Title
Heterogeneity and Unemployment Dynamics

Author
Ahn, Hie Joo

Publication Date
2015

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Heterogeneity and Unemployment Dynamics

A dissertation submitted in partial satisfaction of the requirements for the degree
Doctor of Philosophy

in

Economics

by

Hie Joo Ahn

Committee in charge:

Professor James D. Hamilton, Chair
Professor Thomas H. Baranga
Professor Marjorie Flavin
Professor Gordon H. Hanson
Professor Ronghui Xu

2015

The dissertation of Hie Joo Ahn is approved, and it is acceptable in quality and form for publication on microfilm:

Chair

University of California, San Diego

2015

DEDICATION

To my parents, Youngwook Ahn and Youngsook Choi

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

In my writing dissertation, I have incurred many debts which I am glad to have this opportunity to acknowledge. First of all, I am deeply grateful to professor James D. Hamilton for his invaluable guidance, inspiring comments, and unlimited encouragement throughout my graduate studies at the University of California, San Diego. I learned from him not only how to do research in economics but also never to become discouraged or to give up on the path of pursuing a dream. I am tremendously honored to be one of the students of the advisor whose priority is the success of students and who himself always shows to students the work-ethic and humbleness of a great scholar. At the end of this journey, I became to think that I would like to help others as he does and give back to the world what I have learned from him, just because he have given me so much of his time, effort and energy. Without his support, none of my academic achievement would have been possible.

I am also indebted to professor Marjorie Flavin and Gordon Hanson who has guided and inspired me in the path of developing the dissertation. Their priceless advice have been vital sources of completing this dissertation. I would like to thank professor Thomas Baranga, Johannes Wieland, David Lagakos, Julie Cullen, Gordon Dahl and Roger Gordon for their insightful comments and advice.

In addition, I would like to show my gratitude to the teachers in Seoul National University who led me to pursue a Ph.D degree in economics. Professor Suk-Chul Yoon is the teacher who first suggested me to become an academic scholar. Professor Jaibeom Kim and Sung Hwan Oh helped me tremendously to begin this journey back in Korea. Without them, this amazing journey would have not been even possible.

I would like to thank my colleagues and friends in UCSD, including Ling Shao, Lawrence Schmidt, Myungkyu Shim, Sung Je Byun, Fan Xia, Heea Jung, Aeimit Lakdawala, Albert Yoon, Min Sung Kim, Rafael Acevedo and Suzi Harlow. Their help and support have been invaluable to me. Lastly, I will not forget my old and new office mates - Troy Kravitz, Aislinn Bohren, Jarrod Welch, Hyeong-Ho Moon, Claudio

Labanca, Yaein Baek and Victoria Xie.

I am very grateful to the colleagues and friends that I worked with at the Bank of Korea. I began to learn macroeconomics, monetary policy and econometrics from the great economists there. In particular, I would like to thank economist Kwangmin Park, Seungwon Kim, Sangdai Ryoo, Sosang Moon and Heedon Kang for their invaluable guidance. While learning from and working with them, I became interested in empirical macroeconomics. Most of all, I am greatly grateful to my dear friends, Jiwon Park, Sehyun Ahn, Won Shim and Jongsang Lee for their support and encouragement throughout this journey.

Most of all, my highest appreciation goes to my parents, Youngwook Ahn and Youngsook Choi, for their love and endless support. They are the reason I have survived my journey at UCSD and are the source of my everyday energy. I am also sincerely grateful to my family - my brother, Yangmo, sister in law, Seungyoun and my grand mother. I dedicate this dissertation, the six years of my life, to my parents.

Looking back and going forward, I have the three best luck in my life. The first is to be the daughter of my parents. The second is to be a friend and a colleague of the amazing people that I have met in this journey. The last one is to be one of the students of professor James D. Hamilton.

Chapter 1 is coauthored with James D. Hamilton. Chapter 1 is in preparation for submission. I thank professor James D. Hamilton for his permission to use the co-authored paper in my dissertation Chapter. Chapter 2 and 3 are in preparation for submission.

VITA

2004	B. S. in Business Administration *Magna cum laude*, Seoul National University, Seoul, Korea
2010	M. A. in Economics, University of California, San Diego
2015	Ph. D. in Economics, University of California, San Diego

ABSTRACT OF THE DISSERTATION

Heterogeneity and Unemployment Dynamics

by

Hie Joo Ahn
Doctor of Philosophy in Economics

University of California, San Diego, 2015

Professor James D. Hamilton, Chair

This dissertation consists of three papers about unemployment dynamics. The first chapter is ”Heterogeneity and unemployment dynamics”, the second chapter is ”The role of observed and unobserved heterogeneity in the duration of unemployment spells” and the last chapter is ”Forecasting unemployment using Dynamic Model Adaptation”.

The first chapter develops new estimates of flows into and out of unemployment that allow for unobserved heterogeneity across workers as well as direct effects of unemployment duration on unemployment-exit probabilities. Unlike any previous paper in this literature, we develop a complete dynamic statistical model that allows us to measure the contribution of different shocks to the short-run, medium-run, and long-run variance of unemployment as well as to specific historical episodes. We find that changes in the inflows of newly unemployed are the key driver of economic recessions and identify an increase in permanent job loss as the most important factor.

Then the second chapter explores the role of observed and unobserved hetero-

geneity in explaining both cross-sectional differences across individuals in the duration of unemployment as well as changes in the average duration of unemployment over time. Using CPS micro data I construct for each month the number of individuals who have been looking for work for 1 month, the number looking for work for 2-3 months, the number looking for 4-6 months, and so on, for people grouped according to a variety of observable characteristics. I use a dynamic accounting identity to infer from these vector-valued time series changes in inflows and outflows of different unobserved types of workers within a given observed category. I propose new strategies to explicitly quantify the contribution of unobserved heterogeneity to unemployment duration in the aggregate as well as across individuals. Unobserved heterogeneity explains about one third of the aggregate dispersion in ongoing duration spells of unemployment and 40% of the cross-sectional dispersion in completed duration spells over the 1980-2013 period. The compositional shift of unobserved types is a crucial factor raising the mean duration in progress during the Great Recession. By contrast, observed heterogeneity makes only a minor contribution to either cross-sectional or time-series variation.

The last chapter proposes a new method of combining forecasts based on the recent performance of out-of-sample forecasts for forecasting the U.S. unemployment rate. At every period, a forecaster chooses a single model of which the recent out-of-sample forecasts yields the smallest squared error among a given set of forecasting models to make multiple-period ahead forecasts. The proposed combination method produces more accurate forecasts than existing model averaging methods and the Greenbook forecasts.

Chapter 1

Heterogeneity and Unemployment Dynamics

Abstract. This paper develops new estimates of flows into and out of unemployment that allow for unobserved heterogeneity across workers as well as direct effects of unemployment duration on unemployment-exit probabilities. Unlike any previous paper in this literature, we develop a complete dynamic statistical model that allows us to measure the contribution of different shocks to the short-run, medium-run, and long-run variance of unemployment as well as to specific historical episodes. We find that changes in the inflows of newly unemployed are the key driver of economic recessions and identify an increase in permanent job loss as the most important factor.

1.1 Introduction

What accounts for the sharp spike in the unemployment rate during recessions? The answer traditionally given by macroeconomists was that falling product demand leads firms to lay off workers, with these job separations a key driver of economic downturns. That view has been challenged by Hall (2005) and Shimer (2012), among others, who argued that cyclical fluctuations in the unemployment rate are instead primarily driven by declines in the job-finding rates for unemployed workers.

This debate has become particularly important for understanding the Great Recession and its aftermath. In June 2011– two years into the recovery– the unemployment rate still stood at 9.1%, higher than the peak in any postwar recession other than 1982. Even more troubling, the average duration of those unemployed at that time was 40 weeks, about twice the highest value reached in any month over 1947-2005. Of those workers who had been unemployed for less than one month in June 2011, only 57% were still unemployed the next month. By contrast, of those who had been unemployed for more than 6 months as of June 2011, 93% were still unemployed the following month.[1]

This fact that the long-term unemployed find jobs or leave the labor force more slowly than others is a strikingly consistent feature in the postwar data, and could be fundamental for understanding the respective contributions of unemployment inflows and outflows during recessions. For example, workers who lose their jobs due to involuntary permanent separation may have a more difficult time finding new jobs than people who quit voluntarily (Bednarzik, 1983; Fujita and Moscarini, 2013). If the number of involuntary separations increases during a recession, it could show up as what other researchers have interpreted as a fall in the job-finding rate and increase in the duration of unemployment even if the key driver of the recession was the increase in involuntary separations.

The phenomenon that unemployment exit rates fall with the duration of unemployment has been widely studied, with explanations falling into two broad categories. One possibility is that the experience of being unemployed for a longer period of time directly changes the characteristics of a fixed individual. Following van den

[1] The values for p^1_{t+1} and $p^{7.+}_{t+1}$ were calculated from

$$p^1_{t+1} = \frac{U^1_t - U^2_{t+1}}{U^1_t},\ p^{7.+}_{t+1} = \frac{U^{7.+}_t - (U^{7.+}_{t+1} - U^7_{t+1})}{U^{7.+}_t}$$

for U^n_t the number unemployed with duration n months at t. The reported series are seasonally adjusted with X-12-ARIMA.

Berg and van Ours (1996) we will refer to this possibility as ”genuine duration dependence”. Individuals lose human capital the longer they are unemployed (Acemoglu, 1995; Ljungqvist and Sargent, 1998), employers may statistically discriminate against those who have been unemployed for longer (Eriksson and Rooth, 2014; Kroft, Lange, and Notowidigdo, 2013), and individuals may search less the longer they have been unemployed (Faberman and Kudlyak, 2014). We will refer to such negative genuine duration dependence, that is, a condition where a longer period spent in unemployment directly reduces the probability of finding a job, as ”unemployment scarring.” Another possibility is positive genuine duration dependence. For example, the longer a person has been unemployed, the more willing they may be to accept a low-paying job or simply to drop out of the labor force. Katz and Meyer (1990a,b) argued that such effects may become important as unemployment benefits become exhausted. We will refer to the possibility that the probability of exiting unemployment increases as a consequence of a longer duration of unemployment as ”motivational” effects.

A quite different explanation for the differences in unemployment exit probabilities across the different duration categories is that there are important differences across job-seekers from the very beginning, arising for example from differences in the reason the individuals left their previous job or in differences in ex ante abilities or motivation across workers. The longer an individual is observed to have been unemployed, the greater the chance that the individual is a member of a group whose unemployment exit probabilities were low to begin with. That such cross-sectional heterogeneity might be important for the question studied by Hall and Shimer was recognized as far back as Darby, Haltiwanger, and Plant (1986), who argued that heterogeneity accounted for falling job-finding rates during recessions in a manner consistent with the traditional macroeconomic interpretation of recessions. A number of researchers have tried to investigate this hypothesis by looking at differences across job seekers in observable characteristics such as demographics, education, industry, occupation, geographical region, and reason for unemployment. Baker (1992), Shimer (2012), and Kroft, Lange, Notowidigdo, and Katz (2014) found that such variables contributed little to variation over time in long-term unemployment rates, while Aaronson, Mazumder and Schechter (2010), Bachmann and Sinning (2012), Barnichon and Figura (2013), Hall (2014), and Hall and Schulhofer-Wohl (2014) documented important differences across observable characteristics. Elsby, Michaels and Solon (2009) found that incorporating observable heterogeneity reduced the imputed role of cyclical variation in unemployment exit rates.

However, no two individuals with the same coarse observable characteristics are in fact identical. It seems undeniable that a given pool of unemployed individuals that conditions on any set of observed characteristics is likely to become increasingly represented by those with lower ex ante exit probabilities the longer the period of time for which the individuals have been unemployed. Most of the above studies assume that conditional on observable characteristics, unemployed individuals are identical in terms of their transition probabilities into and out of unemployment. The result is that the imputed exit probabilities are determined solely from the most recent labor force statistics as if every month was a new steady state of the economy, not taking into account the fact that each individual has a unique history of unemployment. This approach misses a key feature of economic recessions and unemployment dynamics. Once one acknowledges heterogeneity across workers, the pool of those looking for work at a given point in time– and therefore the exit rates for individuals in that group– depends on the specific history of conditions whereby those individuals came to be unemployed. This means that more information than the current month's labor force statistics is necessary to account for the different histories of unemployed individuals and thus to credibly analyze the contributions of the inflows and outflows.

A large literature has attempted to separate genuine duration dependence from cross-sectional heterogeneity based on observable covariates for unemployed workers (Heckman and Singer, 1984) and the difference between calendar time and individual duration (van den Berg and van Ours, 1996). Our approach is closest to that in Hornstein (2012) who used dynamic accounting identities to track directly the way the characteristics of the pool of unemployed workers with unobserved cross-sectional heterogeneity would depend on the previous history. Hornstein used a minimum-distance estimation with identification achieved by smoothing penalties and only considered negative genuine duration dependence. By contrast, our paper provides a completely specified dynamic model that allows for both time variation in unobserved cross-sectional differences in worker characteristics as well as nonmonotonic genuine duration dependence.

Our approach offers a number of other advantages over previous studies. We provide a statistical framework for generating variance decompositions as well as historical decompositions of observed changes in unemployment over any subsample. In doing so we resolve a key shortcoming in much of the previous literature. Most previous studies used correlations between unemployment and the steady-state unemployment rate predicted by either inflows or outflows to draw conclusions about how much of the variation in unemployment is due to each factor. However, the unemployment rate is highly serially correlated and possibly nonstationary. What do we even mean by its variance, and how do we distinguish between the contribution to this variance of short-term versus long-term influences? Previous studies often addressed these issues by using some kind of detrending procedures. By contrast, our paper develops a complete statistical model with nonstationary driving processes, which as a by-product generates a forecast of unemployment at any horizon in the future. Since the forecast error at any specified horizon has a stationary distribution and well defined mean squared error whether or not the underlying process is nonstationary, as in den Haan (2000) we can calculate the fraction of the variance in unanticipated changes in unemployment over any horizon that is attributable to the various shocks in the model. This allows us to measure the dynamic contributions of different factors to unemployment and allows us to make very clear statements about the importance for short-run, medium-run, and long-run dynamics as well as over specific historical episodes. This is one of the key innovations of our approach and is entirely new to this literature.

In Section 1 we introduce the data that we will use in this analysis based on the number of job-seekers each month who report they have been looking for work at various search durations. We describe the accounting identities that will later be used in our full dynamic model and use average values of observable variables over the sample to explain the intuition for how such duration data can be used to separately identify cross-sectional heterogeneity and genuine duration dependence. We also use these calculations to illustrate why cross-sectional heterogeneity appears to be more important than genuine duration dependence in terms of explaining the broad features of these data.

In Section 2 we extend this framework into a full dynamic model in which we postulate the existence of two types of workers at any given date. Type H workers have a higher ex ante probability of exiting unemployment than type L workers, and all workers are also subject to potential scarring or motivational effects. Our model postulates that the number of newly unemployed individuals of either type, as well as the probability for each type of exiting the pool of unemployed at each date, evolve over time according to unobserved random walks. We show how one can calculate the likelihood function for the observed unemployment data and an inference about each of the state variables at every date in the sample using an extended Kalman filter.

Empirical results are reported in Section 3. Broken down in terms of inflows versus outflows, we find that variation over time in the inflows of the newly unemployed are more important than outflows from unemployment in accounting for errors in predicting aggregate unemployment at all horizons. Broken down in terms of types of workers, inflow and outflow probabilities for type L workers are more important than those for type H workers, and account for 90% of the uncertainty in predicting unemployment 2 years ahead. In recessions since 1990, shocks to the inflows of type L workers were the most important cause of rising unemployment during the recession. We find a non-monotonic contribution of genuine duration dependence, with scarring effects dominating up to 1 year but motivational effects apparent for those unemployed longer than a year.

We also highlight the key features of the data that lead us to these conclusions. At the end of recessions since 1990, the number of newly unemployed began to decline even as the total number of unemployed continued to rise (see Figure 1.1). In terms of the dynamic accounting identities, this must mean that either there was an increase in the inflows of type L workers or a change in the outflow probabilities for either group. In the first case, the effects would not show up in the 4-6 month category until 4 months later, in the 7-12 month category until 7 months later, and so on. If it was a change in the type H outflow probabilities, it would show up immediately in the shorter duration groups but not in the longer duration groups since the latter have few individuals of this type left, while a change in type L outflow probabilities would show up immediately and most dramatically as a change in the longer-duration groups. During the later stages of the Great Recession, for example, the changes primarily follow the first pattern, suggesting that an increased inflow of type L job-seekers was the most important development, though we find some evidence that changes in outflow probabilities for type L individuals also contributed to the slow recovery.

We offer interpretations of our findings in Section 4 by relating our estimated series to those available from other sources. We conclude that a key difference between type L and type H workers is the circumstances under which they left their previous job. Our imputed series for newly unemployed type L workers behaves very similarly to separate measures of the number of new job-seekers who were involuntarily separated from their previous job for a reason other than what was described as a temporary layoff. We conclude that, consistent with the traditional interpretation of business cycles, the key reason that unemployment spikes during recessions is a change in the circumstances under which individuals lose their jobs.

In Section 5 we investigate the robustness of our approach to various alternative specifications, including alternative methods to account for the change in the CPS questionnaire in 1994, allowing for correlation between the innovations of the underlying structural shocks in our model, and the possible effects of time aggregation. While such factors could produce changes in some of the details of our inference, our overall conclusions (summarized in Section 6) appear to be quite robust.

1.2 Observable implications of heterogeneity

The Bureau of Labor Statistics reports for each month t the number of Americans who have been unemployed for less than 5 weeks. Our baseline model is specified at the monthly frequency, leading us to use the notation U_t^1 for the above BLS-reported magnitude, indicating these individuals have been unemployed for 1 month or less as of month t. BLS also reports the number who have been unemployed for between 5 and 14 weeks (or 2-3 months, denoted $U_t^{2.3}$), 15-26 weeks ($U_t^{4.6}$) and longer than 26 weeks ($U_t^{7.+}$). We also used the raw CPS micro data from which these aggregates were constructed to break down the last group further into those unemployed with duration 7-12 months ($U_t^{7.12}$) and those with longer than 1 year ($U_t^{13.+}$).[2]

The data used in our analysis are graphed in Figure 1.1. Our purpose in this paper is to explore what variation in these duration-specific components U_t^x across time can tell us about unemployment dynamics. Our focus will be on the following question– of those individuals who are newly unemployed at time t, what fraction will still be unemployed at time $t + k$? We presume that the answer to this question depends not just on aggregate economic conditions over the interval $(t, t + k)$ but also on the particular characteristics of those individuals. Let w_{it} denote the number of people of type i who are newly unemployed at time t, where we interpret

$$U_t^1 = \sum_{i=1}^{I} w_{it}. \tag{1.1}$$

[2] See Appendix for further details of data construction.

We define $P_{it}(k)$ as the fraction of individuals of type i who were unemployed for one month or less as of date $t-k$ and are still unemployed and looking for work at t. Note that in order for someone to have been unemployed for 2-3 months at time t, they either must have been newly unemployed at $t-1$ and still looking for a job at t, or they were newly unemployed at $t-2$ and still looking at $t-1$ and t:

$$U_t^{2.3} = \sum_{i=1}^{I} \left[w_{i,t-1} P_{i,t}(1) + w_{i,t-2} P_{i,t}(2) \right]. \tag{1.2}$$

Likewise

$$U_t^{4.6} = \sum_{i=1}^{I} \sum_{k=3}^{5} \left[w_{i,t-k} P_{i,t}(k) \right] \tag{1.3}$$

$$U_t^{7.12} = \sum_{i=1}^{I} \sum_{k=6}^{11} \left[w_{i,t-k} P_{i,t}(k) \right] \tag{1.4}$$

$$U_t^{13.+} = \sum_{i=1}^{I} \sum_{k=12}^{47} \left[w_{i,t-k} P_{i,t}(k) \right] \tag{1.5}$$

where following Hornstein (2012) we terminate the calculations after 4 years of unemployment.

To get some intuition about what observation of the U_t^x aggregates can tell us about w_{it} and $P_{it}(k)$, we consider in this section some simple time-invariant examples. Suppose that none of the above magnitudes depended on time. How much could we learn from the average values of U^x? As a first simplest case, suppose that everyone was identical before they became unemployed, and how long they have been unemployed had no consequences for the probability of finding a job next month. In other words, our first case assumes that the fraction of unemployed individuals at time $t-1$ who are still unemployed at t is some constant p regardless of the date or the individual's circumstances. For this special case we would have

$$P_{it}(k) = p^k \text{ for all } i, t. \tag{1.6}$$

In this case, equations (1.1) and (1.2) imply

$$U^{2.3} = U^1(p + p^2).$$

Thus under the above assumptions, just by observing the average number of people newly unemployed and the average number unemployed of duration 2-3 months, we could obtain an estimate of p:

$$U^{2.3}/U^1 = p + p^2. \tag{1.7}$$

Table 1.1 reports that on average over 1976-2013, there were 3,210 thousand Americans whose duration of unemployment is less than 5 weeks and 2,303 thousand reporting an unemployment spell that so far had continued for 2 or 3 months. Equation (1.7) would then imply an estimate $\hat{p} = 0.484$ for the average fraction of unemployed individuals who would still be unemployed one month later.

However, note that these same homogeneity assumptions would also imply

$$U^{4.6}/U^1 = (p^3 + p^4 + p^5) \tag{1.8}$$

If indeed $\hat{p} = 0.484$, equation (1.8) would predict a value for $U^{4.6}$ of 625, whereas we see in Table 1.1 that the actual value is 1,238. If workers really were all identical, we would expect to see far fewer individuals whose unemployment spells lasted longer than 3 months than we do in the data. The indicated conclusion is that those individuals who have been unemployed for 3 months on average have different characteristics (and a lower probability of finding a job next month) than the typical worker who has only been unemployed for 1 month.

Consider next a generalization of the above special case in which there is still no heterogeneity across workers and no aggregate variation ($w_{it} = w$ for all i and t), but we do allow for genuine duration dependence arising from factors referred to in the introduction as scarring or motivational effects. Specifically, suppose that the fraction of unemployed individuals who had been unemployed for τ months as of the previous month who are still unemployed in the current month is given by some function $p(\tau)$. Unemployment scarring would correspond to $p(\tau)$ being an increasing function of τ, while if the motivational effect dominates, $p(\tau)$ would be a decreasing function of τ. In this case (1.6) generalizes to

$$P_{i,t}(k) = p(1)p(2)\cdots p(k),$$

and (1.7) and (1.8) become

$$U^{2.3}/U^1 = p(1) + p(1)p(2) \tag{1.9}$$

$$U^{4.6}/U^1 = p(1)p(2)p(3)[1 + p(4) + p(4)p(5)]. \tag{1.10}$$

Suppose we were willing to choose a parametric form for the function $p(\tau)$ as in Katz and Meyer (1990b):

$$p(\tau) = \exp\{-\exp[x + d(\tau - 1)]\} \quad \text{for} \quad \tau = 1, 2, 3, ... \tag{1.11}$$

One benefit of this functional form is that $p(\tau)$ is guaranteed to be between 0 and 1 for any values of x, d, or τ, a feature that will be helpful when we get to a generalization of this set-up in the following section in which we will allow for variation in x over time. A negative value for the parameter d would correspond to unemployment scarring whereas $d > 0$ would represent a motivational effect. Substituting (1.11) into (1.9) and (1.10) produces a system of 2 equations which we can solve numerically for x and d as functions of the observed values for $U^{2.3}/U^1$ and $U^{4.6}/U^1$ given in Table 1.1. The solution turns out to be $x = -0.252$ and $d = -0.296$. The negative value for d is supportive of the unemployment scarring hypothesis, consistent with the inference above that it is not possible to reconcile the relative values of U^1, $U^{2.3}$, and $U^{4.6}$ without some kind of heterogeneity.

The problem with relying purely on genuine duration dependence is seen if we try to use the inferred values for the function $p(\tau)$ to estimate the value for $U^{7.12}$ and $U^{13.+}$. These turn out to be 1,235 and 5,617, respectively. Note in particular that this predicted value for $U^{13.+}$ is far larger than the observed value of 636. Any "unemployment scarring" that is operating on workers who have been unemployed for longer than 12 months seems to be very different from that experienced by those unemployed for only 2-5 months. One possibility is that the functional form (1.11) is misspecified. However, another possibility worth exploring is that there are important ex-ante differences between individuals, with some likely to get a job more quickly than others. As a result of these ex-ante differences, when one looks at a given pool of workers who have been unemployed for τ months, a larger fraction of the pool is going to be accounted for by those with lower job-finding probabilities the larger the value of τ.

To illustrate how this could work, suppose there are $I = 2$ types of workers, which we will label type H and type L in anticipation of the normalization that type L workers have a lower probability of exiting unemployment. With cross-sectional heterogeneity but no genuine duration dependence, equation (1.6) becomes

$$P_{it}(k) = p_i^k \text{ for all } t. \tag{1.12}$$

Substituting (1.12) into (1.1) through (1.4) gives a system of 4 equations which we can solve for the 4 unknowns (w_H, w_L, p_H, p_L) as functions of the observed averages $(U^1, U^{2.3}, U^{4.6}, U^{7.12})$. The solution turns out to be $w_L = 679$, $w_H = 2,531$, $p_L = 0.848$ and $p_H = 0.360$. The type H workers comprise a very high fraction, 78.8%, of the initial pool of unemployed U^1. But because they are more likely to be the ones who find jobs quickly, there are fewer type H workers included in group $U^{2.3}$ and even fewer in $U^{4.6}$ and $U^{7.12}$. This changing composition can account for the feature of the data that a specification without cross-sectional heterogeneity would attribute to unemployment scarring.

We can also use these values for (w_H, w_L, p_H, p_L) in equation (1.5) to get a predicted value for $U^{13.+}$ of 614, not far from the observed value of 636. These calculations suggest that cross-sectional heterogeneity is a more promising potential explanation of unemployment dynamics than genuine duration dependence.

Finally, we note that it is possible to estimate a model that allows for both cross-section heterogeneity and genuine duration dependence. Suppose we generalize (1.11) to

$$p_i(\tau) = \exp\{-\exp[x_i + d(\tau - 1)]\} \quad \text{for} \quad \tau = 1, 2, 3, ... \tag{1.13}$$

for $i = H$ or L. Equations (1.1)-(1.5) then give us a system of 5 equations in the 5 unknowns (w_H,w_L, x_H, x_L, d). The solutions turn out to be $w_L = 683$, $w_H = 2,528$, $p_L(1) = 0.846$, $p_H(1) = 0.360$ and $d = -0.003$. These estimates allow little role for genuine duration dependence, with the slightly negative value for d now implying that unemployment scarring may be more important than motivation. However, this effect is quite tiny: the probability of exiting unemployment goes down less than 0.001 as the duration of unemployment increases by 1 month.

The above calculations demonstrate that given parametric assumptions, it is possible to come up with estimates of the relative importance of cross-sectional heterogeneity and genuine duration dependence in explaining why some individuals remain unemployed for so long. However, the examples discussed so far were quite limited in that we assumed that all parameters were constant over time. More generally, the observed values of U_t^x for some particular t could tell us about the portions and probabilities for different types of workers at that date if we knew something about the prior history. By assuming that the magnitudes of $w_{Ht}, w_{Lt}, x_{Ht}, x_{Lt}$ evolve gradually over time, we can use observations of $U_t^1, U_t^{2.3}, U_t^{4.6}, U_t^{7.12}, U_t^{13.+}$ in a nonlinear state-space model to form an inference about the changing values of $w_{Ht}, w_{Lt}, x_{Ht}, x_{Lt}$ and separately infer the contribution of time-invariant genuine duration dependence, as we demonstrate in the next section.

1.3 Dynamic formulation

The previous section discussed a static example in order to illustrate how cross-sectional heterogeneity and genuine duration dependence can be identified from observed reports of unemployment duration. However, our main interest lies in the contribution of the two types of heterogeneity to unemployment dynamics. Here we set up a state-space model where the dynamic behavior of the observed vector $y_t = (U_t^1, U_t^{2.3}, U_t^{4.6}, U_t^{7.12}, U_t^{13.+})'$ is determined as a nonlinear function of latent dynamic variables– the inflows and outflow probabilities for unemployed individuals with unobserved heterogeneity. Due to the nonlinear nature of the resulting model, we draw inference on the latent variables using the extended Kalman filter.

1.3.1 State-space representation

We assume smooth variation over time for the latent variables of interest, $w_{Ht}, w_{Lt}, x_{Ht}, x_{Lt}$, with each assumed to follow an unobserved random walk, e.g.,

$$w_{Ht} = w_{H,t-1} + \epsilon_{Ht}^{w}.$$

A random walk is a flexible and parsimonious way of modeling time-varying latent variables. As in the previous steady-state example, we consider 4 years to be the maximum duration.[3] Suppose that we observe the elements of y_t with measurement error

[3] Allowing a different maximum duration of unemployment, for instance, 3 years, does not change the results significantly.

$r_t = (r_t^1, r_t^{2.3}, r_t^{4.6}, r_t^{7.12}, r_t^{13.+})'$. The measurement equations are thus written as follows,

$$U_t^1 = \sum_{i=H,L} w_{it} + r_t^1 \tag{1.14}$$

$$U_t^{2.3} = \sum_{i=H,L} [w_{i,t-1}P_{i,t}(1) + w_{i,t-2}P_{i,t}(2)] + r_t^{2.3} \tag{1.15}$$

$$U_t^{4.6} = \sum_{i=H,L} \sum_{k=3}^{5} [w_{i,t-k}P_{i,t}(k)] + r_t^{4.6} \tag{1.16}$$

$$U_t^{7.12} = \sum_{i=H,L} \sum_{k=6}^{11} [w_{i,t-k}P_{i,t}(k)] + r_t^{7.12} \tag{1.17}$$

$$U_t^{13.+} = \sum_{i=H,L} \sum_{k=12}^{47} [w_{i,t-k}P_{i,t}(k)] + r_t^{13.+} \tag{1.18}$$

where

$$P_{i,t}(j) = p_{i,t-j+1}(1)p_{i,t-j+2}(2)...p_{i,t}(j). \tag{1.19}$$

We assume that for type i workers who have already been unemployed for τ months as of time $t-1$, the fraction who will still be unemployed at t is given by

$$p_{i,t}(\tau) = \exp[-\exp(x_{i,t} + d_\tau)] \quad \text{for} \quad \tau = 1, 2, 3, ... \tag{1.20}$$

where d_τ determines the nature of genuine duration dependence experienced by an unemployed individual with duration of unemployment τ months and x_{it} is a time-varying magnitude influencing the unemployment exit probability for all workers of type i regardless of their duration. Like the inflows w_{LT} and w_{Ht}, we assume that the parameters x_{Lt} and x_{Ht} governing outflow probabilities also follow a random walk. Note that because we have assumed that the genuine-duration dependence effects as summarized by d_τ are time-invariant and that the type-specific effects x_{it} evolve smoothly over time, it is possible to estimate a different value for the parameter d_τ for each τ. We investigated a number of different specifications for d_τ and found the best fit using linear splines at $\tau = 6$ and $\tau = 12$ which we use for the baseline analysis:

$$d_\tau = \begin{cases} \delta_1(\tau - 1) & \text{for } \tau < 6 \\ \delta_1[(6-1)-1] + \delta_2[\tau - (6-1))] & \text{for } 6 \leq \tau < 12 \\ \delta_1[(6-1)-1] + \delta_2[(12-1)-(6-1)] + \delta_3[\tau - (12-1)] & \text{for } 12 \leq \tau. \end{cases} \tag{1.21}$$

Positive δ_j for $j = 1, 2, 3$ imply motivational effects while negative values imply unemployment scarring over the relevant duration ranges.

We can arrive at the likelihood function for the observed data $\{y_1, ..., y_T\}$ by assuming that the vector of measurement errors r_t are independent Normal, where R_1, $R_{2.3}$, $R_{4.6}$, $R_{7.12}$ and $R_{13.+}$ are the standard deviations of $r_t^1, r_t^{2.3}, r_t^{4.6}$, $r_t^{7.12}$ and $r_t^{13.+}$ respectively:

$$r_t \sim N(0, R)$$

$$\underbrace{R}_{5\times 5} = \begin{bmatrix} R_1^2 & 0 & 0 & 0 & 0 \\ 0 & R_{2.3}^2 & 0 & 0 & 0 \\ 0 & 0 & R_{4.6}^2 & 0 & 0 \\ 0 & 0 & 0 & R_{7.12}^2 & 0 \\ 0 & 0 & 0 & 0 & R_{13.+}^2 \end{bmatrix}.$$

Let ξ_t be the vector $(w_{Lt}, w_{Ht}, x_{Lt}, x_{Ht})'$ and $\epsilon_t = (\epsilon_{Lt}^w, \epsilon_{Ht}^w, \epsilon_{Lt}^x, \epsilon_{Ht}^{x\prime})'$. Our assumption that the latent factors evolve as random walks would be written as

$$\underbrace{\xi_t}_{4\times 1} = \xi_{t-1} + \underbrace{\epsilon_t}_{4\times 1} \tag{1.22}$$

$$\underbrace{\epsilon_t}_{4\times 1} \sim N(\underbrace{0}_{4\times 1}, \underbrace{\Sigma}_{4\times 4})$$

$$\underbrace{\Sigma}_{4\times 4} = \begin{bmatrix} (\sigma_L^w)^2 & 0 & 0 & 0 \\ 0 & (\sigma_H^w)^2 & 0 & 0 \\ 0 & 0 & (\sigma_L^x)^2 & 0 \\ 0 & 0 & 0 & (\sigma_H^x)^2 \end{bmatrix}.$$

In Section 5 we will also report results for a specification in which the shocks are allowed to be contemporaneously correlated.

Since the measurement equations (1.14)-(1.18) are a function of $\{\xi_t, \xi_{t-1}, ..., \xi_{t-47}\}$, the state equation should describe the joint distribution of ξ_t's from $t-47$ to t, where I and 0 denote a (4×4) identity and zero matrix, respectively:

$$\underbrace{\begin{bmatrix} \xi_t \\ \xi_{t-1} \\ \xi_{t-2} \\ \vdots \\ \xi_{t-46} \\ \xi_{t-47} \end{bmatrix}}_{192\times 1} = \underbrace{\begin{bmatrix} \underbrace{I}_{4\times 4} & \underbrace{0}_{4\times 4} & 0 & 0 & ... & 0 & 0 & 0 \\ I & 0 & 0 & 0 & ... & 0 & 0 & 0 \\ 0 & I & 0 & 0 & ... & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & ... & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & ... & I & 0 & 0 \\ 0 & 0 & 0 & 0 & ... & 0 & I & 0 \end{bmatrix}}_{192\times 192} \underbrace{\begin{bmatrix} \xi_{t-1} \\ \xi_{t-2} \\ \xi_{t-3} \\ \vdots \\ \xi_{t-47} \\ \xi_{t-48} \end{bmatrix}}_{192\times 1} + \underbrace{\begin{bmatrix} \underbrace{\epsilon_t}_{4\times 1} \\ \underbrace{0}_{4\times 1} \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}}_{192\times 1}. \quad (1.23)$$

1.3.2 Estimation

Our system takes the form of a nonlinear state space model in which the state transition equation is given by (2.9) and observation equation by (1.14)-(1.18) where $P_{i,t}(j)$ is given by (1.19) and $p_{i,t}(\tau)$ by (1.20). Our baseline model has 12 parameters to estimate, namely the diagonal terms in the variance matrices Σ and R and the parameters governing genuine duration dependence, $\delta_{1,}$ δ_2 and δ_3. Because the observation equation is nonlinear in x_{it}, the extended Kalman filter can be used to form the likelihood function for the observed data $\{y_1, ..., y_T\}$ and form an inference about the unobserved latent variables $\{\xi_1, ..., \xi_T\}$, as detailed in Appendix. Inference about historical values for ξ_t provided below correspond to full-sample smoothed inferences, denoted $\hat{\xi}_{t|T}$.

1.4 Results for the baseline specification

We estimated parameters for the above nonlinear state-space model using seasonally adjusted monthly data on $y_t = (U_t^1, U_t^{2.3}, U_t^{4.6}, U_t^{7.12}, U_t^{13.+})'$ for $t =$ January 1976 through December 2013. Figure 1.2 plots smoothed estimates for $p_{i,t}(1)$, the probability that a newly unemployed worker of type i at $t-1$ will still be unemployed at t. These average 0.34 for type H individuals and 0.81 for type L individuals. The probabilities of type H remaining unemployed rise during the early recessions but are less cyclical in the last two recessions. By contrast, the continuation probabilities for type L individuals rise in all recessions and continued to rise after the end of the last 3 recessions. The gap between the two probabilities increased significantly over the last 20 years.

Figure 1.3 plots inflows of individuals of each type into the pool of newly unemployed. Type H workers constitute 76% on average of the newly unemployed. Inflows of both types increase during recessions. New inflows of type H workers declined immediately at the end of every recession, but inflows of type L workers continued to rise after the recessions of 1990-91 and 2001 and were still at above-average levels 3 years after the end of the Great Recession. This changing behavior of type L workers' inflows appears to be another important characteristic of jobless recoveries. The Great Recession is unique in that the inflows of type L workers as well as the continuation probabilities reached higher levels than any earlier dates in our data set.

The combined implications of these cyclical patterns are summarized in Figure 1.4. Before the Great Recession, the share in total unemployment of type L workers fluctuated between 30% and 50%, falling during expansions and rising during and after recessions. But during the Great Recession, the share of type L workers skyrocketed to over 80%. The usual recovery pattern of a falling share of type L workers has been very slow in the aftermath of the Great Recession.

While the inflows of type H workers show a downward trend since the 1980's, those of type L workers exhibit an upward trend. This difference in the low frequency movements of the two series provides a new perspective on the secular decrease in the inflows to unemployment and the secular rise in the average duration of unemployment. Abraham and Shimer (2002) and Aaronson, Mazumder and Schechter (2010) showed that the substantial rise in average duration of unemployment between mid-1980 and mid-2000 can be explained by the CPS redesign, the aging of the population and the increased labor force attachment of women. Bleakley, Ferris and Fuhrer (1999) concluded that the downward trend in inflows can be explained by reduced churning during this period. Figure 1.3 shows that the downward trend in the inflows is mainly driven by type H workers. The increased share of type L inflows contributed to the rise in the average duration of unemployment since the 1980's. This suggests that unobserved heterogeneity is important in accounting for low frequency dynamics in the labor market as well as those for business cycle frequencies.

Table 1.2 provides parameter estimates for our baseline model. We find a value for δ_1, the parameter that governs genuine duration dependence for unemployment durations less than 6 months, that is near zero and statistically insignificant. The estimate of δ_2 (applying to individuals unemployed for more than 5 months and less than 1 year) is statistically significant and negative. The negative sign is consistent with the scarring hypothesis– the longer someone from either group has been unemployed, provided the duration has been 11 months or less, the more likely it is that person will be unemployed next month. On the other hand, we find a statistically significant positive value for δ_3 (unemployment lasting for a year and over). Once someone has been unemployed for more than a year, it becomes more likely as more months accumulate that they will either find a job or exit the labor force in any given month, consistent with what we have labeled motivational effects. This non-monotonic behavior of genuine duration dependence is displayed graphically in Figure 1.5.

Although the values of δ_2 and δ_3 are statistically significant, they play a relatively minor role compared to ex ante heterogeneity in accounting for differences in exit probabilities by duration of unemployment. As seen in Panel B of Figure 1.5, our estimates of genuine duration dependence imply relatively modest changes in continuation probabilities for type L workers for most horizons. And while the implications for long-horizon continuation probabilities for type H workers may appear more significant, they are empirically irrelevant, since the probability that type H workers would be unemployed for more than 12 months is so remote ($0.5^{12} = 2.4 \times 10^{-4}$).

1.4.1 Variance decomposition

Many previous studies have tried to summarize the importance of different factors in determining unemployment by looking at correlations between the observed unemployment rate and the steady-state unemployment rate predicted by each factor of interest alone; see for example Fujita and Ramey (2009) and Shimer (2012). One major benefit of our framework is that it delivers a much cleaner answer to this question in the form of variance decompositions.

Variance decomposition is a familiar method in linear VARs for measuring how much each shock contributes to the mean squared error (MSE) of an s-period-ahead forecast of a magnitude of interest.[4] Here we focus on forecasts of the total number of people unemployed. In a linear VAR, both the MSE and the portion attributable to each component are functions of population parameters that depend on the horizon s but not the date, and the sum of the contributions of each of the factors exactly equals the overall MSE.

