. Additional controls include age, gender, race, marital status, education, union status, reason for displacement, and tenure and earnings at lost job. Industries are categorized into 14 categories. Year refers to year of displacement.

FIGURE A.2: SURVIVAL AND EXIT RATES - ALTERNATIVE BINS



(A) SURVIVAL RATE

(B) EXIT RATE

(C) SURVIVAL RATE

(D) EXIT RATE

Notes: Short notice refers to a notice of 1-2 months and long notice refers to a notice of more than 2 months. 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.

TABLE A.3: UNEMPLOYMENT INSURANCE TAKE-UP

Unemployment Duration	Observations	Recieved UI Benefits
0 Weeks	532	0.05
1-12 Weeks	1175	0.45
>12 Weeks	1180	0.81

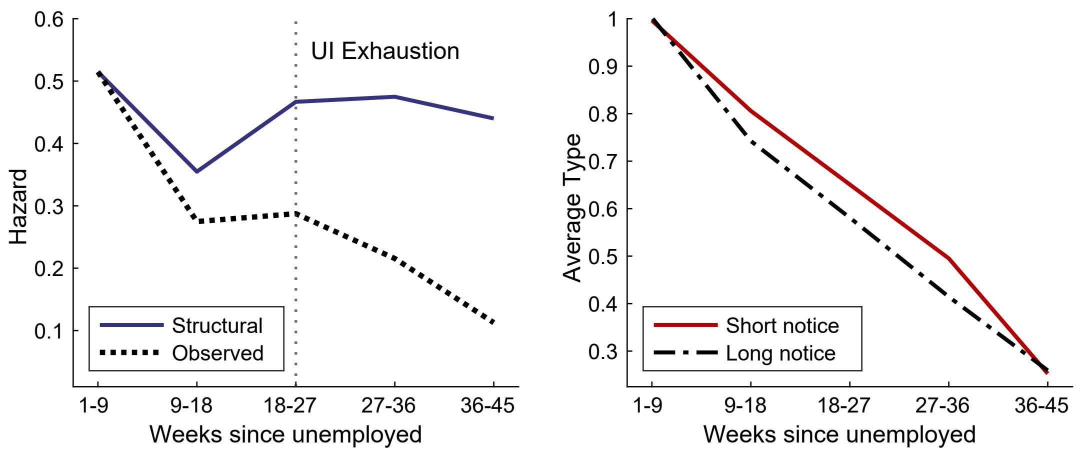
Notes: This table reports the percentage of individuals in the baseline sample who reported receiving UI benefits by the duration of unemployment.

FIGURE A.3: TIMING OF BENEFIT EXHAUSTION

Notes: The figure shows the percentage of individuals who exhausted their UI benefits in each interval. The sample is restricted to individuals in the baseline sample who reported receiving UI benefits.

hazard. Moreover, mimicking the baseline estimates, the structural hazard increases more than the average up until UI exhaustion and remains constant thereafter.

FIGURE A.4: DATA AND ESTIMATES - 9 WEEK INTERVALS



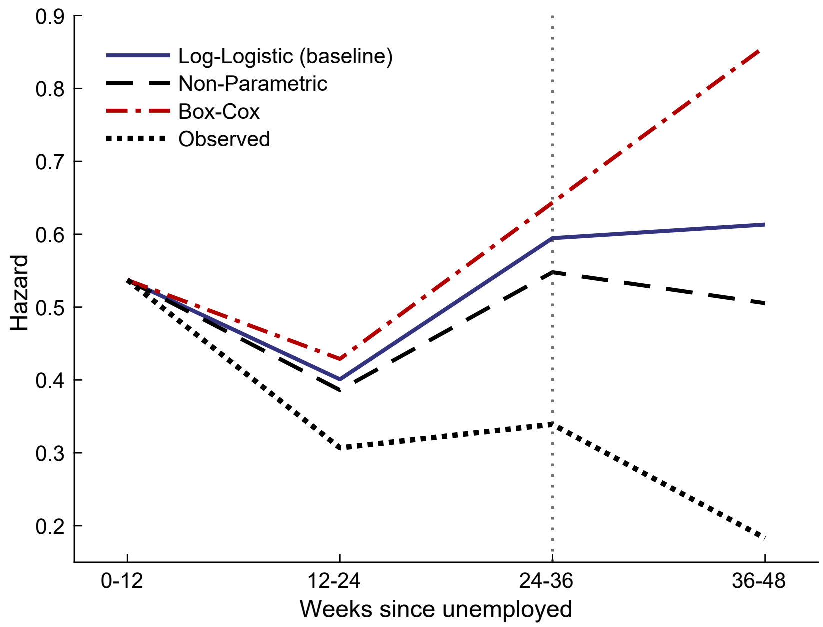
(A) STRUCTURAL HAZARD

(B) AVERAGE TYPE

Notes: Solid line in panel A presents the estimates for the structural hazard from the Mixed Hazard model. Dotted line presents the observed job-finding rate from the data. Panel B presents the implied average-type at each duration for those with short and long notice.

The Mixed Hazard model specified in Section 1.2 is non-parametrically identified. Figure A.5 presents non-parametric estimates for the structural hazard along with the baseline estimates that assume a log-logistic functional form for the hazard (eq. (1.2)). We can see that qualitatively the non-parametric hazard is similar to the baseline estimates. In the literature it is common to impose a Weibull or a Gompertz hazard. However, I choose the log-logistic because it allows the hazard to be non-monotonic. In Figure A.5, I also present estimates assuming the Box-Cox functional form, given by $\psi(d) = \exp\left[\frac{\alpha d^{\beta}-1}{\beta}\right]$. With $\beta \to 0$, this converges to the Weibull hazard, with $\beta = 1$ it is equal to Gompertz, and $\beta = 0$ implies constant hazard. The estimates from this

FIGURE A.5: ESTIMATES WITH DIFFERENT FUNCTIONAL FORMS

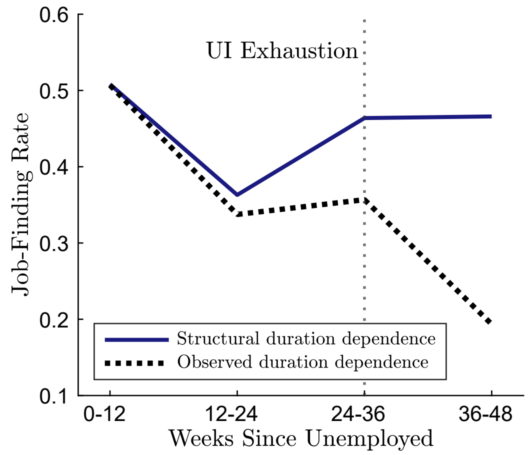


Notes: Figure presents estimates for the structural hazard from the Mixed Hazard model under alternative parametric assumptions. Dotted line presents the observed job-finding rate from the data.