In our case we have the simple system for the latent (4×1) vector

$$\xi_{t+1} = \xi_t + \epsilon_{t+1}$$

[4] See for example Hamilton (1994a, Section 11.5).

from which

$$\begin{aligned}\xi_{t+s} &= \xi_t + \epsilon_{t+1} + \epsilon_{t+2} + \epsilon_{t+3} + ... + \epsilon_{t+s} \\ &= \xi_t + u_{t+s}.\end{aligned}$$

Letting $y_t = (U_t^1, U_t^{2.3}, U_t^{4.6}, U_t^{7.12}, U_t^{13.+})'$ denote the (5×1) vector of observations for date t, our model implies that in the absence of measurement error
$y_t = h(\xi_t, \xi_{t-1}, \xi_{t-2}, ..., \xi_{t-47})$ where $h(\cdot)$ is a known nonlinear function. Hence

$$y_{t+s} = h(u_{t+s} + \xi_t, u_{t+s-1} + \xi_t, ..., u_{t+1} + \xi_t, \xi_t, \xi_{t-1}, ..., \xi_{t-47+s}).$$

We can take a first-order Taylor expansion of this function around $u_{t+j} = 0$ for $j = 1, 2, ..., s$,

$$y_{t+s} \simeq h(\xi_t, ..., \xi_t, \xi_t, \xi_{t-1}, ..., \xi_{t-47+s}) + \sum_{j=1}^{s} [H_j(\xi_t, \xi_t, ..., \xi_t, \xi_t, \xi_{t-1}, ..., \xi_{t-47+s})] u_{t+j}$$

for $H_j(\cdot)$ the (5×4) matrix associated with the derivative of $h(\cdot)$ with respect to its jth argument. Using the definition of u_{t+j}, this can be rewritten as

$$y_{t+s} \simeq c_s(\xi_t, \xi_{t-1}, ..., \xi_{t-47+s}) + \sum_{j=1}^{s} [\Psi_{s,j}(\xi_t, \xi_{t-1}, ..., \xi_{t-47+s})] \epsilon_{t+j} \tag{1.24}$$

for $\Psi_{s,j}(\cdot)$ a known (5×4)-valued function of $\xi_t, \xi_{t-1}, ..., \xi_{t-47+s}$. The MSE associated with an s-period-ahead forecast of y_{t+s} is then

$$\begin{aligned}
& E(y_{t+s}-\hat{y}_{t+s|t})(y_{t+s}-\hat{y}_{t+s|t})' \\
= & \sum_{j=1}^{s}[\Psi_{s,j}(\xi_t,\xi_{t-1},...,\xi_{t-47+s})]\Sigma[\Psi_{s,j}(\xi_t,\xi_{t-1},...,\xi_{t-47+s})]' \\
= & \sum_{j=1}^{s}\sum_{m=1}^{4}\Sigma_m[\Psi_{s,j}(\xi_t,\xi_{t-1},...,\xi_{t-47+s})e_m][\Psi_{s,j}(\xi_t,\xi_{t-1},...,\xi_{t-47+s})e_m]'
\end{aligned}$$

for e_m column m of the (4×4) identity matrix and Σ_m the row m, column m element of Σ. Thus the contribution of innovations of type L worker's inflows (the first element of $\epsilon_t=(\epsilon^w_{L,t},\epsilon^w_{H,t},\epsilon^x_{L,t},\epsilon^x_{H,t})'$) to the MSE of the s-period-ahead linear forecast error of total unemployment, $\mathbf{1}'y_t$, is given by

$$\mathbf{1}'\sum_{j=1}^{s}\Sigma_1[\Psi_{s,j}(\xi_t,\xi_{t-1},...,\xi_{t-47+s})e_1][\Psi_{s,j}(\xi_t,\xi_{t-1},...,\xi_{t-47+s})e_1]'\mathbf{1} \tag{1.25}$$

where $\mathbf{1}$ denotes a (5×1) vector of ones. Note that as in the constant-parameter linear case, the sum of the contributions of the 4 different structural shocks would be equal to the MSE of an s-period-ahead linear forecast of unemployment in the absence of measurement error. However, in our case the linearization is taken around time-varying values of $\{\xi_t,\xi_{t-1},...,\xi_{t-47+s}\}$. We can evaluate equation (1.25) at the smoothed inferences $\{\hat{\xi}_{t|T},\hat{\xi}_{t-1|T},...,\hat{\xi}_{t-47+s|T}\}$ and then take the average value across all dates t in the sample. This gives us an estimate of the contribution of the type L worker's inflows to unemployment fluctuations over a horizon of s months:

$$\begin{aligned}
q_{s,1} \;=\; & T^{-1}\sum_{t=1}^{T}\mathbf{1}'\sum_{j=1}^{s}\Sigma_1[\Psi_{s,j}(\hat{\xi}_{t|T},\hat{\xi}_{t-1|T},...,\hat{\xi}_{t-47+s|T})e_1] \\
& [\Psi_{s,j}(\hat{\xi}_{t|T},\hat{\xi}_{t-1|T},...,\hat{\xi}_{t-47+s|T})e_1]'\mathbf{1}.
\end{aligned}$$

Consequently $q_{s,1}/\sum\limits_{m=1}^{4}q_{s,m}$ would be the ratio of the first factor's contribution to unemployment volatility at horizon s.

Figure 1.6 shows the contribution of each factor to the mean squared error in predicting overall unemployment as a function of the forecasting horizon. If one is trying to forecast unemployment one month ahead, uncertainty about future inflows of type H and type L workers are equally important. However, the farther one is looking into the future, the more important becomes uncertainty about what is going to happen to type L workers. If one is trying to predict one or two years into the future, the single most important source of uncertainty is inflows of new type L workers, followed by uncertainty about their outflows. Much of the MSE associated with a 2-year-ahead forecast of unemployment comes from not knowing when the next recession will begin or the current recession will end. For this reason, the MSE associated with 2-year-ahead forecasts is closely related to what some researchers refer to as the "business cycle frequency" in a spectral decomposition. If we are interested in the key factors that change as the economy moves into and out of recessions, inflows and outflows for type L workers are most important. We will provide additional evidence on this point in Section 3.2.

The last panel of Figure 1.6 breaks these contributions separately into inflows and outflows. Both inflows and outflows are important. However, the uncertainty about future inflows are more important in accounting for the error we would make in predicting total unemployment, accounting for more than 60% of the MSE throughout the forecasting horizon.

1.4.2 Historical decomposition

A separate question of interest is how much of the realized variation over some historical episode came from particular structural shocks. In case of a linear VAR, we can decompose the historical time path for y between some date t and $t + s$ into the component that would have been predicted at time t and the part that is due to innovations in each of the shocks. A similar approach can be adopted in our case. The smoothed inferences satisfy

$$\hat{\xi}_{t+s|T} = \hat{\xi}_{t|T} + \hat{\epsilon}_{t+1|T} + \hat{\epsilon}_{t+2|T} + \hat{\epsilon}_{t+3|T} + ... + \hat{\epsilon}_{t+s|T}$$

where $\hat{\epsilon}_{t+s|T} = \hat{\xi}_{t+s|T} - \hat{\xi}_{t+s-1|T}$. For any date $t+s$ we then have the following model-inferred value for the number of people unemployed:

$$\mathbf{1}' h(\hat{\xi}_{t+s|T}, \hat{\xi}_{t+s-1|T}, \hat{\xi}_{t+s-2|T}, ..., \hat{\xi}_{t+s-47|T}).$$

For an episode starting at some date t, we can then calculate

$$\mathbf{1}' h(\hat{\xi}_{t|T}, \hat{\xi}_{t|T}, \hat{\xi}_{t|T}, ..., \hat{\xi}_{t|T}, \hat{\xi}_{t-1|T}, ..., \hat{\xi}_{t+s-47|T}).$$

This represents the path that unemployment would have been expected to follow between t and $t+s$ as a result of initial conditions at time t if there were no new shocks between t and $t+s$. Given this path for unemployment that is implied by initial conditions, we can then isolate the contribution of each separate shock between t and $t+s$. Using the linearization in equation (1.24) allows us to represent the realized deviation from this path in terms of the contribution of individual historical shocks:

$$y_{t+s} \simeq c_s(\hat{\xi}_{t|T}, \hat{\xi}_{t-1|T}, ..., \hat{\xi}_{t-47+s|T}) + \sum_{j=1}^{s} [\Psi_{s,j}(\hat{\xi}_{t|T}, \hat{\xi}_{t-1|T}, ..., \hat{\xi}_{t-47+s|T})] \hat{\epsilon}_{t+j|T}. \quad (1.26)$$

From the above equation, we get a contribution for example of $\epsilon^w_{L,t+1}, \epsilon^w_{L,t+2}, ..., \epsilon^w_{L,t+s}$ (the shocks to w_L between $t+1$ and $t+s$) to the deviation between the level of unemployment at $t+s$ from the value predicted on the basis of initial conditions at t:

$$\mathbf{1}' \sum_{j=1}^{s} [\Psi_{s,j}(\hat{\xi}_{t|T}, \hat{\xi}_{t-1|T}, ..., \hat{\xi}_{t-47+s|T})] e_1 \hat{\epsilon}'_{t+j|T} e_1.$$

Figure 1.7 shows the contribution of each component to the realized unemployment rate in the last five recessions. In each panel, the solid line (labeled U_{base}) gives the change in the unemployment rate relative to the value at the start of the episode that would have been predicted on the basis of initial conditions. Typically an increase in the inflow of type L workers (whose contribution to total unemployment is indicated by the starred red curves) is the most important reason that unemployment rises during a recession. A continuing increase of these inflows even after the recession was over was an important factor in the jobless recoveries from the 1990 and 2001 recessions. During the Great Recession, outflows as well as the inflows of type L workers were an important driver of the rise in the unemployment, with the sustained deterioration of outflow probabilities the main contributor to the slow recovery of the unemployment rate.[5]

Figure 1.8 provides some intuition about the features of the observed data that cause us to draw these conclusions. We focus on the period after October 2008, when the economy rapidly deteriorated. Our inference in Figure 1.7 concluded that the key factor accounting for the rise in unemployment at this time was an increase in the inflow of type L workers, with a secondary contribution of an increase in the continuation probabilities for type L workers. The actual subsequent observed paths for $U_t^1, ..., U_t^{13.+}$ are given by the black circled lines in Figure 1.8. All five measures increased substantially after October 2008. What do we observe in the way these five series increased that leads us to lay the blame mostly on an increase in w_t^L and to a lesser degree an increase in p_t^L?

The forecasts for each series based on our model's inference $\hat{\xi}_{t|T}$ about conditions as of October 2008 are indicated by the solid red lines. Since we treat inflows as random walks, the model forecast for U_t^1 is a horizontal line. The fact that the number

[5]Because of the length and severity of the recession of 2007-2009, the linearization (1.26) around the November 2007 values on which the last panel is based becomes poorer as we try to predict values for 2010. This is why the "U_{all}" line in the last panel falls below the actual path of unemployment in the case of this recession. As a robustness check, we also calculated the exact nonlinear contribution of each component in isolation of the others to the actual observed unemployment rate and the picture is very similar. The advantage of the linear decomposition is that the sum of the individual contributions exactly equals the aggregate, whereas the same is not true in a nonlinear dynamic representation.

of newly unemployed went up must mean that there was some change in either w_{Lt} or w_{Ht}. Our model also treats continuation probabilities as random walks, which means, given our inferred mix of type L and type H workers as of October 2008, our model predicted that $U_t^{2.3}$ would rise for 3 months before flattening out, $U_t^{4.6}$ would rise for 6 months before flattening, and so on. If there was an unanticipated change in subsequent exit probabilities $p_{i,t+s}$ this would show up as higher values for $U_t^{2.3}$ than predicted over the next 3 months, higher values for $U_t^{4.6}$ over the next 6 months, and so on. There is some modest evidence of this in the subsequent paths of $U_{t+s}^{7.12}$ and $U_{t+s}^{13.+}$, but only $p_{L,t+s}$ could matter for these, since $p_{H,t+s}$ matters very little for unemployment durations longer than 6 months. Hence the primary reason that we see the rise in $U_{t+s}^{2.3}$ after $s = 3$ months and in $U_{t+s}^{4.6}$ after $s = 6$ months must have been an increase in inflows of type L workers during months $t + 1, ..., t + s$.

We also plot as the starred fuschia lines in Figure 1.8 the implied paths for U_{t+s}^{x} conditional on both $\hat{\xi}_{t|T}$ and on type L inflows $\hat{w}_{L,t+s|T}$. These inflows alone can account for most of what was observed. If we allow also for a more minor contribution of deteriorating type L outflow probabilities (that is, also condition on $\hat{p}_{L,t+s|T}$) we can explain the observed behavior of all 5 series quite well.

Our results also offer a new perspective on an emerging debate about the causes of the fall in average unemployment exit probabilities and increase in very long spells of unemployment observed in the Great Recession. Hall (2014) argued that the explanation is a compositional shift of jobseekers toward types with low exit probabilities, for instance permanent job losers. By contrast, recent studies by Bachman and Sinning (2012) and Kroft, Lange, Notowidigdo and Katz (2014) concluded that compositional changes played little role. We add a new factor that none of these studies considered, which is the possibility of changes in the inflows of workers with unobserved heterogeneity, and find that it provides an additional reason to favor Hall's interpretation. Our estimates suggest that not only growing inflows of type L workers but also their declining exit probabilities gradually changed the composition of the pool of unemployed as seen in Figure 1.4, and that this is the primary reason that the composite exit probabilities for the pool declined during the recession. After the recession was over, the number of type L workers remained high mainly due to the sustained high continuation probabilities of type L workers despite the decreased inflows of type L workers.

1.5 Who are the type L workers?

Shimer (2012) concluded that the most important potential source of heterogeneity across different workers could be differences in the reasons the individuals became unemployed. He found that the job-finding probability of job losers on temporary layoff is higher than that of other job losers and the fraction of unemployment represented by job losers not on temporary layoff exhibits clear counter-cyclicality. Darby, Haltiwanger and Plant (1986) argued that counter-cyclicality in the average unemployment duration mainly comes from the increased inflow of prime-age workers suffering permanent job loss who are likely to have low job-finding probabilities. Bednarzik (1983) also noted that permanently separated workers are more likely to experience a long duration of unemployment, while Fujita and Moscarini (2013) showed that the unemployed who are likely to experience long-term unemployment spells tend to be those who are not recalled to work by their previous employers.

Panel A of Figure 1.9 breaks down people looking for work in terms of the reason they came to be unemployed. Dark bars describe the share of people who have been looking for work for less than one month by reason and white bars the share of those who have been looking for more than 6 months by reason. Permanent job losers and job losers on temporary layoff each account for about one fifth of new entrants into the pool of unemployed. By contrast, those on temporary layoff account for less than 3% of the unemployed with duration longer than 6 months, while around half of the long-term unemployed are accounted for by permanent job losers. This means that the unemployment exit probabilities of permanent job losers are much lower than those of job losers on temporary layoff.

Panel B of Figure 1.9 plots the inflows to unemployment by reason. Both the inflows of permanent job losers and those on temporary layoff exhibit counter-cyclicality. They rise as the recession begins and fall as the recession ends. In Panel C of Figure 1.9 we compare our estimate of the number of newly unemployed type L workers to the number of those newly unemployed who gave permanent separations from their previous job as the reason[6]. The two series were arrived at using different data and different methodologies but exhibit remarkably similar dynamics. By contrast, our series for newly unemployed type L workers does not look much like any of the other series in Panel B. Panel D compares the total number of those unemployed who gave permanent separation as the reason to our estimate of the total number of unemployed type L workers, for which the correspondence is even more striking.

Recall from Figures 1.2 and 1.3 that the overwhelming majority of newly unemployed individuals are able to find a new job quickly, and from Figure 1.4 that the longer an expansion continues, the more the pool of unemployed individuals consists of those we have labeled as type H. These features seem related to the well-known observation that in normal times there is a tremendous amount of churning in the labor market, with millions of workers entering and exiting the unemployment pool every month even as the overall unemployment rate remains low— see for example, Davis, Faberman and Haltiwanger (2006). Lazear and Spletzer (2012) showed using micro data from JOLTS that churning is procyclical, with quits accounting for the major part of it. However, our measure of type H inflows often rises during recessions. It is clear that in addition to normal churning arising from those who quit their job voluntarily, unemployment due to temporary layoffs is another important part of what we have characterized as type H unemployment. Temporary layoffs rise during recessions, but insofar as many of these individuals often return to their old jobs relatively quickly, our

[6]Permanent separations include permanent job losers and persons who completed temporary jobs. The separate series, permanent job losers and persons who completed temporary jobs, are publicly available from 1994, but their sum (permanent separations) is available back to 1976.

procedure is likely assigning most of those on temporary layoff to type H rather than type L.

Within any categorization based on observable characteristics there are still important differences across individuals. For example, within the "permanently separated" category, many workers do end up being recalled to their old positions (Fujita and Moscarini, 2013), and such individuals are likely be included in our type H designation. On the other hand, some of the individuals in every reported BLS category may have a history of low performance or poor interpersonal and communication skills[7] and would be categorized in our approach as type L. Although allowing for unobserved heterogeneity within any given group of common observed characteristics seems critical for this kind of study, our conclusion is that the single most important distinction between the latent classes of workers identified by our approach arises from the circumstances under which the individuals came to be unemployed, with permanently separated workers likely accounting for the majority of our type L workers. Normal churning of the labor market and temporary layoffs appear to be a big part of what we are capturing with our type H designation, with many permanently separated workers and labor force entrants who are hired as replacement workers likely also included in our H group.

A separate paper by Ahn (2014) provides further evidence in support of this interpretation. Ahn (2014) allows for both observed and unobserved heterogeneity by fitting models like the one developed here to subsets of workers sorted based on observable characteristics. She replaced our observation vector y_t based on aggregate unemployment numbers with $y_{jt} = (U_{jt}^{1}, U_{jt}^{2.3}, U_{jt}^{4.6}, U_{jt}^{7.12}, U_{jt}^{13.+})'$ where $U_{jt}^{2.3}$ for example denotes the number of workers with observed characteristic j who have been unemployed for 2-3 months, the idea being that within the group j there are new in-

[7]ManpowerGroup's 2013 Talent Shortage Survey showed that there is growing shortage of interpersonal skills. Firms reported that a lack of interpersonal skills like communication, collaboration and creativity, and a disregard for punctuality, appearance and flexibility are important problems among the entry-level job candidates.

flows (w_{jHt} and w_{jLt}) and outflows (p_{jHt} and p_{jLt}) of two unobserved types of workers. Of particular interest for the present discussion are the results when j corresponds to one of the 5 reasons for why the individual was looking for work. Panel A of Figure 1.10 displays Ahn's estimated values for new inflows of type L workers for each of the categories as well as the sum $\sum_{j=1}^{5} \hat{w}_{jLt|T}$. Our series $\hat{w}_{Lt|T}$ inferred from aggregate data is also plotted again for comparison. The sum of micro estimates is very similar to our aggregate estimates, and the individual micro components reveal clearly that those we have described as type L workers primarily represent a subset of people who were either permanently separated from their previous job or are looking again for work after a period of having been out of the labor force.

Ahn (2014) also calculated the models' inferences about the total number of type L individuals in any given observable category j who were unemployed in month t. These are plotted in Panel B of Figure 1.10. Here the correspondence between the aggregate inference and the sum of the micro estimates is even more compelling, as is the conclusion that type L unemployed workers represent primarily a subset of those permanently separated from their old jobs or re-entering the labor force.

1.6 Robustness checks

Here we examine how our conclusions would change under a number of alternative specifications, including changes in the unemployment measures used, alternative specifications of genuine duration dependence, possible correlations among the shocks, and reformulation of the model in terms of weekly rather than a monthly frequency.

1.6.1 Accounting for the structural break in the CPS survey

As noted in Appendix, a redesign in the CPS survey in 1994 introduced a structural break with which any user of these data has to deal. Our baseline estimates adjusted the unemployment duration data using differences between rotation groups 1 and

5 and groups 2-4 and 6-8 in the the CPS micro data. Here we summarize how our results would change if we were to instead use the adjustment employed by Hornstein (2012).

Table 1.3 summarizes the implications of alternative specifications for what we see as the most important conclusions that emerge from our baseline analysis. The table breaks down the MSE of a forecast of the overall level of unemployment at 3-month, 1-year, and 2-year forecast horizons into the fraction of the forecast error that is attributable to various shocks. Column 1 gives the numbers implied by our baseline specification and highlights our key conclusion that inflows account for more than half the variance at all horizons. Inflows of type L workers are most important but the outflows of type L workers and the inflows of type H workers are also crucial at a 3-month horizon. At a 1- or 2-year horizon, shocks to inflow and outflow probabilities for type L workers are the most important factors.

Column 2 of Table 1.3 reports the analogous variance decompositions when we instead use Hornstein's data adjustment as described in Appendix. This produces very little change in these numbers. In column 3 we use only data subsequent to the redesign in 1994 making no adjustment to the reported BLS figures. This reduces the estimated contribution of inflows of type L workers at shorter horizons, but preserves our main finding that for business-cycle frequencies, changes for type L workers account for most of the fluctuations in unemployment, with changes in type L inflows accounting for about half the variance of unemployment at the 2-year horizon. We obtained similar results using the full data set from 1976-2013 with no adjustments for the 1994 redesign (column 4). We also found that the non-monotonic pattern in the genuine duration dependence is preserved regardless of data adjustment methods.

Note that although we report the likelihood and Schwarz's (1978) Bayesian criterion in rows 2 and 3 of Table 1.3, the values for columns 2-4 are not comparable with the others due to a different definition of the observable data vector y_t.

1.6.2 Alternative specifications for genuine duration dependence

Our baseline specification assumed that a single parameter δ_1 described genuine duration dependence for any worker unemployed for less than 6 months. We also estimated a model in which each of the observed duration categories (2-3 months, 4-6 months, 7-12 months, and greater than 12 months) was characterized by a different genuine duration parameter, replacing (1.21) with

$$d_\tau = \begin{cases} \delta_1^A(\tau - 1) & \text{for } \tau < 3 \\ \delta_1^A(3-2) + \delta_1^B(\tau - 2) & \text{for } 3 \leq \tau < 6 \\ \delta_1^A(3-2) + \delta_1^B(5-2) + \delta_2(\tau - 5) & \text{for } 6 \leq \tau < 12 \\ \delta_1^A(3-2) + \delta_1^B(5-2) + \delta_2(11-5) + \delta_3(\tau - 11) & \text{for } 12 \leq \tau. \end{cases}$$

Adding this additional parameter δ_1^B results in only a trivial improvement in the likelihood function and virtually no change in any of the variance decompositions, as seen in column 5 of Table 1.3.

1.6.3 Allowing for correlated shocks

Our baseline specification assumed that the shocks to w_{Lt}, w_{Ht}, p_{Lt} and p_{Ht} were mutually uncorrelated. It is possible to generalize this in a parsimonious way by allowing a factor structure to the innovations, $\varepsilon_t = \lambda F_t + u_t$, where $F_t \sim N(0,1)$, λ is a (4×1) vector of factor loadings, and u_t is a (4×1) vector of mutually uncorrelated idiosyncratic components with variance matrix $E(u_t u_t') = Q$:

$$E(\varepsilon_t \varepsilon_t') = \lambda\lambda' + Q$$

$$Q = \begin{bmatrix} (q_H^w)^2 & 0 & 0 & 0 \\ 0 & (q_L^w)^2 & 0 & 0 \\ 0 & 0 & (q_H^x)^2 & 0 \\ 0 & 0 & 0 & (q_L^x)^2 \end{bmatrix}.$$

In this case the variance decomposition (1.25) becomes

$$\begin{aligned} & E(y_{t+s} - \hat{y}_{t+s|t})(y_{t+s} - \hat{y}_{t+s|t})' \\ = & \sum_{j=1}^{s} [\Psi_{s,j}(\xi_t, \xi_{t-1}, ..., \xi_{t-47+s})](\lambda\lambda' + Q)[\Psi_{s,j}(\xi_t, \xi_{t-1}, ..., \xi_{t-47+s})]' \\ = & \sum_{j=1}^{s} [\Psi_{s,j}(\xi_t, \xi_{t-1}, ..., \xi_{t-47+s})]\lambda\lambda'[\Psi_{s,j}(\xi_t, \xi_{t-1}, ..., \xi_{t-47+s})]' \\ + & \sum_{j=1}^{s}\sum_{m=1}^{4} Q_m[\Psi_{s,j}(\xi_t, \xi_{t-1}, ..., \xi_{t-47+s})e_m][\Psi_{s,j}(\xi_t, \xi_{t-1}, ..., \xi_{t-47+s})e_m]' \end{aligned}$$

for Q_m the row m, column m element of Q. Because the factor F_t has an effect on all four components, it is not possible to impute the term involving $\lambda\lambda'$ to any one of the four shocks individually. However, we can calculate the portion of the MSE that is attributable to this aggregate factor along with those of each of the individual idiosyncratic shocks in u_t. This is reported in column 6 of Table 1.4, and variance decompositions are plotted in Figure 1.11. The aggregate factor by itself accounts for 58% of the MSE of a 3-month-ahead forecast of unemployment, and inflows and outflows of type H workers account for another 19%. The aggregate factor is strongly correlated with outflows of type L workers. If we isolate the idiosyncratic component of each shock that is uncorrelated with the other three, shocks to inflows of type L workers account for only a quarter of the 3-month-ahead forecast error and almost 1/3 of the 2-year-ahead forecast error. There is essentially no role for the idiosyncratic component of outflows of type L workers, since changes in these outflows are so highly correlated with the other three shocks. This suggest that the probability of exiting unemployment of type L workers is closely related to an aggregate shock. Considering that the share of type L workers in unemployment is importantly driven by their outflows, it implies that the compositional change of unemployment can be interpreted as an aggregate phenomenon that is core to the dynamics of economic recessions.

1.6.4 Time aggregation

Focusing on monthly transition probabilities understates flows into and out of unemployment since someone who loses their job in week 1 of a month but finds a new job in week 2 would never be counted as having been unemployed. Shimer (2012) argued that this time-aggregation bias would result in underestimating the importance of outflows in accounting for cyclical variation in unemployment, and Fujita and Ramey (2009), Shimer (2012) and Hornstein (2012) all formulated their models in continuous time.

On the other hand, Elsby, Michaels and Solon (2009) questioned the theoretical suitability of a continuous-time conception of unemployment dynamics, asking if it makes any sense to count a worker who loses a job at 5:00 p.m. one day and starts a new job at 9:00 a.m. the next as if they had been unemployed at all. We agree, and think that defining the central object of interest to be the fraction of those newly unemployed in month t who are still unemployed in month $t + k$, as in our baseline model, is the most useful way to pose questions about unemployment dynamics. Nevertheless, and following Kaitz (1970), Perry (1972), Sider (1985), Baker (1992), and Elsby, Michaels and Solon (2009) we also estimated a version of our model formulated in terms of weekly frequencies as an additional check for robustness.

We can do so relatively easily if we make a few simplifying assumptions. We view each month t as consisting of 4 equally-spaced weeks and assume that in each of these weeks there is an inflow of w_{it} workers of type i, each of whom has a probability $p_{it}(0) = \exp[-\exp(x_{it})]$ of exiting unemployment the following week. This means that for those type i individuals who were newly unemployed during the first week of month t, $w_{it}[p_{it}(0)]^3$ are still unemployed as of the end of the month. Thus for the model interpreted in terms of weekly transitions, equation (1.14) would be replaced by

$$U_t^1 = \sum_{i=H,L} \{w_{it} + w_{it}[p_{it}(0)] + w_{it}[p_{it}(0)]^2 + w_{it}[p_{it}(0)]^3\} + r_t^1.$$

Likewise (1.15) becomes

$$U_t^{2.3} = \sum_{i=H,L} \sum_{s=1}^{4} \left\{w_{i,t-1}[p_{i,t-1}(1)]^{8-s} + w_{i,t-2}[p_{i,t-2}(2)]^{12-s}\right\} + r_t^{2.3}$$

for $p_{it}(\tau)$ given by (1.20)-(1.21) for $\tau = 1, 2$. Note that although this formulation is conceptualized in terms of weekly inflow and outflows w_i and p_i, the observed data y_t are the same monthly series used in our other formulations, and the number of parameters is the same as for our baseline formulation.

The weekly formulation achieves a slightly lower value for the likelihood function and, as seen in Table 1.3, does not change our substantive conclusions.

1.7 Conclusion

People who have been unemployed for longer periods than others have dramatically different probabilities of exiting unemployment, and these relative probabilities change significantly over the business cycle. Even when one conditions on observable characteristics, unobserved differences across people and the circumstances under which they came to be unemployed are crucial for understanding these features of the data.

We have shown how the time series of unemployment levels by different duration categories can be used to infer inflows and outflows from unemployment for workers characterized by unobserved heterogeneity. In contrast to other methods, our approach uses the full history of unemployment data to summarize inflows and outflows from unemployment and allows us to make formal statistical statements about how much of the variance of unemployment is attributable to different factors as well as identify the particular changes that characterized individual historical episodes.

In normal times, around three quarters of those who are newly unemployed find jobs quickly. But in contrast to the conclusions of Hall (2005) and Shimer (2012), we find that more than half the variance in unemployment comes from shocks to the number of newly unemployed, and a key feature of economic recessions is newly unemployed individuals who have significantly lower job-finding probabilities. Our inferred values for the size of this group exhibit remarkably similar dynamics to separate measures of the number of people who permanently lose their jobs. We conclude that recessions are characterized by a change in the circumstances under which people become unemployed that makes it harder for them to find new jobs.

1.8 Acknowledgements

Chapter 1 is coauthored with James D. Hamilton. Chapter 1 is in preparation for submission.

1.9 Figures and Tables

1.9.1 Figures

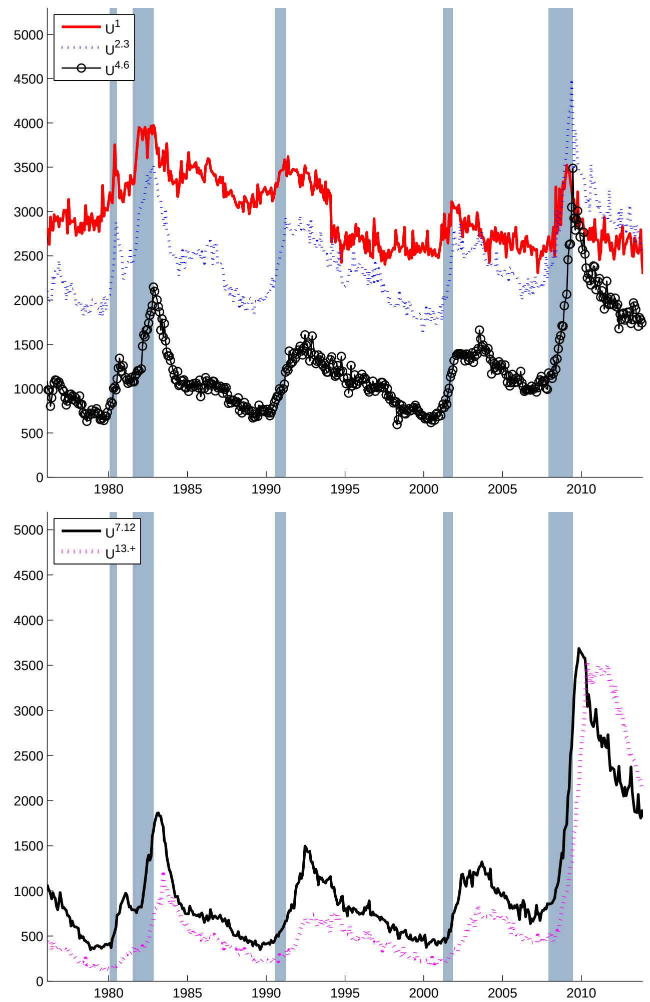

Figure 1.1: Number of unemployed individuals (in thousands) by duration of time they have already been unemployed as of the indicated date.
Panel A: those unemployed 1 month, 2-3 months, and 4-6 months. Panel B: those unemployed 7-12 months and more than 12 months.

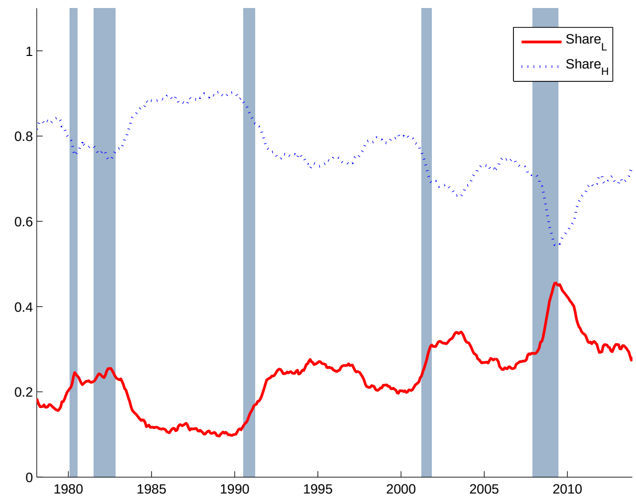

Figure 1.2: Probability that a newly unemployed worker of each type will still be unemployed the following month, $\hat{p}_{it|T}(1)$ for $i = L, H$.

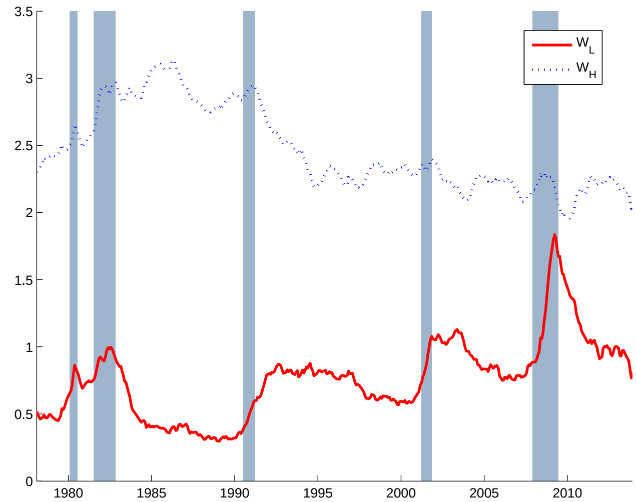

Figure 1.3: Number of newly unemployed workers of each type, $\hat{w}_{it|T}$ for $i = L, H$.

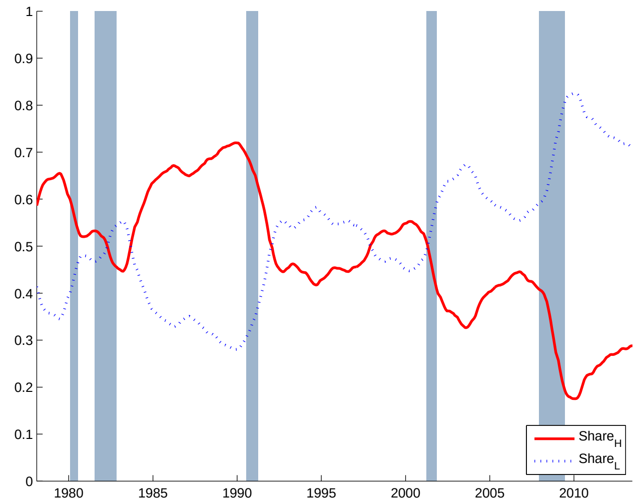

Figure 1.4: Share of total unemployment accounted for by each type of worker

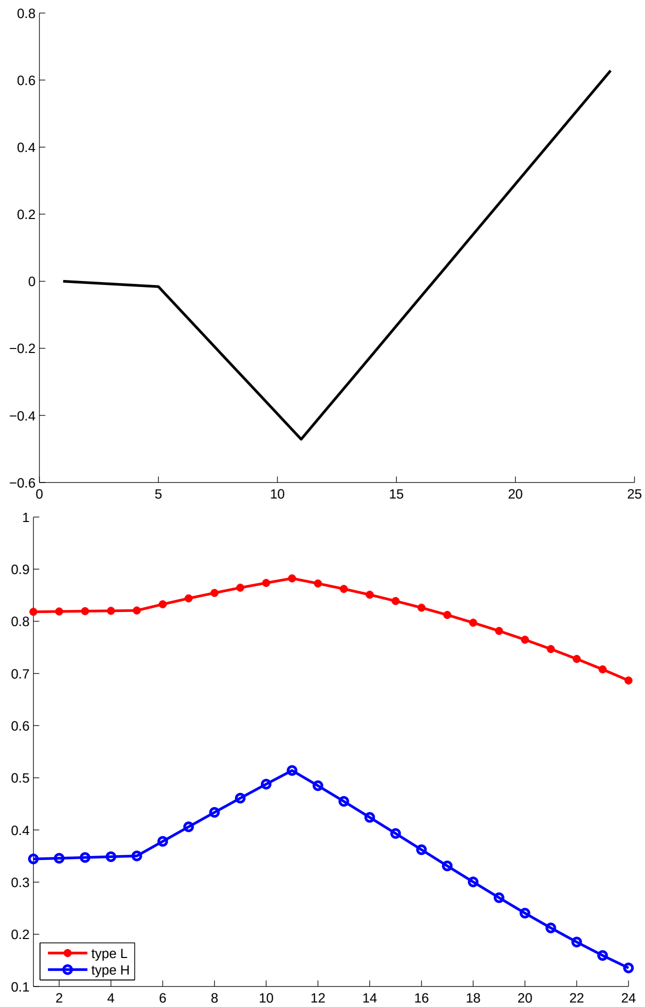

Figure 1.5: Estimates of genuine duration dependence.
Panel A: plot of d_τ as a function of τ (months spend in unemployment). Panel B: average continuation probabilities of type H and type L workers as a function of duration of unemployment.

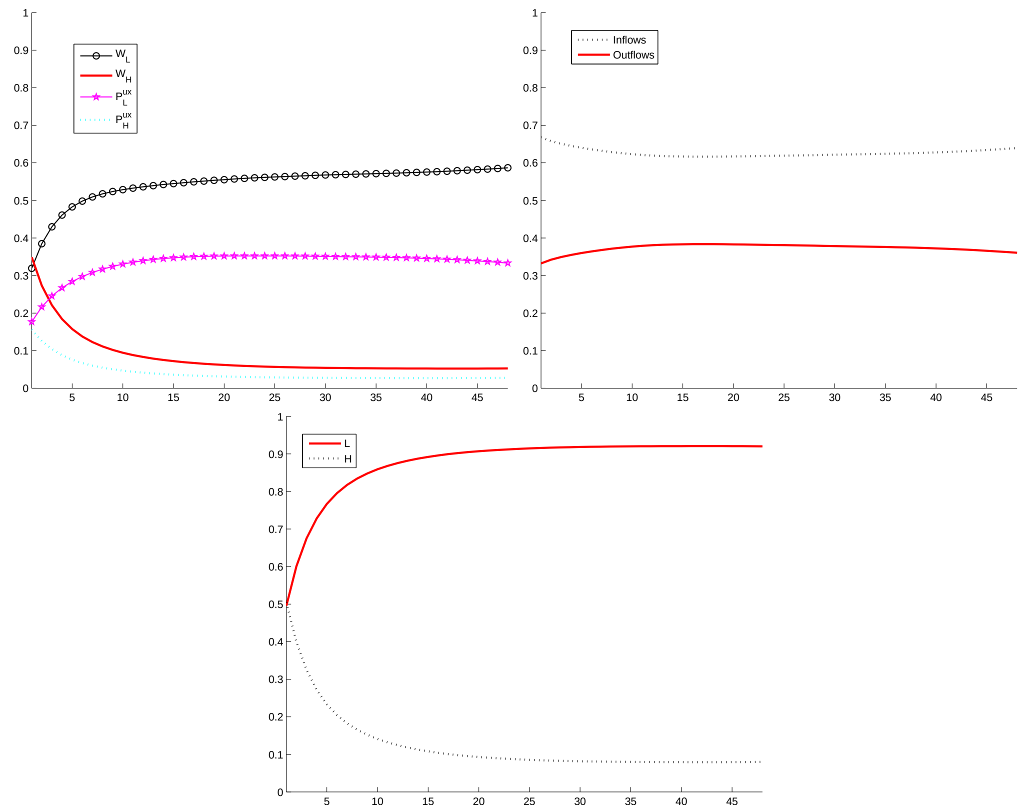

Figure 1.6: Fraction of variance of error in forecasting total unemployment at different horizons attributable to separate factors.
Horizontal axis: number of months ahead s for which the forecast is formed. Panel A: contribution of each of the factors $\{w_{Ht}, w_{Lt}, x_{Ht}, x_{Lt}\}$ separately. Panel B: combined contributions of $\{w_{Ht}, x_{Ht}\}$ and $\{w_{Lt}, x_{Lt}\}$. Panel C: combined contributions of $\{w_{Ht}, w_{Lt}\}$ and $\{x_{Ht}, x_{Lt}\}$.

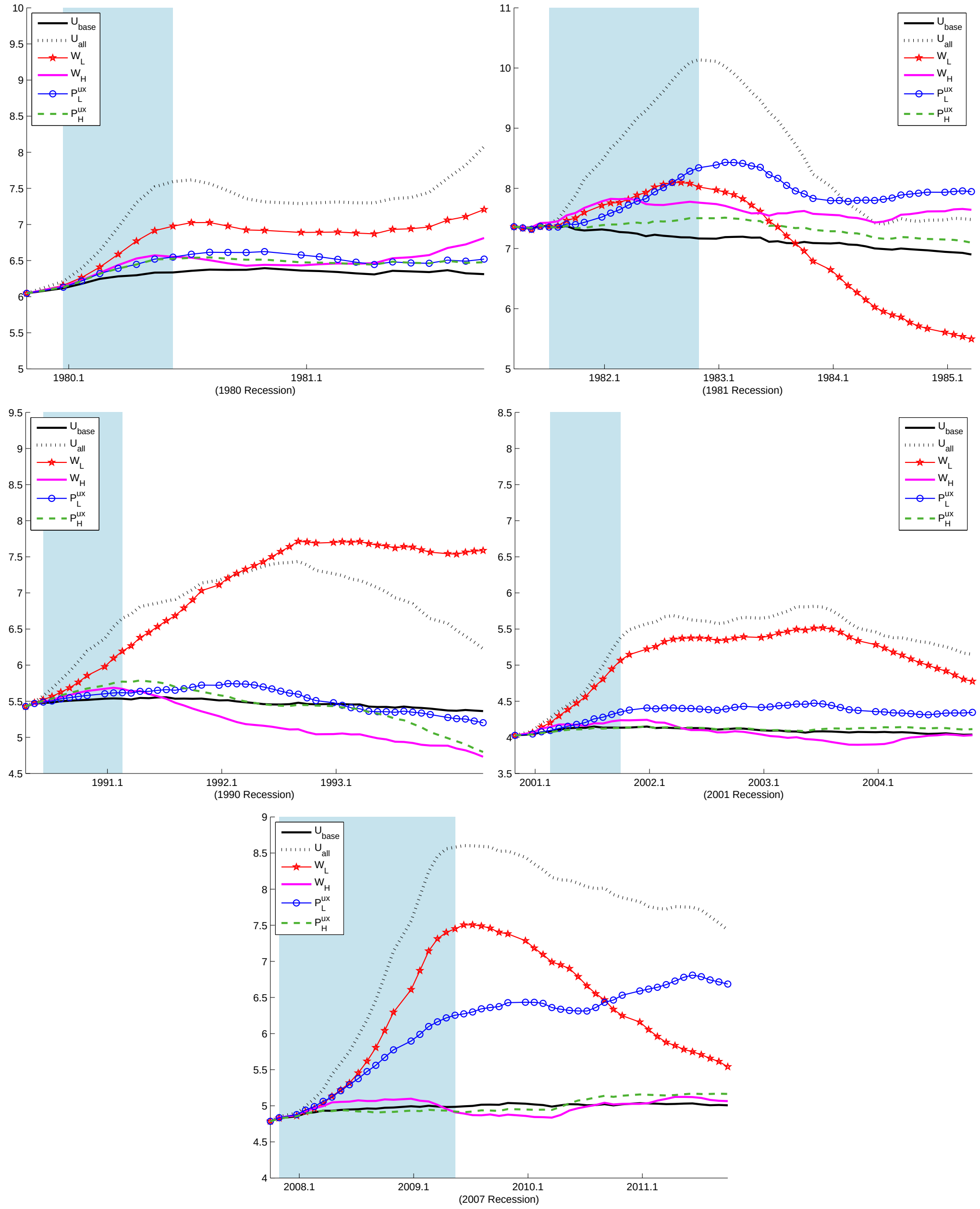

Figure 1.7: Historical decompositions of of U.S. recessions

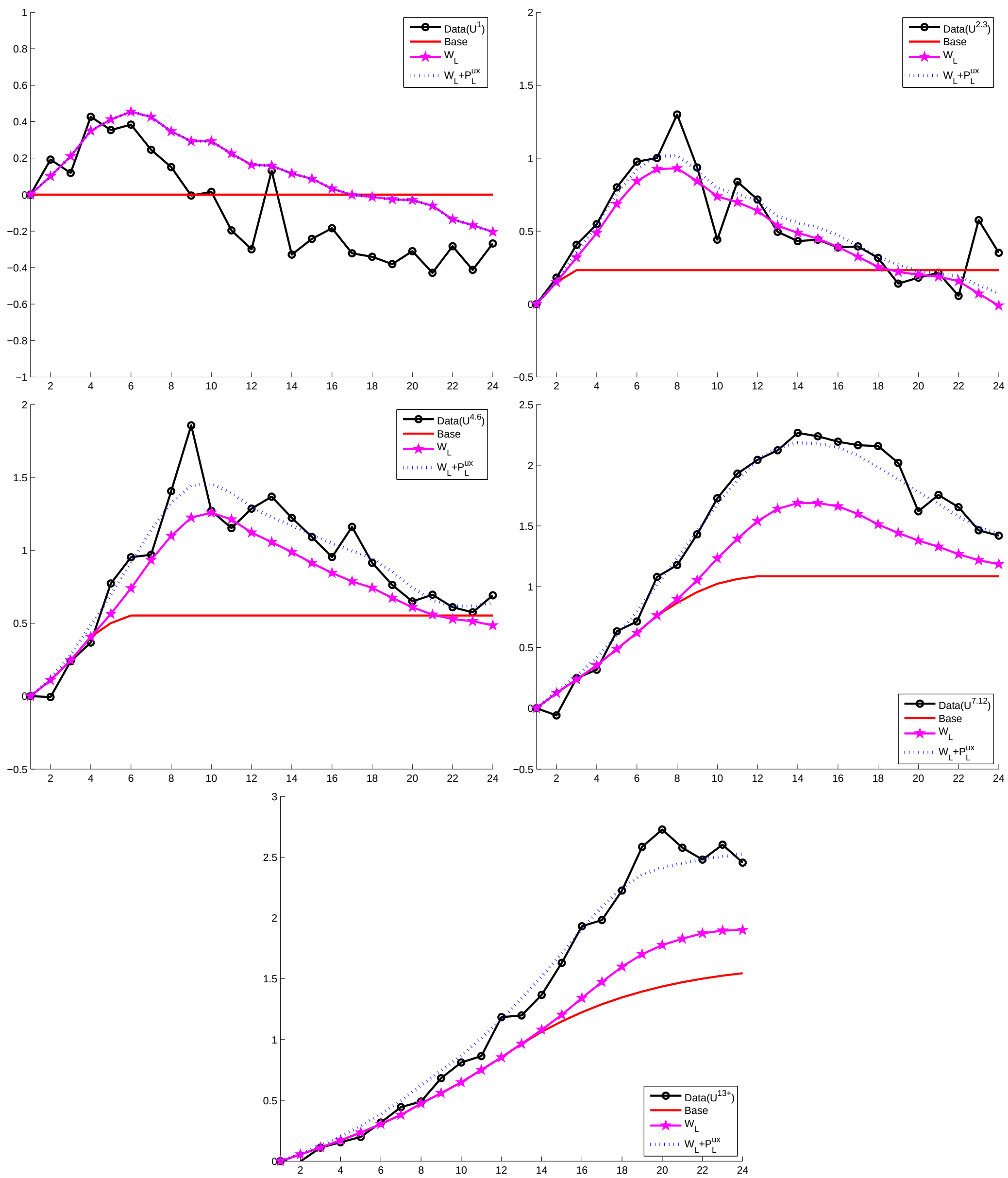

Figure 1.8: Data and forecasts for $U_t^1, \ldots, U_t^{13.+}$ during the Great Recession. Horizontal axis: number of months ahead s for which the forecast is formed in October 2008. Vertical axis: percentage point deviation from the initial values at October 2008.

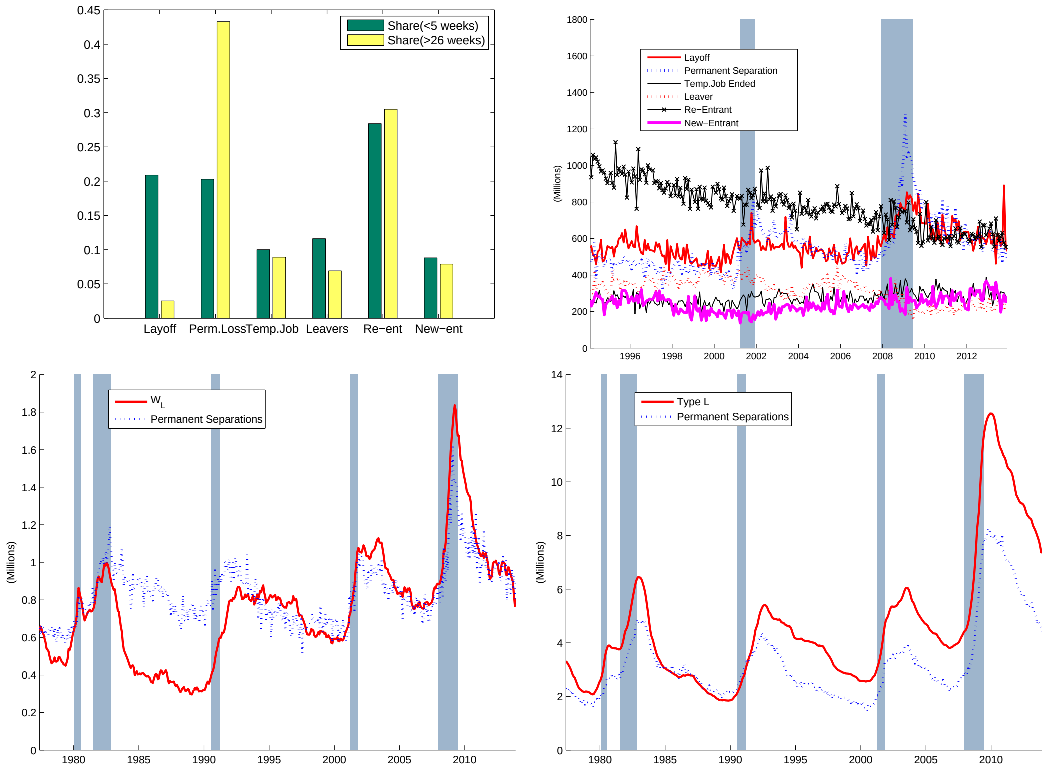

Figure 1.9: Reason for unemployment

Panel A: share of unemployment by reason (1994-2013 average). Panel B: inflows to unemployment by reason for unemployment. Panel C: inflows of type L workers compared with workers newly unemployed due to permanent job loss or end of a temporary job. Panel D: total numbers of unemployed type L workers compared to total numbers of unemployed due to permanent job loss or end of temporary job.

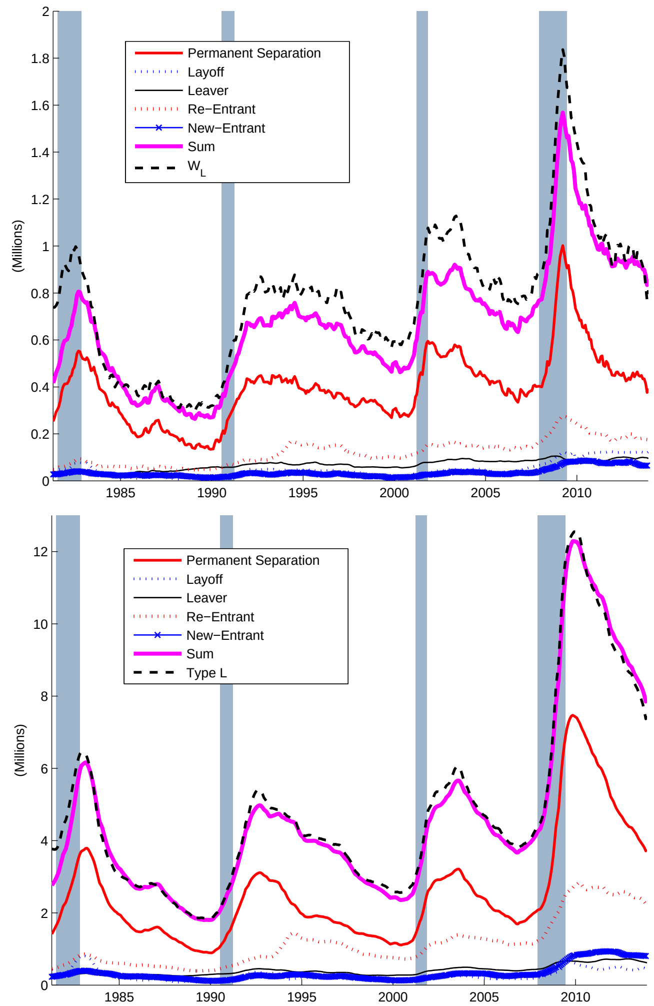

Figure 1.10: Inflows and total numbers of type L workers by reason of unemployment. Panel A: number of type L individuals who are newly unemployed by reason of unemployment along with the sum across reasons (thick fuchsia) and inference based on uncategorized aggregate data (dashed black). Panel B: number of type L workers who have been unemployed for any duration by reason of unemployment along with the sum across reasons (thick fuchsia) and inference based on uncategorized aggregate data (dashed black). Source: Ahn (2014).

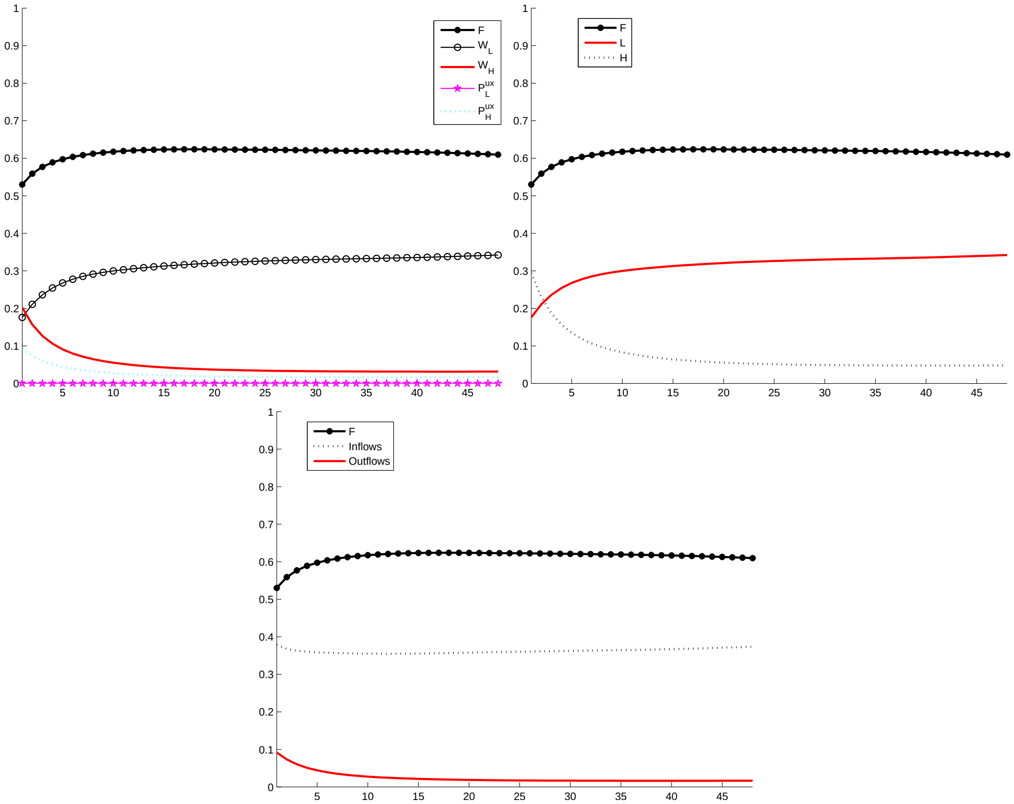

Figure 1.11: Fraction of variance of error in forecasting total unemployment at different horizons attributable to separate factors in the model with correlated errors.
Horizontal axis: number of months ahead s for which the forecast is formed. Panel A: contribution of the aggregate factor F_t along with the idiosyncratic components of $\{w_{Ht}, w_{Lt}, x_{Ht}, x_{Lt}\}$ separately. Panel B: combined contributions of idiosyncratic components of $\{w_{Ht}, x_{Ht}\}$ and $\{w_{Lt}, x_{Lt}\}$ along with aggregate factor F_t. Panel C: combined contributions of idiosyncratic components of $\{w_{Ht}, w_{Lt}\}$ and $\{x_{Ht}, x_{Lt}\}$ along with aggregate factor F_t.

1.9.2 Tables

Table 1.1: Average number unemployed by duration of unemployment (in thousands, 1976-2013)

U^{1}	$U^{2.3}$	$U^{4.6}$	$U^{7.12}$	$U^{13.+}$
3,210	2,303	1,238	1,050	636

Table 1.2: Parameter estimates for the baseline model

White (1982) quasi-maximum-likelihood standard errors in parentheses.

σ_L^w	0.0446*** (0.0043)	R_1	0.0977*** (0.0058)	δ_1	-0.0040 (0.0163)
σ_H^w	0.0465*** (0.0060)	$R_{2.3}$	0.0760*** (0.0043)	δ_2	-0.0759*** (0.0258)
σ_L^x	0.0445*** (0.0049)	$R_{4.6}$	0.0776*** (0.0067)	δ_3	0.0846*** (0.0292)
σ_H^x	0.0218*** (0.0029)	$R_{7.12}$	0.0590*** (0.0049)		
		R_{13+}	0.0366*** (0.0027)		
No. of Obs.		456			
Log-Likelihood		2,401.13			

1.10 Appendices

Measurement issues and seasonal adjustment

The seasonally adjusted numbers of people unemployed for less than 5 weeks, for between 5 and 14 weeks, 15-26 weeks and for longer than 26 weeks are published by the Bureau of Labor Statistics. To further break down the number unemployed for longer than 26 weeks into those with duration between 27 and 52 weeks and with longer than 52 weeks, we used seasonally unadjusted CPS microdata publicly available at the NBER website (http://www.nber.org/data/cps_basic.html). Since the CPS is a probability sample, each individual is assigned a unique weight that is used to produce the aggregate data. From the CPS microdata, we obtain the number of unemployed whose duration of unemployment is between 27 and 52 weeks and the number longer than 52 weeks. We seasonally adjust the two series using X-12-ARIMA,[8] and calculated the ratio of those unemployed 27-52 weeks to the sum. We then multiplied this ratio by the published BLS seasonally adjusted number for individuals who had been unemployed for longer than 26 weeks to obtain our series $U_t^{7.12}$.[9]

An important issue in using these data is the redesign of the CPS survey in 1994. Before 1994, individuals were always asked how long they had been unemployed. After the redesign, if an individual is reported as unemployed during two consecutive months, then her duration is recorded automatically as the sum of her duration last month and the number of weeks between the two months' survey reference periods. Note that if an individual was unemployed during each of the two weeks surveyed, but worked at a job in between, that individual would likely self-report a duration of unemployment

[8] An earlier version of this paper dealt with seasonality by taking 12-month moving averages and arrived at similar overall results to those presented in this version. As a further check on the approach used here, we compared the published BLS seasonally adjusted number for those unemployed with duration between 15 and 26 weeks to an X-12-ARIMA-adjusted estimate constructed from the CPS microdata, and found the series to be quite close.

[9] This adjustment is necessary because the published number for unemployed with duration longer than 26 weeks is different from that directly computed from the CPS microdata, although the difference is subtle. The difference arises because the BLS imputes the numbers unemployed with different durations to various factors, e.g., correction of missing observations.

Table 1.3: Comparison of variance decomposition across different models

Column (1) Baseline model, (2) Alternative data set, (3) Post 94 data set, (4) Unadjusted data set, (5) Unrestricted GDD, (6) Correlated shocks, (7) Weekly frequency. SIC calculated as minus twice the log likelihood plus number of parameters k times log of sample size ($T = 456$). Note that likelihood and SIC for columns 2-4 are not comparable with the others because the data on y_t are different. F denotes the aggregate factor

		(1)	(2)	(3)	(4)	(5)	(6)	(7)
No. of param.		12	12	12	12	13	16	12
Log-Likelihood		2401.13	2310.23	1156.56	2432.27	2401.27	2412.88	2399.60
SIC		-4728.79	-4546.99	-2247.53	-4791.07	-4722.95	-4727.80	-4725.73
3 month	F	-	-	-	-	-	0.577	-
	w_L	0.430	0.422	0.251	0.166	0.431	0.236	0.395
	w_H	0.221	0.249	0.230	0.380	0.220	0.127	0.149
	p_L	0.245	0.258	0.270	0.190	0.246	0.000	0.217
	p_H	0.104	0.071	0.249	0.264	0.103	0.060	0.239
	Inflows	0.651	0.670	0.481	0.546	0.651	0.363	0.544
	L group	0.675	0.680	0.521	0.356	0.677	0.236	0.613
2 year	F	-	-	-	-	-	0.623	-
	w_L	0.561	0.522	0.437	0.501	0.561	0.325	0.528
	w_H	0.057	0.064	0.068	0.118	0.057	0.034	0.041
	p_L	0.352	0.395	0.408	0.286	0.353	0.000	0.357
	p_H	0.029	0.019	0.087	0.095	0.029	0.017	0.075
	Inflows	0.619	0.586	0.505	0.619	0.618	0.360	0.569
	L group	0.913	0.917	0.845	0.787	0.914	0.325	0.884

to be less than 5 weeks before the redesign, but the duration would be imputed to be a number greater than 5 weeks after the redesign.

As suggested by Elsby, Michaels and Solon (2009) and Shimer (2012) we can get an idea of the size of this effect by making use of the staggered CPS sample design. A given address is sampled for 4 months (called the first through fourth rotations, respectively), not sampled for the next 8 months, and then sampled again for another 4 months (the fifth through eighth rotations). After the 1994 redesign, the durations for unemployed individuals in rotations 2-4 and 6-8 are imputed, whereas those in rotations 1 and 5 are self-reported, just as they were before 1994. For those in rotation groups 1 and 5, we can calculate the fraction of individuals who are newly unemployed and compare this with the total fraction of newly unemployed individuals across all rotations. The ratio of these two numbers is reported in Panel A of Figure A1, and averaged 1.15 over the period 1994-2007 as reported in the second row of Table A1. For comparison, the ratio averaged 1.01 over the period 1989-1993, as seen in the first row. This calculation suggests that if we want to compare the value of U_t^1 as calculated under the redesign to the self-reported numbers available before 1994, we should multiply the former by 1.15. This is similar to the adjustment factors of 1.10 used by Hornstein (2012), 1.154 by Elsby, Michaels and Solon (2009), 1.106 by Shimer (2012), and 1.205 by Polivka and Miller (1998).

For our study, unlike most previous researchers, we also need to specify which categories the underreported newly unemployed are coming from. Figure A1 reports the observed ratios of rotation 1 and 5 shares to the total for the various duration groups, with average values summarized in Table A1. One interesting feature is that under the redesign, the fraction of those with 7-12 month duration from rotations 1 and 5 is very similar to that for other rotations, whereas the fraction of those with 13 or more months is much lower.[10] Based on the values in Table A1, we should scale up the estimated

[10]One possible explanation is digit preference– an individual is much more likely to report having been unemployed for 12 months than 13 or 14 months. When someone in rotation 5 reports they have been unemployed for 12 months, BLS simply counts them as such, and if they are still unemployed the following month, BLS imputes to them a duration of 13 months. The imputed number of people 13

values for U_t^1 and scale down the estimated values of $U_t^{2.3}$ and $U_t^{13.+}$ relative to the pre-1994 numbers. The values for $U_t^{4.6}$ and $U_t^{7.12}$ seem not to have been affected much by the redesign. Our preferred adjustment for data subsequent to the 1994 redesign is to multiply U_t^1 by 1.15, $U_t^{2.3}$ by 0.87, $U_t^{13.+}$ by 0.77, and leave $U_t^{4.6}$ and $U_t^{7.12}$ as is. We then multiplied all of our adjusted duration figures by the ratio of total BLS reported unemployment to the sum of our adjusted series in order to match the BLS aggregate exactly.

Hornstein (2012) adopted an alternative adjustment, assuming that all of the imputed newly unemployed came from the $U^{2.3}$ category. He chose to multiply U_t^1 by 1.10 and subtract the added workers solely from the $U_t^{2.3}$ category. As a robustness check we also report results using Hornstein's proposed adjustment in Section 5.1, as well as results using no adjustments at all.

An alternative might be to to use the ratios for each t in Figure A1 rather than to use the averages from Table A1. However, as Shimer (2012) and Elsby, Michaels and Solon (2009) mentioned, such an adjustment would be based on only about one quarter of the sample and thus multiplies the sampling variance of the estimate by about four, which implies that noise from the correction procedure could be misleading in understanding the unemployment dynamics.

Table 1.4: Average ratio of each duration group's share in the first/fifth rotation group to that in total unemployment

	U^1	$U^{2.3}$	$U^{4.6}$	$U^{7.12}$	$U^{13.+}$
1989-1993	1.01	1.01	0.96	1.02	0.97
1994-2007	1.15	0.87	0.95	1.05	0.77

months and higher is significantly bigger than the self-reported numbers, just as the imputed number of people with 2-3 months appears to be higher than self-reported.

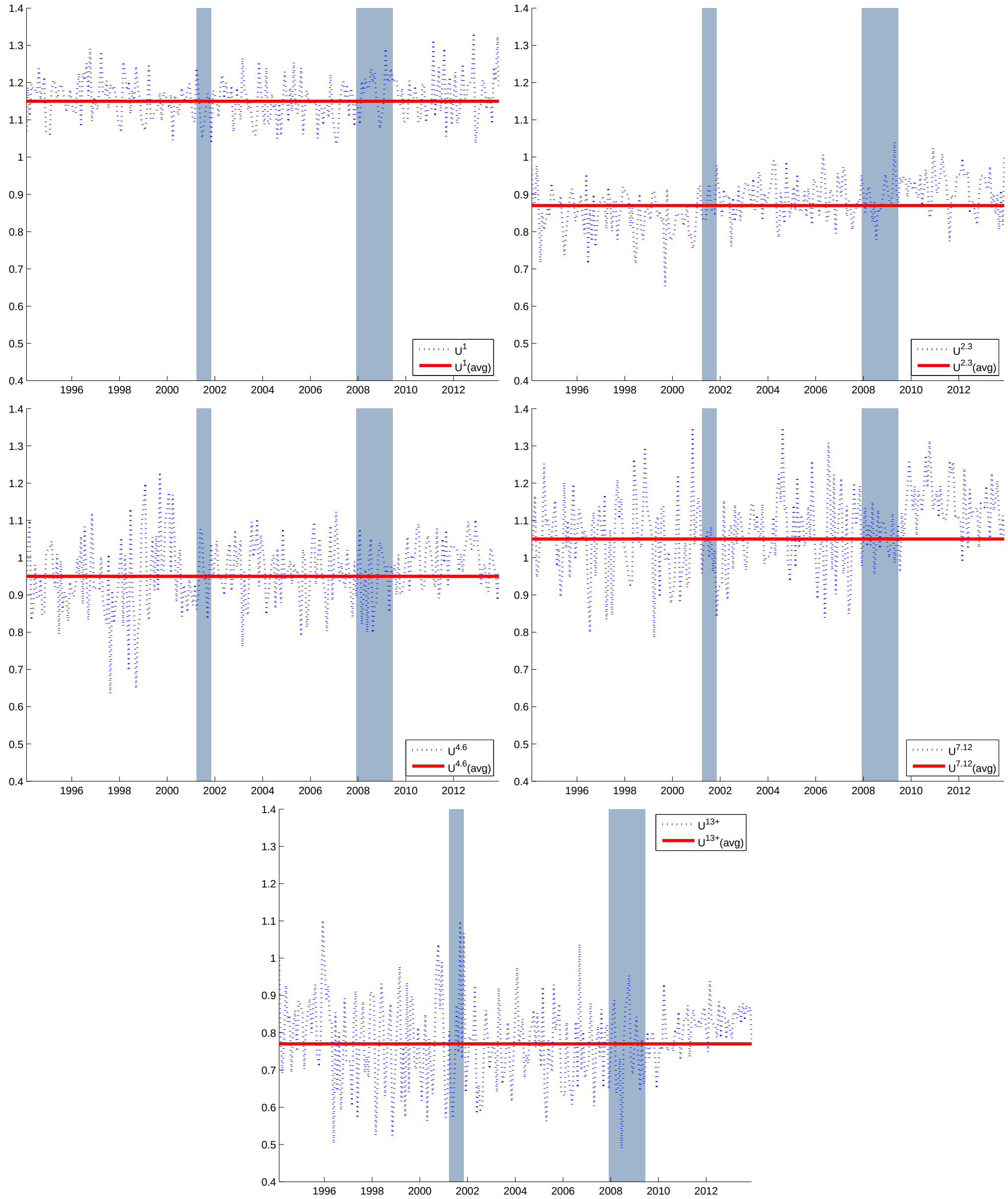

Figure 1.12: Ratio of each duration group's share in the first and fifth rotation groups to that in all rotation groups

Estimation Algorithm

The system (2.9) and (1.14)-(1.18) can be written as

$$x_t = Fx_{t-1} + v_t$$

$$y_t = h(x_t) + r_t$$

for $x_t = (\xi'_t, \xi'_{t-1}, ..., \xi'_{t-47})'$, $E(v_t v'_t) = Q$, and $E(r_t r'_t) = R$. The function $h(.)$ as well as elements of the variance matrices R and Q depend on the parameter vector $\theta = (\delta_1, \delta_2, \delta_3, R_1, R_{2.3}, R_{4.6}, R_{7.12}, R_{13+}, \sigma^w_L, \sigma^w_H, \sigma^x_L, \sigma^x_H)'$. The extended Kalman filter (e.g., Hamilton, 1994b) can be viewed as an iterative algorithm to calculate a forecast $\hat{x}_{t+1|t}$ of the state vector conditioned on knowledge of θ and observation of $Y_t = (y'_t, y'_{t-1}, ..., y'_1)'$ with $P_{t+1|t}$ the MSE of this forecast. With these we can approximate the distribution of y_t conditioned on Y_{t-1} as $N(h(\hat{x}_{t|t-1}), H'_t P_{t|t-1} H_t + R)$ for $H_t = \partial h(x_t)/\partial x'_t | x_t = \hat{x}_{t|t-1}$ from which the likelihood function associated with that θ can be calculated and maximized numerically. The forecast of the state vector can be updated using

$$\hat{x}_{t+1|t} = F\hat{x}_{t|t-1} + FK_t(y_t - h(\hat{x}_{t|t-1}))$$

$$K_t = P_{t|t-1} H_t (H'_t P_{t|t-1} H_t + R)^{-1}$$

$$P_{t+1|t} = F(P_{t|t-1} - K_t H'_t P_{t|t-1})F' + Q.$$

A similar recursion can be used to form an inference about x_t using the full sample of available data, $\hat{x}_{t|T} = E(x_t | y_T, ..., y_1)$ and these smoothed inferences are what are reported in any graphs in this paper.

Hamilton(1994) mentions that it is desirable to use smoothed estimates because smoothed estimates incorporate all the available information up to the end of the sample to help improve inference on the historical values that the state vector takes. The sequence of smoothed estimates, $\{\hat{\xi}_{t|T}\}_{t=1}^{T}$, is calculated as follows.

$$\hat{\xi}_{t|T} = \hat{\xi}_{t|t} + J_t(\hat{\xi}_{t+1|T} - \hat{\xi}_{t+1|t})$$

for $t = T-1, T-2, ..., 1$, where $J_t = P_{t|t} F' P^{-1}_{t+1|t}$. The corresponding mean squared errors are found by iterating on in reverse order for $t = T-1, T-2, ..., 1$

$$P_{t|T} = P_{t|t} + J_t(P_{t+1|T} - P_{t+1|t})J_t'.$$

The initial variables are constructed from the following methods. Since the CPS micro data is not publicly available before 1976, we made predictions of the number unemployed with duration 7-12 months and that with duration 13 months and over from January 1972 to December 1975 by multiplying the average shares of the number unemployed with duration 7-12 months and 13 months and over in 1976 to the number unemployed with duration longer than 26 weeks each month.

With the data from February.1972 to January.1976, we estimate the steady state values for the four latent variables, $w_{Ht}, w_{Lt}, x_{Ht}, x_{Lt}$, each month using the set of equations (1)-(5) in Section 1, by presetting the estimated steady state value for the genuine duration dependence parameter for the entire sample period from January 1976 to December 2013. The estimates of initial variables enter into the initial vector $\xi_t's$ for $t = -46, -45, ..., 1$. By setting large diagonal elements of $P_{1|0}$,the particular value of initial variables in $\hat{x}_{1|0}$ has little influence on any of the results.

PZ Algorithm

Maximization of the likelihood function $\sum_{t=1}^{T} \log f(y_t|Y_{t-1})$ is made difficult by non-convexity and multimodality of the likelihood surface. We developed a new algorithm, which we call a PZ algorithm, which helped considerably in the estimation. The procedure of PZ Algorithm is as follows.

1. Divide the parameters into two sets or more. For illustrative purpse, the parameters are grouped into two sets, set A(θ_A) and set B(θ_B).

2. Fix θ_A at a set of random starting values $\theta_A^{(0)}$ and then search set B parameters yielding maximum likelihood from starting values, $\theta_B^{(0)}$. Let the estimated set B parameters in this step be $\hat{\theta_B}^{(0)}$ and the value of likelihood function be $L_B^{(0)}$.

3. Fix θ_B at $\hat{\theta_B}^{(0)}$ and then estimate set A parameters yielding maximum likelihood from starting values, $\theta_A^{(0)}$. Let the estimated set A parameters in this step be $\hat{\theta_A}^{(0)}$ and the value of likelihood function be $L_A^{(0)}$.

4. Fix θ_A at $\hat{\theta_A}^{(0)}$ and then search set B parameters yielding maximum likelihood from starting values, $\hat{\theta_B}^{(0)}$. Let the estimated set B parameters in this step be $\hat{\theta_B}^{(1)}$ and the value of likelihood function be $L_B^{(1)}$.

5. Fix θ_B at $\hat{\theta_B}^{(1)}$ and then estimate set A parameters yielding maximum likelihood from starting values, $\hat{\theta_A}^{(0)}$. Let the estimated set A parameters in this step be $\hat{\theta_A}^{(1)}$ and the value of likelihood function be $L_A^{(1)}$.

6. Repeat these processes until $L_B^{(n)}$ and $L_A^{(n)}$, $\hat{\theta_A}^{(n)}$ and $\hat{\theta_A}^{(n-1)}$ and $\hat{\theta_B}^{(n)}$ and $\hat{\theta_B}^{(n-1)}$ converge to each other.

We use two algorithms, Newton-Raphson and pattern search for each search step. First we find an optimum using Newton-Raphson method. Treating the parameters found by Newton-Raphson method as starting values, we again apply pattern search, a global optimization algorithm which does not use derivatives in finding an optimum. Search algorithms using derivatives such as Newton-Raphson are often weak at finding a global optimum and the performance heavily depends on how close the starting values are to the global optimum, when the likelihood surface is highly non-convex. Nonetheless, we use Newton-Raphson in the first stage, because it helps the search process to be faster and performs better than using pattern search alone.

To check whether the estimated parameters are global maximum, we generated data of which the data generating process follows the baseline state space model. Given the artificial data, we estimated the baseline state space model with randomly drawn starting values by using PZ Alrogithm. The estimated parameters and the likelihood values are identical across different starting values that we tried. This confirms us that PZ Algorithm is successful in estimating the set of parameters which are the global maximum.

To compare the performance of PZ Algorithm to other popularly used numerical search algorithms such as Newton-Raphson and Genetic algorithm, we estimated the model with the same generated data and the same sets of starting values using different algorithms. For instance, with the randomly generate 10 sets of starting values, we experimented whether each solver finds the set of parameters which yields the likelihood value 6108.3 which is the global maximum. While PZ Algorithm succeeded in finding the global maximum with all the 10 sets of starting values, Newton-Raphson method found the global maximum with five sets of starting values out of ten and Genetic algorithm which does not depend on starting values in the search procedure did not succeed in finding the global maximum at all. See the following figure. However, the range of log-likelihood values found by Newton-Raphson method is much larger than that of Genetic algorithm.

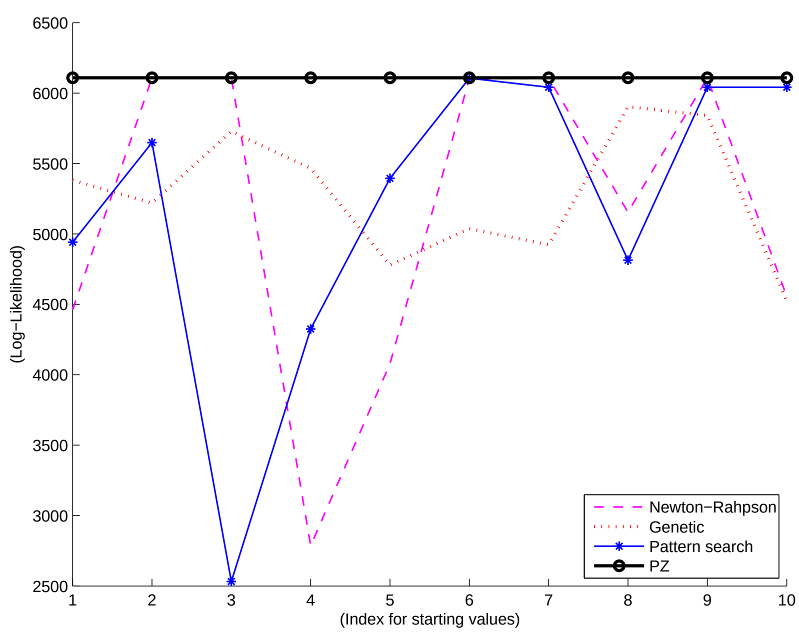

Figure 1.13: Comparison of performance

Data correction: Hornstein (2012)'s method

Hornstein (2012) is different from ours in the way to adjust the CPS redesign in 1994. Hornstein (2012) decreased the size of those who have been unemployed for 5 to 14 weeks by 10% of the short-term unemployed to compensate the inrease of the number unemployed for less than 5 weeks. He claims that the results are not sensitive with respect to the choice of adjustment factor. However, Hornstein (2012)'s adjustment method assumes that only unemployed individuals with duration of unemployment 5-14 weeks have intervening short employment spells between the two reference periods.

To see how a different correction method might change the result, we estimated the model using the dataset corrected with Hornstein (2012)'s method. The parameter estimates are in Table O1. The estimated series and the contribution of each factor to the unemployment rate are documented in Figure O1-O5. Overall, the estimation results are similar to our baseline estimation result.

Genuine Duration Dependence

In this section, we check the validity of various restrictions on the genuine duration dependence parameters. We consider four alternative specification for the genuine duration dependence. The most unrestricted model (Model 4) allows $\delta_1, \delta_2, \delta_3, \delta_4$ to be different from each other. The resrictions on the other three models are summarized below.

	Restriction
Model 1	$\delta_1 = \delta_2 = \delta_3 = \delta_4$
Model 2	$\delta_1 = \delta_2$, $\delta_3 = \delta_4$
Model 3	$\delta_1 = \delta_2$
Model 4	No Restriction

The various likelihood ratio tests, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), suggest that Model 3 is more probable to minimize the infomation loss than the alternatives. Therefore, we adopt Model 3 as the baseline model. The estimation and test results are found in the following table.

Table 1.5: Parameter estimates

White (1982) quasi-maximum-likelihood standard errors in parentheses. Likelihood ratio test(LRT) statistics are computed with Model (4) to be the unrestricted model. (df) denotes degree of freedom.