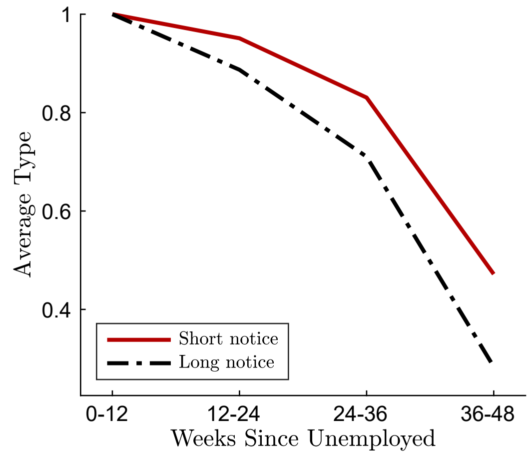
specification result in an increasing hazard.

Finally, Figure A.6 presents the estimates on a sample of individuals who were laid off not as a result of a plant closure. Corresponding standard errors are presented in Table A.6. As we can see from this table that even for this sample, we can reject that the individual hazard declines after UI exhaustion.

FIGURE A.6: EXCLUDING PLANT CLOSURES



(A) STRUCTURAL HAZARD



(B) AVERAGE TYPE

Notes: Figure represents estimates from the Mixed Hazard model excluding individuals laid off due to plant closures. Dotted line in panel A presents the observed job-finding rate from the data. Panel B presents the implied average-type at each duration for those with short and long notice. Standard errors are presented in Table A.4.

TABLE A.4: ROBUSTNESS: EXCLUDE PLANT CLOSURES

		Full Sample		Exclude Plant Closures	
	Description	Estimate	SE	Estimate	SE
$\psi(1)$	Structural hazard: 0-12 weeks	0.54	0.01	0.51	0.01
$\psi(2)$	Structural hazard: 12-24 weeks	0.40	0.06	0.36	0.08
$\psi(3)$	Structural hazard: 24-36 weeks	0.59	0.11	0.46	0.13
$\psi(4)$	Structural hazard: 36-48 weeks	0.61	0.12	0.47	0.16

Notes: Table presents baseline estimates from the Mixed Hazard model alongside estimates from a sample that excludes individuals laid off due to plant closures.

TABLE A.5: ROBUSTNESS: RECESSION AND NORMAL YEARS

	Description	Normal Years Estimate	Normal Years SE	Recession Years Estimate	Recession Years SE
$\psi(1)$	Structural hazard: 0-12 weeks	0.55	0.01	0.50	0.02
$\psi(2)$	Structural hazard: 12-24 weeks	0.41	0.08	0.37	0.10
$\psi(3)$	Structural hazard: 24-36 weeks	0.60	0.14	0.63	0.15
$\psi(4)$	Structural hazard: 36-48 weeks	0.61	0.16	0.67	0.16
μ_2	Second moment	1.17	0.12	1.33	0.18
μ_3	Third moment	1.36	0.59	1.75	0.96
μ_4	Fourth moment	1.51	2.07	2.12	3.70

Notes: Table presents estimates from the Mixed Hazard model separately for years 2001-2002 and 2007-2009 and the rest of the years.

TABLE A.6: ROBUSTNESS: TENURE AND AGE QUARTILES

	Description	Age Estimate	Age SE	Tenure Estimate	Tenure SE
$\psi(1)$	Structural hazard: 0-12 weeks	0.47	0.01	0.47	0.01
$\psi(2)$	Structural hazard: 12-24 weeks	0.33	0.05	0.36	0.05
$\psi(3)$	Structural hazard: 24-36 weeks	0.49	0.11	0.55	0.18
$\psi(4)$	Structural hazard: 36-48 weeks	0.51	0.12	0.57	0.18

Notes: Table presents estimates from the Mixed Hazard model averaged over age and tenure quartiles.

A.4 Generalization

The main identification result in the paper relies on two crucial assumptions: (i) notice length is independent of worker type, and (ii) the structural hazard after the initial period is identical regardless of notice length. In this section, I generalize the identification result and show that it is possible to identify structural duration dependence and the moments of heterogeneity distribution as long as we know how the structural hazard after the initial period as well as the distribution of heterogeneity varies across workers with different notice lengths. In particular, consider two lengths of notice and define κ_d as the difference between the d^{th} moment of ν conditional on l' and l as follows

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

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,

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

Now if for some $\bar{D}$ we know κ_d for $d = 1, ..., \bar{D}$ and γ_d for $d = 2, ..., \bar{D}$, we can identify the first $\bar{D}$ structural hazards and moments of type distribution for each notice length up to scale. [3] To see why this is the case, we can extend the discussion below Theorem 1.

[3]Alternatively, we could know κ_d for $d = 2, ..., \bar{D}$ and γ_d for $d = 1, ..., \bar{D}$. Also, in theory, the choice of defining γ_d and κ_d as a ratio or a difference does not matter as we can write the proof of identification in

For notice length l, at $d=1$ and $d=2$, we have

$$
\begin{aligned}
\tilde{h}(1|l) &= \psi_l(1)\mu_{l,1} \\
\tilde{h}(2|l) &= \psi_l(2)\mu_{l,1}\left(\frac{1-\tilde{h}(1|l)(\mu_{2,l}/\mu_{1,l}^2)}{1-\tilde{h}(1|l)}\right)
\end{aligned}
$$

As before, if we knew the extent of heterogeneity across workers i.e. the variance of ν amongst l notice individuals we would be able to infer structural duration dependence $\psi_l(2)/\psi_l(1)$ from observed duration dependence $\tilde{h}(2|l)/\tilde{h}(1|l)$. Now, we also observe the hazard conditional on notice length l', which is given by

$$
\tilde{h}(2|l') = \gamma_2\psi_l(2)(\mu_{l,1}+\kappa_1)\left(\frac{1-\tilde{h}(1|l')\big((\mu_{2,l}+\kappa_2)/(\mu_{1,l'}+\kappa_1)^2\big)}{1-\tilde{h}(1|l')}\right)
$$

So now if we compare $\tilde{h}(2|l')$ to $\tilde{h}(2|l)$, as before, the difference between the two depends on $\mu_{2,l}$, however, now it also depends on γ_2, κ_1, and κ_2. So if we know γ_2, κ_1, and κ_2, we can still back out $\mu_{2,l}$. The intuition for the result is that we know how the structural hazards for different notice lengths at $d=2$ should vary if there was no heterogeneity. Then if we observe the structural hazards being different over and above what we would expect with no heterogeneity, we can attribute that to the presence of heterogeneity.