	Model 1	Model 2	Model 3	Model 4
σ_L^w	0.0399***	0.0403***	0.0446***	0.0451***
	(0.0039)	(0.0039)	(0.0043)	(0.0048)
σ_H^w	0.0463***	0.0469***	0.0465***	0.0465***
	(0.0062)	(0.0062)	(0.0060)	(0.0058)
σ_L^x	0.0458***	0.0462***	0.0445***	0.0439***
	(0.0052)	(0.0055)	(0.0049)	(0.0053)
σ_H^x	0.0191***	0.0191***	0.0218***	0.0221***
	(0.0026)	(0.0026)	(0.0029)	(0.0030)
R_1	0.0987***	0.0986***	0.0977***	0.0975***
	(0.0058)	(0.0058)	(0.0058)	(0.0058)
$R_{2.3}$	0.0770***	0.0776***	0.0760***	0.0759***
	(0.0045)	(0.0046)	(0.0043)	(0.0043)
$R_{4.6}$	0.0831***	0.0847***	0.0776***	0.0774***
	(0.0067)	(0.0071)	(0.0067)	(0.0071)
$R_{7.12}$	0.0579***	0.0564***	0.0590***	0.0593***
	(0.0050)	(0.0050)	(0.0049)	(0.0050)
R_{13+}	0.0363***	0.0359***	0.0366***	0.0367***
	(0.0026)	(0.0026)	(0.0027)	(0.0027)
δ_1	0.0061***	-0.0215	-0.0040	0.0069
	(0.0024)	(0.0155)	(0.0163)	(0.0653)
δ_2	-	-	-	-0.0240
	-	-	-	(0.0475)
δ_3	-	0.0097**	-0.0759***	-0.0760**
	-	(0.0043)	(0.0258)	(0.0306)
δ_4	-	-	0.0846***	0.0840**
	-	-	(0.0292)	(0.0383)
BIC	-4698.26	-4700.13	-4728.79	-4722.95
LRT	43.06	35.06	0.28	-
(df)	3	2	1	-
No. of Obs.	456	456	456	456
Log-likelihood	2379.74	2383.74	2401.13	2401.27

Estimation Results with Hornstein (2012)'s Unemployment Measures

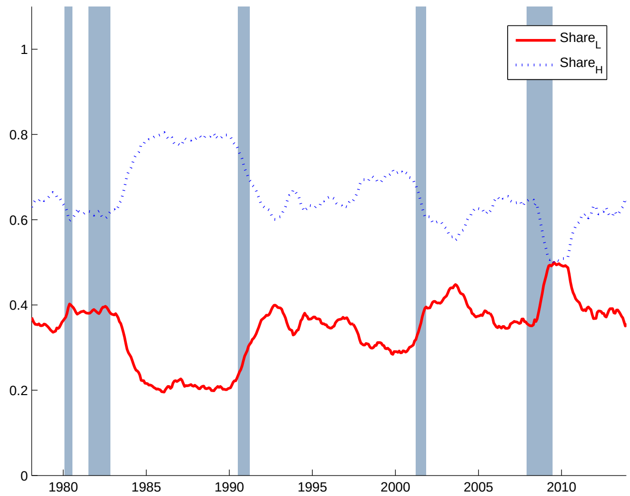

Figure 1.14: Probability that a newly unemployed worker of each type will still be unemployed the following month

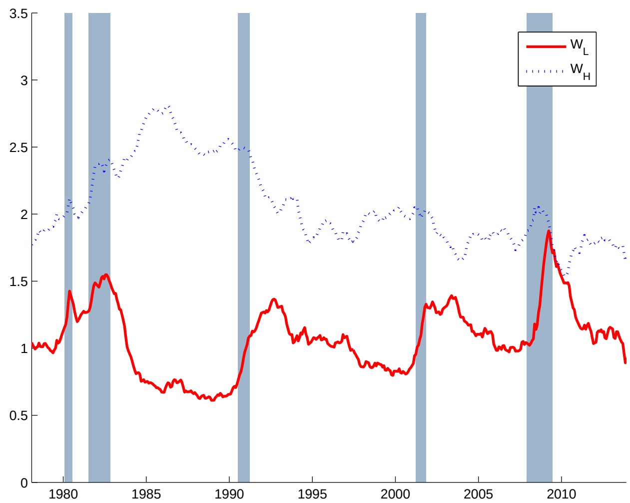

Figure 1.15: Number of newly unemployed workers of each type

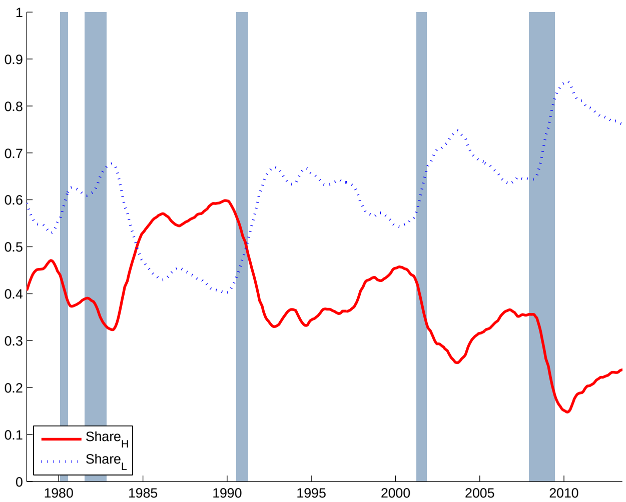

Figure 1.16: Share of total unemployment accounted for by each type of worker

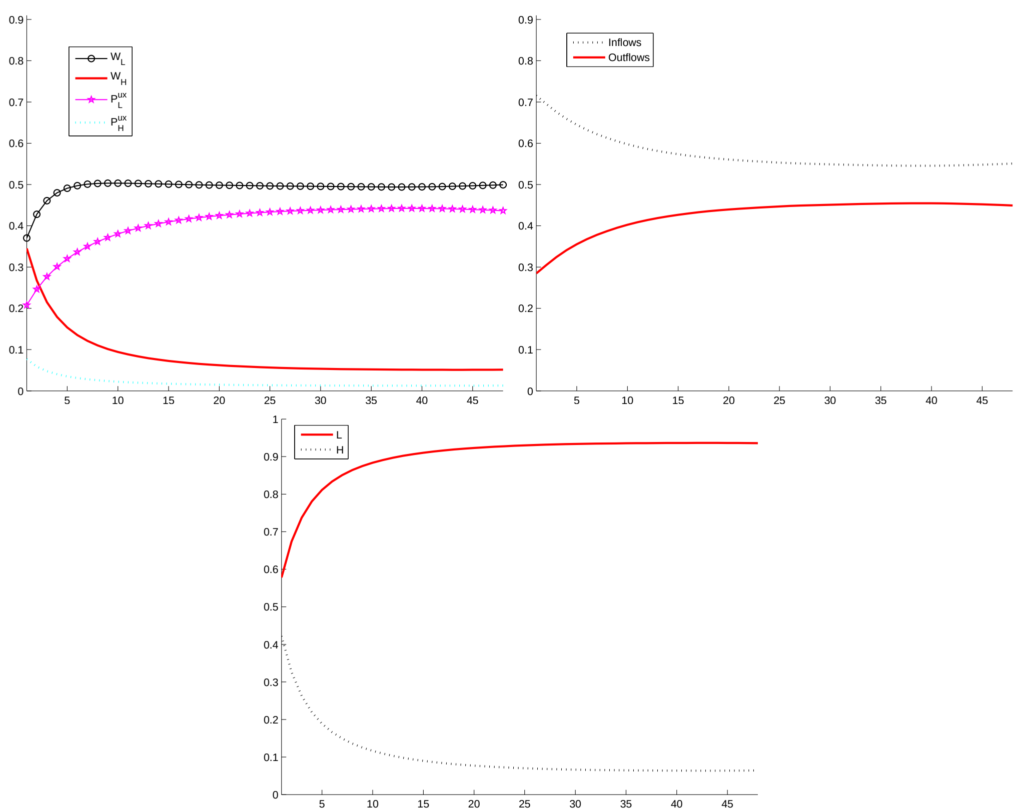

Figure 1.17: Fraction of variance of error in forecasting total unemployment at different horizons attributable to separate factors

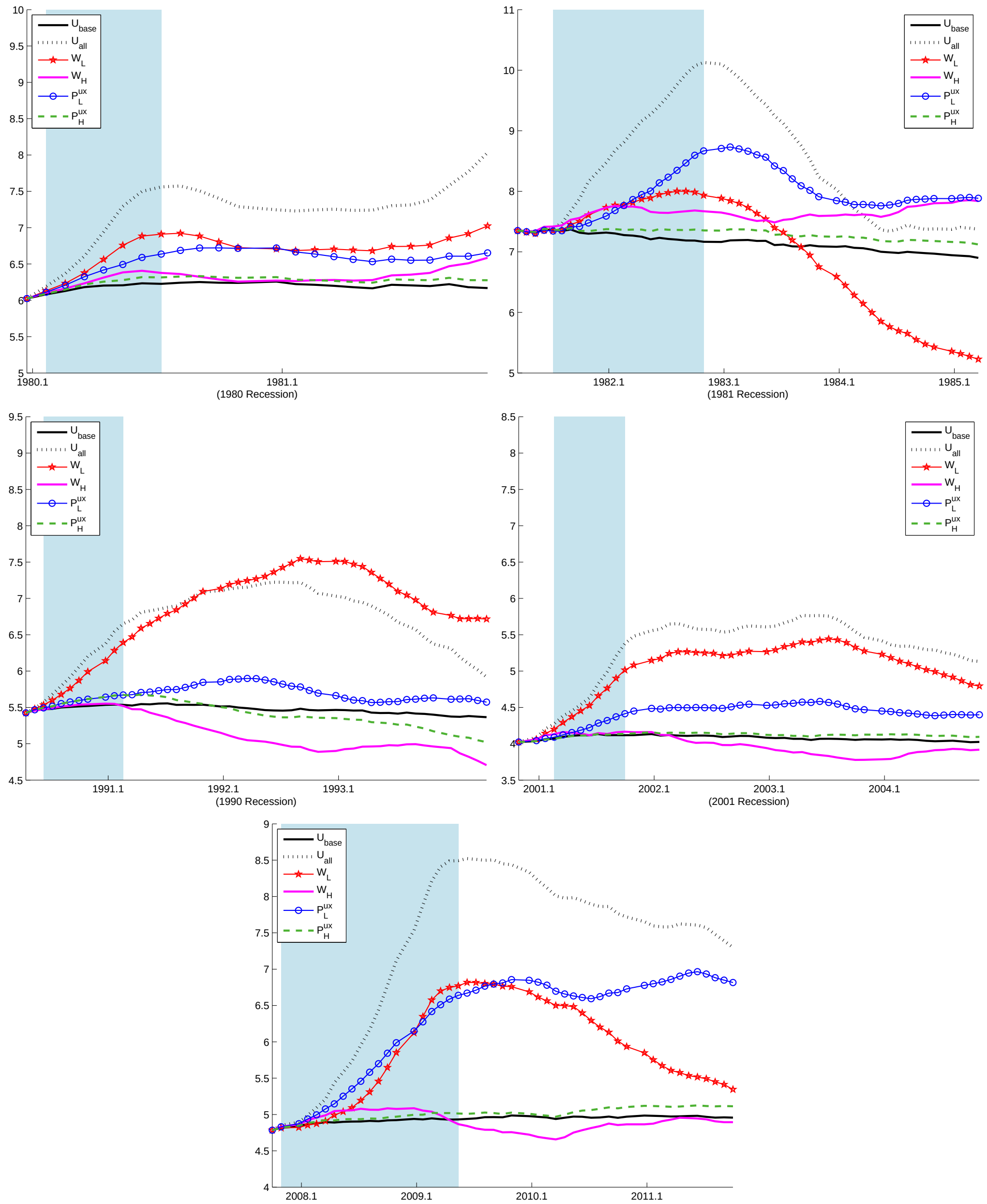

Figure 1.18: Historical decompositions of of U.S. recessions

Estimation Result of Model with Factor Structure

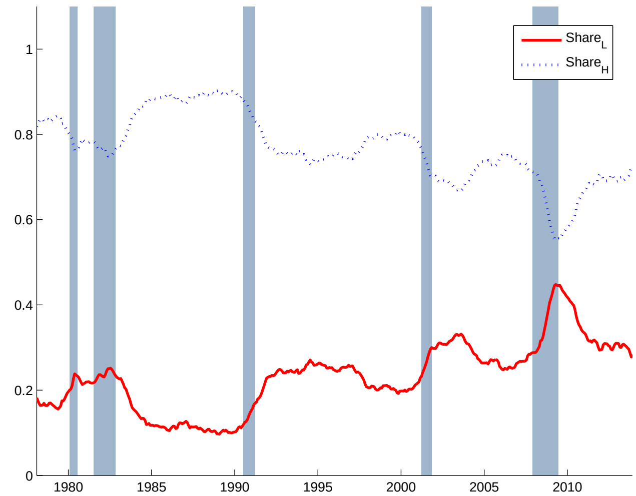

Figure 1.19: Probability that a newly unemployed worker of each type will still be unemployed the following month

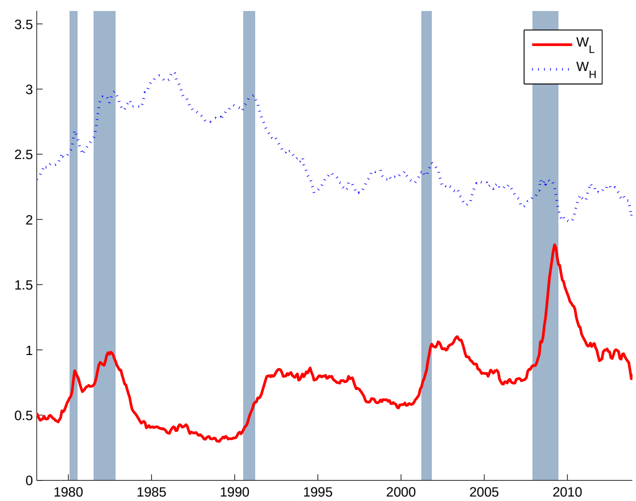

Figure 1.20: Number of newly unemployed workers of each type

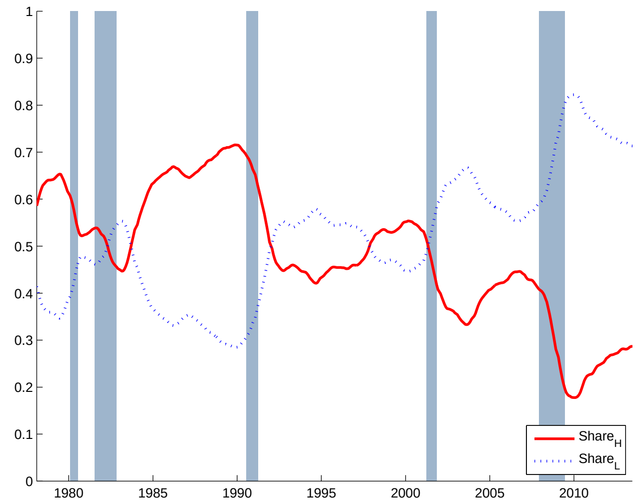

Figure 1.21: Share of total unemployment accounted for by each type of worker

Estimation Result of Model with Weekly Transition

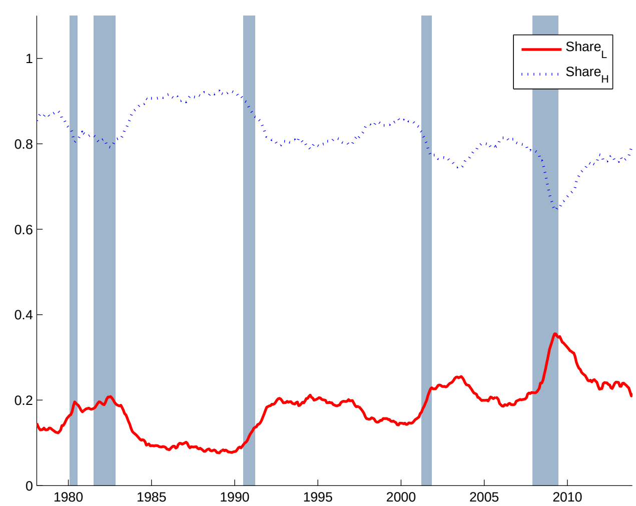

Figure 1.22: Probability that a newly unemployed worker of each type will still be unemployed the following month

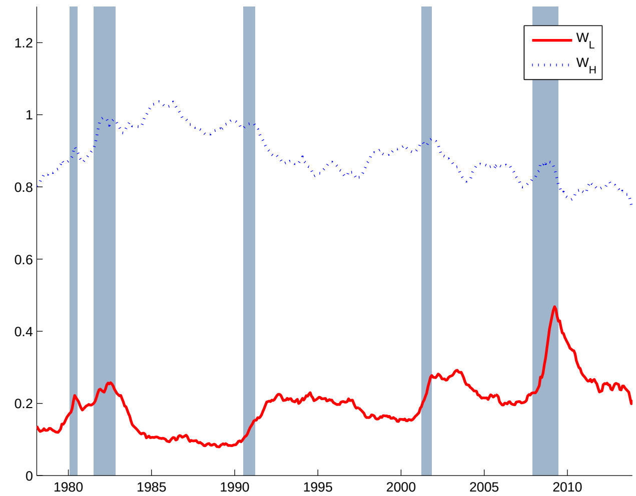

Figure 1.23: Number of newly unemployed workers of each type

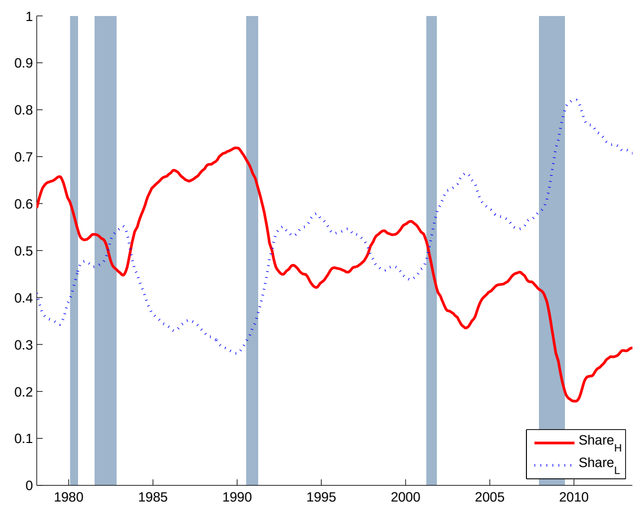

Figure 1.24: Share of total unemployment accounted for by each type of worker

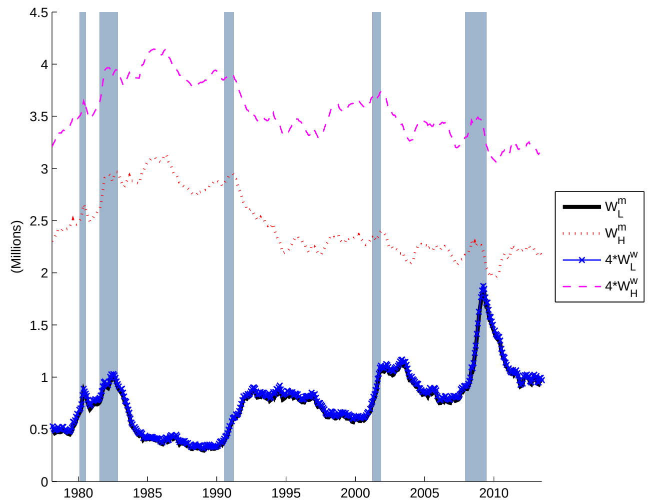

Figure 1.25: Comparison of Inflows

$4 * W_L^w$ and $4 * W_H^w$ denote the total inflows implied by weekly inflows of type L and type H group. W_L^m and W_H^m are the inflows estimated from the baseline model.

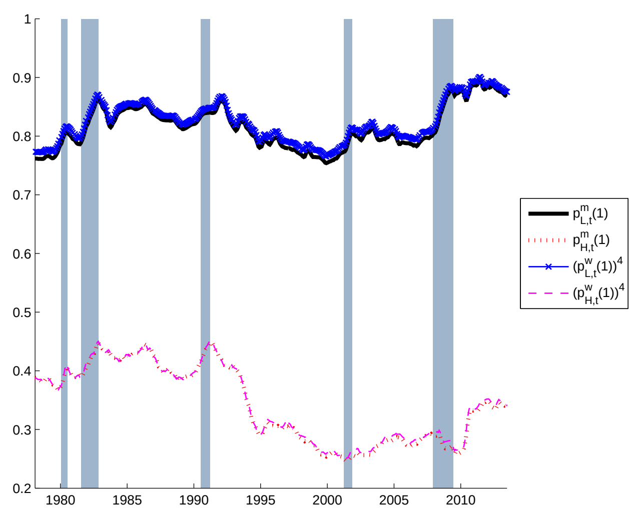

Figure 1.26: Comparison of Continuation Probabilities

$p_L^w(1)^4$ and $p_H^w(1)^4$ denote the monthly continuation probabilities imputed from the weekly continuation probabilities. $p_L^m(1)$ and $p_H^m(1)$ are the continuation probabilities estimated from the baseline model.

Chapter 2

The Role of Observed and Unobserved Heterogeneity in the Duration of Unemployment Spells

Abstract. This paper explores the role of observed and unobserved heterogeneity in explaining both cross-sectional differences across individuals in the duration of unemployment as well as changes in the average duration of unemployment over time. I use a dynamic accounting identity to infer from these vector-valued time series changes in inflows and outflows of different unobserved types of workers within a given observed category. I propose new strategies to explicitly quantify the contribution of unobserved heterogeneity to unemployment duration in the aggregate as well as across individuals. Unobserved heterogeneity explains about one third of the aggregate dispersion in ongoing duration spells of unemployment and 40% of the cross-sectional dispersion in completed duration spells over the 1980-2013 period. The compositional shift of unobserved types is a crucial factor raising the mean duration in progress during the Great Recession. By contrast, observed heterogeneity makes only a minor contribution to either cross-sectional or time-series variation.

2.1 Introduction

The Great Recession has had a major impact on labor market in the United States. In September 2014 - five years into the recovery – the unemployment rate fell below 6% from its peak of 10% in October 2009. However, the average duration of unemployment continued to rise after the end of the recession peaking 41 weeks in December 2011 and has remained elevated. The staggered recovery of average duration of unemployment has been at the center of interest of economists and policy makers. To identify the source of the rise in unemployment duration and its slow recovery, economists have looked at the relation between observable characteristics of unemployed individuals and their duration of unemployment to understand the labor market during recessions.[1] Looking at observable characteristics of the unemployed, such as gender, age, education, occupation, industry, or reason for unemployment, we find that the duration of unemployment rises significantly for almost every observable characteristic during the Great Recession and compositional shifts in the unemployment of different groups explain little of the observed increase in average duration of unemployment. This observation seems to suggest that the rise in unemployment duration is driven by aggregate factors that affect many workers in a similar way. Is this analysis an accurate portrait of the reality?

Figure 2.1 displays some key evidence that has been overlooked in understanding the duration of unemployment in the aggregate. Of the men with some college education who had been unemployed for less than one month as of June 2012, only 58% were still unemployed the next month. In contrast, 92% of the men with some college education who had been unemployed for more than 6 months as of June 2012 were still unemployed the following month. This fact that within a disaggregated group

[1] See for example Aaronson, Mazumder, and Schechter (2010), Elsby, Hobijn, and Sahin (2010) Estevao and Tsounta (2011), Bachmann and Sinning (2012), Davis and von Wachter (2012), Sahin, Song, Topa and Violante (2012), Shimer (2012), Barnichon and Figura (2013), Hall and Schulofer-Wohl (2014), Hall (2014), Kroft, Lange, Notowidigdo, and Katz (2014) and Krueger, Cramer and Cho (2014).

of unemployed individuals the long-term unemployed find jobs or leave the labor force more slowly than others is a strikingly consistent feature in the data and could be crucial for understanding the recent path of the average duration of unemployment.

What accounts for this phenomenon? The only logical inference is that there are important differences on average between the people sharing the same observable characteristics who had been unemployed for 6 months and those who had only been unemployed for 1 month in June 2012, with something about circumstances of the former group making it more likely they would still be unemployed in July 2012 compared to individuals in the latter group.

What accounts for this phenomenon? The only logical inference is that there are important differences on average between the people sharing the same observable characteristics who had been unemployed for 6 months and those who had only been unemployed for 1 month in June 2012, with something about circumstances of the former group making it more likely they would still be unemployed in July 2012 compared to individuals in the latter group.

No two people with the same observed characteristics are actually identical, no matter how fine a gradation of observed categories we use from available data. If some individuals newly unemployed at time t with a given observed characteristic j have an intrinsically lower probability of exiting unemployment than others, those individuals will necessarily make up a larger fraction of those with observed characteristic j who have been unemployed for s months at time $t + s$ as a necessary consequence of dynamic sorting. Consequently, the unemployment exit probabilities of long-term unemployed individuals become lower than those of short-term unemployed individuals as we observe in Figure 2.1. A quite different explanation for the phenomenon is that the experience of being unemployed for a longer period of time directly changes the characteristics of a fixed individual. This possibility is often referred to as genuine duration dependence following Van den Berg and van Ours (1996). Employers may statistically discriminate against those who have been unemployed for longer (Eriksson and Rooth, 2014; Kroft, Lange, and Notowidigdo, 2013) or individuals lose human capital the longer they are unemployed (Acemoglu, 1995; Ljungqvist and Sargent, 1998). I will refer to the condition where a longer period spent in unemployment directly reduces the probability of finding a job as "unemployment scarring." Another possibility is that the longer a person has been unemployed, the more likely they may be to accept a low-paying job or simply drop out of the labor force. Katz and Meyer (1990a,b) argued that such effects may become important as unemployment benefits become exhausted. I will refer to the possibility that the probability of exiting unemployment increases as a consequence of a longer duration of unemployment as "motivational" effects.

Suppose we postulate unobserved heterogeneity within a group of people who all have the same observed characteristics. Some of those newly unemployed at time t with observed characteristic j may be of an unobserved type H with a high probability p_{jt}^{H} of exiting unemployment next month, while others of type L have a lower probability p_{jt}^{L}. We can infer something about changes in inflows or exit probabilities for the two unobserved types by examining the joint panel dynamics of $\{U_{jt}^{1}, U_{jt}^{2.3}, U_{jt}^{4.6}, U_{jt}^{7.12}, U_{jt}^{13.+}\}_{t=1}^{T}$ by making use of the dynamic accounting identity that tells us that those who have been unemployed for 2 or 3 months at time t must have become newly unemployed either at $t-1$ or $t-2$. For example, if there is an increase in the fraction of newly unemployed type L workers at time t with no changes in exit probabilities p_{jt}^{L}, this would first show up in the form of an increase in $U_{j,t+1}^{2.3}$ and $U_{j,t+2}^{2.3}$, with the increase in $U_{j,t+s}^{4.6}$ not showing up until $t+3$, $U_{j,t+s}^{7.12}$ only after 6 months, and so on. By postulating that changes in the inflows and outflows for each unobserved type follow a random walk, Ahn and Hamilton (2014) showed using aggregate unemployment data that a nonlinear state-space model could be used to infer changes in these inflows and outflows as well as estimate the effect from genuine duration dependence. In this paper, I apply a related idea to data disaggregated on the basis of a variety of observed characteristics j to study the role of observed and unobserved heterogeneity in the duration of unemployment.

This paper explores unobserved cross-sectional heterogeneity within people grouped on the basis of various observable characteristics. The identification of unobserved cross-sectional individual difference is achieved from the dynamic accounting identity of unemployment introduced in Ahn and Hamilton (2014). The key feature of dynamic accounting identity is that it allows us to directly observe the distribution of unemployment duration not only of a group with observable characteristic but also of unobserved types residing in the group at time t. Using this feature, we can analyze how much observed and unobserved heterogeneity contributed to the aggregate distribution of unemployment. I propose new strategies to analyze the role of observed and unobserved heterogeneity in the variance and the mean of unemployment duration at each point in time. No previous study has attempted to analyze the role of unobserved heterogeneity within a disaggregated group in the cross-sectional dispersion and time-series dynamics of unemployment duration in the aggregate and this paper is the first to quantify the contribution of unobserved heterogeneity.

Most of the previous studies[2] in the micro-econometric literature assumed that unobserved heterogeneity does not vary over time and analyzed its static effect on the individual exit probability from unemployment.[3] However, changes in labor demand or institutional factors have different effects on individuals with different unobserved attributes. This suggests that unobserved heterogeneity has dynamic features. The approach of this paper is distinct from existing studies in that it allows time variation in unobserved heterogeneity by modeling the observed data as coming from a mixture of two distributions that each vary over time. This enables us to directly infer how inflows of newly unemployed and exit probabilities for currently unemployed type H and L individuals who share the same observed characteristics vary over time. This paper is the first to consider the dynamic behavior of unobserved heterogeneity within groups of in-

[2]See for example Heckman and Singer (1984a,b), Katz and Meyer (1990a,b), Acemoglu (1995), van den Berg and van Ours (1996), Ljungqvist and Sargent (1998), van den Berg and van der Klaauw (2001) Abbring, van den Berg and van Ours (2002), Kroft, Lange and Notowidigso (2013), Krueger and Mueller (2010, 2012), Eriksson and Rooth (2014) and Faberman and Kudlyak (2014).

[3]To my knowledge, Botosaru (2013) is the only one to propose an identification strategy to analyze the dynamic effect of unobserved heterogeneity in individual unemployment hazard.

dividuals sharing the same detailed observed characteristics using a fully specified statistical model. To incorporate fine gradation of individual characteristics, I constructed a unique dataset of unemployment duration using CPS micro data. This paper differs from Ahn and Hamilton (2014) and Hornstein (2012) in the following ways. Ahn and Hamilton (2014) considered dynamic features in unobserved heterogeneity only in aggregate level data and did not allow possible time variation in genuine duration dependence. Hornstein (2012) used minimum-distance estimation in applying the dynamic accounting identity to disaggregated data including age, industry and occupation but with identification incompletely achieved by smoothing penalties. Furthermore, the model not only captures non-monotonic patterns in genuine duration dependence but also allows time variation in genuine duration dependence which is new to the literature. I allow the nature of genuine duration dependence to change over time depending on the generosity of UI benefits. Unlike previous studies where relatively shorter ranges of duration are considered, I model the full range of durations which could exist in the economy.

I find the striking feature that in every group of unemployed individuals with the same observable characteristic there exists substantial heterogeneity in the duration of unemployment both in normal and recessionary periods, which is key to the counter-cyclical fluctuations of unemployment duration of each group. Across various groups the difference between type H and L average duration of unemployment is between 2 and 6 months on average. The share of type L workers in the unemployment of each group exhibits strong counter-cyclical fluctuations, which drives the rise in unemployment duration of the group during recessions. In the aggregate, unobserved types within each group of unemployed individuals with observable characteristic explains about 30% of variance of unemployment duration. In addition, I find that the compositional shift of unobserved types is a crucial factor raising the average duration of unemployment during the Great Recession. By contrast, observed heterogeneity plays a very limited role in explaining both time variation and cross-sectional dispersion of unemployment duration. Particularly, I find that changes in unobserved heterogeneity among newly unemployed individuals are important accounting for about one third of the rise in the average duration of unemployment during the Great Recession and the dispersion of completed duration spells across individuals.

One might argue that during the recent recession, the average duration of unemployment rose dramatically not due to unobserved heterogeneity but because of the deterioration of job finding probability accompanied by unemployment scarring. According to this view, individuals stay unemployed longer due to the substantial fall in the job finding probability and the more of the long-term unemployed become less likely to find a job even further as they become exposed to unemployment scarring. The core difference between this view and the conclusion of this paper lies in how much newly unemployed individuals can drive the rise in the long-term unemployment, a crucial driver of the average duration of unemployment. I highlight the key feature of the data which lead me to the conclusion discussed in details in Section 2.2. By the mechanical feature of U^7, those who have been unemployed for seven consecutive months, changes in type L inflows at time t would show up 6 months later in $U^{7.12}$ at time $t+6$. Likewise, newly unemployed individuals at time t would begin to influence 12 months later in $U^{13.+}$ at time $t+12$. Meanwhile, changes in the continuation probabilities at time t would have immediate effect on $U^{7.12}$ and $U^{13.+}$, since they affect the transitions from unemployment of all the unemployed individuals whose duration of unemployment is six months and over. This suggest that when we forecast $U^{7.12}$ and $U^{13.+}$ with the information of the inflows and continuation probabilities up to t, the errors made in forecasting $U^{7.12}$ between t and $t+6$ and in forecasting $U^{13.+}$ between t and $t+12$ are driven by unanticipated changes in the continuation probabilities. If the exit probability from unemployment is a crucial factor in the development of average duration of unemployment, we will observe large forecast errors, but if it is the inflows of workers with low unemployment exit probabilities, we will observe small forecast errors. Changes in $U^{7.12}$ and $U^{13.+}$ during the later part of the Great Recession observed in prime age men with associate degree or some college education show the latter pattern serving as the evidence for the importance of unobserved heterogeneity.

In Section 1, I describe the data used in this study, which are the numbers of individuals who share a given set of observable characteristics and who report they have been looking for work at various search durations. I show that if the variables governing inflows and exit probabilities follow a random walk, a dynamic accounting identity can be used in a nonlinear state-space model to calculate the likelihood function for the observed data and form an inference about the number of unemployed individuals of each type along with their exit probabilities at each date t in the sample.

Section 2 reports estimation results. I find that different groups show different patterns of genuine duration dependence and that unemployed individuals are likely to postpone the exits from unemployment as they receive the UI benefits for a longer time but to accelerate after they exhaust the UI benefits. I also demonstrate that unobserved heterogeneity is crucial in understanding the low frequency dynamics of unemployment duration and incidence of unemployment in addition to observed demographical changes and the supply of highly educated labor force in the past decades. Type H inflows decreased over time, while type L inflows exhibited a slow upward trend.

In Section 3, I analyze the contribution of observed and unobserved heterogeneity of unemployed individuals to the aggregate distribution of unemployment duration. I demonstrate how one can decompose the variance of unemployment duration in each point in time into the component explained by observed individual characteristics and unobserved types. In addition, I show that how much the compositional change of observed and unobserved heterogeneity in the unemployment contributed to the rise in average duration of unemployment.

Section 4 looks at the extent to which heterogeneity of newly unemployed individuals contributed to the duration of unemployment. I demonstrate how one can use the inflows and continuation probabilities of different groups to measure the contribution of observed and unobserved heterogeneity in the inflows of unemployed individuals to the cross-sectional dispersion and the time series variation of unemployment duration.

In Section 5, I further consider different sets of individual characteristics to check the robustness of the conclusion including reason for unemployment, industry, occupation and other detailed individual categories. The main result is robust regardless of the observable characteristics that are considered.

2.2 Model and Estimation

Ahn and Hamilton (2014) demonstrated that conventional models of unemployment in which unemployed individuals are identical and changes in unemployment are determined by inflows and outflows of homogeneous individuals fail to explain the observed distribution of unemployment duration. They show that the observed distribution of unemployment duration can be explained by allowing both cross-sectional heterogeneity among unemployed individuals and genuine duration dependence. For the cross-sectional heterogeneity, they assume that two groups - type H and L - exist in the unemployment. The inflows and continuation probabilities evolve over time according to unobserved random walks. Meanwhile, the genuine duration dependence is assumed to be deterministic. Given the set of assumptions, Ahn and Hamilton (2014) showed how the inflows and continuation probabilities of two groups at time t can be uncovered using the dynamic accounting identity. The dynamic accounting identity is cast in a state space model where the observed numbers of individuals unemployed for 1 month, 2-3 months, 4-6 months, 7-12 months and over 1 year are expressed into functions of current and past values of inflows and continuation probabilities. I use a variant of the state space model proposed in Ahn and Hamilton (2014) to estimate the continuation probabilities and the fraction of individuals of each of the two unobserved types within an unemployed group with observed characteristics.

2.2.1 Model

The dynamic accounting identity is applied to the observed numbers of unemployed individuals with a certain observable characteristic j whose duration of unemployment is 1 month, 2-3 months, 4-6 months, 7-12 months and longer than 1 year, $y_{jt} = (U^{1}_{jt}, U^{2.3}_{jt}, U^{4.6}_{jt}, U^{7.12}_{jt}, U^{13.+}_{jt})'$. Suppose that these numbers are observed with measurement errors $r_{jt} = (r^{1}_{jt}, r^{2.3}_{jt}, r^{4.6}_{jt}, r^{7.12}_{jt}, r^{13.+}_{jt})'$. Assume that there are two types of workers - type H and L - in group j. Let w^{H}_{jt} be the number of people of type H who are newly unemployed at time t and w^{L}_{jt} be that of type L newly unemployed individuals, where we interpret

$$U^{1}_{jt} = w^{H}_{jt} + w^{L}_{jt} + r^{1}_{jt}. \tag{2.1}$$

I assume smooth variations over time for w^{H}_{jt} and w^{L}_{jt} with each assumed to follow an unobserved random walk, e.g.,

$$\begin{aligned} w^{H}_{jt} &= w^{H}_{j,t-1} + \epsilon^{Hw}_{jt} \\ w^{L}_{jt} &= w^{L}_{j,t-1} + \epsilon^{Lw}_{jt} \end{aligned}$$

where ϵ^{Hw}_{jt} and ϵ^{Lw}_{jt} are the innovation terms which drive the dynamics of w^{H}_{jt} and w^{L}_{jt}. A random walk is a flexible and parsimonious way of modeling time-varying latent variables. In addition, a random walk on parameters is often used as a general approach that can pick up structural changes (Baumeister and Peersman, 2013). Random walk specifications allow the inflows and the continuation probabilities to track structural breaks in the duration data of CPS, which might come from the 1994 redesign of the questionnaire, changes in the definition of words, changes in the classification of industries and occupations and so on and then to just move on after the break to adapt to whatever comes next.[4]

[4]I do not adjust the number of individuals unemployed for 1 month after the CPS redesign in 1994

I define $P^z_{jt}(k)$ as the fraction of individuals of type z who were unemployed for one month or less as of date $t-k$ and are still unemployed and looking for work at time t. Note that in order for someone to have been unemployed for 2-3 months at time t, they either must have been newly unemployed at time $t-1$ and looking for a job at t, or they were newly unemployed at $t-2$ and still looking at $t-1$ and t. Thus $U^{2.3}_{jt}$ can be written as follows

$$U^{2.3}_{jt} = \sum_{z=H,L} \left[w^z_{j,t-1} P^z_{jt}(1) + w^z_{j,t-2} P^z_{jt}(2) \right] + r^{2.3}_{jt}. \tag{2.2}$$

Likewise $U^{4.6}_{jt}, U^{7.12}_{jt}$ and $U^{13.+}_{jt}$ are

$$U^{4.6}_{jt} = \sum_{z=H,L} \sum_{k=3}^{5} \left[w^z_{j,t-k} P^s_{jt}(k) \right] + r^{4.6}_{jt} \tag{2.3}$$

$$U^{7.12}_{jt} = \sum_{z=H,L} \sum_{k=6}^{11} \left[w^z_{j,t-k} P^z_{jt}(k) \right] + r^{7.12}_{jt} \tag{2.4}$$

$$U^{13.+}_{jt} = \sum_{z=H,L} \sum_{k=12}^{47} \left[w^z_{j,t-k} P^z_{jt}(k) \right] + r^{13.+}_{jt}, \tag{2.5}$$

where I terminate the calculations after 4 years of unemployment following Hornstein (2012) and Ahn and Hamilton (2014).

I assume that for unemployed individuals in group j who have already been unemployed for τ months as of time $t-1$, the fraction who will still be unemployed at t is given by

$$p^z_{jt}(\tau) = \exp[-\exp(x^z_{jt} + d^{g_t}_{j\tau})] \quad \text{for} \quad \tau = 1, 2, 3, ... \tag{2.6}$$

where x^z_{jt} is a time-varying magnitude influencing the unemployment exit probability

in this paper. Previous studies increased the number of individuals unemployed for 1 month, because it can affect the contribution of inflows and outflows to unemployment dynamics. Since the main focus of this paper is not to analyze the role of inflows and outflows to unemployment dynamics but to investigate the role of unobserved heterogeneity in shaping the distribution of unemployment duration, I take the duration data as it is.

for all workers of type z regardless of their duration. I also assume that x_{jt}^{H} and x_{jt}^{L} evolve as random walks as follows

$$\begin{aligned} x_{jt}^{H} &= x_{j,t-1}^{H} + \epsilon_{jt}^{Hx} \\ x_{jt}^{L} &= x_{j,t-1}^{L} + \epsilon_{jt}^{Lx}. \end{aligned}$$

I allow the magnitude of genuine duration dependence, $d_{j\tau}^{g_t}$, to depend not only on the duration of unemployment τ but also on conditions governing eligibility for unemployment insurance during month t as captured by the value of g_t. For months t when eligibility for most individuals would expire after $\tau = 6$ months we set g_t = 0, whereas during periods of extended elibility we use g_t = E. We suppose that genuine duration dependence in the first regime is governed by the parameters δ_{j1}^{0}, δ_{j2}^{0}, and δ_{j3}^{0} whereas in the second regime they become δ_{j1}^{E}, δ_{j2}^{E}, and δ_{j3}^{E}. The value of $\delta_{jh}^{g_t}$ governs the genuine duration dependence for someone who has been unemployed for τ months at a time when the regime is g_t according to

$$d_{j\tau}^{g_t} = \begin{cases} \delta_{j1}^{g_t}(\tau - 1) & \text{for } \tau < 6 \\ \delta_{j1}^{g_t}[(6-1)-1] + \delta_{j2}^{g_t}[\tau - (6-1))] & \text{for } 6 \leq \tau < 12 \\ \delta_{j1}^{g_t}[(6-1)-1] + \delta_{j2}^{g_t}[(12-1)-(6-1)] + \delta_{j3}^{g_t}[\tau - (12-1)] & \text{for } 12 \leq \tau \end{cases} .$$

Because we have assumed that the genuine-duration dependence effects as summarized by $d_{j\tau}^{0}$ and $d_{j\tau}^{E}$ are time-invariant and that the type-specific effects x_{jt}^{z} evolve smoothly over time, it is possible to estimate a different value for the parameter $d_{j\tau}^{0}$ and $d_{j\tau}^{E}$ for each τ. Positive $\delta_{jh}^{g_t}$ for $h = 1, 2, 3$ imply motivational effects while negative values imply unemployment scarring over the relevant duration ranges. The reason of allowing the size of genuine duration dependence to vary depending on the unemployment insurance situation is as follows. Krueger and Mueller (2008) showed that unemployed individuals search harder for a job as their duration gets longer while they receive the UI benefits. Meyer (1990) and Katz and Meyer (1990a,b) found that unemployed individuals accelerate exiting unemployment status right before they exhaust their unemployment insurance benefits. These studies suggest the possibility that the relationship between exit probabilities and unemployment duration can change depending on the maximum duration of UI benefits. As the maximum duration of UI benefits are extended, unemployed individuals might reduce their search effort[5], postpone unemployment exits to receive the UI benefits, but quickly leave the labor force or accept a low-paying jobs due to discouragement after they exhaust UI benefits. An extension of eligibility would automatically be implemented providing an additional 13 weeks of eligibility beyond the usual 26 weeks in any state whose unemployment rate exceeds 6.5%, and an additional 20 weeks if the unemployment rate exceeds 8.0%.[6] Since the UI policies are different across states, we have to consider state-by-state differences over time to rigorously study the effect of the extension of UI benefits. However, since the analysis is beyond the scope of this paper, I simply assumed that $g_t = 0$ whenever the national unemployment rate is below 6.5% while $g_t = E$ for any months in which the national unemployment rate is above 6.5%.[7]

The fraction of individuals of type z who were unemployed for one month or less as of date $t-k$ and are still unemployed and looking for work at time t, $P^z_{jt}(k)$, can be written as a product of monthly fractions $p^z_{j,t-j+h}(h)$ for $h = 1, 2, ..., j$ as follows

[5]In the longitudinal analysis of unemployed workers in New Jersey in 2009 and 2010, Krueger and Mueller (2011) found that the amount of time that jobless workers devoted to job search declined by 1.5 minutes with each additional week of unemployment. This suggests that unemployed individuals reduce search effort when the UI benefits become more generous.

[6]The extended benefit program is triggered when a state's insured unemployment rate (IUR), the ratio of insured unemployed workers to the total employment, or total unemployment rate (TUR) reach certain levels. All states must pay up to 13 weeks of extended benefits if the IUR for the previous 13 weeks is at least 5% and is 120% of the average of the rates for the same 13-week period in each of the two previous years. There are two other optional thresholds that states may choose (states may choose one, two, or none). If the state has chosen a given option, it would provide the following. Option 1: an additional 13 weeks of benefits if the state's IUR is at least 6%, regardless of previous years' averages. Option 2: an additional 13 weeks of benefits if the state's TUR is at least 6.5% and is at least 110% of the state's average TUR for the same 13 weeks in either of the previous two years; an additional 20 weeks of benefits if the state's TUR is at least 8% and is at least 110% of the state's average TUR for the same 13 weeks in either of the previous two years (Whittaker and Isaacs, 2014).

[7]I also estimated versions of the model with δ_{jh} constant across months with similar results.

$$P^z_{jt}(j) = p^z_{j,t-j+1}(1)p^z_{j,t-j+2}(2)...p^z_{jt}(j). \tag{2.7}$$

We can arrive at the likelihood function for the observed data $\{y_{j1}, ..., y_{jT}\}$ by assuming that the vector of measurement errors r_{jt} is independent Normal,

$$r_{jt} \sim N(0, R_j),$$

$$\underbrace{R_j}_{5\times5} = \begin{bmatrix} (R^1_j)^2 & 0 & 0 & 0 & 0 \\ 0 & (R^{2.3}_j)^2 & 0 & 0 & 0 \\ 0 & 0 & (R^{4.6}_j)^2 & 0 & 0 \\ 0 & 0 & 0 & (R^{7.12}_j)^2 & 0 \\ 0 & 0 & 0 & 0 & (R^{13.+}_j)^2 \end{bmatrix},$$

where R^1_j, $R^{2.3}_j$, $R^{4.6}_j$, $R^{7.12}_j$ and $R^{13.+}_j$ are the standard deviations of $r^1_{jt}, r^{2.3}_{jt}, r^{4.6}_{jt}, r^{7.12}_{jt}$ and $r^{13.+}_{jt}$ respectively. Let ξ_{jt} be the vector $(w^L_{jt}, w^H_{jt}, x^L_{jt}, x^H_{jt})'$ and $\epsilon_{jt} = (\epsilon^{Lw}_{jt}, \epsilon^{Hw}_{jt}, \epsilon^{Lx}_{jt}, \epsilon^{Hx}_{jt})'$. Our assumption that the latent factors evolve as random walks would be written as[8]

$$\underbrace{\xi_{jt}}_{4\times1} = \xi_{j,t-1} + \underbrace{\epsilon_{jt}}_{4\times1} \tag{2.8}$$

$$\underbrace{\epsilon_{jt}}_{4\times1} \sim N(\underbrace{0}_{4\times1}, \underbrace{\Sigma_j}_{4\times4})$$

$$\underbrace{\Sigma_j}_{4\times4} = \begin{bmatrix} (\sigma^w_{jL})^2 & 0 & 0 & 0 \\ 0 & (\sigma^w_{jH})^2 & 0 & 0 \\ 0 & 0 & (\sigma^x_{jL})^2 & 0 \\ 0 & 0 & 0 & (\sigma^x_{jH})^2 \end{bmatrix}.$$

Since the measurement equations (2.1)-(2.5) are a function of $\{\xi_{jt}, \xi_{j,t-1}, ..., \xi_{j,t-47}\}$, the state equation should describe the joint distribution of ξ_{jt}'s from $t-47$ to t, where

[8] The shock could be contemporaneously correlated and can be captured with a factor structure of Σ_j. Ahn and Hamilton (2014) showed that imposing a factor structure does not change the result.