Theorem 2. *For some l, l', define $\kappa_d = \mu_{l',d} - \mu_{l,d}$ and $\gamma_d = \psi_{l'}(d)/\psi_l(d)$. 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 $\{\psi_l(d), \psi_{l'}(d)\}_{d=1}^{\bar{D}}$ and*

any case. Here, I define κ_d as a difference but γ_d as a ratio as it makes it easy to vary these parameters when I look at how the estimates change with respect to these parameters.

the conditional moments of the type distribution $\{\mu_{l,d},\mu_{l',d}\}_{d=1}^{\bar{D}}$ *are identified up to a scale from* $\{G(d|l),G(d|l')\}_{d=1}^{\bar{D}}$.

Proof. First note that we can rewrite $eq.$ (1.1) as follows,

$$g(d|l)=\psi_l(d)\sum_{k=1}^{d}c_k(\boldsymbol{\psi_{l,d-1}})\mu_{l,k} \tag{A.8}$$

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

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

Now we can prove the statement of the theorem by induction. First note that the statement is true for $\bar{D}=1$. To see this, note that

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

We will normalize $\mu_{l,1}=1$. Then we can solve for $\psi_l(1)=g(1|l)$ and $\psi_{l'}(1)=\frac{g(1|l')}{1+\kappa_1}$.

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

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

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

$$
\mu_{l,d} = \frac{(-1)^d}{\prod_{s=1}^{d-1}\psi_l(s)}\left[\sum_{k=1}^{d-1} c_k(\boldsymbol{\psi_{l,d-1}})\mu_{l,k} - \frac{g(d|l)}{\psi_l(d)}\right]
$$

Using the fact that $\mu_{l',d} = \mu_{l,d} + \kappa_d$ and plugging that in the expression for $g(d|l')$, we get

$$
\begin{aligned}
g(d|l') &= \psi_{l'}(d)\left[\sum_{k=1}^{d-1} c_k(\boldsymbol{\psi_{l',d-1}})\mu_{l',k} + (-1)^{d-1}\mu_{l',d}\prod_{s=1}^{d-1}\psi_{l'}(s)\right] \\
&= \psi_{l'}(d)\left[\sum_{k=1}^{d-1} c_k(\boldsymbol{\psi_{l',d-1}})\mu_{l',k} + (-1)^{d-1}\kappa_d\prod_{s=1}^{d-1}\psi_{l'}(s) - \Gamma_{d-1}\left(\sum_{k=1}^{d-1} c_k(\boldsymbol{\psi_{l,d-1}})\mu_{l,k} - \frac{g(d|l)}{\psi_l(d)}\right)\right] \\
&= \psi_{l'}(d)\left[\sum_{k=1}^{d-1} c_k(\boldsymbol{\psi_{l',d-1}})\mu_{l',k} + (-1)^{d-1}\kappa_d\Psi_{l'}(d-1) - \Gamma_{d-1}\sum_{k=1}^{d-1} c_k(\boldsymbol{\psi_{l,d-1}})\mu_{l,k}\right] + \Gamma_d g(d|l)
\end{aligned}
$$

where $\Gamma_d = \prod_{s=1}^{d}\gamma_s$ and $\Psi_l(d) = \prod_{s=1}^{d}\psi_l(s)$. Then we can solve for,

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

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

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

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

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

A.4.1 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 the 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 the 2nd, 3rd, and 4th central moment, the variance, skewness, and kurtosis, for the two groups are identical. The non-central moments would be impacted by scale changes, so all four κ_ds 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 for 25 equidistant values for κ_1 in the interval $[-0.04, 0.07]$. [4] κ_2, κ_3 and κ_4 are defined as above. Residuals from this exercise are presented in panel B of Figure A.7. In panel A, I present the estimates for structural duration dependence for three different values of κ. From here, we can see that for positive values of κ_1 our baseline estimate of structural

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

hazard is unchanged. However, in the case that individuals with longer notices would be less likely to exit unemployment in any case, our baseline estimate overestimates the increase in structural hazard. This is because, we attribute all the difference between the hazard at later durations to be due to heterogeneity, but in the case when $\kappa_1 < 0$, some of this difference is simply due to individuals with longer notice lengths being composed of workers who are less likely to find a job at any point in time. However, since the minimizing value of κ_1 is slightly above 0, this exercise points towards the robustness of the baseline estimates. [5]

2. Allow structural hazards after the first period to vary

Now as in the baseline estimation, I assume notice length to be independent of worker type. But now we will 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 $d = 1,..,\bar{D}$ and $\gamma_d = \gamma$ for $d = 2,..,\bar{D}$. I estimate the model for 25 equidistant values for γ in the interval $[0.95, 1.2]$. Results from this exercise are presented in Figure A.8. The results point towards the structural hazard being slightly greater for individuals with a longer notice even beyond the first 12 weeks. As we can see from panel A of Figure A.8, this suggests that the baseline estimates might be underestimating the role of dynamic selection. The reason for this is similar as before, in the case that the hazard for long notice workers is higher even beyond the initial period, the gap between the long and short notice average exit rates due to composition would be greater than

[5] While I did not formally prove that κ_1 is identified, due to assuming the log-logistic hazard, the model had one free parameter which is in probably enabling us to be able to pin down κ_1.

what we assumed in the baseline estimation.

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

I create a 20×20 grid for values of $\kappa \in [-0.15, 0.15]$ and $\gamma \in [0.95, 1.20]$. I reestimate the model for each point in the grid. Panel A of Figure A.9 presents the residuals for different values in the grid. While, panel B of Figure A.9 presents estimates at the maximizing values.

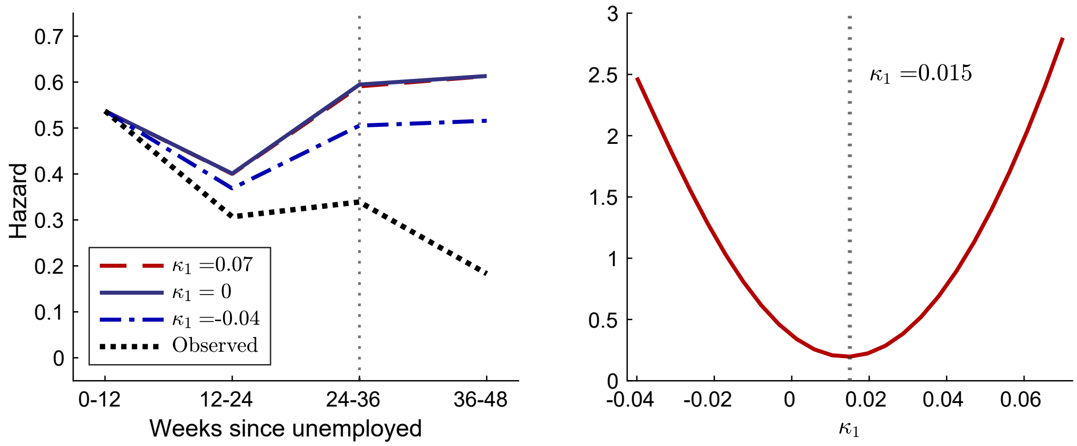
A.5 Search Model

A.5.1 Calibration Details

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), assume that costs of job search are given $c(s) = \theta s^{1+\rho}/(1+\rho)$. I set $\rho = 1$ and $\theta = 50$. [6] Table A.7

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

FIGURE A.7: ALLOW AVERAGE TYPE TO VARY

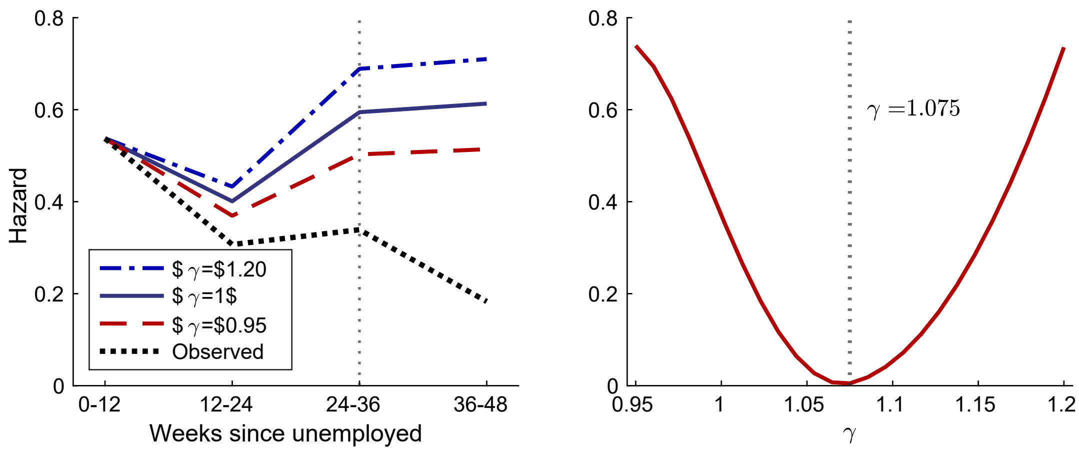


(A) STRUCTURAL HAZARD

(B) RESIDUALS

Notes: The figure presents results from estimation of the more general Mixed Hazard model specified in A.4. The model is estimated assuming that the structural hazards after the first 12 weeks are equalized for two lengths of notice. However, 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.

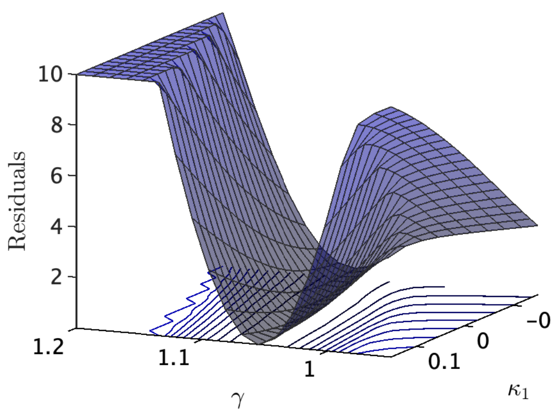
FIGURE A.8: ALLOW STRUCTURAL HAZARDS AFTER THE FIRST PERIOD TO VARY



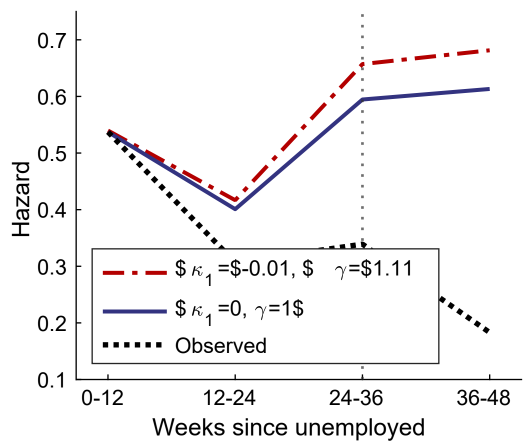
(A) STRUCTURAL HAZARD (B) RESIDUALS

Notes: The figure presents results from estimation of the more general Mixed Hazard model specified in A.4. The model is estimated assuming that notice lengths are independent of a worker's type. However, the structural hazards after the initial period for individuals with different lengths of notice are allowed to vary according to the parameter γ. Panel A presents the residuals from GMM estimation for different values of γ.

FIGURE A.9: ALTERNATIVE ASSUMPTIONS ON STRUCTURAL HAZARDS AND HETEROGENEITY DISTRIBUTION



(A) RESIDUALS



(B) STRUCTURAL HAZARD

Notes: The figure presents results from estimation of the more general Mixed Hazard model specified in A.4. 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 hazards after the initial period for individuals with different lengths of notice are allowed to vary according to the parameter γ. Panel A presents the residuals from GMM estimation for different values of κ_1 and γ.

summarizes the calibration parameters.