I and 0 denote a (4×4) identity and zero matrix, respectively:

$$\underbrace{\begin{bmatrix} \xi_{jt} \\ \xi_{j,t-1} \\ \xi_{j,t-2} \\ \vdots \\ \xi_{j,t-46} \\ \xi_{j,t-47} \end{bmatrix}}_{192\times1} = \underbrace{\begin{bmatrix} \underbrace{I}_{4\times4} & \underbrace{0}_{4\times4} & 0 & 0 & ... & 0 & 0 & 0 \\ I & 0 & 0 & 0 & ... & 0 & 0 & 0 \\ 0 & I & 0 & 0 & ... & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & ... & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & ... & I & 0 & 0 \\ 0 & 0 & 0 & 0 & ... & 0 & I & 0 \end{bmatrix}}_{192\times192} \underbrace{\begin{bmatrix} \xi_{j,t-1} \\ \xi_{j,t-2} \\ \xi_{j,t-3} \\ \vdots \\ \xi_{j,t-47} \\ \xi_{j,t-48} \end{bmatrix}}_{192\times1} + \underbrace{\begin{bmatrix} \underbrace{\epsilon_{jt}}_{4\times1} \\ \underbrace{0}_{4\times1} \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}}_{192\times1}. \quad (2.9)$$

2.2.2 Estimation

Our system takes the form of a nonlinear state space model in which the state transition equation is given by (2.9) and observation equation by (2.1)-(2.5) where $P^z_{jt}(\tau)$ is given by (2.7) and $p^z_{jt}(\tau)$ by (2.6). Our baseline model has 15 parameters to estimate, namely the diagonal terms in the variance matrices Σ_j and R_j and the parameters governing genuine duration dependence, δ^0_{j1}, δ^0_{j2}, δ^0_{j3}, δ^E_{j1}, δ^E_{j2} and δ^E_{j3}.

Because the observation equation is nonlinear in the latent variables of interest, the extended Kalman filter can be used to form the likelihood function for the observed data $\{y_{j1}, ..., y_{jT}\}$ and form an inference about the unobserved latent variables $\{\xi_{j1}, ..., \xi_{jT}\}$,. Inference about historical values for ξ_{jt} provided below correspond to full-sample smoothed inferences, denoted $\hat{\xi}_{jt|T}$.

The system of equations is estimated with maximum likelihood. The maximization of the likelihood function is difficult by non-convexity and multimodality of the likelihood surface. To tackle this difficulty, I used PZ algorithm that helped considerably in estimation.

2.2.3 Data

This section briefly describes the data used in this study. I constructed from CPS micro data the numbers of people who have observed characteristic j and have been unemployed for one month, 2-3 months, 4-6 months, 7-12 months, and longer than one year.[9] The observations for month t are collected in the vector $y_{jt} = (U^1_{jt}, U^{2.3}_{jt}, U^{4.6}_{jt}, U^{7.12}_{jt}, U^{13.+}_{jt})'$ for t running from January 1976 through December 2013.[10] For the baseline model each value of j summarizes the individual's age (broken down into the 3 categories 16-24, 25-44, or over 45), 3 categories of education (high school graduates or less, some college or associate degree, or college degree) and gender (male or female). This would generate a total of 18 possible categories, though given limited numbers in some cells, I pooled the second two education categories together for men or women in the age group 16-24, and also pooled the two higher education categories together for women over age 45, for a total of 15 different values that j can take on.[11]

[9]The CPS microdata used for the construction of numbers unemployed with duration less than 5 weeks, between 5 and 14 weeks, between 15 and 26 weeks, between 27 and 52 weeks and longer than 52 weeks by each individual characteristics are publicly available at the NBER website (http://www.nber.org/data/cps_basic.html). Since the CPS is a probability sample, each individual is assigned a unique weight that is used to produce the aggregate data series.

[10]It is well known that the CPS redesign in 1994 understates the size of individuals unemployed for 1 month and could have subsequently affected the size of longer duration groups after 1994. I do not take into account the possible effect of CPS redesign in 1994 on the distribution of unemployment duration for two reasons. First, the goal of this paper is to explain the observed distribution of unemployment duration and the observed share of long-term unemployment. Although the data is observed with possible measurement errors, correcting the measurement errors is beyond the scope of this paper. Second, the main interest of this paper is not about the relative importance of inflows and outflows in the unemployment dynamics, in which the correction of short-term unemployment for the CPS redesign in 1994 could be important as mentioned by Elsby, Michaels and Solon (2009) and Shimer (2012).

[11]The fifteen groups that are considered for the baseline case: (1) Men/Age 16-24/High school graduates and less than high school, (2) Men/Age 16-24/ Some college, associate degree and college graduates, (3) Men/Age 25-44/ High school graduates and less than high school, (4) Men/ Age 25-44/ Some college and associate degree (5) Men/ Age 25-44/ College graduates, (6) Men/ Age 45 and over/ High school graduates and less than high school, (7) Men/ Age 45 and over/ Some college and associate degree, (8) Men/ Age 45 and over/ College graduates, (9) Women/Age 16-24/High school graduates and less than high school, (10) Women/Age 16-24/ Some college, associate degree and college graduates, (11) Women/Age 25-44/ High school graduates and less than high school, (12) Women/ Age 25-44/ Some college and associate degree (13) Women/ Age 25-44/ College graduates, (14) Women/ Age 45 and over/ High school graduates and less than high school, (15) Women/ Age 45 and over/ Some college, associate degree and college graduates.

For extensions beyond the baseline case, I repeated the analysis with y_{jt} broken down by 5 different reasons for unemployment, by 8 different industries, or by 6 different occupations.[12] The numbers of people who report having been unemployed for only one month (U^1), the number who have been unemployed for 2-3 months $(U^{2.3})$, 4-6 months $(U^{4.6})$, 7-12 months $(U^{7.12})$ and more than 12 months $(U^{13.+})$ for the baseline case are plotted in Figures 2.2 and 2.3.[13] The series rise during a recession and fall back down after the recession is over. However, the cyclical patterns change after the 1990 recession. Before 1990, the number of unemployed individuals recovered quickly after recessions are over. After the 1990 recession, those numbers continued to rise after the recession was over in most of the groups and began to fall much later after the end of recession. In particular, $U^{7.12}$ and $U^{13.+}$ continued to go up for several years after the end of recessions and recovered very slowly, which is the key feature of the jobless recovery.

The shares of long-term unemployment are documented in Figures 2.4 and 2.5. For most groups, the fraction continued to rise after the end of recession. In every group, the share of long-term unemployment reached a record high in the year 2011.

In addition, each group exhibits a different trend. There are three demographic changes that generate low-frequency movements particularly in the inflows: the aging population, the increased supply of labor force with higher education and the increased labor force participation and job attachment of women. Abraham and Shimer (2002) show the number of women who re-enter the labor force to look for a job decreased

[12] I consider five categories for reason for unemployment: (1) temporary layoff, (2) permanent separations (including other separation and temporary job ended), (3) job leavers, (4) re-entrants, (5) new entrants, eight categories for industry: (1) agriculture, forestry, fishing, farming and mining, (2) construction, (3) manufacturing, (4) wholesale and retail trade, (5) transportation, utilities and information, (6) Finance, (7) Service, (8) Public administration and six categories for occupation: (1) management, business and financial occupations, (2) professional and related occupations, (3) sales and related occupations, (4) office and adminstrative support occupations, (5) service occupations, (6) construction, extraction, installation, maintenance, repair, production, transportation and material moving ocupations. I used the same detailed occupation classification which Jaimovich and Siu (2012) used to construct their series available from January 1983.

[13] Seasonally adjusted series are used for the estimation. However, I plotted 12 month moving averages so that it is easier for readers to observe the patterns of five series.

over time, as it became less likely for them to leave the labor force from employment.

Interestingly, the increased labor force attachment of women and the aging population seem to interact with the increased supply of labor force with higher education. In the top panels of Figures 2.2 and 2.3, we observe a downward trend in the inflows of young men or women with no college experience. This trend is typically not seen in older or more educated individuals. Particularly, the downward trend in the inflows of women with lower education suggests that they may have been more likely in the past to quit their job and leave the labor force due to marriage or raising children in the past. The increased inflows of workers aged 45 and over are mainly found among those with education attainment higher than a high school diploma. The aging population seems to be closely associated with the changes in education level of the society when it comes to the dynamics of unemployment and unemployment duration. The aggregate number of newly unemployed shows a secular decline because the downward trends outweigh the upward trends.

2.3 Empirical results

In this section, I report smoothed estimates for the inflows and unemployment continuation probabilities of type H and L individuals, the patterns of genuine duration dependence and the share in the unemployment of type H and L individuals in each group. With the estimates, I recover the distribution of unemployment duration of type H and L individuals in group j each point in time by calculating the number of those unemployed for τ months at time t of type H and L in group j, $U_{jt}^{H\tau}$ and $U_{jt}^{L\tau}$ for $\tau = 1, 2, 3, ...$

The smoothed estimates for the continuation probabilities of types H and L individuals are plotted in Figures 2.6 and 2.7. The stark feature is that in every group of unemployed individuals with the same observable characteristic there exists substantial heterogeneity in terms of the continuation probability both in normal and recessionary periods. Average type L continuation probabilities (reported in row 2 of Tables 3 and 4) are between 0.75 and 0.97. Average type H continuation probabilities (row 3) are between 0.35 and 0.54. In addition, the continuation probability of each group is counter-cyclical but at the same time exhibits its own dynamics. The Great Recession is distinct from other recessionary episodes in that the continuation probabilities of both types reached the highest levels of the sample period during the recession and remain elevated even after the end of the recession for most groups. This might reflect the insufficient demand for workers of both types. However, in some groups, such as men aged 25 and over with high school diploma or less, type H continuation probabilities recovered close to the pre-recession levels but type L probabilities still remain elevated.

Figures 2.8 and 2.9 plot smoothed estimates for the inflows of type H and L individuals within each group j. Over the full sample, type L individuals make up a small portion in the inflows and represent on average 10-40% of the newly unemployed across all groups, as seen in the first rows of Tables 3 and 4. During recessions, inflows of both types H and L rise during recessions but the share of type L workers goes up in the inflows. This is because type L inflows exhibit stronger counter-cyclicality than type H inflows. During the Great Recession, inflows of type L workers reached the highest level and took the largest share in newly unemployed individuals for most groups. It is also interesting to observe that the cyclical fluctuations of type L inflows began to show a quite different pattern from the past after the 1990 recession. While type L inflows decreased right after the recession was over in the 1980's, type L inflows either continued to rise or recovered slowly after the end of recession since the 1990's, which is a new feature of the jobless recovery. Consequently, the type L share in the inflows continued to remain elevated even after the end of recession and it took longer for the share of type L workers to go back to the pre-recession level. This feature is more prominent among women than men.

Looking at the individual group in detail, type L inflows of every group rises during recessions as shown in Figure 2.10. The composition of each group in total type L inflows shows that there is no distinct group which mainly drives the cyclical fluctuation of type L inflows. This suggests that those who have particular low exit probabilities are found in most of the groups and that the number of them in each group rises during recessions. If we consider the composition of broad category such as education and age as shown in Figure 2.11, we still do not see any group of particular characteristics driving the rise in the total type L inflows and not observe clear counter-cyclical changes of composition in the total type L inflows like the disaggregated case.

Among the individual characteristics to be considered in the baseline case, the education attainment of high school diploma or less seems to be the most important single characteristic of newly unemployed type L workers. As documented in Table 2.15, unemployed individuals with high school diploma or less take more than 50% of type L inflows. Type L individuals tend to be less educated than a typical member of the labor force, but this is primarily driven by the fact that less-educated individuals make up a higher portion of the newly unemployed. People with high school diploma or less make up the smaller portion of newly unemployed type L individuals than they do of newly unemployed individuals. The overall evidence suggests that type L individuals cannot be uniquely associated with any fixed age-gender-education characteristic, although they are important component of the inflows.

Tables 2.1 and 2.4 report parameter estimates for the 15 datasets, with the patterns of genuine duration dependence implied for each group shown in Figure 2.12. Some of the parameters that govern the genuine duration dependence are not statistically significant. Nonetheless, there are three interesting features. First, each group displays different patterns of genuine duration dependence.[14] Second, non-monotonic patterns are observed among most groups. In the range of duration up to 6 months, the

[14]Van den Berg and van Ours (1996) also found different patterns in the genuine duration dependence across various demographic groups. They found motivational effects for black workers unlike white male who are more likely to be exposed to unemployment scarring. They explain that strong anticipation of unemployment benefits exhaustion is the reason for the motivational effects. This paper is different from theirs in that I consider non-monotonic genuine duration dependence and used the data with more detailed individual characteristics.

coefficients $\delta_{j1}^{g_t}$ can be either positive or negative.[15] Unemployed individuals become less likely to exit the unemployment pool between 6 and 11 months as the duration of unemployment gets longer, but they are more likely to do so from 1 year.[16] Third, when the UI benefits are likely to be extended beyond 6 months, $\delta_{j2}^{g_t}$ becomes slightly smaller or more negative, but $\delta_{j3}^{g_t}$ tends to become stronger after 1 year than during normal times in many groups. In other words, unemployed individuals become more likely to postpone giving up job search while they receive the extended UI benefits, but to give up job search much faster after the UI benefits are exhausted.[17]

2.3.1 Distribution of unemployment duration

With the estimates, we can calculate the number of those unemployed for τ months at time t of type H and L in group j, $U_{jt}^{H\tau}$ and $U_{jt}^{L\tau}$. This allows us to directly observe the distribution of unemployment duration of type H and L individuals in group j at t. Figures 2.11Fraction of variance of error in forecasting total unemployment at different horizons attributable to separate factors in the model with correlated errors. Horizontal axis: number of months ahead s for which the forecast is formed. Panel A: contribution of the aggregate factor F_t along with the idiosyncratic components of $\{w_{Ht}, w_{Lt}, x_{Ht}, x_{Lt}\}$ separately. Panel B: combined contributions of idiosyncratic components of $\{w_{Ht}, x_{Ht}\}$ and $\{w_{Lt}, x_{Lt}\}$ along with aggregate factor F_t. Panel C: combined contributions of idiosyncratic components of $\{w_{Ht}, w_{Lt}\}$ and $\{x_{Ht}, x_{Lt}\}$ along with aggregate factor F_t. and Figure 2.12 plot the distribution of unemployment

[15] Krueger and Mueller (2008) showed that unemployed individuals search harder for a job as they stay unemployed longer until 6 months in the unemployment when they exhaust the UI benefits. At the same time, they become less attractive to potential employers as their duration of unemployment gets longer. Kroft, Notowidigdo and Katz (2013) showed from a controlled experiment that the likelihood of receiving a callback for an interview significantly decline during the first eight months in the unemployment. This implies that an unemployed individual could become more or less likely to exit unemployment depending on the characteristics of individuals, as they stay longer unemployed until 6 months in the unemployment.

[16] Possible long-term consequences of unemployment scarring (e.g., Paul Oyer, 2006) or extension of UI benefits (e.g., Schmieder, von Wachter and Benderm, 2012) are not considered in this paper.

[17] Rothstein (2012) and Farber and Valletta (2013) showed the extended UI benefits during the Great Recession played a limited role in the rise of unemployment rate and duration of unemployment.

duration of different groups and type H and L distribution within each group averaged over December 2007 and December 2013.

In each group the difference in the shape of distribution is substantial between type H and L workers. While type H distribution attenuates sharply as the duration increases, type L distribution attenuates slowly contributing the duration distribution of group j to have a fat tail. There also exists difference in the dispersion of distribution among the groups of unemployed individuals with different observable characteristics. More flat distributions are observed in older and highly educated job seekers. However, the difference is fairly small compared to that between type H and L distributions.

Using the distributions, we can calculate the average unemployment duration of each type z for $z = H, L$ in group j from

$$D_{jt}^{z} = \frac{\sum_{\tau=1}^{48} \tau U_{jt}^{z\tau}}{\sum_{\tau=1}^{48} U_{jt}^{z\tau}},$$

and the average unemployment duration of group j from the following

$$D_{jt} = \frac{\sum_{z=H,L} \sum_{\tau=1}^{48} \tau U_{jt}^{z\tau}}{\sum_{z=H,L} \sum_{\tau=1}^{48} U_{jt}^{z\tau}}.$$

The duration of unemployment of type H and L individuals of each group is plotted in Figures 2.16 and 2.17. Type H and L workers' durations of unemployment differ substantially in normal and recessionary periods. Average type L duration is between 5 and 10 months and average type H duration is between 2 and 3 months. While type H unemployment duration exhibits weak counter-cyclicality, the unemployment duration of type L workers shows strong counter-cyclical fluctuations. The unemployment duration of each group comove with its type L duration suggesting that type L duration is the key determinant of the unemployment duration of each group.

The share of type L workers in each unemployment group j also exhibits strong counter-cyclicality as shown in Figures 2.18 and 2.19. This is driven by the rises in both the inflows and continuation probabilities of type L workers during recessions. Due to the record high levels of inflows and continuation probabilities, the share of type L workers reached the highest level in the majority of groups during the Great Recession. Except for a few groups[18], the shares of type L workers remain elevated compared to the pre-recession levels in the year 2013, 4 years into the recovery. This suggests that the compositional shift of unobserved heterogeneity within each group can be a crucial factor driving the rise in unemployment duration as well as the sluggish recovery of it. The dramatic rise in type L share as well as type L duration commonly found in most of the groups explains why we observe the sharp increase in the average duration of unemployment and the share of long-term unemployment in every corner of the economy as discussed in previous studies. (Hall, 2014; Kroft, Lange, Notowidigdo and Katz, 2014; Krueger, Cramer and Cho, 2014)

2.3.2 Identification of inflows and outflows

One might argue that the average duration of unemployment rose dramatically not due to unobserved heterogeneity but because of the deterioration of job finding probability accompanied by unemployment scarring. According to this view, individuals stay unemployed longer due to the substantial fall in the job finding probability and to unemployment scarring. The core difference between this view and the conclusion of this paper lies in how much differences across newly unemployed individuals contribute to the rise in the long-term unemployment, a crucial driver of the average duration of unemployment. Here, I highlight the key feature of the data which leads me to my conclusion.

The intuition of this exercise as follows. Changes in type L inflows at time t would show up 6 months later in $U^{7.12}$ at time $t+6$, because only those who have been

[18] Men/Age 45 and over/Associate degree or some college and Women/Age 25-44/College graduates.

unemployed for six consecutive months after they initially become unemployed would be counted in U^7. Likewise, newly unemployed individuals at time t would begin to influence $U^{13.+}$ 12 months later at time $t+12$. By contrast, changes in the continuation probabilities at time t would have an immediate effect on $U^{7.12}$ and $U^{13.+}$, since they affect the transitions from unemployment of all the unemployed individuals. Therefore, the direct and clear litmus test for the importance of continuation probabilities is to see the errors made in forecasting $U^{7.12}$ between t and $t+6$ and in forecasting $U^{13.+}$ between t and $t+12$ based on the information of the inflows and continuation probabilities up to time t. Large forecast errors suggest the importance of type L continuation probabilities, while small forecast errors imply the crucial role in the type L inflows.

As a representative example, I first consider men aged 25-44 years with some college education or an associate degree. The forecasts for each observed series based on our smoothed inferences on the inflows and continuation probabilities about conditions as of September 2008 are indicated by the solid red lines in Figure 2.19. Since we treat inflows as random walks, the model forecast for U^1_{jt} is a horizontal line. The fact that the number of newly unemployed went up over the next 10 months implies that there was some change in either w^L_{jt} or w^H_{jt}. Our model also treats continuation probabilities as random walks, which means, given our inferred mix of type L and H workers as of September 2008, our model predicted that $U^{4.6}_{jt}$ would rise for 6 months before flattening out, $U^{7.12}_{jt}$ would rise for 12 months before flattening, and so on. If there was an unanticipated change in subsequent exit probabilities for individuals of type L, this would show up as higher values for $U^{7.12}_{jt}$ than predicted over the next 6 months, higher values for $U^{13.+}_{jt}$ at least over the next 12 months, and so on. Only $p^L_{j,t+s}$ could matter for these, because $p^H_{j,t+s}$ plays a very small role in unemployment durations longer than 6 months. The observed values for $U^{7.12}_{jt}$ and $U^{13.+}_{jt}$ are only slightly higher than one would forecast based on conditions as of September 2008, meaning that most of the increase in the duration of unemployment for this demographic group was due to increased inflows of type L workers into the pool of unemployed rather than a change in their unemployment continuation probabilities. This is the feature of the data that causes our model to infer relatively little change in continuation probabilities for this demographic group during the later part of the recession (as seen in the fourth panel of Figure 2.6) but a big increase in inflows of type L individuals (Figure 2.8). The starred fuchsia line in Figure 2.19 shows the implied path for $U^x_{j,t+s}$ if we condition only on subsequent inflows of type L individuals but none of the other shocks. Moreover, we see from the starred fuchsia lines how type L inflows determined the differential peaks for each duration group and why the peaks of each duration group were delayed by the difference in the durations between two adjacent groups.

However, different demographic groups sometimes exhibited different patterns. Figure 21 shows the analogous data and forecasts for women aged 25-44 years with some college education or an associate degree. U^1_{jt} for this demographic group does not exhibit a clear increase and neither does $U^{2.3}_{jt}$. However, $U^{4.6}_{jt}, U^{7.12}_{jt}$ and $U^{13.+}_{jt}$ all rise during the period of interest and the size of increase becomes bigger in the longer-term unemployment. This implies that the rise in the long-term unemployment is not driven mainly by the inflows. Unlike the previous case, the observed values of $U^{7.12}_{jt}$ is significantly higher than the corresponding forecast made on November 2008 over the next 12 months. Likewise, the observed values for $U^{13.+}_{jt}$ are substantially larger than the predicted values over the next 16 months. In this group, the rise in average duration of unemployment is mainly explained by type L continuation probabilities as seen in the fourth panel of Figure 2.7.

These two examples show that the crucial factors driving the rise in long-term unemployment are different across groups. However, it is notable the pattern in Figure 20 is often the one seen for other demographic groups. Composition of change in the inflows appears to be critical in explaining the rise in average duration of unemployment during the Great Recession.

2.3.3 Unobserved heterogeneity and trends in the duration of unemployment

Aside from cyclical properties, different demographic groups display quite different low frequency movements in their inflows and continuation probabilities. One of the important features in the labor market in the last three decades is that the average duration of unemployment showed an upward trend in spite of the secular decrease in the incidence of unemployment as discussed in the previous studies[19]. My empirical results provide a new explanation for the observed trends in the labor market, and establish that unobserved heterogeneity is crucial in understanding the low frequency dynamics of unemployment duration and the incidence of unemployment.

The upward trend in the inflows of individuals age 45 and over seen in the last rows of Figures 2.3 and 2.4 is driven by both type H and L workers. In contrast, the downward trends in the inflows of men age 16-24 whose education attainment is less than and equal to high school diploma, in those of women age 16-24 who are not college graduates and in those of women age 25-44 who are high school graduates or less are mainly driven by the secular decrease of type H inflows. The level in type L inflows either stays at the same level or even increases over time. In other words, the downward trends in the inflows of these groups have been concentrated in the group of people who are likely to exit unemployment status relatively quickly and are not accompanied by the decreases in the individuals who are likely to stay unemployed longer. This suggests that the demographic changes and the increased education level in the society have asymmetric effects on type H and L workers. It further implies that unobserved heterogeneity in the inflows is critical in understanding the upward trend in the average duration of unemployment accompanied by a downward trend in the number of newly unemployed individuals.

One of the factors that drove a secular rise in the average duration of unem-

[19] See for example Abraham and Shimer (2002) and Aaronson, Mazumder and Schechter (2010).

ployment of women is the upward trend in their type H continuation probabilities.[20] This might show another aspect of increased labor force attachment of women. Rather than leaving the labor force when they lose their jobs, women might have become more likely to continue searching for a job longer. In sum, these features show that unobserved heterogeneity is crucial in understanding the low frequency dynamics in the U.S. labor market.

2.4 Heterogeneity and aggregate unemployment duration across all workers

So far I have investigated the role of unobserved heterogeneity within a group of workers who all share the same observed characteristics in determining the unemployment duration of that particular group. In this section, I broaden the view from individual group to the consequences for the overall labor force, asking three different questions. (1) For the total group of all individuals who become newly unemployed during a particular month t, how much of the differences across those individuals in how long it will take before they complete their unemployment spell can be explained by differences across individuals in observed characteristics and unobserved types? (2) For the total group of all individuals who are unemployed during a particular month t, how much of the differences across those individuals in how long they have been unemployed so far can be explained by differences across individuals in observed characteristics and unobserved types? (3) How much of the changes over time in the average length of time that unemployed individuals have been looking for work be explained by changes over time in observed characteristics and unobserved types?

[20] This pattern is also found in type H continuation probabilities of re-entrants and new entrants to the labor force.

2.4.1 Heterogeneity in the cross-sectional dispersion

2.4.2 Heterogeneity and the cross-section dispersion of completed unemployment spells

I first analyze how much of the differences across newly unemployed individuals in any given month in how long it will take before they complete their unemployment spell can be explained on the basis of their observed characteristics and unobserved types. I propose to answer this question by measuring how much knowledge of observed characteristics or unobserved types of newly unemployed individuals could reduce the mean squared error of the predicted length of time before that individual will exit unemployment.

Consider first an econometrician who has data on the observed characteristics O_m of a newly unemployed individual m but does not know the unobserved type Q_m of the individual. Conditional on knowing that $O_m = j$ (individual m has observed characteristic j), the econometrician would know that the current unemployment exit probability for the individual is q_{jt}. If this exit probability is expected to persist[21], the number of months that individual m is going to be unemployed before finding a job, n_{mt}, would have a geometric distribution with mean given by $E(n_{mt}|O_m = j) = 1/q_{jt}$, expected square $E(n_{mt}^2|O_m = j) = (2/q_{jt}^2) - (1/q_{jt})$, and variance $Var(n_{mt}|O_m = j) = (1/q_{jt}^2) - (1/q_{jt})$. If f_{jt} denotes the fraction of those newly unemployed at time t with observed characteristic j, then the average squared error for predicting realized duration across the pool of individuals who became newly unemployed at time t that would be made by an econometrician who made optimal use of observed characteristics

[21]This is likely to be a very good approximation since x_{jt} follows a random walk and allows simple closed-form expressions for all the following formulas. Note I also abstract from genuine duration dependence for purposes of calculating these summary statistics. A generalization of these results that takes into account potential time variation in future q_{jt}^z and genuine duration dependence can be obtained by simulating the model forward from any date t. This more general simulation method is the one that was used to produce Figure 22. Analytical closed-form expressions for the simpler case are used in the text in order to communicate the motivation and intuition behind these calculations.

would be given by

$$\sum_{j=1}^{J} f_{jt}[(1/q_{jt}^2) - (1/q_{jt})].$$

Compare this with the average squared error in predicting unemployment duration that would be made by an econometrician who did not make use of information about the observed characteristics of individual m. The predicted duration in this case would just be the unconditional mean,

$$\begin{aligned} E(n_{mt}) &= \sum_{j=1}^{J} f_{jt} E(n_{mt}|O_m = j) \\ &= \sum_{j=1}^{J} f_{jt}(1/q_{jt}), \end{aligned}$$

with associated MSE

$$\begin{aligned} & E(n_{mt}^2) - [E(n_{mt})]^2 \\ = & \sum_{j=1}^{J} f_{jt} E(n_{mt}^2|O_m = j) - \left[\sum_{j=1}^{J} f_{jt} E(n_{mt}|O_m = j)\right]^2 \\ = & \sum_{j=1}^{J} f_{jt}[(2/q_{jt}^2) - (1/q_{jt})] - \left[\sum_{j=1}^{J} f_{jt}(1/q_{jt})\right]^2 \\ = & \left\{\sum_{j=1}^{J} f_{jt}[(1/q_{jt}^2) - (1/q_{jt})]\right\} + \left\{\sum_{j=1}^{J} f_{jt}(1/q_{jt}^2) - \left[\sum_{j=1}^{J} f_{jt}(1/q_{jt})\right]^2\right\} \quad (2.10) \end{aligned}$$

The left-hand side of (2.10) gives the MSE in predicting unemployment duration if we make no use of observed characteristics. The first term in the right-hand side is the MSE if we make use of observed characteristics, while the second term is the amount by which observed characteristics contribute to the dispersion of unemployment duration across those individuals who become newly unemployed at time t. In other words, the first term in the right-hand side is the average dispersion within each group and the second term captures the dispersion explained by the difference among observed groups. Equation (2.10) corresponds to the familiar decomposition

$$Var(n_{mt}) = \underbrace{E[Var(n_{mt}|O_m = j)]}_{\substack{\text{heterogeneity} \\ within \\ \text{observed characteristics.}}} + \underbrace{Var[E(n_{mt}|O_m = j)]}_{\substack{\text{heterogeneity} \\ between \\ \text{observed characteristics}}}. \tag{2.11}$$

The first term in the right hand side of equation (3.2) captures the component of variance that can be explained by heterogeneity within observed characteristics and the second term is the variance explained by heterogeneity between observed characteristics.

Next suppose that the econometrician further has information on the individual's unobserved type z represented by $Q_m = z$ for $z = H$ or L. Conditional on knowing that $O_m = j$ and $Q_m = z$, the econometrician would know that the current unemployment exit probability for the individual is q_{jt}^z. The mean squared error we would make in predicting person m's duration of unemployment would be given by

$$E(n_{mt}^2|Z_m = z, O_m = j) - [E(n_{mt}|Z_m = z, O_m = j)]^2 = [1/(q_{jt}^z)^2] - (1/q_{jt}^z).$$

If f_{jt}^z denotes the fraction of those newly unemployed at time t with observed characteristic j and unobserved type z, then the average squared error for predicting realized duration across the pool of individuals who became newly unemployed at time t that would be made by the econometrician who used only the observed characteristics would be given by

$$\sum_{z=H,L} f_{jt}^z[(1/(q_{jt}^z)^2 - (1/q_{jt}^z)]. \tag{2.12}$$

Using equation (2.12), we can further decompose the mean squared error for predicting the completed duration spells of individuals with observed characteristic j

into the MSE if we also know the unobserved type within group j as follows:

$$Var(n_{mt}|O_m = j) = \qquad (2.13)$$
$$\left\{\sum_{z=H,L} f_{jt}^z[(1/(q_{jt}^z)^2 - (1/q_{jt}^z)]\right\} + \left\{\sum_{z=H,L} f_{jt}^z[1/(q_{jt}^z)^2] - [\sum_{z=H,L} f_{jt}^z(1/q_{jt}^z)]^2\right\}.$$

The first term on the right-hand side is the average dispersion within each type and the second term captures the dispersion explained by the difference between the two types. Equation (**??**) corresponds to the familiar decomposition

$$Var(n_{mt}|O_m = j) = \underbrace{E[Var(n_{mt}|Z_m = z, O_m = j)]}_{\substack{\text{heterogeneity} \\ within \\ \text{unobserved types}}} + \underbrace{Var[E(n_{mt}|Z_m = z, O_m = j)]}_{\substack{\text{heterogeneity} \\ between \\ \text{unobserved types}}}. \qquad (2.14)$$

In other words, the dispersion in completed duration spells of individuals in group j is decomposed into the average variance explained by heterogeneity within unobserved types and the variance accounted for by heterogeneity between unobserved types.

By plugging equation (**??**) into equation (2.10), we have the full decomposition result for the mean squared error of individual duration spells

$$\begin{aligned} & E(n_{mt}^2) - [E(n_{mt})]^2 \\ = & \left\{\sum_{j=1}^{J} f_{jt}(1/q_{jt}^2) - [\sum_{j=1}^{J} f_{jt}(1/q_{jt})]^2\right\} \\ & + \sum_{j=1}^{J} f_{jt}\left\{\sum_{z=H,L} f_{jt}^z[1/(q_{jt}^z)^2] - [\sum_{z=H,L} f_{jt}^z(1/q_{jt}^z)]^2\right\} \\ & + \sum_{j}^{J} f_{jt}\left\{\sum_{z=H,L} f_{jt}^z[(1/(q_{jt}^z)^2 - (1/q_{jt}^z)]\right\}. \end{aligned} \qquad (2.15)$$

This corresponds to

$$
\begin{aligned}
Var(n_{mt}) \quad = \quad & \underbrace{Var[E(n_{mt}|O_m = j)]}_{\substack{\text{heterogeneity}\\ between \\ \text{observed characteristics}}} + \underbrace{E_{O_m}\{Var[E_{Z_m}(n_{mt}|Z_m = z, O_m = j)]\}}_{\substack{\text{heterogeneity}\\ between \\ \text{unobserved types}}} \\
& + \underbrace{E_{O_m}\{E_{Z_m}[Var(n_{mt}|Z_m = z, O_m = j)]\}}_{\substack{\text{idiosyncratic}\\ \text{component}}}. \qquad (2.16)
\end{aligned}
$$

The first component in equation (2.16) is the variance that is accounted for by the difference in expected duration spells across individuals with different observed characteristics. The second term is the amount explained by differences in the average completed duration spells of type H or L workers. The last term is the remaining mean squared error, resulting from idiosyncratic differences across individuals that are not captured by either observed characteristics or unobserved types.

So far I have assumed that the exit probability of each group remains constant at q^z_{jt} for $t + 1$, $t + 2$,... However, the exit probabilities change both as a result of innovations in $x^z_{j,t+s}$ as well as genuine duration dependence. Generalizations of the three terms in equation (2.16) are easy enough to calculate by simulating the dynamic model given the conditioins holding as of any given date t in the sample.

Figure 22 displays the result of this decomposition for every month in the sample. If we knew a newly unemployed individual's unobserved type, we can reduce on average 40% of the mean squared error in forecasting his or her duration spells before leaving the unemployment status. By contrast, knowing an individual's observed characteristics only helps to reduce the MSE of forecasts by around 5% throughout the sample period. This result suggests that unobserved heterogeneity is crucial in accounting for the cross-sectional dispersion of completed duration spells, while observed heterogeneity plays a limited role in explaining differences in the duration of unemployment across different individuals. After the Great Recession, the contribution of unobserved heterogeneity rose to 60% and stayed elevated until the end 2013.

I also repeated the above calculations using alternative characterizations of observable characteristics, replacing the $j = 1, ..., 15$ different age-gender-education categories with $j = 1, ..., 5$ categories based on reason for unemployment.[22] The reason given for unemployment is slightly more successful in predicting differences in unemployment duration across individuals. For example, reason for unemployment accounts for 5% of the MSE on average over 2010-2011, compared with only 2% for the 15 gender-age-education categories. I also constructed groups based on 6 different observed occupations or on 8 different industry groups. Both of these explain less of the variance than does reason for unemployment. The overall conclusion from Figure 22 appears to be quite robust – unobserved types are much more important in explaining the differences across individuals in unemployment duration than are any differences in observable characteristics.

2.4.3 Heterogeneity and the cross-section dispersion of ongoing unemployment spells

A second way to analyze the role of heterogeneity is to consider for the current pool of unemployed individuals how long each individual reports they have been looking for a job so far. Let d_{mt} be the number of months that an individual m has been unemployed at time t. Suppose that an econometrician makes use of an individual m's observed characteristic O_m. Conditional on knowing $O_m = j$, the econometrician would know the average duration of unemployment of those unemployed who share the same observable characteristic j, $E(d_{mt}|O_m = j)$, and their variance of unemployment duration, $Var(d_{mt}|O_m = j)$. Then, the aggregate variance of unemployment duration is decomposed as follows:

$$Var(d_{mt}) = \underbrace{E[Var(d_{mt}|O_m = j)]}_{\substack{\text{heterogeneity} \\ within \\ \text{observed characteristics.}}} + \underbrace{Var[E(d_{mt}|O_m = j)]}_{\substack{\text{heterogeneity} \\ between \\ \text{observed characteristics}}}. \tag{2.17}$$

[22]Reasons for unemployment are represented by $j = 1$ for temporary layoff, $j = 2$ for permanent separation, $j = 3$ for job leavers, $j = 4$ for re-entrants to the labor force, and $j = 5$ for new entrants.

The first term in the right hand side of equation (2.17) captures the component of variance that can be explained by heterogeneity within observed characteristics and the second term is the variance explained by heterogeneity between observed characteristics.

Next suppose that the econometrician further has information on the individual's unobserved type z represented by $Q_m = z$ for $z = H$ or L. Conditional on knowing that $O_m = j$ and $Q_m = z$, the econometrician would know the average duration of unemployment of those unemployed whose observable characteristic is j and unobserved type is z, $E(d_{mt}|O_m = j, Q_m = z)$, and their variance of unemployment duration, $Var(d_{mt}|O_m = j, Q_m = z)$. The variance of the number of months that individuals with observable characteristic j have been unemployed can then be further decomposed into

$$Var(d_{mt}|O_m = j) = \underbrace{E[Var(d_{mt}|Z_m = z, O_m = j)]}_{\substack{\text{heterogeneity} \\ \textit{within} \\ \text{unobserved types}}} + \underbrace{Var[E(d_{mt}|Z_m = z, O_m = j)]}_{\substack{\text{heterogeneity} \\ \textit{between} \\ \text{unobserved types}}}. \tag{2.18}$$

In other words, the variance of individual unemployment duration of group j is decomposed into the average variance explained by heterogeneity within unobserved types and the variance accounted for by heterogeneity between unobserved types.

By plugging equation (2.18) into equation (2.17), we have the full decomposition result for the mean squared error of individual duration spells as follows.

$$\begin{aligned} Var(d_{mt}) &= \underbrace{Var[E(d_{mt}|O_m = j)]}_{\substack{\text{heterogeneity} \\ \textit{between} \\ \text{observed characteristics}}} + \underbrace{E_{O_m}\{Var[E_{Z_m}(d_{mt}|Z_m = z, O_m = j)]\}}_{\substack{\text{heterogeneity} \\ \textit{between} \\ \text{unobserved types}}} \\ &\quad + \underbrace{E_{O_m}\{E_{Z_m}[Var(d_{mt}|Z_m = z, O_m = j)]\}}_{\substack{\text{idiosyncratic} \\ \text{component}}}. \end{aligned} \tag{2.19}$$

The first component in equation (2.19) is the variance that is accounted for by the difference in the average duration across individuals with different observable characteristics. The second term is the amount explained by differences in the average duration of workers of type H or L. The last term is the remaining dispersion, resulting from idiosyncratic differences across individuals that are not captured by either observed characteristics or unobserved types.

Figure 23 displays the result of this decomposition for every month in the sample. Panel A shows that the aggregate variance of ongoing unemployment duration rises during recessions and continues to go up after the end of recession. The variance of unemployment duration accounted for by unobserved types exhibits the similar pattern. It increased substantially after the Great Recession was over and reached the highest level of the sample period. Panel B shows the share of variance accounted for by unobserved heterogeneity and observed heterogeneity. Unobserved heterogeneity accounts for around 30% of the variance of unemployment duration in the aggregate, while observed heterogeneity plays a limited role explaining less than 10% of the dispersion thoughout the 1980-2013 period. This suggests that unobserved heterogeneity is important in the cross-sectional dispersion of ongoing duration spells in the economy and is consistent with the conclusion of previous studies that it is hard to identify the observed individual characteristics which make some individuals stay unemployed longer than others.

2.4.4 Heterogeneity and time-series variation in the average duration of ongoing unemployment spells

Whereas the previous subsection explored factors influencing the cross-section dispersion of the length of time that unemployed individuals had spent up to month t in looking for a job, in this subsection I look at changes over time in the average length of time that currently unemployed individuals have spent searching for a job. There are two factors that drive the rise in the average duration of unemployment during a recession. The first factor is the compositional shift towards groups with traditionally longer duration of unemployment. The second factor is the rise in individual duration of unemployment. I analyze how much each factor contributed to the rise in observed duration of unemployment during the Great Recession by taking into account both observed and unobserved heterogeneity.

To quantify how much compositional shifts could have explained the rise in the average duration of unemployment, I hold the unemployment duration of each group and each unobserved type fixed at the value of December 2007, and investigate how much the compositional changes of individuals with observable characteristics and unobserved types can explain the overall rise in the average duration of unemployment. By first allowing the share of observed groups to vary over time with the share of unobserved types fixed and then further allowing the share of both observed groups and unobserved types to vary over time, we can separately analyze how much the compositional change of observed groups and unobserved types contribute to the rise in average duration of unemployment.

The average duration of unemployment is written as follows

$$D_t^{full} = \sum_{j=1}^{J} F_{jt}(f_{jt}^H D_{jt}^H + f_{jt}^L D_{jt}^L).$$

where F_{jt} is the fraction of group j in the unemployment, f_{jt}^z and D_{jt}^z are the fraction and the unemployment duration of type z workers in group j at time t for $z = H, L$.

Consider first the case in which D_{jt}^z and f_{jt}^z are fixed and only F_{jt} varies over time. Let $\bar{D}_j^z$ and $\bar{f}_j^z$ be the value of unemployment duration and the fraction of type z workers in group j at time t = 2007:M12 for $z = H, L$. Then the change in the average duration of unemployment due to changes in the composition of groups based on observed characteristics, $D_t^{com\_o}$, is

$$D_t^{com\_o} = \sum_{j=1}^{J} F_{jt}(\bar{f}_j^H \bar{D}_j^H + \bar{f}_j^L \bar{D}_j^L).$$

This series is plotted as the green line with x's in Figure 24. The compositional changes of groups with observed characteristics play no role in the rise of average unemployment duration.

Consider second the case in which only D^z_{jt} is fixed at the value of $t = 2007$:M12 and the fraction of unemployed individuals with observed characteristic j and unobserved type H and L in group j varies over time. Then the average duration of unemployment driven by changes in the composition of both unobserved types and observed groups, $D_t^{com\_ou}$, is

$$D_t^{com\_ou} = \sum_{j=1}^{J} F_{jt}(f_{jt}^H \bar{D}_j^H + f_{jt}^L \bar{D}_j^L).$$

This series is plotted as the dashed red line in Figure 24. The compositional change of unobserved heterogeneity drives most of the rise in unemployment duration during the Great Recession. It continues to play an important role in raising the average duration of unemployment after the end of the recession accounting for the half of the rise in unemployment duration between December 2007 and March 2010. Unlike the conclusion of previous studies[23] which find that compositional changes explain very little of the development of the average duration of unemployment during the Great Recession, I find that compositional shift is an important factor once we consider unobserved heterogeneity.

Lastly, I consider the case in which only the unemployment duration of each type in group j is allowed to vary over time, while the fractions are fixed at the value of time $t = 2007$:M12. Then the average duration of unemployment driven by changes in the duration of each type in group j is

$$D_t^{dur} = \sum_{j=1}^{J} \bar{F}_j(\bar{f}_j^H D_{jt}^H + \bar{f}_j^L D_{jt}^L).$$

Changes in the duration of each type drive the rise in average duration of unemployment in the later part of Great Recession and after the recession was over. Along with the compositional change, the rise in unemployment duration of unobserved types also

[23] See for example Kroft, Lange, Notowidigdo and Katz (2013).

played an important role in the path of unemployment duration after the recent recession.

2.5 Analysis based on alternative observable characteristics

In this section I make use not just of the baseline breakdown of unemployed workers by gender, age, and education but also a number of other observable characteristics to try to get a richer description of some of the observed attributes that type L workers have in common.