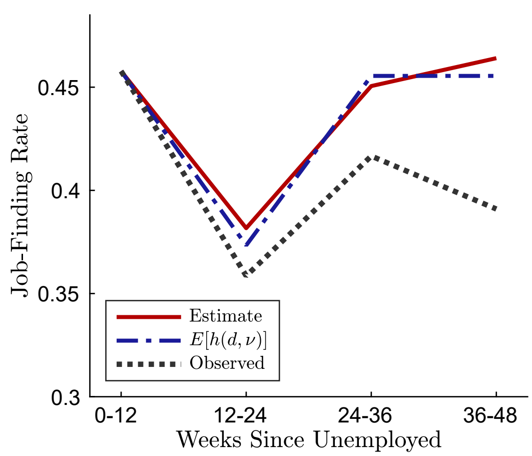
TABLE A.7: CALIBRATION PARAMETERS

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

A.5.2 Simulation

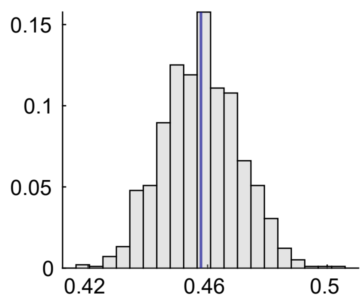
In this section, I simulate data from the search model presented in Section 1.6. To incorporate multiple notice period, I let the offer rate in the first period be different for long (L) and short (S) notice individuals. I set $\nu_H = 1$, $\nu_L = 0.5$ and $\pi = 0.5$, $\delta_L(1) = 1.25$, $\delta_S(1) = 1$, and $\delta(d) = 0.95$ for $d = 2, 3, 4$. Rest of the parameters are set as specified in Table A.7. I assume there are 2500 individuals, half of whom receive the L length notice. I simulate data on job-finding rates for this model 1000 times. Average of estimates for the structural hazard is presented in Figure A.10, while the distribution of the estimates is presented in Figure A.11.

FIGURE A.10: SIMULATION: AVERAGE ESTIMATE

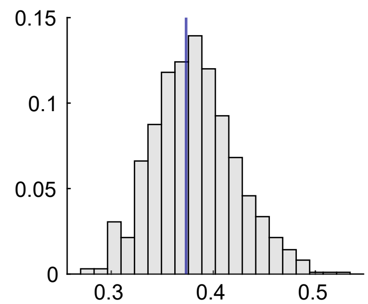


Notes: The figure presents average estimate from 1000 simulations of the search model. The dashed line presents the implied structural duration dependence $\mathbb{E}[h(d|\nu)]$ from this model. While, the dotted line presents the observed structural duration dependence $\mathbb{E}[h(d|\nu)|D \geq d]$ from this model.

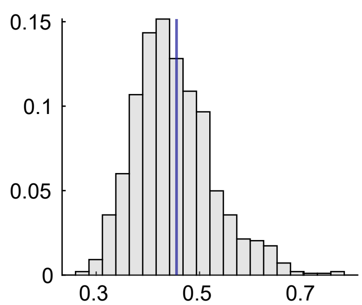
FIGURE A.11: SIMULATION: DISTRIBUTION OF ESTIMATES



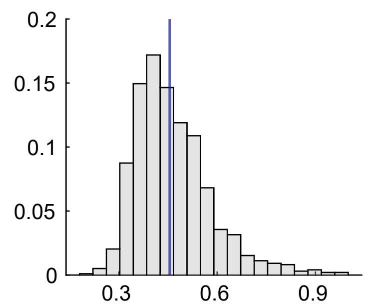
(A) $\psi(1)$



(B) $\psi(2)$



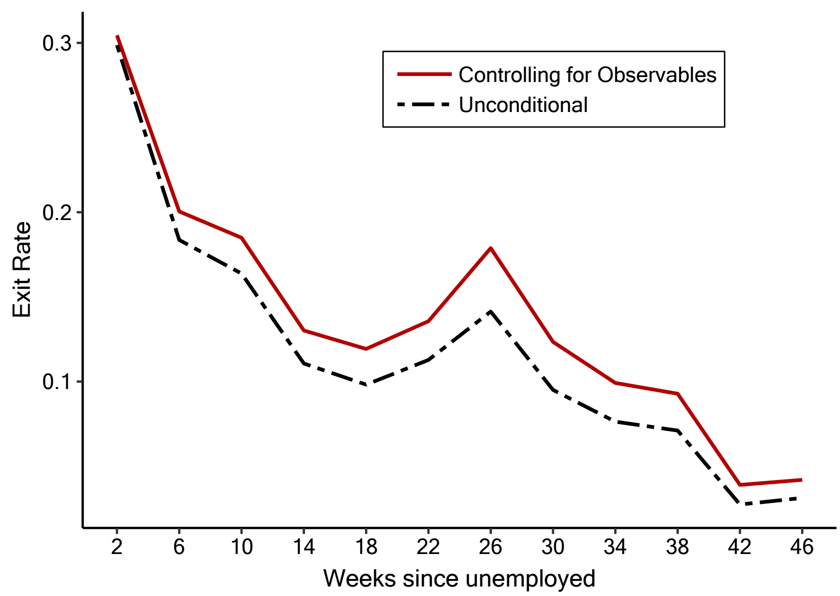
(C) $\psi(3)$



(D) $\psi(4)$

Notes: The figure presents distribution of estimates of structural duration dependence on simulated data from the search model. The blue line represents the true value for each structural hazard.

FIGURE A.12: COX PROPORTIONAL HAZARD MODEL



Notes: The figure presents estimates of the structural hazard from the Cox proportional hazard model (`coxph` in R). Sample consists of 29,775 individuals from the DWS for years 1996-2020 who worked full-time at their previous employer and did not expect to be recalled to their last job. Observations with missing values on unemployment duration are excluded. Observable characteristics controlled for include age, gender, race, education, marital status, reason for displacement, union status, years of tenure and earnings at last job, year of displacement, industry, and state of residence.

APPENDIX B

UNDERSTANDING THE RACIAL EMPLOYMENT GAP: THE ROLE OF SECTORAL SHIFTS

B.1 Proofs and Derivations

B.1.1 Proof of Proposition 1

Proof. Worker i chooses to search for employment opportunities in the location with the highest expected utility, such that

$$l_i^* = \underset{l}{\operatorname{argmax}}\{u_{lg}\varepsilon_{li}\}$$

where $u_{lg} = w_{lg} l_{lg}$.

Probability that an individual chooses local labor market l' is given by:

$$
\begin{aligned}
\frac{P_{l'g}}{P_g} &= \mathbb{E}_{\varepsilon_{l'}}[\Pr(u_{l'g}\varepsilon_{l'} > u_{lg}\varepsilon_l) \ \ \forall l \neq l'] \\
&= \int_0^{\infty} \exp\left[-\sum_{l\neq l'}\left(\frac{u_{l'g}\varepsilon_{l'}}{u_{lg}}\right)^{-1/\kappa_g}\right] f(\varepsilon_l')d\varepsilon_l' \\
&= \int_0^{\infty} \exp\left[-\left(u_{l'g}^{-1/\kappa_g}\sum_{l\neq l'} u_{lg}^{1/\kappa_g}+1\right)\varepsilon_{l'}^{-1/\kappa_g}\right] 1/\kappa_g \varepsilon^{-1/\kappa_g - 1} d\varepsilon_l' \\
&= \frac{u_{lg}^{1/\kappa_g}}{\sum_l u_{lg}^{1/\kappa_g}}
\end{aligned}
$$

Derivation uses the fact that if ε is distributed Fréchet with $F(\varepsilon) = exp(-\varepsilon^{-1/\kappa})$, then $f(\varepsilon) = 1/\kappa\varepsilon^{1/\kappa-1}\exp(-\varepsilon^{1/\kappa})$.

Plugging in equilibrium values of w_{lg} and l_{lg}, we can find equilibrium expected utility from location l as follows:

$$
u_{lg}^* = \omega_{lg}A_{lg} \times \gamma_{lg}\left[\frac{(1-\omega_{lg})\gamma_{lg}}{F_{lg}}\right]^{\frac{\alpha_g}{1-\alpha_g}} A_{lg}^{\frac{\alpha_g}{1-\alpha_g}} = c_{lg}A_{lg}^{\frac{1}{1-\alpha_g}}
$$

where $c_{lg} = \omega_{lg}\gamma_{lg}\left[\frac{(1-\omega_{lg})\gamma_{lg}}{F_{lg}}\right]^{\frac{\alpha_g}{1-\alpha_g}}$. Plugging in u_{lg}^* in the expression for $P_{l'g}/P_g$ we get the expression specified in the proposition. □

B.1.2 Proof of Proposition 2

Proof.

$$\frac{P^*_{lg}}{P_g} = \frac{\tilde{c}_{lg} A_{lg}^{\frac{1}{\kappa_g(1-\alpha_g)}}}{\sum_l \tilde{c}_{lg} A_{lg}^{\frac{1}{\kappa_g(1-\alpha_g)}}}$$

Taking the log of eq. (2.7) we get,

$$\ln l^*_{lg} = \ln(c_{lg}/\omega_{lg}) + \frac{\alpha_g}{1-\alpha_g} \ln A_{lg}$$

Now if we total differentiate this expression, we get

$$\frac{\partial l^*_{lg}}{l^*_{lg}} = \frac{\alpha_g}{1-\alpha_g} \frac{\partial A_{lg}}{A_{lg}} \to \hat{l}^*_{lg} = \frac{\alpha_g}{1-\alpha_g} \hat{A}_{lg}$$

Similarly, taking the log of eq. (2.8) we get,

$$\ln P^*_{lg} = \ln \tilde{c}_{lg} + \frac{1}{\kappa_g(1-\alpha_g)} \ln A_{lg} - \ln\left[\sum_l \tilde{c}_{lg} A_{lg}^{\frac{1}{\kappa_g(1-\alpha_g)}}\right]$$

Now again if we total differentiate this expression, we get

$$\hat{P}^*_{lg} = \frac{1}{\kappa_g(1-\alpha_g)} \hat{A}_{lg} - \frac{1}{\sum_l \tilde{c}_{lg} A_{lg}^{\frac{1}{\kappa_g(1-\alpha_g)}}} \cdot \frac{1}{\kappa_g(1-\alpha_g)} \sum_l \tilde{c}_{lg} A_{lg}^{\frac{1}{\kappa_g(1-\alpha_g)}} \frac{\partial A_{lg}}{A_{lg}}$$

Since $\frac{P^*_{lg}}{P_g} = \frac{\tilde{c}_{lg} A_{lg}^{\frac{1}{\kappa_g(1-\alpha_g)}}}{\sum_l \tilde{c}_{lg} A_{lg}^{\frac{1}{\kappa_g(1-\alpha_g)}}}$, we can write the above expression as,

$$\hat{P}^*_{lg} = \frac{1}{\kappa_g(1-\alpha_g)} \hat{A}_{lg} - \frac{1}{\kappa_g(1-\alpha_g)} \sum_l \pi^p_{lg} \hat{A}_{lg}$$

where $\pi^p_{lg} = P^*_{lg}/P_g$. □

B.1.3 Proof of Proposition 3

Proof. First note that, $L_g = \sum_l L_{lg}$ so we can write $\hat{L}_g = \sum_l \pi_{lg} \hat{L}_{lg}$. Since $l_{lg} = L_{lg}/P_{lg}$, we have

$$\hat{L}^*_g = \sum_l \pi_{lg} (\hat{l}_{lg} + \hat{P}_{lg})$$

Now if we plug in the expressions from Proposition 2, we get

$$\hat{L}^*_g = \frac{\alpha_g}{1-\alpha_g} \sum_l \pi_{lg} \hat{A}_{lg} - \frac{1}{\kappa_g(1-\alpha_g)} \sum_l (\pi^p_{lg} - \pi_{lg}) \hat{A}_{lg}$$

Note that,

$$(\pi^p_{lg} - \pi_{lg}) = \pi_{lg} \left(\frac{P_{lg}}{P_g} \cdot \frac{L_g}{L_{lg}} - 1 \right) = \pi_{lg} \left(\frac{l_g}{l_{lg}} - 1 \right)$$

□

B.2 Data

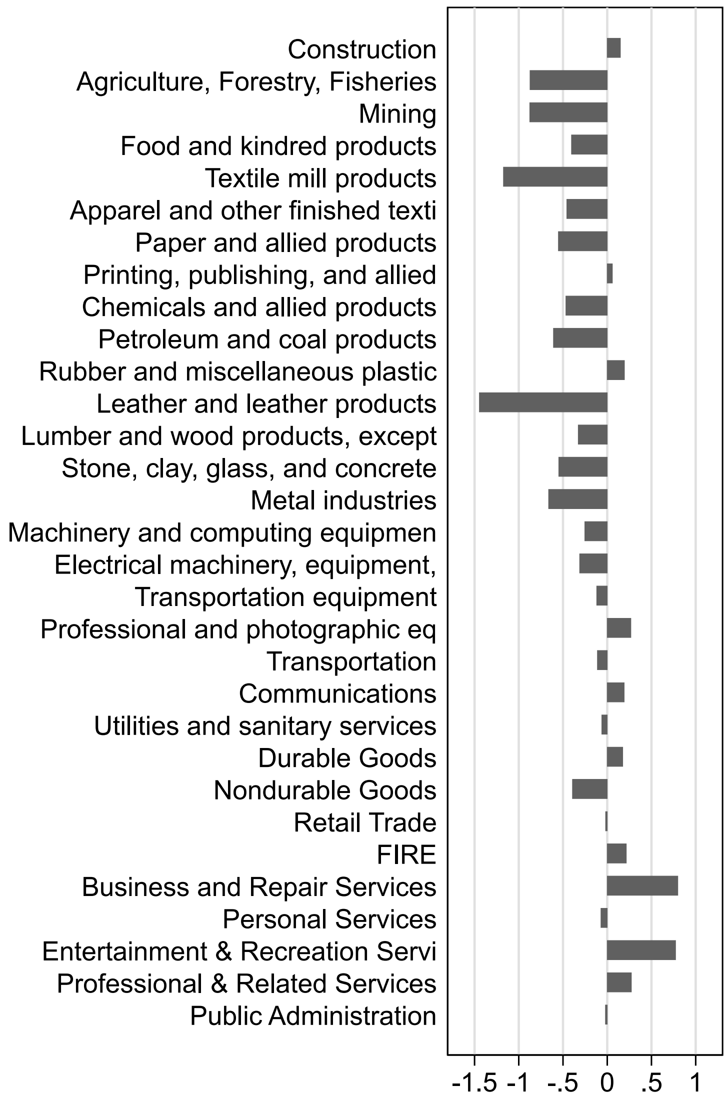
This section contains robustness checks and additional figures left out of the main text. Table B.1 shows the results of our decomposition exercise in section 2.4.2 using both our baseline specification and our specifications where we include additional controls in the regression. Specifically, Panel (A) using the coefficient estimates from Column (1) of Table 2.2, Panel (B) using the coefficient estimates from Column (2) of Table 2.2, and Panel (C) using the coefficient estimates from Column (3) of Table 2.2 to perform the decomposition. Our results are of very similar magnitude across different specifications.