I first ask the question, given a newly unemployed type L individual, what observed characteristics is the person most likely to exhibit? Tables 2.5-2.7 provide summary statistics for the baseline categories of gender, age and education. The key result is summarized in the first row of Table 2.15 - education is the single most important attribute distinguishing newly unemployed type L workers from a typical member of the labor force. For an average month t in the sample, more than half of the newly unemployed type L individuals have a high-school education or less, whereas these individuals make up only 40% of the total labor force. However, this group also accounts for 2/3 of the total number of newly unemployed (see column 2 of Table 2.15). In other words, although type L workers are likely to be less educated than a typical member of the workforce, this can be more than accounted for by the fact that less educated individuals are more likely to become unemployed in the first place.

I next repeated the analysis breaking individuals down instead in terms of one of 8 different industries in which they had previously been employed. Parameter estimates for this data set are reported in Table 2.8-2.9 and summarized in Table 2.10 and the second row of Table 2.15. Thirty-one percent of newly unemployed type L workers are likely to have previously worked in construction or manufacturing, compared to only 17% of the total labor force. But we see in column 2 of Table 2.15 that this again simply reflects the fact that newly unemployed workers of any type are likely to have come disproportionately from this sector.

I also applied our model to 6 different occupation categories, with results reported in Tables 2.11 and 2.12 and the third row of Table 2.15. Again there are no striking differences in observed occupation between newly unemployed type L and type H individuals.

Finally I looked at 5 different reasons for unemployment, with parameter estimates in Table 2.13 and summaries in Table 2.15. Forty-two percent of newly unemployed type L workers are likely to have indicated "permanent separation" as their reason for becoming unemployed, compared with only 25% of newly unemployed type H workers.

To summarize, a newly unemployed type L individual is more likely than a typical member of the labor force to be poorly educated, have previously worked in construction or manufacturing, and to have lost their job as a result of involuntary permanent separation. However, only the last factor appears to be an important observable characteristic that helps distinguish type L individuals from other newly unemployed individuals.

I present these data graphically in Figure 2.26. The contribution of workers who were permanently separated from their previous job to the increase in inflows of type L workers during a typical recession is particularly striking (see Panel E and F of Figure 26). Cyclically related inflows from the construction-manufacturing and wholesale-retail (second left panel) are also dramatic but not unique to the type L individuals.

The importance of permanently separated workers in the type L inflows illustrates a key reason for thinking about unobserved heterogeneity as a dynamic process. Although it is common in the micro literature to think of unobserved heterogeneity as a fixed characteristic of a given worker, these results suggest that the difficulty an individual has in obtaining a job is very much tied to changing economic conditions, for example, skills that are no longer in demand.

No matter which individual characteristics and no matter how fine a gradation of individual characteristics available in the CPS micro data we consider, the same broad conclusion emerges as in our baseline case, and the finding that unobserved heterogeneity is key in accounting for the duration of unemployment is quite robust. The differences across individuals that cause some to spend much longer in unemployment than others are not perfectly captured by the observable characteristics of the unemployed but the evidence suggests that the nature most closely associated to type L individuals is whether he or she is permanently separated from previous employers.

2.6 Conclusion

Average duration of unemployment rises during recessions. If we analyze the unemployment duration by observable characteristics, it is commonly found that the unemployment duration rises during recessions for almost every any observable characteristic. However, even when we condition on observable characteristics within each group, there still exist people who have lower exit probabilities than others. Increased inflows of those who have intrinsically low exit probabilities and the deterioration of their exit probabilities during economic downturns are the major drivers of the rise in average duration of unemployment. I show that changes in unobserved heterogeneity among unemployed individuals are crucial in understanding the rise in average duration of unemployment during the Great Recession as well as the cross-sectional dispersion over the 1980-2013 period.

Given the empirical findings of this paper, one might be tempted to ask what ultimately constitutes the important cross-sectional heterogeneity that is commonly found within any group of individuals who all share the same observed characteristics. No matter how fine a gradation of individual characteristics we consider, there still exists substantial unexplained heterogeneity within each group.

I conclude the paper by briefly discussing some of the policy implications based on the empirical findings. The characteristics that make an individual more likely to remain unemployed for long periods are determined before an individual became unemployed. People with particular attributes or circumstances are likely to stay long in the unemployment regardless of the recovery of other parts in the economy once they become unemployed. This suggests expansionary policy measures might have a limited ability to reduce the unemployment duration, insofar as policies cannot change the intrinsic characteristics of unemployed individuals. It further implies that we need to exert effort in reducing the number of workers separating from firms as well as enhancing the re-employment prospect of job seekers.

2.7 Acknowledgements

Chapter 2 is in preparation for submission.

2.8 Figures and Tables

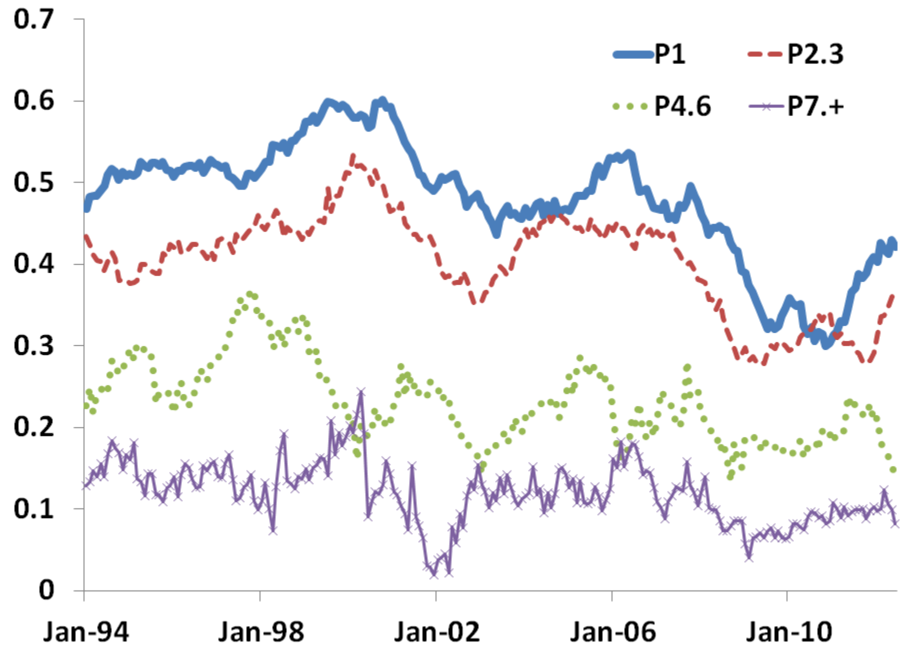

Figure 2.1: Exit probability from unemployment by duration of men with some college education, 1994-2012

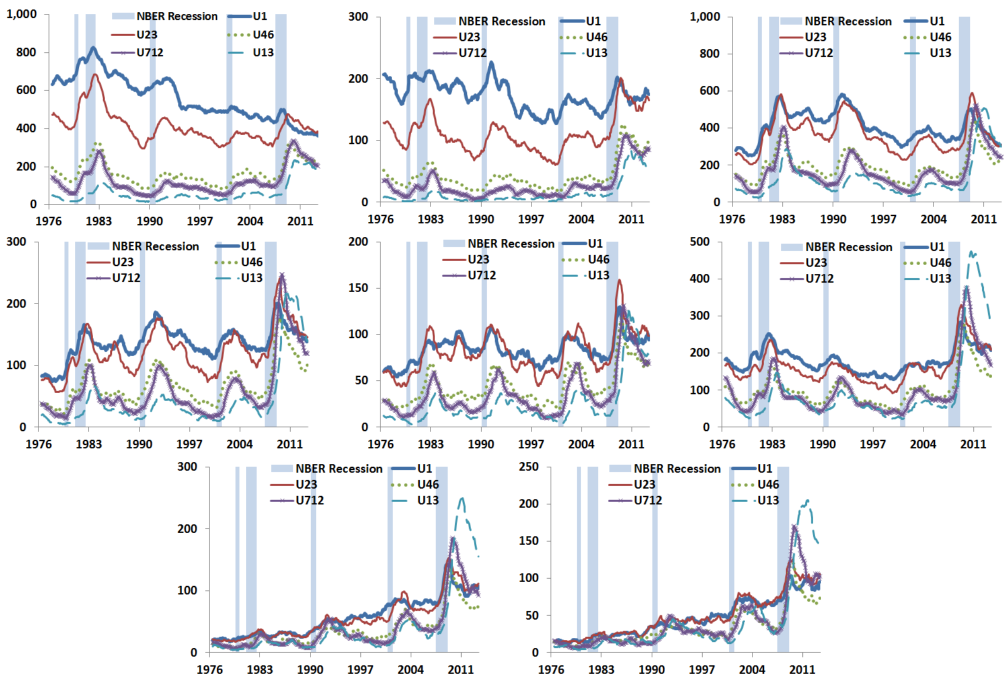

Figure 2.2: Numbers unemployed by duration (male)

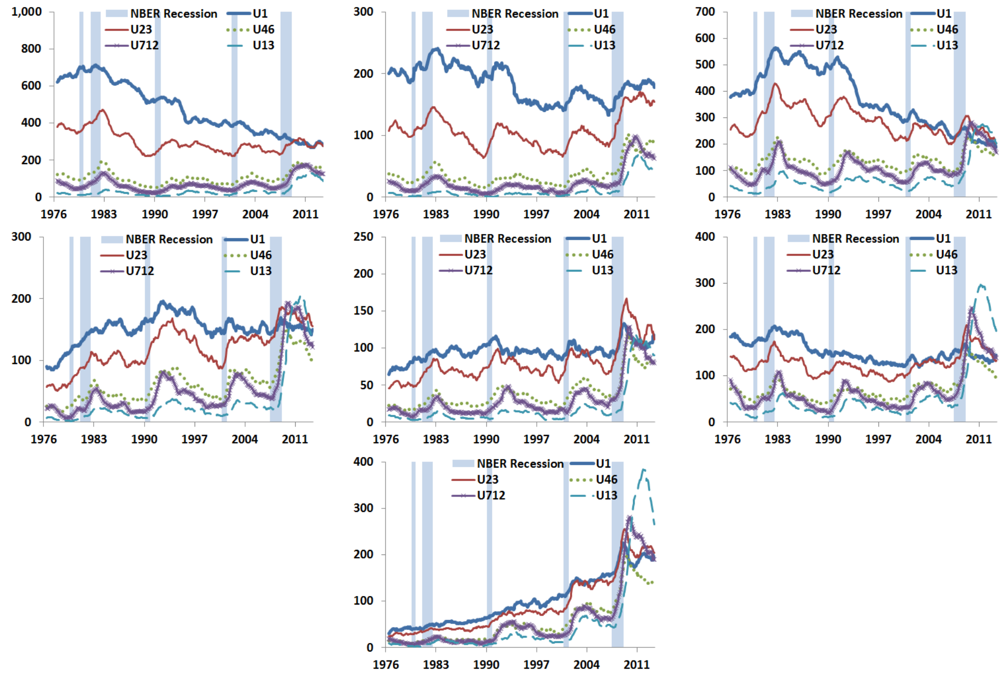

Figure 2.3: Numbers unemployed by duration (female)

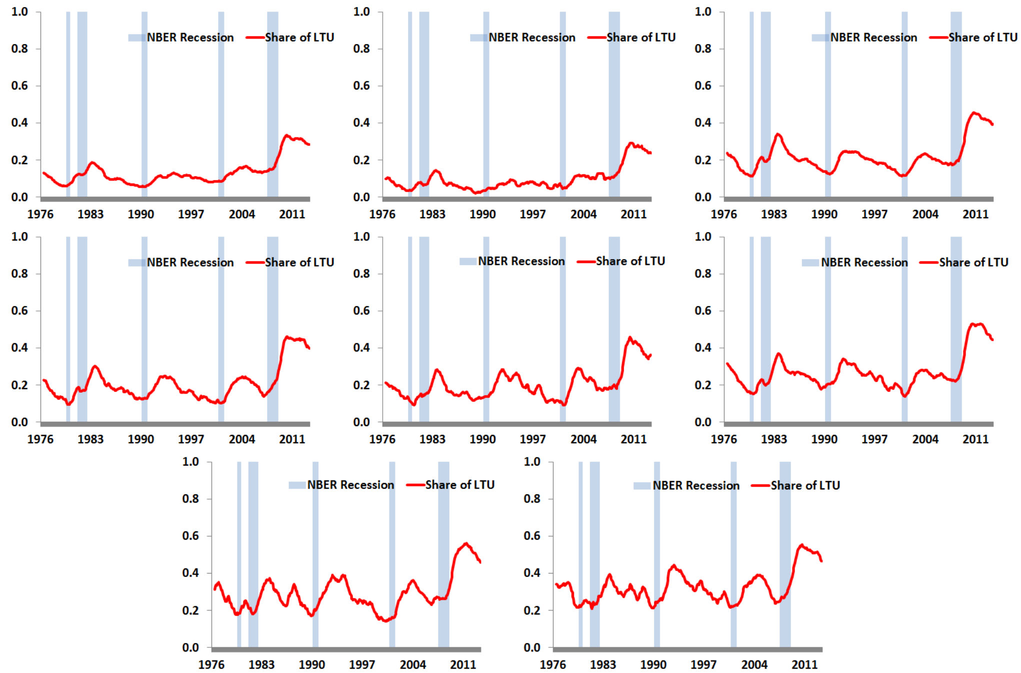

Figure 2.4: Share of long-term unemployment (male)

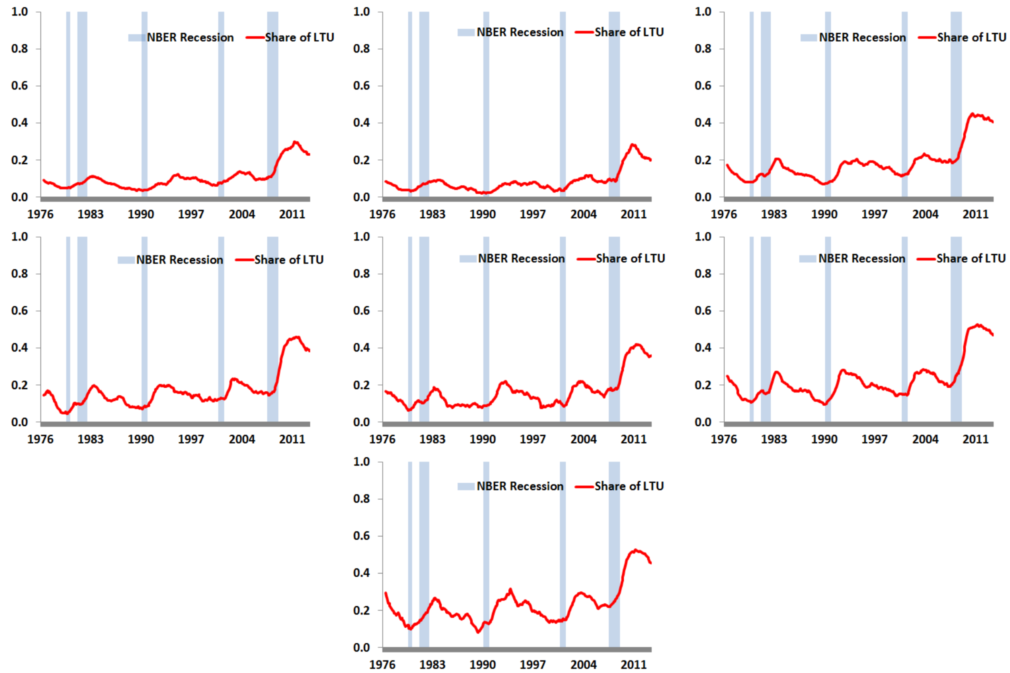

Figure 2.5: Share of long-term unemployment (female)

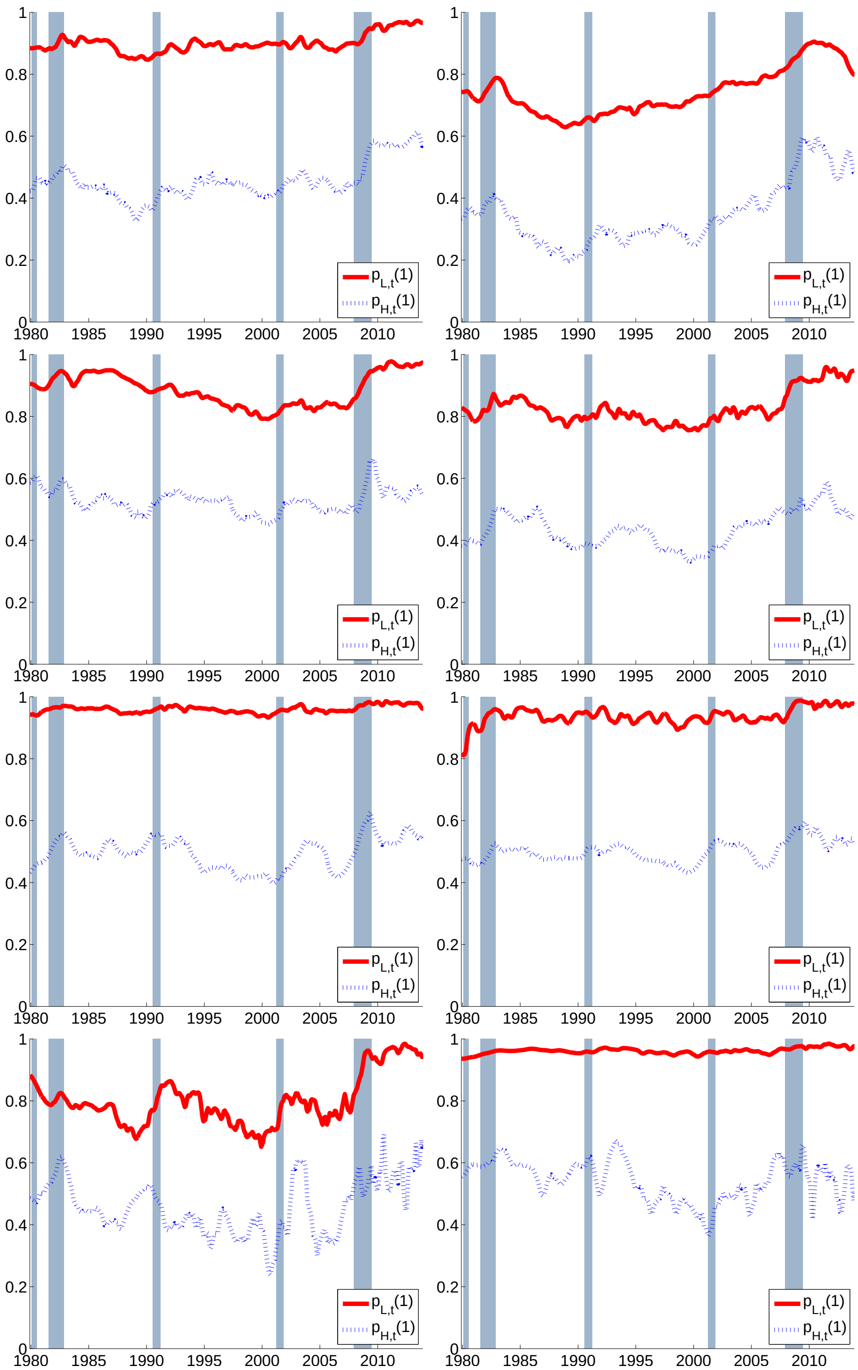

Figure 2.6: Probability that a newly unemployed worker of each type will still be unemployed the following month (male)

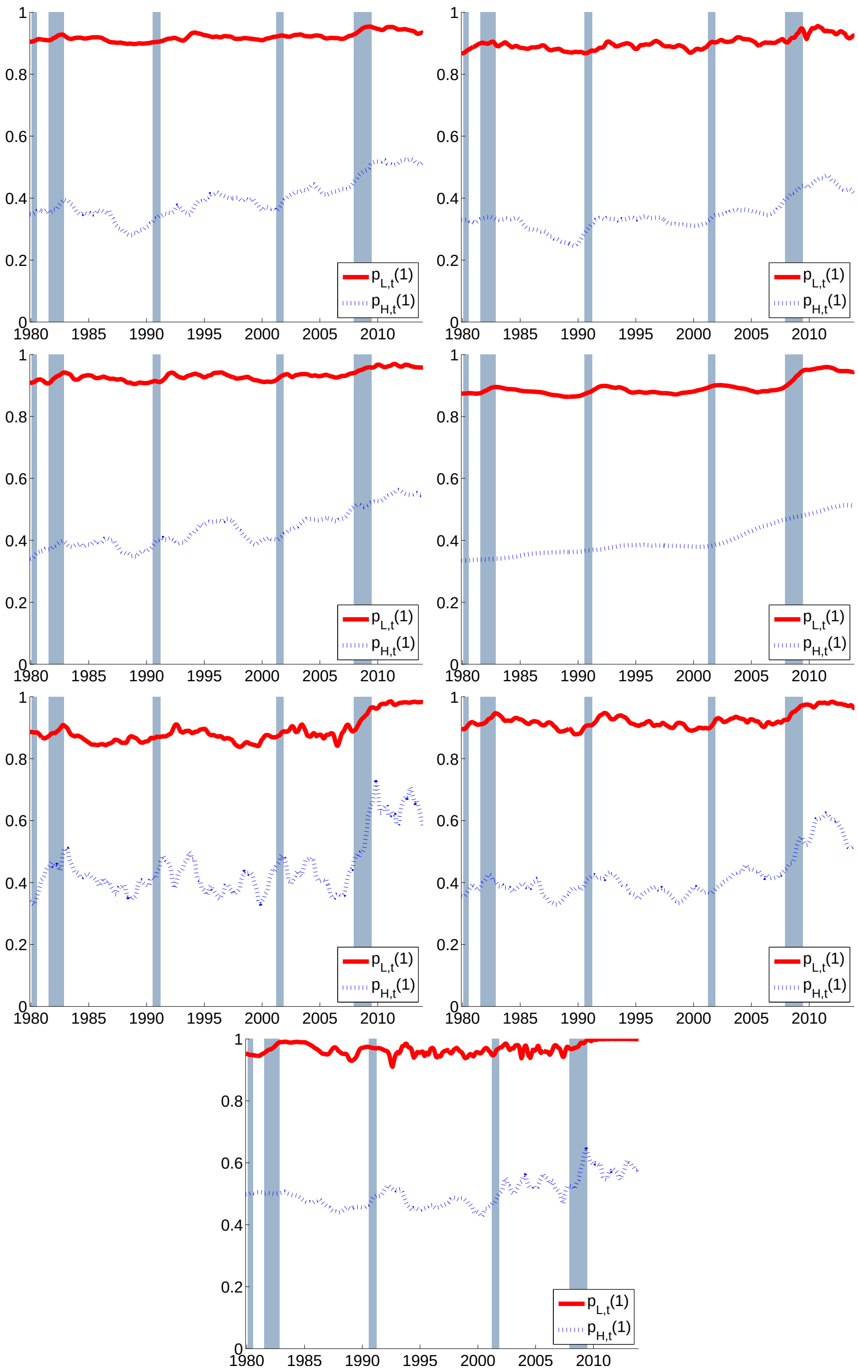

Figure 2.7: Probability that a newly unemployed worker of each type will still be unemployed the following month (female)

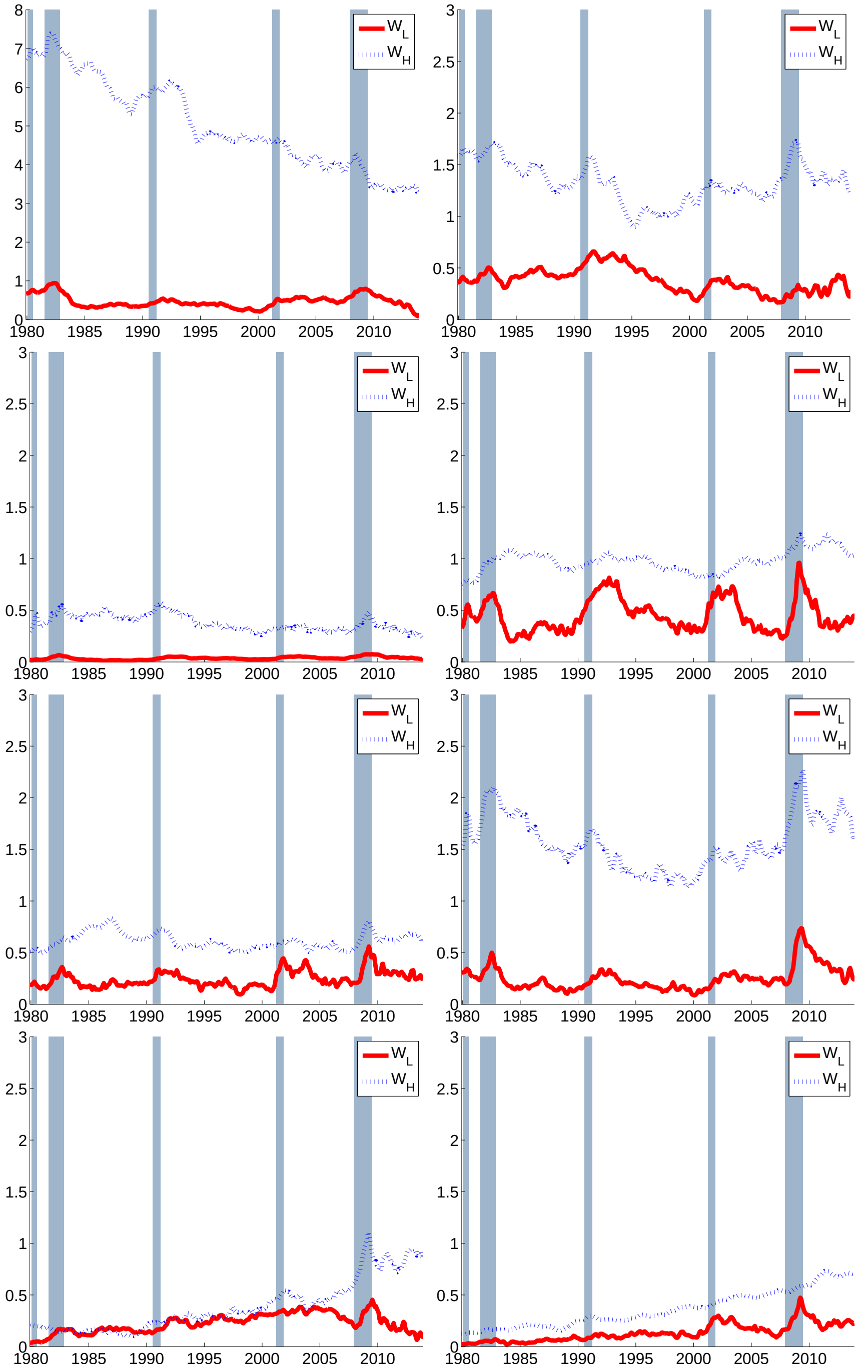

Figure 2.8: Number of newly unemployed workers of each type (male)

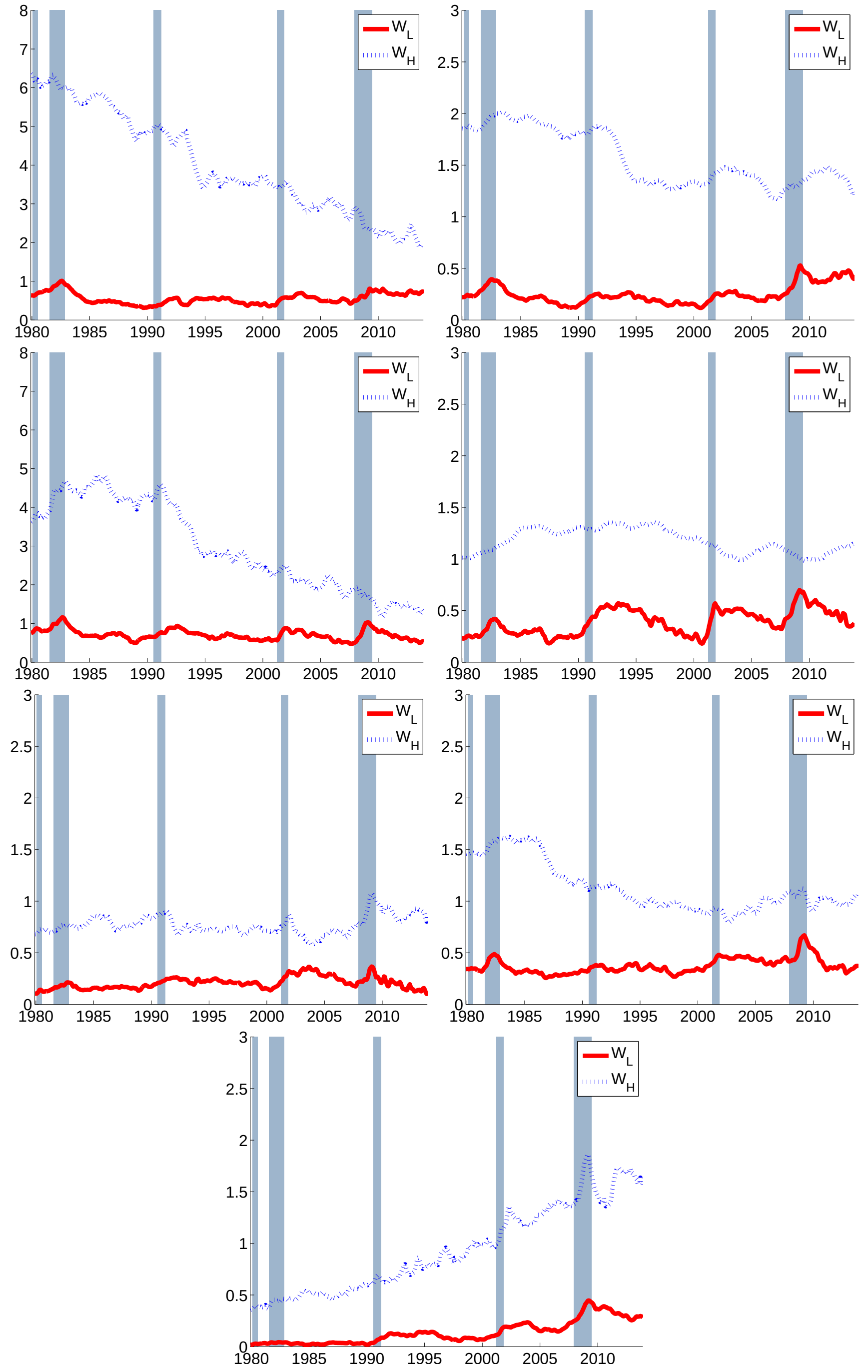

Figure 2.9: Number of newly unemployed workers of each type (female)

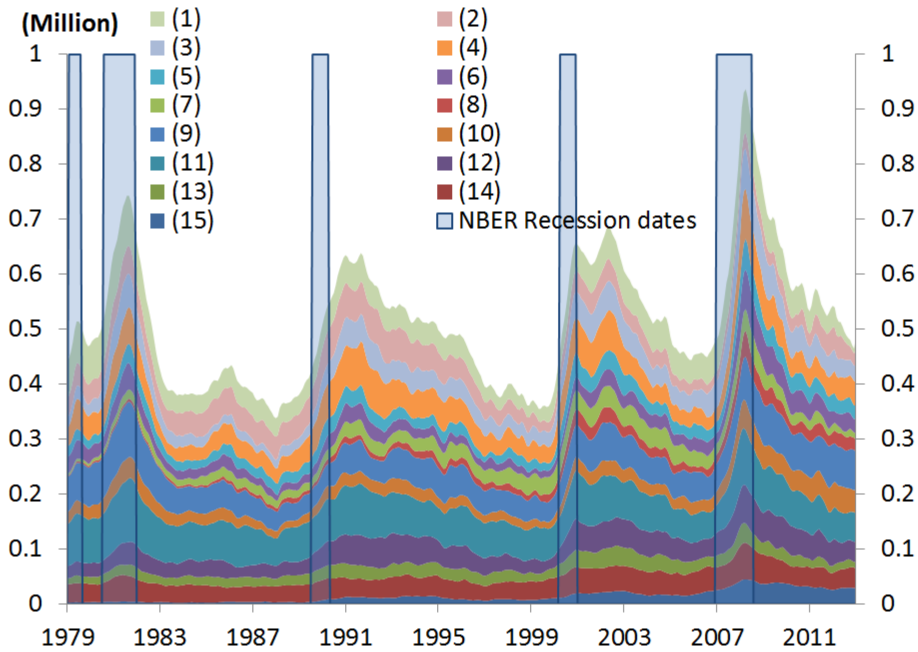

Figure 2.10: Composition of total type L inflows

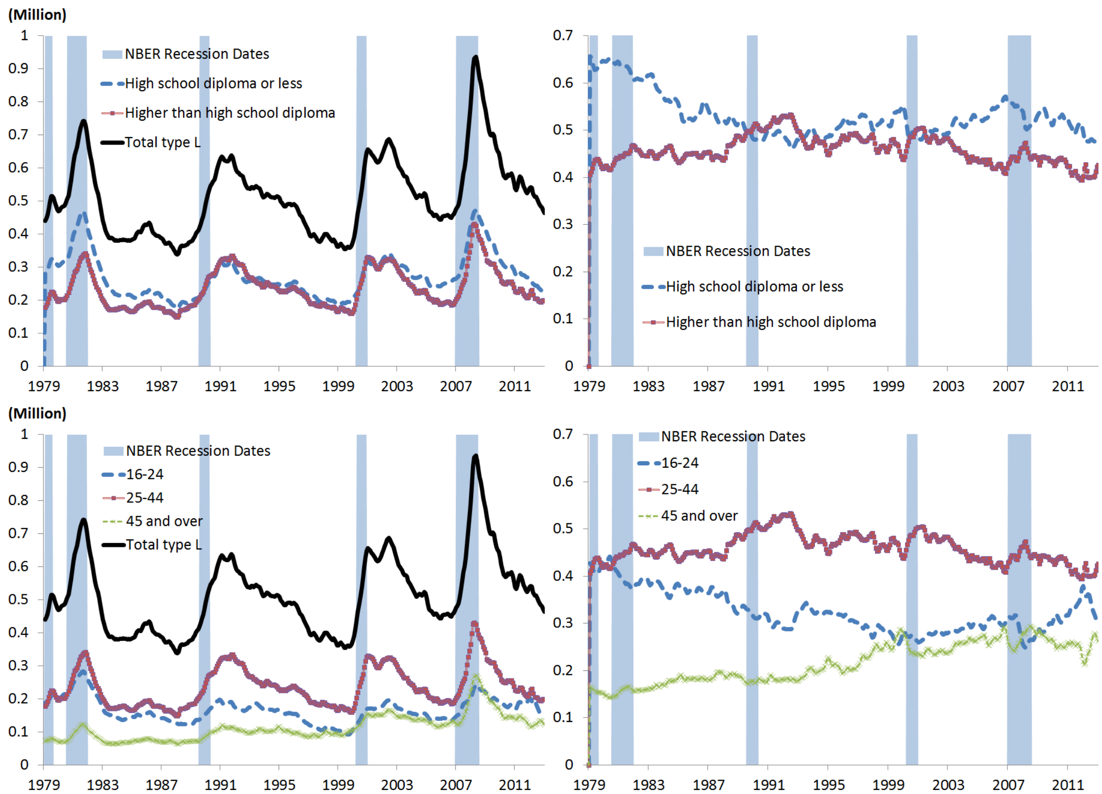

Figure 2.11: Size and share of type L individuals of each group by education and age

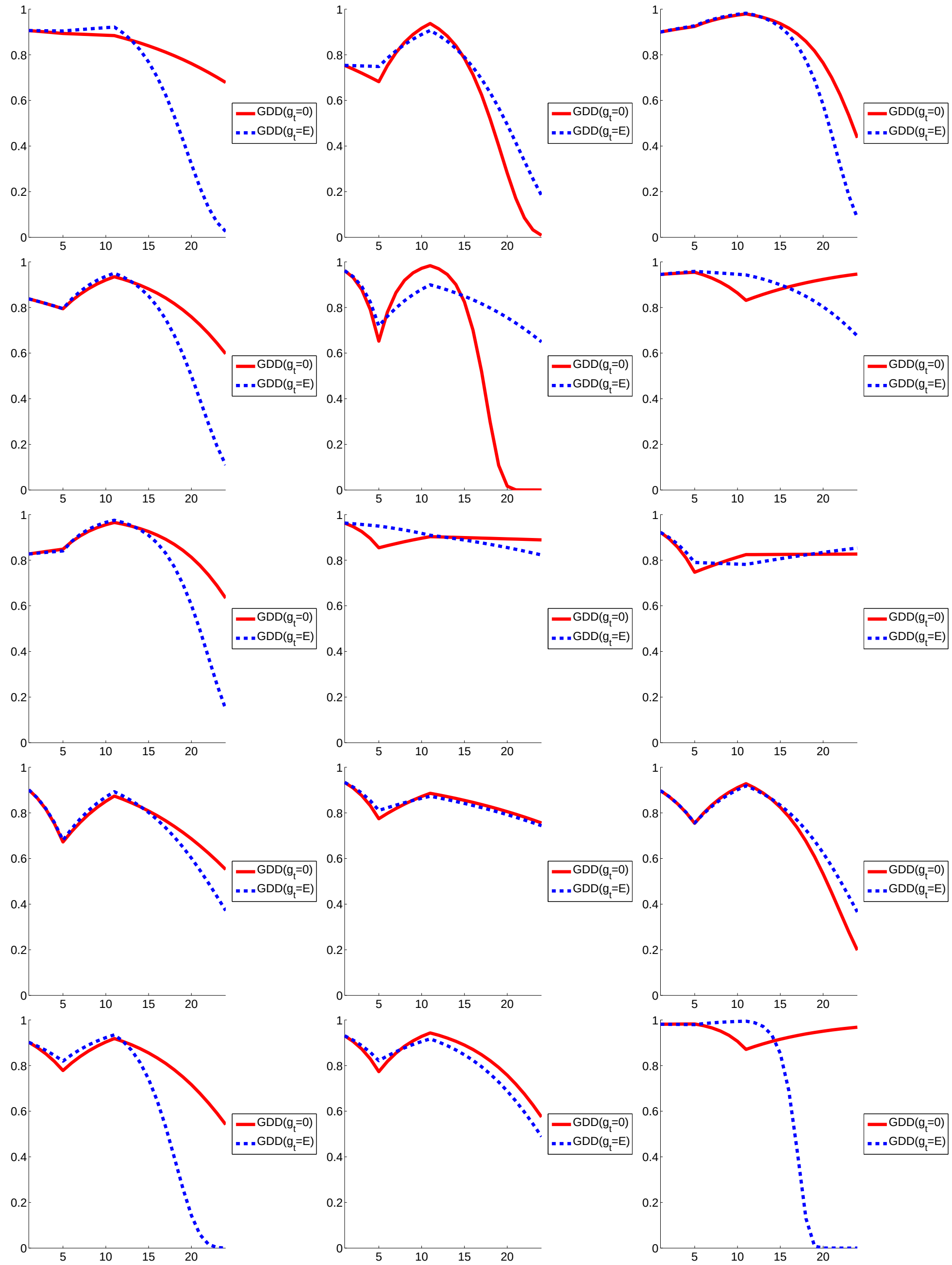

Figure 2.12: Estimates of genuine duration dependence

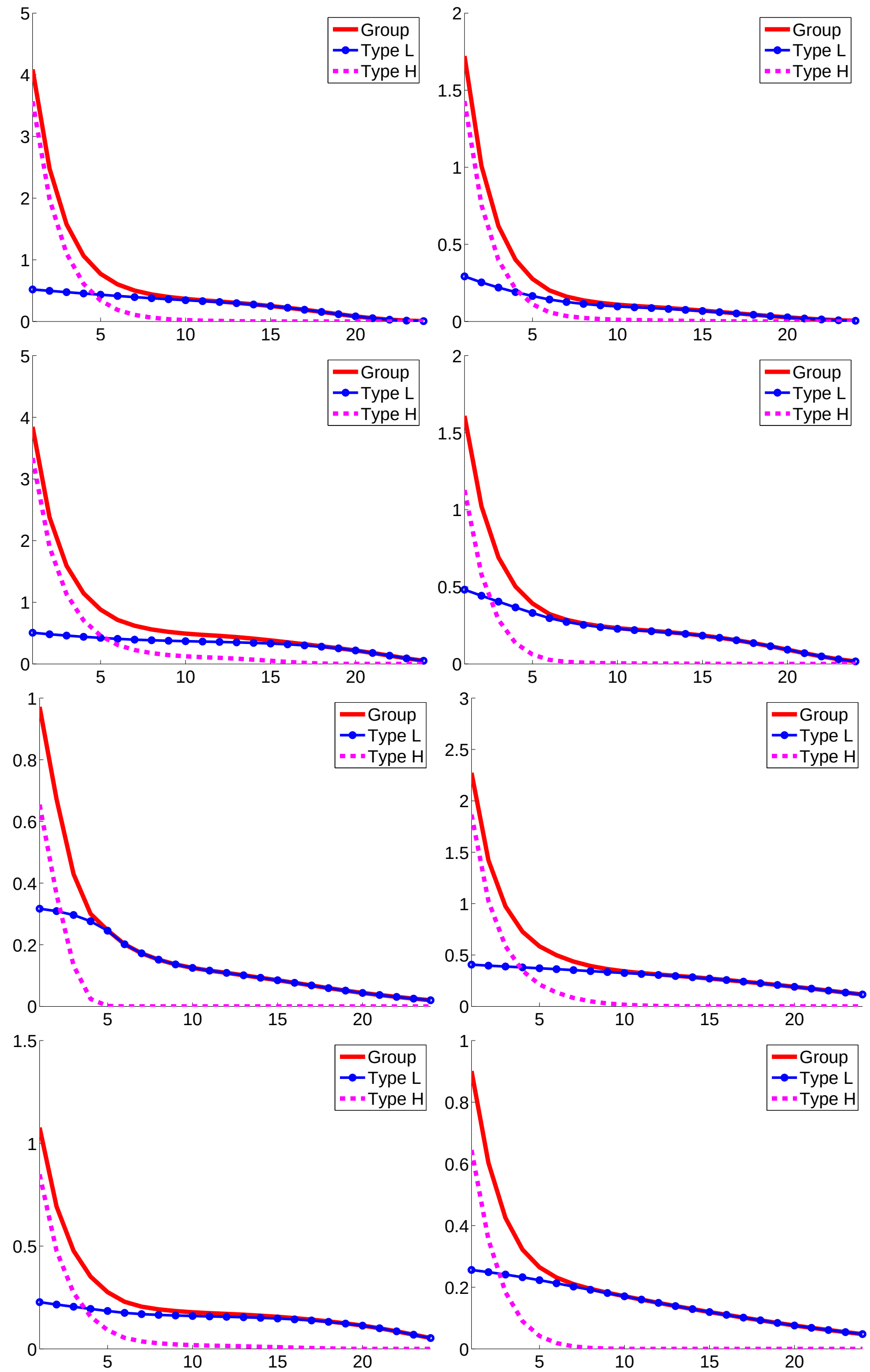

Figure 2.13: Distribution of unemployment duration by each type of worker (male)

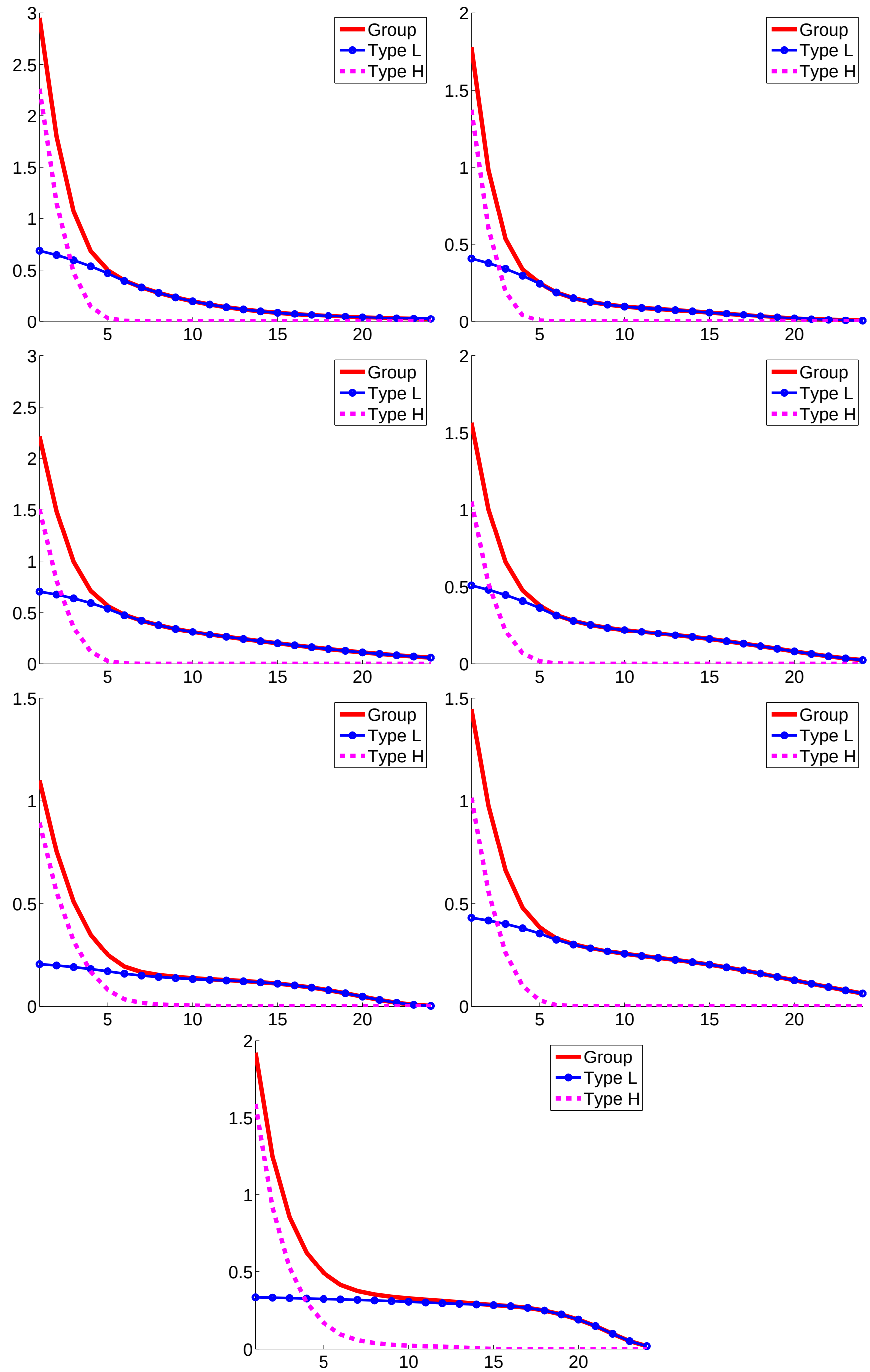

Figure 2.14: Distribution of unemployment duration by each type of worker (female)

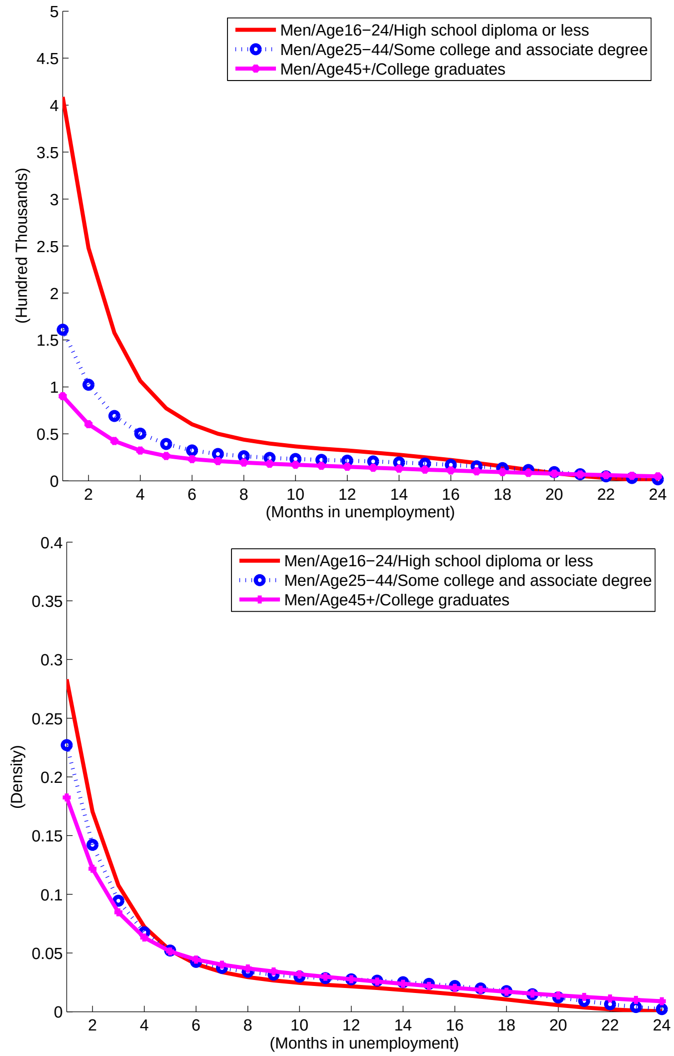

Figure 2.15: Distribution of unemployment duration of (1) men/age 16-24/high school graduates or less, (2) men/age 25-44/some college and associate degree and (3) men/age 45 and over/college graduates.

Average of December 2007-December 2013. X-axis: unemployment duration in months, Y-axis: (Panel A) Number of unemployed individuals in hundred thousands, (Panel B) density.

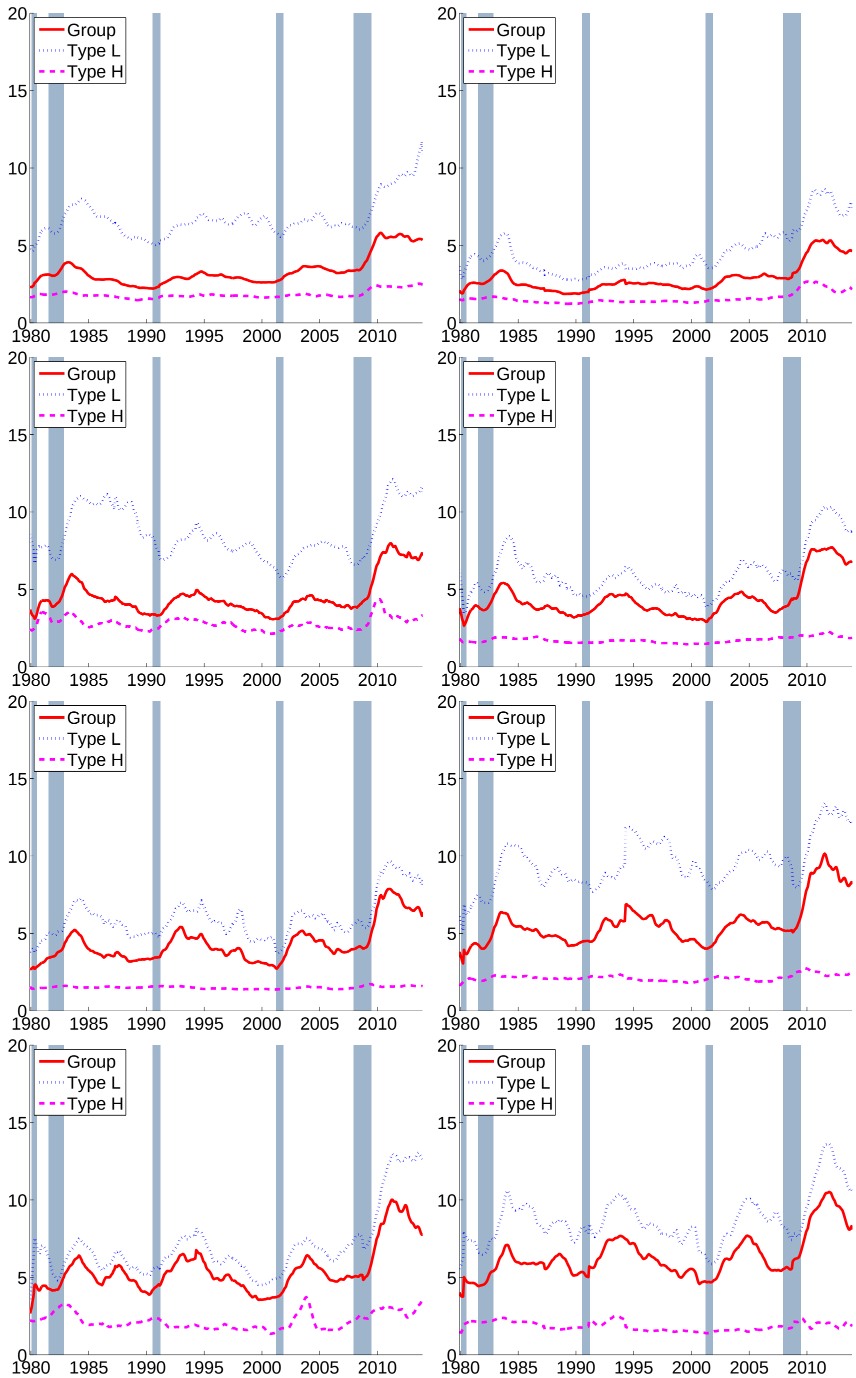

Figure 2.16: Average duration of unemployment by each type of worker in months (male)

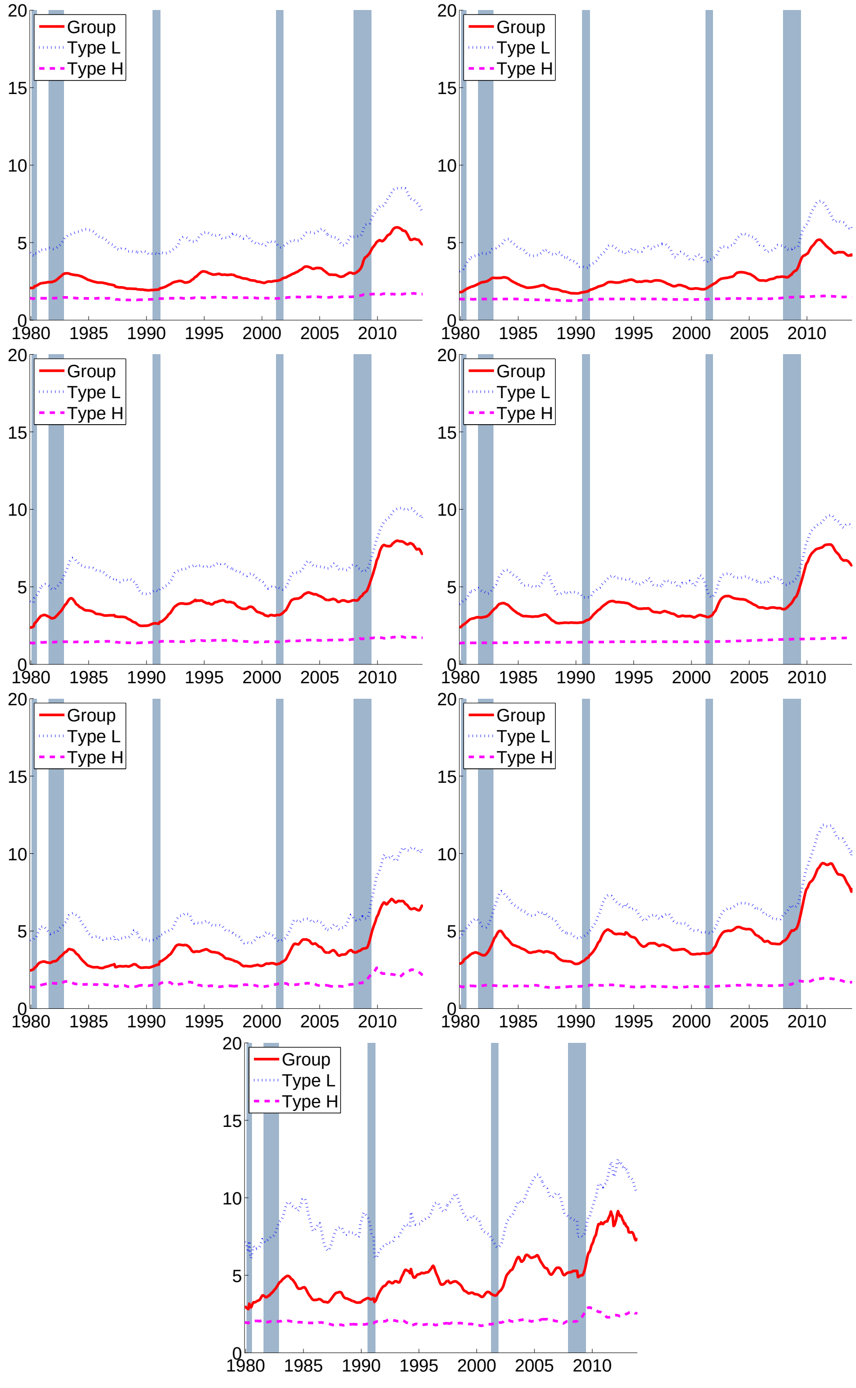

Figure 2.17: Average duration of unemployment by each type of worker in months (female)

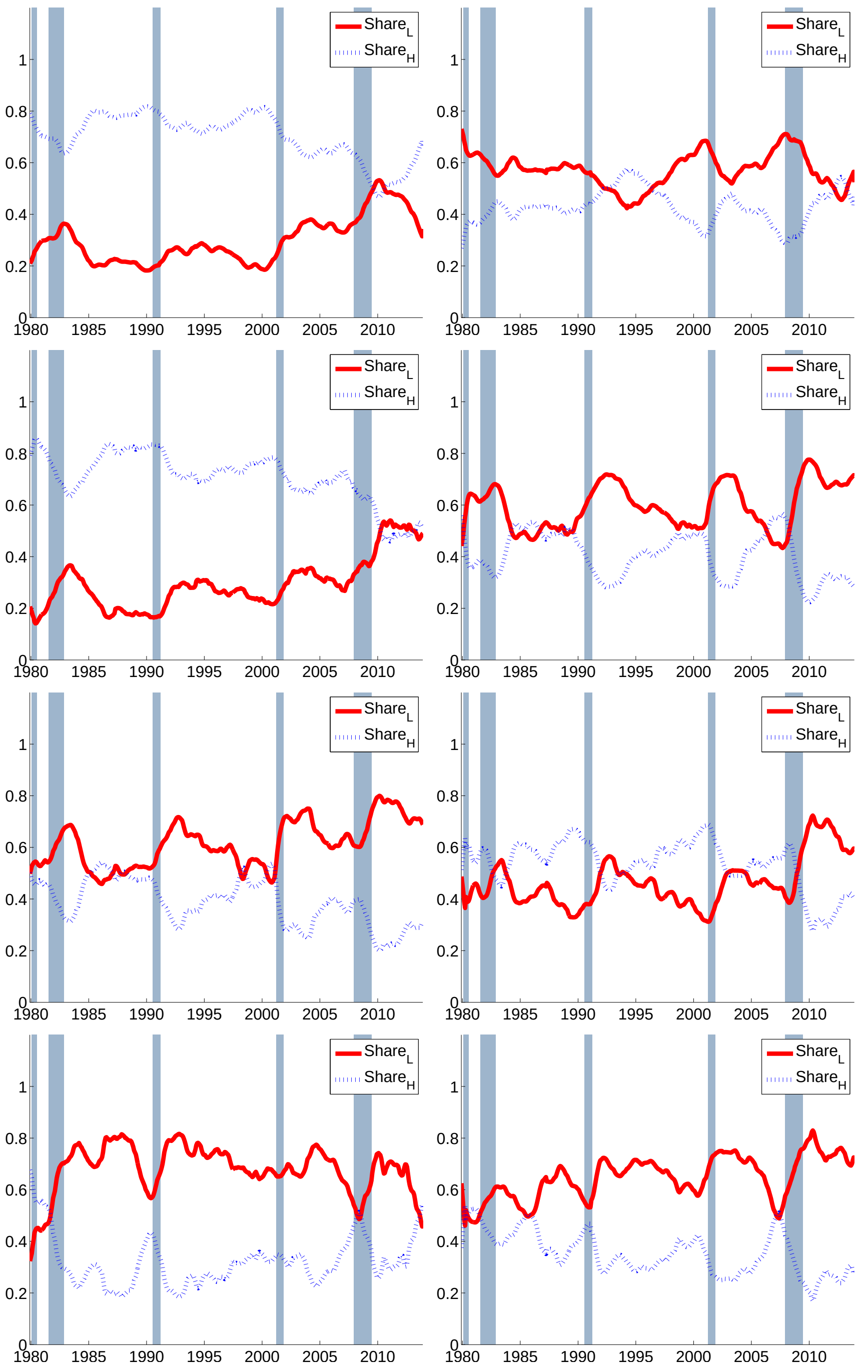

Figure 2.18: Share of total unemployment accounted for by each type of worker (male)

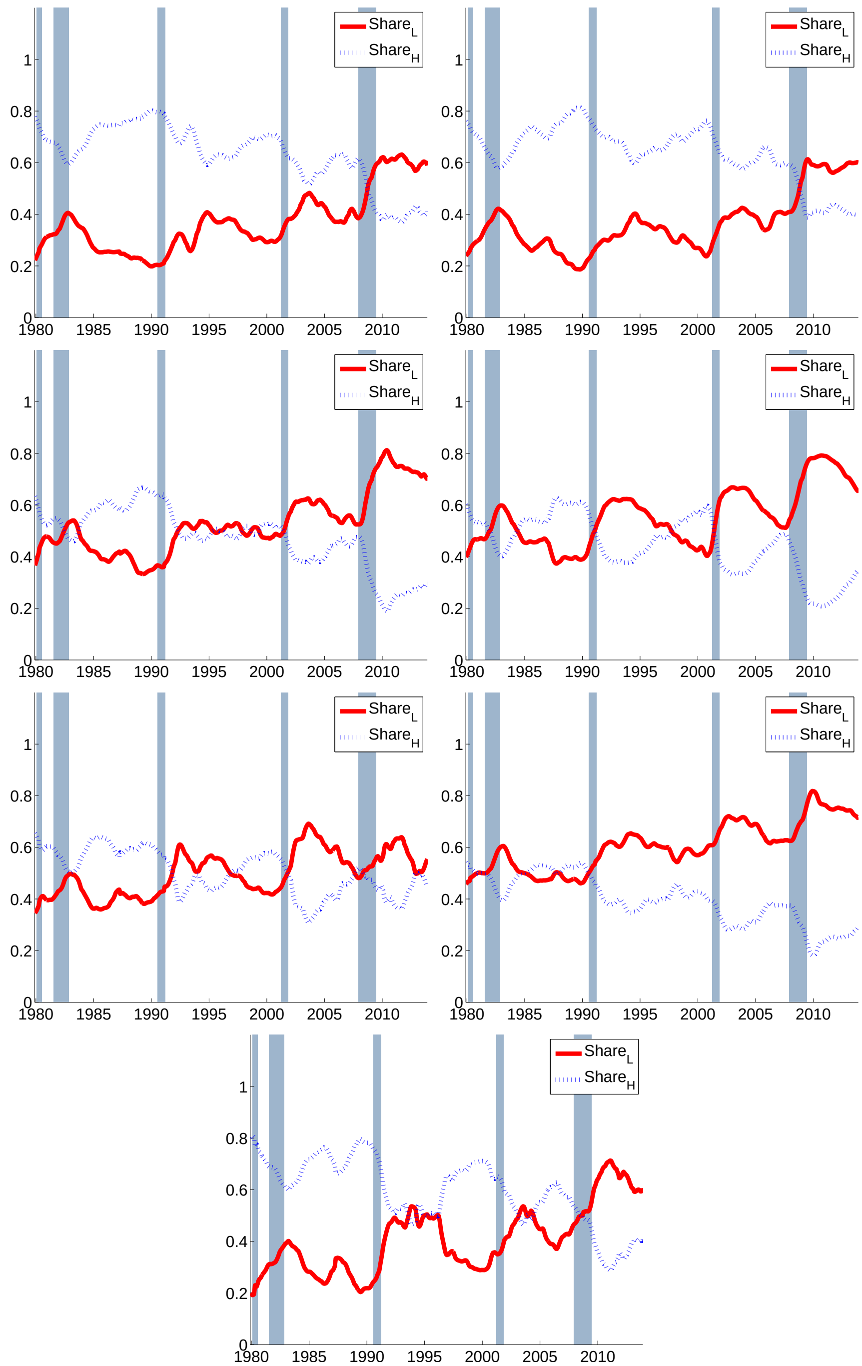

Figure 2.19: Share of total unemployment accounted for by each type of worker (female)

Figure 2.20: Data and forecasts for $U_{jt}^{1}, ..., U_{jt}^{13.+}$ of men/age 25-44/some college and associate degree during the Great Recession.
Horizontal axis: number of months ahead s for which the forecast is formed in September 2008

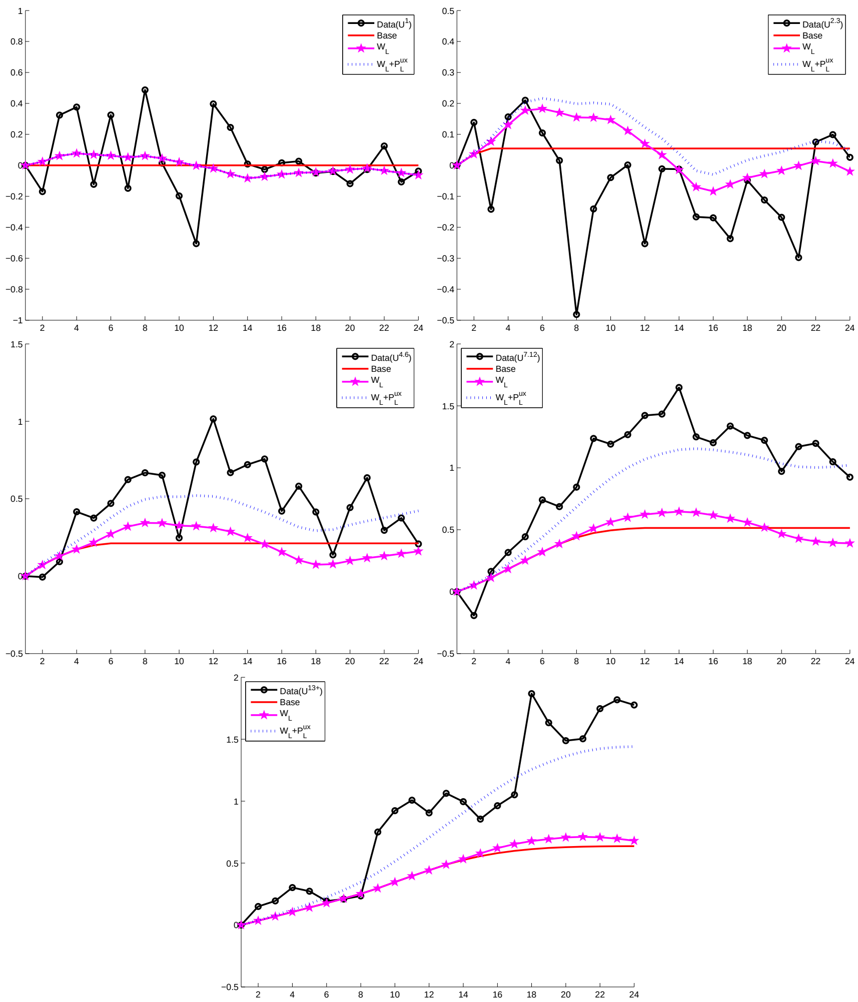

Figure 2.21: Data and forecasts for $U_{jt}^{1}, ..., U_{jt}^{13.+}$ of women/age 25-44/ some college and associate degree during the Great Recession.
Horizontal axis: number of months ahead s for which the forecast is formed in November 2008

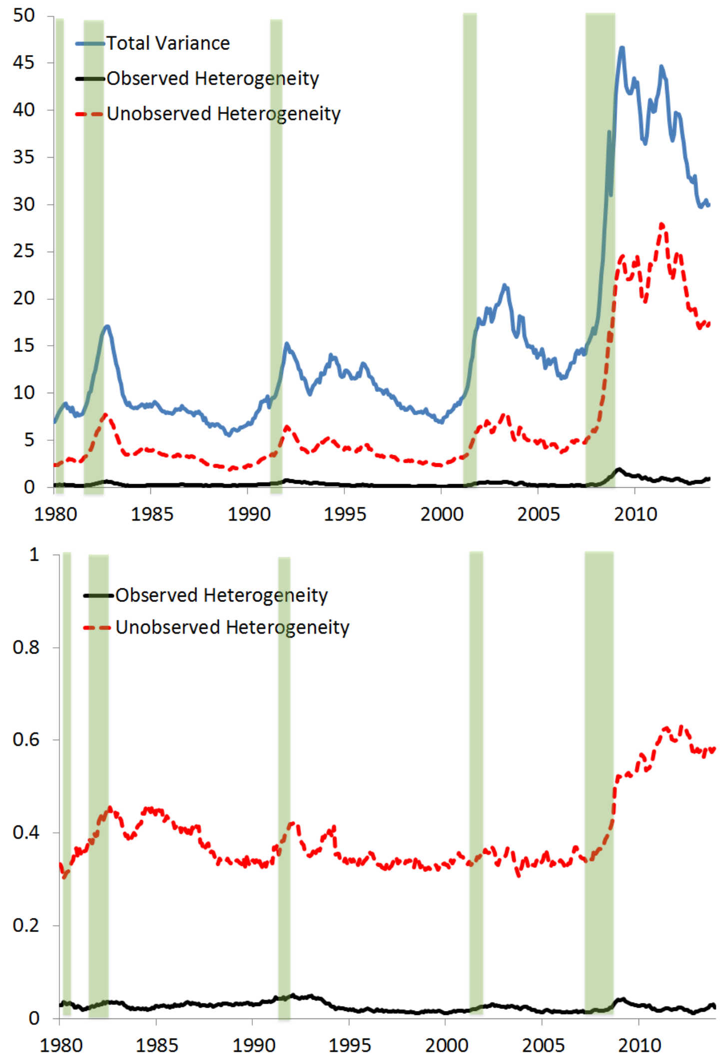

Figure 2.22: Amount of mean squared error in predicting the completed duration spells of unemployment across individuals accounted for by observed and unobserved heterogeneity

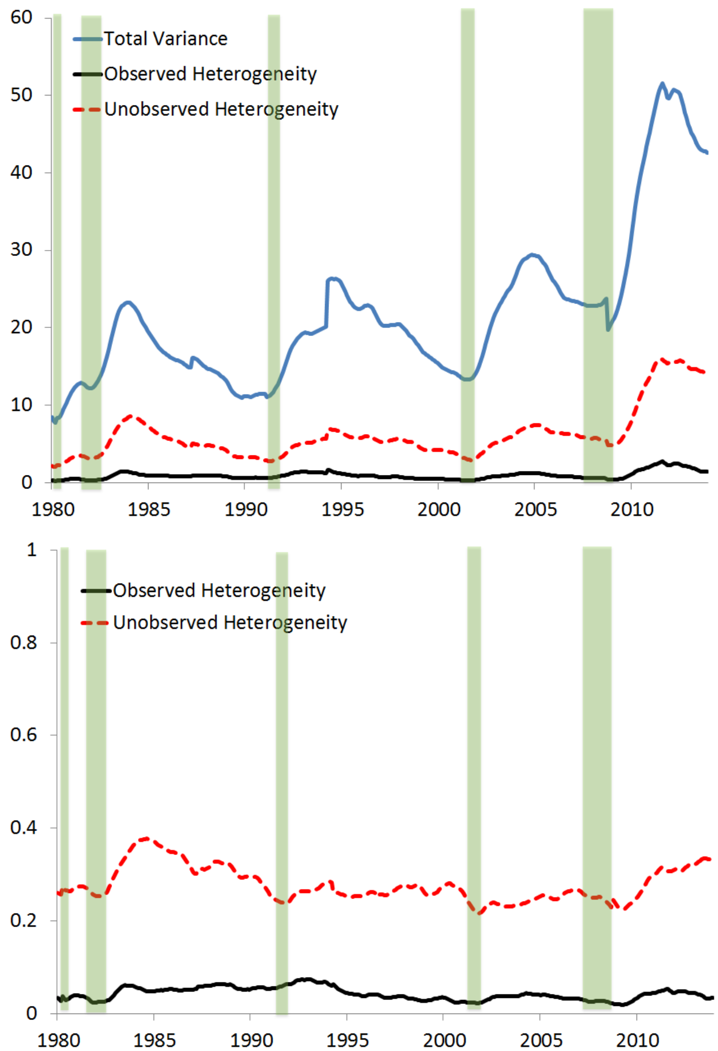

Figure 2.23: Amount of variance of unemployment duration in the aggregate accounted for by observed and unobserved heterogeneity

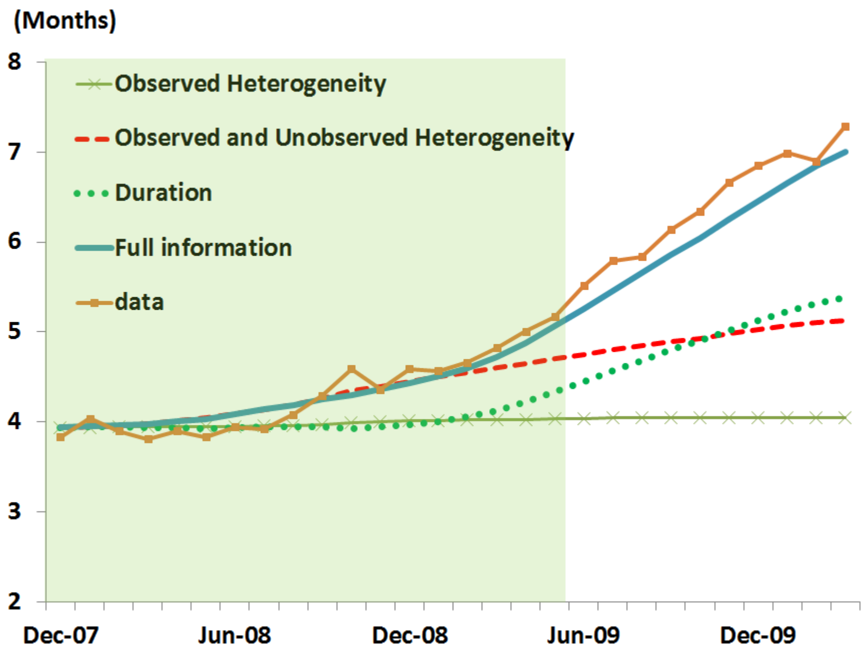

Figure 2.24: Compositional variation and average duration of unemployment during the Great Recession and its recovery phase (December 2007 - March 2010)

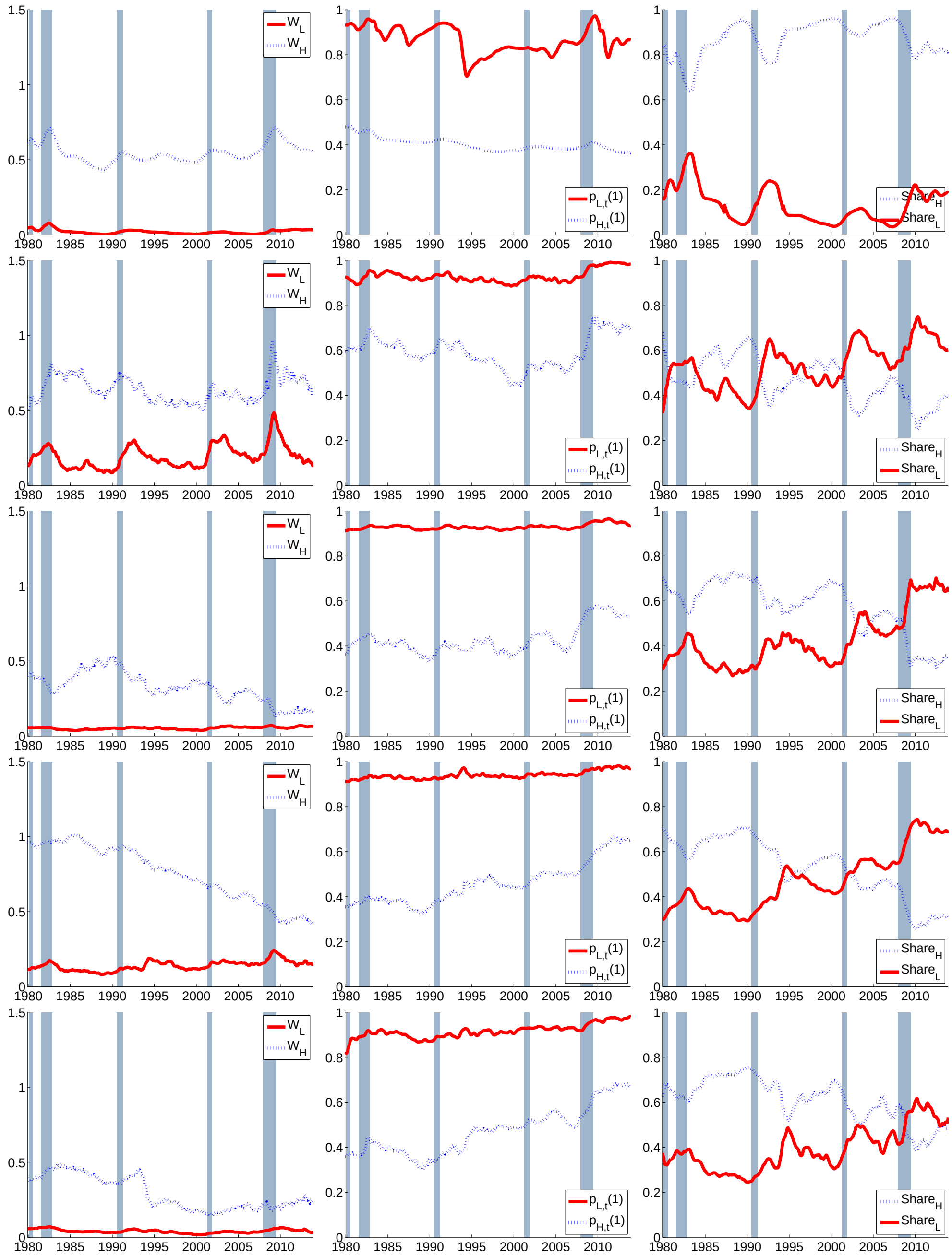

Figure 2.25: Inflows, continuation probabilities and share of type H and L in unemployment by reason for unemployment

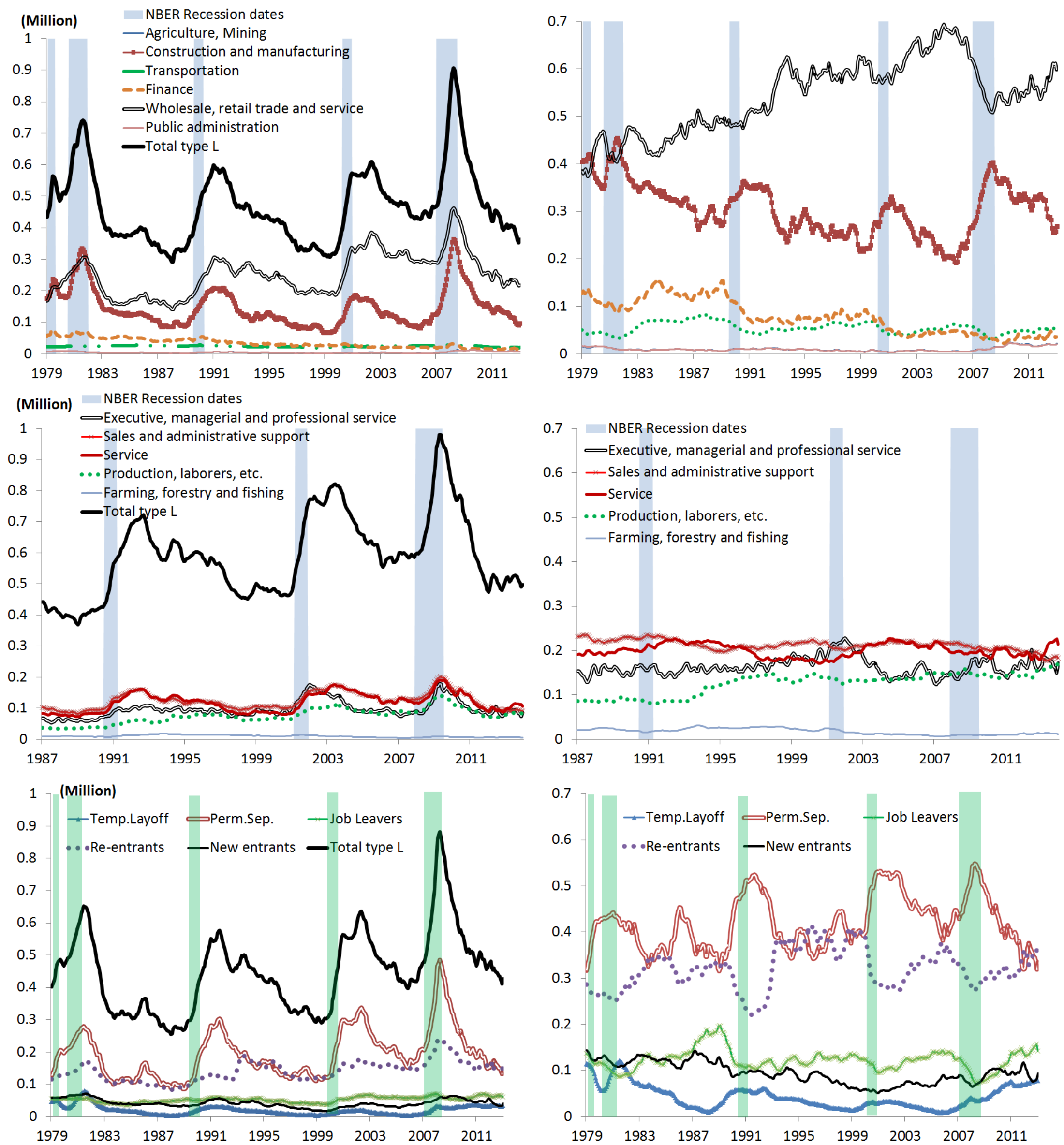

Figure 2.26: Size and share of type L individuals of each group by reason for unemployment, industry and occupation

Table 2.1: Parameter estimates for the baseline model (Male)

White (1982) quasi-maximum-likelihood standard errors in parentheses. (1) Age 16-24/High school graduates and less than high school, (2) Age 16-24/ Some college, associate degree and college graduates, (3) Age 25-44/ High school graduates and less than high school, (4) Age 25-44/ Some college and associate degree.

	(1)	(2)	(3)	(4)
σ^w_{jL}	0.0036**	0.0342	0.0037	0.0533***
	(0.0015)	(0.0210)	(0.0030)	(0.0088)
σ^w_{jH}	0.0106***	0.0540**	0.0222***	0.0339
	(0.0019)	(0.0238)	(0.0028)	(0.0210)
σ^x_{jL}	0.0980*	0.0494**	0.1000***	0.0997
	(0.0525)	(0.0224)	(0.0251)	(0.0974)
σ^x_{jH}	0.0337***	0.0537**	0.0388***	0.0473
	(0.0052)	(0.0262)	(0.0075)	(0.0470)
R^1_j	0.0342***	0.2099***	0.0268***	0.1855***
	(0.0020)	(0.0103)	(0.0020)	(0.0119)
$R^{2.3}_j$	0.0308***	0.1543***	0.0277***	0.1561***
	(0.0015)	(0.0109)	(0.0016)	(0.0155)
$R^{4.6}_j$	0.0186***	0.1011***	0.0215***	0.1218***
	(0.0009)	(0.0069)	(0.0012)	(0.0090)
$R^{7.12}_j$	0.0166***	0.0816***	0.0198***	0.1082***
	(0.0008)	(0.0048)	(0.0011)	(0.0088)
$R^{13.+}_j$	0.0123***	0.0748***	0.0196***	0.0963***
	(0.0008)	(0.0075)	(0.011)	(0.0138)
δ^0_1	0.0334	0.0756	-0.0726	0.0657
	(0.0666)	(0.0877)	(0.0513)	(0.1361)
δ^0_2	0.0150	-0.2960**	-0.2197***	-0.2043
	(0.1122)	(0.1299)	(0.0664)	(0.1749)
δ^0_3	0.0881	0.3305	0.2823***	0.1564
	(0.2034)	(0.2143)	(0.1052)	(0.1611)
δ^E_1	0.0052	0.0058	-0.0826*	0.0649
	(0.0539)	(0.0646)	(0.0423)	(0.0784)
δ^E_2	-0.0346	-0.1805	-0.2413***	-0.2510***
	(0.0953)	(0.1144)	(0.0815)	(0.0862)
δ^E_3	0.2924	0.2193**	0.3794***	0.2901***
	(0.2244)	(0.1062)	(0.1096)	(0.0903)
No.Obs.	409	409	409	409
Log-L	4639.70	1278.74	4404.54	948.68

Table 2.2: Parameter estimates for the baseline model (Male), continued.