Figure B.1 plots the magnitude of our proxies for labor demand shifts at the sectoral level for each sector in our sample. In particular, it shows $\ln(L_{s,2000}) - \ln(L_{s,1970})$, where $L_{s,y}$ is total employment in sector s in year y.

TABLE B.1: EMPLOYMENT RESPONSE VS. POPULATION RESPONSE

	Employment Response (1)	Population Response (2)	Total Response (3)
Panel A: Baseline specification			
Black Men	-0.031	0.006	-0.024
White Men	-0.006	0.006	0.000
Panel B: With additional controls			
Black Men	-0.033	0.006	-0.027
White Men	-0.009	0.008	-0.001
Panel C: With additional controls and state fixed effects			
Black Men	-0.032	0.005	-0.027
White Men	-0.012	0.007	-0.005

FIGURE B.1: INDUSTRY LEVEL SHIFTS



APPENDIX C

THE DISCRETE-TIME MIXED PROPORTIONAL HAZARD MODEL

C.1 Expressions for t=1,2

Normalize $\phi(0)\mu_1 = 1$ and denote $\gamma = \phi(1)/\phi(0)$. Now expanding eq. (3.1) for $t = 1,2$ and $l = 1,0$:

$$g_1(1) = \psi(1)\gamma$$

$$g_0(1) = \psi(1)$$

$$g_1(2) = \psi(2)[\gamma - \gamma^2\psi(1)\lambda_2]$$

$$g_0(2) = \psi(2)[1 - \psi(1)\lambda_2]$$

From the first two equations we can solve for $\psi(1) = g_0(1)$ and $\gamma = g_1(1)/g_0(1)$. Once we plug in those values, we can get use the next two equations to solve for $\psi(2)$ and μ_2 as follows:

$$\psi(2) = \frac{g_0(1)(g_1(2) - \gamma^2 g_0(2))}{g_1(1) - \gamma^2 g_0(1)}, \quad \lambda_2 = \frac{g_1(2) - \gamma g_0(2)}{g_0(1)(g_1(2) - \gamma^2 g_0(2))}$$

As long as $\gamma \neq 1$, expressions for $\psi(2)$ and μ_2 are well defined.

Similarly for $t = 3$,

$$g_1(3) = \psi(3)[\gamma - \psi(1)\gamma^2\lambda_2 - \psi(2)\gamma^2\lambda_2 + \gamma^3\psi(1)\psi(2)\lambda_3]$$

$$g_0(3) = \psi(3)[1 - \psi(1)\lambda_2 - \psi(2)\lambda_2 + \psi(1)\psi(2)\lambda_3]$$

Then we can write:

$$g_1(3) - \gamma^3 g_0(3) = \psi(3)[(\gamma - \gamma^3) - (\psi(1) + \psi(2))(\gamma^2 - \gamma^3)\lambda_2]$$

So

$$\psi(3) = \frac{g_1(3) - \gamma^3 g_0(3)}{(\gamma - \gamma^3) - (\psi(1) + \psi(2))(\gamma^2 - \gamma^3)\lambda_2}$$

We can plug in $\lambda_2, \gamma, \psi(1), \psi(2)$ and solve for $\psi(3)$. One caveat is that for the above term to be well defined we need the denominator to be non-zero. So we need that

$$(1 - \gamma^2) \neq (\psi(1) + \psi(2))\gamma(1 - \gamma)\lambda_2$$

or

$$1 - \psi(1)\lambda_2 - \psi(2)\lambda_2 \neq -1/\gamma$$

If parameters are appropriately bounded, then $\psi(3)$ will be identified. To solve for λ_3:

$$\frac{g_1(3)}{g_0(3)} = \frac{\gamma - \psi(1)\gamma^2\lambda_2 - \psi(2)\gamma^2\lambda_2 + \gamma^3\psi(1)\psi(2)\mu_3}{1 - \psi(1)\lambda_2 - \psi(2)\lambda_2 + \psi(1)\psi(2)\mu_3}$$

In which case,

$$\lambda_2 = \frac{g_0(3)(\gamma - \psi(1)\gamma^2\lambda_2 - \psi(2)\gamma^2\lambda_2) - g_1(3)(1 - \psi(1)\lambda_2 - \psi(2)\lambda_2)}{\psi(1)\psi(2)(g_1(3) - \gamma^3 g_0(3))}$$

Bibliography

Abbring, J. H. (2012). Mixed Hitting-Time Models. *Econometrica 80*(2), 783–819.

Adão, R., C. Arkolakis, and F. Esposito (2019, February). General Equilibrium Effects in Space: Theory and Measurement. Technical Report 25544, National Bureau of Economic Research, Inc.

Alvarez, F., K. Borovickova, and R. Shimer (2021). Consistent Evidence on Duration Dependence of Price Changes. Technical report.

Alvarez, F. E., K. Borovičková, and R. Shimer (2016, April). Decomposing Duration Dependence in a Stopping Time Model. Technical Report w22188, National Bureau of Economic Research.

Amior, M. and A. Manning (2018, July). The Persistence of Local Joblessness. *American Economic Review 108*(7), 1942–1970.

Autor, D. H., D. Dorn, and G. H. Hanson (2013, October). The China Syndrome: Local Labor Market Effects of Import Competition in the United States. *American Economic Review 103*(6), 2121–2168.

Baker, M. and A. Melino (2000). Duration dependence and nonparametric heterogeneity: A Monte Carlo study. *Journal of Econometrics 96*(2), 357–393.

Bartik, T. (1991). Who Benefits from State and Local Economic Development Policies? Books from Upjohn Press, W.E. Upjohn Institute for Employment Research.

Batistich, M. K. and T. N. Bond (2019). Stalled Racial Progress and Japanese Trade in the 1970s and 1980s. pp. 60.

Bayer, P. and K. K. Charles (2018, August). Divergent Paths: A New Perspective on Earnings Differences Between Black and White Men Since 1940. *The Quarterly Journal of Economics 133*(3), 1459–1501.

Blanchard, O. J. and L. F. Katz (1992). Regional Evolutions. *Brookings Papers on Economic Activity 23*(1), 1–76.

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.

Bound, J. and R. Freeman (1992). What Went Wrong? The Erosion of Relative Earnings and Employment Among Young Black Men in the 1980s. *The Quarterly Journal of Economics 107*(1), 201–232.

Bound, J. and H. J. Holzer (1993). Industrial Shifts, Skills Levels, and the Labor Market for White and Black Males. *The Review of Economics and Statistics 75*(3), 387–396.

Bound, J. and H. J. Holzer (2000). Demand Shifts, Population Adjustments, and Labor Market Outcomes during the 1980s. *Journal of Labor Economics 18*(1), 20–54.

Boustan, L. P. (2017). INTRODUCTION. In *Competition in the Promised Land*, Black Migrants in Northern Cities and Labor Markets, pp. 1–13. Princeton University Press.

Brinch, C. N. (2007, April). Nonparametric Identification of the Mixed Hazards Model with Time-Varying Covariates. *Econometric Theory 23*(2), 349–354.

Brown, C. Black-White Earnings Ratios Since the Civil Rights Act of 1964: The Importance of Labor Market Dropouts. *The Quarterly Journal of Economics.*

Burdett, K. and T. Vishwanath (1988, October). Declining Reservation Wages and Learning. *The Review of Economic Studies 55*(4), 655–665.

Card, D. and A. B. Krueger (1992, February). School Quality and Black-White Relative Earnings: A Direct Assessment*. *The Quarterly Journal of Economics 107*(1), 151–200.

Chodorow-Reich, G. (2020, June). Regional data in macroeconomics: Some advice for practitioners. *Journal of Economic Dynamics and Control 115*, 103875.

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.

Davis, S. J., R. J. Faberman, and J. C. Haltiwanger (2013, May). The Establishment-Level Behavior of Vacancies and Hiring *. *The Quarterly Journal of Economics 128*(2), 581–622.

DellaVigna, S., J. Heining, J. F. Schmieder, and S. Trenkle (2020, April). Evidence on Job Search Models from a Survey of Unemployed Workers in Germany. Working Paper 27037, National Bureau of Economic Research.

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.

Dix-Carneiro, R. and B. K. Kovak (2017, October). Trade Liberalization and Regional Dynamics. *American Economic Review 107*(10), 2908–2946.

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.

Ganong, P. and P. Noel (2019, July). Consumer Spending during Unemployment: Positive and Normative Implications. *American Economic Review 109*(7), 2383–2424.

Glaeser, E., J. Scheinkman, and A. Shleifer (1995). Economic Growth in a Cross-Section of Cities. *Journal of Monetary Economics 36*(1), 117–143.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020, August). Bartik Instruments: What, When, Why, and How. *American Economic Review 110*(8), 2586–2624.

Gonzalez, F. M. and S. Shi (2010). An Equilibrium Theory of Learning, Search, and Wages. *Econometrica 78*(2), 509–537.

Gould, E. D. (2020, March). Torn Apart? The Impact of Manufacturing Employment Decline on Black and White Americans. *The Review of Economics and Statistics*, 1–46.

Hahn, J. (1994). The Efficiency Bound of the Mixed Proportional Hazard Model. *The Review of Economic Studies 61*(4), 607–629.

Han, A. and J. A. Hausman (1990). Flexible parametric estimation of duration and competing risk models. *Journal of Applied Econometrics 5*(1), 1–28.

Hansen, L. P. (1982). Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica 50*(4), 1029–1054.

Hausman, J. A. and T. Woutersen (2014). Estimating a semi-parametric duration model without specifying heterogeneity. *Journal of Econometrics 178*(P1), 114–131.

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.

Horowitz, J. L. (1999). Semiparametric Estimation of a Proportional Hazard Model with Unobserved Heterogeneity. *Econometrica 67*(5), 1001–1028.

Hurst, E., Y. Rubinstein, and K. Shimizu (2021, July). Task-Based Discrimination. Working Paper 29022, National Bureau of Economic Research.

Ishwaran, H. (1996, August). Identifiability and rates of estimation for scale parameters in location mixture models. *The Annals of Statistics 24*(4), 1560–1571.

Jaeger, D. A., J. Ruist, and J. Stuhler (2018, February). Shift-Share Instruments and the Impact of Immigration. Working Paper 24285, National Bureau of Economic Research.

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.

Kim, R. and J. Vogel (2021, March). Trade Shocks and Labor Market Adjustment. *American Economic Review: Insights 3*(1), 115–130.

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.

Lang, K. and J.-Y. K. Lehmann (2012, December). Racial Discrimination in the Labor Market: Theory and Empirics. *Journal of Economic Literature 50*(4), 959–1006.

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.

Meyer, B. D. (1990). Unemployment Insurance and Unemployment Spells. *Econometrica 58*(4), 757–782.

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.

Neal, D. (2005, January). Why Has Black-White Skill Convergence Stopped? Technical Report 11090, National Bureau of Economic Research, Inc.

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.

Ridder, G. (1990). The Non-Parametric Identification of Generalized Accelerated Failure-Time Models. *The Review of Economic Studies 57*(2), 167–181.

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.

Ruggles, S., S. Flood, S. Foster, R. Goeken, J. Pacas, M. Schouweiler, and M. Sobek (2021). IPUMS USA: Version 11.0.

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.

Smith, J. P. and F. R. Welch (1977). Black/White Male Earnings and Employment: 1960-70. In *NBER Chapters*, pp. 233–302. National Bureau of Economic Research, Inc.

Topalova, P. (2010, October). Factor Immobility and Regional Impacts of Trade Liberalization: Evidence on Poverty from India. *American Economic Journal: Applied Economics 2*(4), 1–41.

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.

Wilson, W. J. (1987). *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy, Second Edition.* University of Chicago Press.