White (1982) quasi-maximum-likelihood standard errors in parentheses. (5) Age 25-44/ College graduates, (6) Age 45 and over/ High school graduates and less than high school, (7) Age 45 and over/ Some college and associate degree, (8) Age 45 and over/ College graduates.

	(5)	(6)	(7)	(8)
σ^w_{jL}	0.0322***	0.0323***	0.0269***	0.0209***
	(0.0047)	(0.0067)	(0.0071)	(0.0080)
σ^w_{jH}	0.0293***	0.0670***	0.0365**	0.0148***
	(0.0061)	(0.0130)	(0.0161)	(0.0049)
σ^x_{jL}	0.1285**	0.1887***	0.1530*	0.1000
	(0.0560)	(0.0668)	(0.0783)	(0.0704)
σ^x_{jH}	0.0559***	0.0331***	0.1253***	0.1000
	(0.0190)	(0.0099)	(0.0429)	(0.0783)
R^1_j	0.5399***	0.1824***	0.1117***	0.1128***
	(0.0326)	(0.0090)	(0.0091)	(0.0070)
$R^{2.3}_j$	0.6112***	0.1769***	0.0881***	0.0925***
	(0.0368)	(0.0108)	(0.0053)	(0.0058)
$R^{4.6}_j$	0.5976***	0.1360***	0.0806***	0.0928***
	(0.0325)	(0.0078)	(0.0052)	(0.0079)
$R^{7.12}_j$	0.5399***	0.1287***	0.0843***	0.0900***
	(0.0292)	(0.0080)	(0.0053)	(0.0060)
$R^{13.+}_j$	0.6112***	0.1328***	0.0862***	0.0744***
	(0.0311)	(0.0125)	(0.0069)	(0.0066)
δ^0_1	0.5976***	-0.0487	-0.0357	0.3549**
	(0.0433)	(0.0375)	(0.0384)	(0.1437)
δ^0_2	-0.5399***	0.2297***	-0.2570***	-0.0723
	(0.1257)	(0.0521)	(0.0748)	(0.1098)
δ^0_3	0.6112***	-0.0933	0.1970**	0.0111
	(0.1460)	(0.0878)	(0.0773)	(0.0601)
δ^E_1	0.5328***	-0.0703**	-0.0255	0.0776
	(0.0616)	(0.0312)	(0.0392)	(0.1270)
δ^E_2	-0.1890***	0.0537	-0.3212***	0.1008
	(0.0381)	(0.0648)	(0.0621)	(0.1619)
δ^E_3	0.1080***	0.1458	0.3330***	0.0553
	(0.0348)	(0.1247)	(0.0897)	(0.0836)
No.Obs.	409	409	409	409
Log-L	1501.00	597.67	1665.35	1688.86

Table 2.3: Parameter estimates for the baseline model (Female)

White (1982) quasi-maximum-likelihood standard errors in parentheses. (1) Age 16-24/High school graduates and less than high school, (2) Age 16-24/ Some college, associate degree and college graduates, (3) Age 25-44/ High school graduates and less than high school, (4) Age 25-44/ Some college and associate degree.

	(1)	(2)	(3)	(4)
σ^w_{jL}	0.0037	0.0212***	0.0430***	0.0296***
	(0.0031)	(0.0056)	(0.0049)	(0.0049)
σ^w_{jH}	0.0104	0.0340***	0.0998***	0.0221***
	(0.0076)	(0.0063)	(0.0169)	(0.0070)
σ^x_{jL}	0.0473***	0.0857***	0.0629***	0.0251***
	(0.0152)	(0.0323)	(0.0124)	(0.0055)
σ^x_{jH}	0.0235	0.0232***	0.0234***	0.0084***
	(0.0167)	(0.0070)	(0.0056)	(0.0019)
R^1_j	0.0132	0.2113***	0.2920***	0.1867***
	(0.0171)	(0.0104)	(0.0157)	(0.0089)
$R^{2.3}_j$	0.0100**	0.1446***	0.2269***	0.1595***
	(0.0047)	(0.0071)	(0.0119)	(0.0082)
$R^{4.6}_j$	0.3195	0.0976***	0.1852***	0.1116***
	(0.2681)	(0.0056)	(0.0091)	(0.0055)
$R^{7.12}_j$	0.0686	0.0779***	0.1536***	0.1100***
	(0.1003)	(0.0059)	(0.0080)	(0.0074)
$R^{13.+}_j$	0.0012***	0.0598***	0.1228***	0.0876***
	(0.0002)	(0.0043)	(0.0065)	(0.0069)
δ^0_1	0.3195	0.3345**	0.3289***	0.2398
	(0.3481)	(0.1381)	(0.0820)	(0.3424)
δ^0_2	-0.0686	-0.1791	-0.1246*	-0.2210
	(0.4011)	(0.1150)	(0.0671)	(0.4144)
δ^0_3	-0.0012	0.1140	0.0644	0.2364
	(0.1252)	(0.1787)	(0.0746)	(0.5619)
δ^E_1	0.2664***	0.3275**	0.2793***	0.2378
	(0.0682)	(0.1290)	(0.0675)	(0.2399)
δ^E_2	0.0076	-0.2037*	-0.0723	-0.1985**
	(0.0497)	(0.1186)	(0.0445)	(0.0774)
δ^E_3	-0.0340	0.1665	0.0599	0.1901***
	(0.0306)	(0.1558)	(0.0547)	(0.0583)
No.Obs.	409	409	409	409
Log-L	5181.76	1467.28	213.88	1106.08

Table 2.4: Parameter estimates for the baseline model (Female), continued.

White (1982) quasi-maximum-likelihood standard errors in parentheses. (5) Age 25-44/ College graduates, (6) Age 45 and over/ High school graduates and less than high school, (7) Age 45 and over/ Some college, associate degree and college graduates.

	(5)	(6)	(7)
σ^w_{jL}	0.0230***	0.0233***	0.0140***
	(0.0066)	(0.0040)	(0.0030)
σ^w_{jH}	0.0352**	0.0345***	0.0464***
	(0.0154)	(0.0123)	(0.0127)
σ^x_{jL}	0.1168	0.0940***	0.4852**
	(0.1085)	(0.0264)	(0.1940)
σ^x_{jH}	0.0799***	0.0418**	0.0445**
	(0.0258)	(0.0167)	(0.0201)
R^1_j	0.1440***	0.1723***	0.1541***
	(0.0070)	(0.0082)	(0.0149)
$R^{2.3}_j$	0.1257***	0.1540***	0.1343***
	(0.0068)	(0.0083)	(0.0074)
$R^{4.6}_j$	0.0835***	0.1191***	0.1154***
	(0.0055)	(0.0060)	(0.0064)
$R^{7.12}_j$	0.0854***	0.1011***	0.1201***
	(0.0067)	(0.0053)	(0.0084)
$R^{13.+}_j$	0.0724***	0.1037***	0.0887***
	(0.0089)	(0.0066)	(0.0098)
δ^0_1	0.2222**	0.3177	-0.0169
	(0.1101)	(0.1966)	(0.0578)
δ^0_2	-0.1802	-0.2466***	0.3424**
	(0.2249)	(0.0890)	(0.1731)
δ^0_3	0.1520	0.1730*	-0.1118
	(0.3123)	(0.1027)	(0.1242)
δ^E_1	0.1647**	0.2508*	0.0205
	(0.0678)	(0.1423)	(0.0390)
δ^E_2	-0.1807***	-0.1368	-0.2272*
	(0.0606)	(0.1385)	(0.1250)
δ^E_3	0.3737**	0.1625	0.8513***
	(0.1742)	(0.1963)	(0.2194)
No.Obs.	409	409	409
Log-L	1538.65	1050.06	1072.82

Table 2.5: Average type L share and continuation probability (Men)

(1) Men/Age 16-24/High school graduates and less than high school, (2) Men/Age 16-24/ Some college, associate degree and college graduates, (3) Men/Age 25-44/ High school graduates and less than high school, (4) Men/Age 25-44/ Some college and associate degree (5) Men/Age 25-44/ College graduates, (6) Men/Age 45 and over/ High school graduates and less than high school, (7) Men/Age 45 and over/ Some college and associate degree, (8) Men/Age 45 and over/ College graduates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Type L share in inflows	0.09	0.22	0.09	0.32	0.28	0.14	0.37	0.27
Type L continuation prob.	0.90	0.74	0.89	0.83	0.96	0.94	0.80	0.96
Type H continuation prob.	0.46	0.35	0.53	0.44	0.49	0.50	0.46	0.54
Share in total L inflows	0.09	0.07	0.07	0.09	0.05	0.05	0.04	0.03
Share in total inflows	0.17	0.05	0.13	0.05	0.03	0.06	0.02	0.02

Table 2.6: Average type L share and continuation probability (Women)

(9)Women/Age 16-24/High school graduates and less than high school, (10) Women/Age 16-24/ Some college, associate degree and college graduates, (11) Women/Age 25-44/ High school graduates and less than high school, (12) Women/Age 25-44/ Some college and associate degree (13) Women/Age 25-44/ College graduates, (14) Women/Age 45 and over/ High school graduates and less than high school, (15) Women/Age 45 and over/ Some college, associate degree and college graduates.

	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Type L share in inflows	0.12	0.14	0.19	0.25	0.21	0.25	0.13
Type L continuation prob.	0.92	0.90	0.93	0.89	0.89	0.92	0.97
Type H continuation prob.	0.40	0.35	0.44	0.40	0.45	0.42	0.50
Share in L inflows	0.11	0.05	0.14	0.08	0.04	0.07	0.03
Share in total inflows	0.14	0.06	0.11	0.05	0.03	0.05	0.04

Table 2.7: Key characteristic of type L inflows

Key characteristic	Fraction in w_L	Fraction in $w_L + w_H$	Fraction in the labor force
Men	49%	52%	53%
Age 25-44	47%	42%	49%
High school or less	53%	66%	40%

Table 2.8: Parameter estimates by industry

White (1982) quasi-maximum-likelihood standard errors in parentheses. (1) Agriculture, forestry, fishing, farming and mining, (2) Construction, (3) Manufacturing, (4) Wholesale and retail trade. The data is available from January 1976.

	(1)	(2)	(3)	(4)
σ^w_{jL}	0.0061***	0.0712	0.0083***	0.0061***
	(0.0015)	(0.0483)	(0.0009)	(0.0007)
σ^w_{jH}	0.0272**	0.0597	0.0271***	0.0212***
	(0.0130)	(0.1331)	(0.0039)	(0.0043)
σ^x_{jL}	0.1000*	0.0594	0.0992***	0.0998***
	(0.0593)	(0.0659)	(0.0292)	(0.0377)
σ^x_{jH}	0.0136	0.0032	0.0352***	0.0300***
	(0.0082)	(0.4281)	(0.0079)	(0.0079)
R^1_j	0.1314***	0.2959***	0.0276***	0.0377***
	(0.0254)	(0.0658)	(0.0020)	(0.0022)
$R^{2.3}_j$	0.1128***	0.2694	0.0252***	0.0305***
	(0.0064)	(0.1968)	(0.0016)	(0.0017)
$R^{4.6}_j$	0.0809***	0.1878***	0.0194***	0.0214***
	(0.0039)	(0.0135)	(0.0009)	(0.0013)
$R^{7.12}_j$	0.0807***	0.1595***	0.0171***	0.0204***
	(0.0049)	(0.0606)	(0.0011)	(0.0011)
$R^{13.+}_j$	0.0614***	0.1209***	0.0177***	0.0168***
	(0.0041)	(0.0172)	(0.0011)	(0.0010)
δ^0_1	-0.0219	0.3457	0.0470***	0.3089***
	(0.0288)	(2.0672)	(0.0179)	(0.1147)
δ^0_2	0.3843***	-0.2494	0.0098	-0.0948*
	(0.0469)	(3.7930)	(0.0351)	(0.0548)
δ^0_3	0.1806	0.1852	0.0536	0.0076
	(0.1231)	(4.7101)	(0.0360)	(0.0475)
δ^E_1	0.0165	0.2980	0.0022	0.2917***
	(0.0300)	(0.7234)	(0.0161)	(0.1084)
δ^E_2	0.3303***	-0.1448	0.1144***	-0.1293***
	(0.0565)	(0.3916)	(0.0273)	(0.0373)
δ^E_3	0.2268***	0.0552	-0.0144	0.1075*
	(0.0866)	(0.2006)	(0.0215)	(0.0619)
No. Obs.	409	409	409	409
Log-L	1811.79	150.56	4390.54	4253.05

Table 2.9: Parameter estimates by industry

White (1982) quasi-maximum-likelihood standard errors in parentheses. (5) Transportation, (6) Finance, (7) Service, (8) Public Administration. The data is available from January 1976.

	(5)	(6)	(7)	(8)
σ^w_{jL}	0.0279	0.0236***	0.0102***	0.0061
	(0.0995)	(0.0048)	(0.0024)	(0.0038)
σ^w_{jH}	0.0505	0.0229***	0.0219***	0.0275*
	(0.0983)	(0.0078)	(0.0057)	(0.0162)
σ^x_{jL}	0.1000	0.0997***	0.1000***	0.1856
	(0.4548)	(0.0370)	(0.0267)	(0.4083)
σ^x_{jH}	0.0591***	0.0499**	0.0367***	0.0140
	(0.0222)	(0.0202)	(0.0088)	(0.0088)
R^1_j	0.1729***	0.1702***	0.0458***	0.1314***
	(0.0158)	(0.0106)	(0.0029)	(0.0253)
$R^{2.3}_j$	0.1763***	0.1256***	0.0365***	0.1128***
	(0.0121)	(0.0065)	(0.0019)	(0.0069)
$R^{4.6}_j$	0.1248***	0.1044***	0.0315***	0.0808***
	(0.0109)	(0.0054)	(0.0025)	(0.0040)
$R^{7.12}_j$	0.1288***	0.0977***	0.0294***	0.0805***
	(0.0202)	(0.0066)	(0.0019)	(0.0048)
$R^{13.+}_j$	0.0970**	0.0583***	0.0211***	0.0594***
	(0.0426)	(0.0038)	(0.0017)	(0.0086)
δ^0_1	0.0295	0.0937	0.0522	-0.0272
	(0.9985)	(0.0578)	(0.0692)	(0.0525)
δ^0_2	0.1879	0.0773*	-0.0532	0.2616
	(1.1510)	(0.0467)	(0.1034)	(0.3497)
δ^0_3	-0.1346***	-0.0862**	0.0348	0.2848
	(0.0699)	(0.0408)	(0.1058)	(0.7181)
δ^E_1	0.1062	0.0370	0.0404	0.0050
	(1.2303)	(0.0401)	(0.0522)	(0.0702)
δ^E_2	-0.0584	0.0701	-0.2210**	0.2062
	(0.7335)	(0.0674)	(0.1031)	(0.2436)
δ^E_3	0.1344	0.0020	0.4094**	0.3510
	(1.0624)	(0.0533)	(0.1593)	(0.6830)
No. Obs.	409	409	409	409
Log-L	856.54	1449.42	3658.93	1821.11

Table 2.10: Average type L share and continuation probability by industry

Only those who report their previous industry are taken into account in computing the share of each group in the total type L inflows. Newly unemployed individuals who does not have previous industry are not considered. (1) Agriculture, forestry, fishing, farming and mining, (2) Construction, (3) Manufacturing, (4) Wholesale and retail trade, (5) Transportation, (6) Finance, (7) Service, (8) Public Administration.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Type L share in inflows	0.09	0.25	0.16	0.18	0.17	0.25	0.16	0.08
Type L continuation prob.	0.99	0.92	0.92	0.93	0.92	0.89	0.87	0.99
Type H continuation prob.	0.55	0.45	0.50	0.45	0.50	0.37	0.46	0.55
Share in L inflows	0.01	0.17	0.14	0.24	0.05	0.07	0.30	0.01

Table 2.11: Parameter estimates by occupation

White (1982) quasi-maximum-likelihood standard errors in parentheses.

	(1)	(2)	(3)	(4)	(5)	(6)
σ^w_{jL}	0.0083***	0.0445***	0.0381***	0.0074***	0.0043***	0.0088
	(0.0018)	(0.0084)	(0.0123)	(0.0013)	(0.0006)	(0.0133)
σ^w_{jH}	0.0100**	0.0459*	0.0868***	0.0119***	0.0052***	0.0391***
	(0.0045)	(0.0268)	(0.0223)	(0.0029)	(0.0017)	(0.0120)
σ^x_{jL}	0.0982**	0.1135***	0.2032	0.1064**	0.0877***	0.0405
	(0.0398)	(0.0359)	(0.2259)	(0.0521)	(0.0197)	(0.0452)
σ^x_{jH}	0.0512***	0.0384***	0.0359**	0.0335***	0.0350***	0.0599**
	(0.0138)	(0.0109)	(0.0153)	(0.0091)	(0.0072)	(0.0271)
R^1_j	0.0312***	0.2938***	0.2608***	0.0406***	0.0315***	0.1573***
	(0.0022)	(0.0150)	(0.0162)	(0.0022)	(0.0017)	(0.0109)
$R^{2.3}_j$	0.0232***	0.2587***	0.2401***	0.0369***	0.0224***	0.1404***
	(0.0014)	(0.0144)	(0.0130)	(0.0019)	(0.0012)	(0.0094)
$R^{4.6}_j$	0.0197***	0.1822***	0.1800***	0.0263***	0.0186***	0.1056***
	(0.0012)	(0.0093)	(0.0123)	(0.0014)	(0.0009)	(0.0085)
$R^{7.12}_j$	0.0218***	0.1699***	0.1605***	0.0248***	0.0170***	0.0729***
	(0.0014)	(0.0136)	(0.0093)	(0.0016)	(0.0014)	(0.0034)
$R^{13.+}_j$	0.0141***	0.1304***	0.1207***	0.0166***	0.0091***	0.0523***
	(0.0011)	(0.0098)	(0.0272)	(0.0017)	(0.0005)	(0.0040)
δ^0_1	0.1199	0.1819**	0.0395	0.0974	0.1322**	-0.0034
	(0.1385)	(0.0726)	(0.1071)	(0.1128)	(0.0570)	(0.1504)
δ^0_2	-0.1700**	-0.2230***	0.1397	0.0292	-0.0457	0.0249
	(0.0708)	(0.0459)	(0.2008)	(0.1085)	(0.0337)	(0.1229)
δ^0_3	0.1292**	0.2329***	-0.1140*	-0.0550	0.0121	-0.0103
	(0.0581)	(0.0534)	(0.0690)	(0.0489)	(0.0353)	(0.1548)
δ^E_1	0.0936	0.1984**	0.0640	0.1178	0.0650	-0.0028
	(0.1020)	(0.0940)	(0.0772)	(0.0774)	(0.0546)	(0.0909)
δ^E_2	-0.1650*	-0.3349**	-0.1555**	-0.1646***	-0.0119	-0.0631
	(0.0895)	(0.1536)	(0.0730)	(0.0510)	(0.0482)	(0.2680)
δ^E_3	0.2286	0.4527*	0.4208**	0.2981*	0.0745	0.0493
	(0.1607)	(0.2447)	(0.1879)	(0.1554)	(0.0525)	(0.4670)
No.Obs.	409	409	409	409	409	409
Log-L	3579.00	93.601	125.00	3195.39	3935.40	1316.61

Table 2.12: Average type L share and continuation probability by occupation

Only those who report their previous occupation are taken into account in computing the share of each group in the total type L inflows. Newly unemployed individuals who does not have previous occupations are not considered. (1) Executive, administrative, managerial occupation and professional and related service occupation, (2) Sales occupation, (3) Administrative support occupation, (4) Service occupation, (5) Precision production, craft, operator, fabricator and laborers, (6) Farming, forestry and fishing.

	(1)	(2)	(3)	(4)	(5)	(6)
Type L share in inflows	0.26	0.24	0.15	0.19	0.21	0.12
Type L continuation prob.	0.88	0.87	0.92	0.90	0.88	0.85
Type H continuation prob.	0.47	0.43	0.51	0.47	0.33	0.50
Share in L inflows	0.23	0.19	0.11	0.28	0.18	0.03

Table 2.13: Parameter estimates by reason for unemployment

White (1982) quasi-maximum-likelihood standard errors in parentheses.

	Temp.Layoff	Perm.Sep.	Job Leavers	Re-entrants	New entrants
σ^w_{jL}	0.0011***	0.0156***	0.0023***	0.0074***	0.0027***
	(0.0001)	(0.0027)	(0.0004)	(0.0010)	(0.0006)
σ^w_{jH}	0.0039***	0.0278***	0.0148***	0.0120***	0.0132***
	(0.0008)	(0.0050)	(0.0018)	(0.0018)	(0.0022)
σ^x_{jL}	0.0774***	0.1000***	0.0491***	0.0999***	0.1000**
	(0.0262)	(0.0284)	(0.0107)	(0.0377)	(0.0407)
σ^x_{jH}	0.0037	0.0490***	0.0327***	0.0274***	0.0346***
	(0.0030)	(0.0095)	(0.0057)	(0.0084)	(0.0082)
R^1_j	0.0376***	0.0498***	0.0280***	0.0494***	0.0324***
	(0.0036)	(0.0030)	(0.0019)	(0.0025)	(0.0019)
$R^{2.3}_j$	0.0323***	0.0446***	0.0199***	0.0400***	0.0208***
	(0.0036)	(0.0030)	(0.0010)	(0.0022)	(0.0011)
$R^{4.6}_j$	0.0190***	0.0384***	0.0152***	0.0268***	0.0138***
	(0.0011)	(0.0035)	(0.0008)	(0.0017)	(0.0007)
$R^{7.12}_j$	0.0152***	0.0388***	0.0129***	0.0268***	0.0150***
	(0.0015)	(0.0037)	(0.0006)	(0.0018)	(0.0009)
$R^{13.+}_j$	0.0087***	0.0263***	0.0109***	0.0218***	0.0114***
	(0.0011)	(0.0022)	(0.0006)	(0.0015)	(0.0009)
δ^0_1	-0.1469***	0.1474*	0.2900**	0.3327***	0.2636***
	(0.0085)	(0.0865)	(0.1134)	(0.1153)	(0.0470)
δ^0_2	0.0586	0.0356	-0.0030	-0.1386***	-0.3119***
	(0.0416)	(0.0965)	(0.1262)	(0.0318)	(0.0853)
δ^0_3	0.1072	-0.0334	-0.0736	0.0741*	0.4072***
	(0.0900)	(0.0526)	(0.0982)	(0.0390)	(0.1185)
δ^E_1	-0.1490	0.1773***	0.2538***	0.3289***	0.2267***
	(0.0947)	(0.0494)	(0.0902)	(0.1083)	(0.0478)
δ^E_2	0.3522	-0.2709***	-0.0053	-0.1447***	-0.1958***
	(0.2372)	(0.0653)	(0.0539)	(0.0459)	(0.0740)
δ^E_3	-0.1354	0.4605***	-0.0237	0.1240	0.2708**
	(0.1642)	(0.1298)	(0.0275)	(0.0810)	(0.1226)
No.Obs.	409	409	409	409	409
Log-L	4469.50	3162.01	5153.11	3783.22	5054.05

Table 2.14: Average type L share and continuation probability by reason for unemployment

(1)TL: Temporary layoff, (2)PS: Permanent separation, (3)JL: Job leavers, (4)RE: Re-entrants, (5)NE: New-entrants

	TL	PS	JL	RE	NE
Type L share in inflows	0.04	0.23	0.14	0.16	0.04
Type L continuation prob.	0.87	0.93	0.93	0.94	0.92
Type H continuation prob.	0.40	0.59	0.43	0.46	0.47
Share in L inflows	0.04	0.42	0.12	0.32	0.09

Table 2.15: Key characteristic of type L inflows

The fraction of individuals whose education attainment is high school diploma and less in the labor force is the average value of 1992-2013.

Key characteristic	Fraction in w_L	in $w_L + w_H$	in labor force
Age-Gender-Education			
- High school diploma or less	53%	66%	40%
Industry			
- Wholesale, retail trade and service	54%	62%	48%
- Construction and manufacturing	31%	30%	17%
Occupation			
- Sales and administrative support	21%	26%	25%
Reason for unemployment			
- Permanent separation	42%	28%	2%

Chapter 3

Forecasting Unemployment using Dynamic Model Adaptation

Abstract. This paper proposes a new method of combining forecasts based on the recent performance of out-of-sample forecasts for forecasting the U.S. unemployment rate. At every period, a forecaster chooses a single model of which the recent out-of-sample forecasts yields the smallest squared error among a given set of forecasting models to make multiple-period ahead forecasts. The proposed combination method produces more accurate forecasts than existing model averaging methods and the Greenbook forecasts.

3.1 Introduction

As noticed by many economists and forecasters, forecasting the unemployment rate is an important but challenging task. Part of the difficulty comes from the fact that unemployment rate exhibits asymmetry in its dynamics depending on the business cycle phase that is quite different from the pattern of other macroeconomic time series; see for example Montgomery, Zarnowitz, Tsay and Tiao (1998) and Hamilton (2005). This implies that forecasting the unemployment rate could be more difficult around business cycle turning points or in large recessions if the forecaster does not adapt to the change in the dynamics of unemployment rate in a timely manner. From a policy maker's perspective, the cost of misforecasts could be high during these periods, since preemptive monetary policy based on misleading forecasts could hamper the quick recovery of the economy from recessions.

The difficulty in forecasting a time series of which the dynamics changes over time has been described as the model instability problem. In the presence of model instability, the best model to predict a time series is likely to change over time. In this aspect, model instability could be a critical issue in forecasting the unemployment rate. Econometricians have considered various ways to tackle the problem such as modelling a break process, a change in regimes, rolling window estimation and so on (Stock and Watson, 2006; Pesaran, Pettenuzzo and Timmermann, 2006; Billio, Casarin, Ravazzolo and van Dijk, 2012; Groen, Paap and Ravazzolo, 2013; Koop and Korobilis, 2012). Meanwhile, in the literature of model averaging, the focus has been on developing optimal methods of forecast combination that would enable us to get around with model instability; see for example Clements and Hendry (2004), Stock and Watson (2004) and Elliott and Timmermann (2005). In spite of the intensive research in this topic, simple averaging has been the most popular method for forecast combination since it performs relatively well compared to other approaches, often known as "forecasting combination puzzle", and is easy to implement in practice.

The model instability problem could be substantially mitigated if we can identify the best local forecasting model that captures the changing dynamics of economic time series in a timely manner and if the performance of the best forecasting model is persistent. Aiolfi and Timmermann (2006) study the persistence in the forecasting performance of various macroeconomic time-series models and claim that we can improve out-of-sample forecasts through forecast combination by taking advantage of the persistence in the relative performance of forecasting models. This implies a new possibility in forecasting the unemployment rate or economic time series subject to model instability. If the relative performance of forecasting models is persistent, the recent out-of-sample performance of forecasting models could serve as an indicator on whether the best local forecasting model changes. Moreover, a forecasting model combination method based on the recent forecasting performance could translate into improvement in out-of-sample forecasts.

In this paper, I first examine the persistence in the relative performance of unemployment rate forecasting models and find that the best model exhibits substantial persistence in out-of-sample forecasts. Based on this finding, I propose *Dynamic Model Adaptation*, a new forecast combination method designed for forecasting a time series which is subject to model instability. The proposed method is that every period a forecaster chooses the model of which the most recent s-period ahead out-of-sample forecasts has the smallest squared errors among a given set of forecasting models and use that model to make multiple-period ahead forecasts. I investigate empirically the out-of-sample forecasting performance of this new combination method in forecasting the unemployment rate in real time, with a large number of models and with its small subset composed of models showing top performance. The performance of proposed method is compared with that of existing forecast combination approach. I find that with the small group of models, changing forecasting models based on the recent performance of 3-quarter-ahead out-of-sample forecast produces more accurate forecasts than other model averaging approaches and than the same approach with more than 100 models. Particularly, the predictive power of this approach becomes stronger during economic recessions. Considering that monetary policy becomes effective from the second quarter after it is implemented and unemployment forecasting is harder during near a economic recession, the proposed method can help the policy makers by providing better forecasts with ease.

The outline of this paper is as follows. In Section 1, I introduce models and data sets used in forecasting the U.S. unemployment rate in real time. Section 2 examines the persistence in the relative out-of-sample performance of best forecasting models. Section 3 introduces Dynamic Model Adaptation and studies the out-of-sample forecasting performance of the proposed method. Section 4 concludes.

3.2 Forecasting the U.S. Unemployment Rate

In this section, I introduce models and data sets used in forecasting the U.S. unemployment rate in real time. The forecasts for each time period t are made based on the common information set that was used by professional forecasters when they made their forecasts for the Survey of Professional Forecasters and the Greenbook.[1] The time series model for h-step-ahead forecasts of the conditional mean of the target variable, y_{t+h}, has the following form

$$y_{t+h} = f_i(x_t; \theta_h^i) + e_{t+h,t}^i$$

where i is an index for the forecasting model. θ_h^i is a vector of unknown parameters and x_t is a vector of predictor variables that are known at time t and may include $y_{t.}$

[1]The first survey conducted in real time was the one for 1990:Q3. The survey's timing is geared to the release of the Bureau of Economic Analysis' advance report of the national income and product accounts. This report is released at the end of the first month of each quarter. The survey questionannaires are sent after this report is released to the public. The survey's questionnaires report recent historical values of the data from the BEA's advance report and the most recent reports of other government statistical agencies. Thus, in submitting their projections, our panelists' information sets include the data reported in the advance report. For the surveys after the 1990:Q2 survey, they have set the deadlines for responses at late in the second to third week of the middle month of each quarter. A complete list of the dates of deadlines for surveys from 1990:Q2 to the present is available on the Philadelphia Fed's website.

$e^i_{t+h,t}$ is an h-step error term. Individual forecasting models typically only use a subset of the elements of x_t. The forecast of y_{t+h} by the ith model is computed as $f_i(x_t; \hat{\theta}^i_{h,t})$, where $\hat{\theta}^i_{h,t}$ is the estimate of θ^i_h given period-t information. $f_i(\cdot)$ can be a either linear or nonlinear forecasting model.

In the first part of this section, I briefly discuss forecasting models. For $f_i(\cdot)$, I consider broadly four classes of models: (1) univariate model, (2) vector autoregressive model, (3) factor augmented vector autoregressive model and (4) nonlinear steady-state model of labor market flows. In the second part, I discuss the methods to estimate $\hat{\theta}^i_{h,t}$ and to generate out-of-sample forecasts, $f_i(x_t; \hat{\theta}^i_{h,t})$.

3.2.1 Alternative Forecasting Models

Univariate Model

The first class of linear models that I consider is ARIMA model (Box, Jenkins and Reinsel, 1994). For the monthly unemployment rate series, u_t, the model can be written as

$$(1 - \phi_1 L - \cdots - \phi_p L^p)(1 - L)^d u_t = c + (1 - \theta_1 L - \cdots - \theta_q L^q)\varepsilon_t$$

where p, d and q are nonnegative integers and c, $\phi' s$ and $\theta' s$ are parameters, L is the lag operator, and $\{\varepsilon_t\}$ is a sequence of i.i.d. random variables with mean 0 and variance σ^2_ε. I assume that

$$\begin{aligned}
\phi(L) &= 1 - \phi_1 L - \cdots - \phi_p L^p \\
\theta(L) &= 1 - \theta_1 L - \cdots - \theta_q L^q
\end{aligned}$$

have no common factors and have all their 0's outside the unit circle. When $d = 0$, u_t is weakly stationary, while it is unit-root non-stationary when $d > 0$. I use ARIMA(2,0,1) and ARIMA(2,1,0) as univariate forecasting models, the two specifications commonly used in the literature (Montgomery et al. 1998; Barnichon and Nekarda, 2012).

Vector Autoregressive Model

To incorporate relevant information other than the unemployment rate into the forecasts, we can use the vector autoregressive (VAR) model. I consider VAR with two lags for forecasting.[2]

$$y_t = c + \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + e_t, \tag{3.1}$$

Specifically, let

$$y_t = [u_t, x_t]'. \tag{3.2}$$

As regressors that go into vector x_t, I consider labor force flow variables such as the separation rate from employment (s_t), the job finding probability (f_t), and leading labor market indicators such as the monthly average of initial claims for unemployment insurance (uic_t) and the help-wanted index (hwi_t).[3] For forecasting, I take logs of s_t, f_t and uic_t and take a log difference of hwi_t.[4]

Given the four possible regressors, $\ln s_t$, $\ln f_t$, $\ln uic_t$ and $d \ln hwi_t$, there are 15 ways to select the variables for x_t, since the autoregressive components are always included in forecasting model. In other words, there are 15 different VAR forecasting models. For example, if we include all the regressors in vector x_t, then

[2] I considered the lag lengths between 1 and 12 months, but found that the VAR model with two lags overall produces more accurate forecasts.

[3] Note that given the timing convention for the flows used in Shimer (2012) and others, the hazard rates enter the VAR lagged by 1 month.

[4] I use this transformation following Barnichon and Nekarda (2012). I considered other specifications but this generates better forecasts than alternatives.

$$x_t = [\ln s_t,\ \ln f_t,\ \ln uic_t,\ d \ln hwi_t]'. \tag{3.3}$$

As alternative measures of labor market flows, I use the exit probabilities from unemployment by duration, particularly the exit probability from unemployment of newly unemployed individuals and of long-term unemployed individuals (those who have been unemployed longer than 6 months) instead of the separation and the job-finding probability.[5]

Factor Augmented Vector Autoregressive Model

Next, I consider factor augmented vector autoregressive (FAVAR) models. Ahn and Hamilton (2014) and Ahn (2014) show that the difference in the exit probabilities from unemployment by duration comes from ex-ante (cross-sectional) and ex-post heterogeneity of unemployed individuals and that ex-ante heterogeneity is important in forecasting the unemployment. To capture the heterogeneity in unemployment exit probabilities in a parsimonious way, I estimate the factors from the exit probabilities of individuals unemployed for 1 month, 2-3 months, 4-6 months and longer than 6 months using principal components. In spite of the possible dynamic features in the factors, Stock and Watson (2002) show that we can estimate the factors using principal components.[6] I consider two factors, since the two factors explain most of the variation in the exit probabilities by duration.

Let f_{1t} and f_{2t} be the first and second factor, respectively. The first principal component turns out to be associated with the mean exit probability which is driven by the aggregate labor market situations and the aggregate economic policies. The second factor turns out to describe the dynamic feature in the dispersion of the exit probabilities by duration of unemployment. We can think of two explanations for the time variation in the second factor. First, there exists compositional variation in the newly unemployed

[5] See Appendix for the details about the exit probabilities by duration.
[6] Alternatively one can use the dynamic factor approach by Forni et al. (2005).

people or the exit probabilities of a particular group changes substantially. Second, the pattern of ex-post heterogeneity such as the speed of unemployment scarring or motivational effects, the tendency to leave the unemployment status more quickly as the duration gets longer, changes (Ahn and Hamilton, 2014; Ahn, 2014). Both have persistent effects on unemployment and are important in predicting the path of the unemployment rate. The two factors summarize not only the aggregate labor market situations but also the ex-ante and ex-post heterogeneity is a parsimonious way. Therefore we can efficiently utilize the information for forecasting the unemployment rate with FAVAR models.

The model is identical to equation (3.1)-(3.3) except that I replace the labor market flow variables with f_{1t} and f_{2t}. To smooth out high frequency movements and seasonally adjust the exit probabilities by duration, I used 6-month moving average, 12-month moving average and X-12-ARIMA. In addition, I also used seasonally unadjusted exit probabilities for the estimation of factors. In total, I consider four possible sets of $[f_{1t}, f_{2t}]$.

Nonlinear Steady-State Model

The last class of models that I consider is a nonlinear steady-state model proposed by Barnichon and Nekarda (2012). They consider information from the labor force flows through the concept of the conditional steady-state unemployment rate. The conditional steady-state unemployment rate is the rate of unemployment that would prevail eventually if the flows into and out of unemployment remain at the current rates. In steady state, these flows are balanced. However, if the inflow rate jumps, for instance, the conditional steady-state unemployment rate also goes up. Without additional shocks, the unemployment rate would rise toward the new steady state. The speed of convergence is also determined by the size of flows.

With this set up, their forecasting model is built upon two components. The first step is forecasting the labor force flows used to determine the steady-state unemployment and the speed at which actual unemployment converges to steady state. In the second step, they feed the forecasted flows into the model of law of motion which describes how the unemployment rate converges to the new steady-state value from the past unemployment rate.

The labor force flow variables, employment separation rate (s_t) and the job finding probability ($f_t)$, are forecasted using a vector autoregression (VAR). They include the help wanted index (hwi_t), the monthly average of initial claims for unemployment insurance (uic_t) and the unemployment rate in the estimation. Specifically let

$$y_t = [\ln s_t, \ln f_t, d \ln u_t, \ln uic_t, d \ln hwi_t]'.$$

Barnichon and Nekarda estimate the following VAR model,

$$y_t = c + \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \varepsilon_t,$$

and compute j-period ahead out-of-sample forecasts of separation and job finding probability, $\hat{s}_{t+j|t}$ and $\hat{f}_{t+j|t}$, respectively.

Lastly, they plug $\hat{s}_{t+j|t}$ and $\hat{f}_{t+j|t}$ into the law of motion and obtain j-period-ahead forecasts of unemployment by iterating forward on the following model:

$$\hat{u}_{t+j|t} = \hat{\beta}_{t+j} \hat{u}^*_{t+j} + (1 - \hat{\beta}_{t+j}) \hat{u}_{t+j-1|t}$$

$$\hat{u}^*_{t+j} = \frac{\hat{s}_{t+j|t}}{\hat{s}_{t+j|t} + \hat{f}_{t+j|t}}$$

$$\hat{\beta}_{t+j} = 1 - \exp\{-(\hat{s}_{t+j|t} + \hat{f}_{t+j|t})\}.$$

3.2.2 Data Set

The seasonally adjusted unemployment rate and the numbers unemployed for duration less than 5 weeks, 5-14 weeks, 15-26 weeks and longer than 26 weeks that are available on a specified vintage date are obtained from Archival Federal Reserve Economic Data (ALFRED).[7]

The inflow and outflow probabilities are calculated with the real time data using Shimer (2012)'s model. The exit probabilities by duration are computed from Current Population Survey(CPS) micro data in real time. Seasonal adjustment for these series is conducted using two methods - 12-month moving average or the estimation of seasonally adjustment components using X-12-ARIMA in real time. Instead of the seasonal adjustment, I also consider smoothing out the series using 6-month moving average. Following Barnichon and Nekarda (2012), I use monthly averages of seasonally adjusted weekly initial unemployment claims and Barnichon's (2010) composite help-wanted index as leading indicators of the labor market. There are no revisions to the print help-wanted index. Real-time data for initial claims are available beginning in June 2009. The maximum absolute variation in the monthly-average level of weekly initial claims since June 2009 is small. The sample period of the forecasting exercises is January 1976 - December 2014, since CPS micro data is publicly available from January 1976.

3.2.3 Estimation and forecasting methods

In total, there are 161 possible monthly forecasts every month. To estimate the parameters of forecasting models, I use both growing-window estimators and 15-year-rolling-window estimators for the estimation of model parameters. For rolling estimation, the initial origin of forecasts begins from January 1992 and ends at Decem-

[7] The vintage dates of labor force statistics are available in ALFRED website (http://alfred.stlouisfed.org/).

ber 2012.[8] For the estimation of univariate models, ARIMA(2,0,1) and ARIMA(2,1,0), I used rolling regression with a fixed window only. Following Barnichon and Nekarda (2012), I also compute the forecasts using their nonlinear steady-state model with constant hazards (s_t and f_t), which does not require rolling regression of any type.[9]

To generate out-of-sample forecasts, I use iterated forecasting, which is to estimate a dynamic model for data observed at the highest available frequency and then simulate the model to generate forecasts at longer horizons. This approach has the following merits compared to the alternative method, direct forecasting, which is to estimate a separate model for each horizon, regressing future realizations on current information.[10] The iterated approach leads to more efficient parameter estimates for a given model, since it includes data recorded at the highest available frequency and thus uses the largest available sample size. In addition, if the model is misspecified due, for instance, to an omitted variable or because of an incorrect lag order, iterating the model multiple steps ahead can reduce existing biases (Marcellino, Stock and Watson, 2006; Pesaran, Pick and Timmermann, 2010).[11]

I generate out-of-sample forecasts up to 24-months ahead. To compare the forecasts with those of Greenbook and Survey of Professional Forecasters (SPF), I compute the averages of monthly forecasts to produce quarterly forecasts.[12] I take this route instead of forecasting quarterly unemployment directly with quarterly models, because monthly series provides valuable additional information for purposes of short-term

[8]Barnichon and Nekarda (2012) estimate their VAR model over a 15-year rolling windows. They mentioned that a rolling window yielded more accurate forecasts than a recursive window in real time. It is likely because of the low-frequency patterns that windows of 15 years were superior to 10- and 20-year windows.

[9]There are 82 models. 79 models are estimated with rolling regression with a fixed window or with expanding window. In this case, I consider each case as a separate model.

[10]Under the indirect approach, the model specificaiton is the same across all forecast horizons. Only the number of iterations changes with the horizon. Under the direct forecasting, both the model specification and estimates can vay across different forecast horizons (Pesaran, Pick and Timmermann, 2010).

[11]Both approaches have advantages and drawbacks. Direct forecasts are less efficient but could be more likely to be robust to model misspecification as they are typically linear projections of current realization on past data.

[12]Survey of Professional Forecasters (SPF) was originally conducted by the American Statistical Association (ASA) and NBER, although it has been conducted by the Federal Reserve Bank of Philadelphia.

forecasting (Montgomery et al.,1998).[13]

Table 3.1 and Table 3.2 report the performance of selected models measured by root-mean-squared-errors (RMSE).[14] In Table 1, for the entire sample period, the RMSE of the nonlinear steady state model with rolling estimators with a fixed window ($BN(f)$) is the smallest among the alternative forecasting models in the current quarter forecast and is even smaller than that of median SPF forecasts. In the longer-horizon forecasts, the FAVAR forecasting models perform better than other models in 4-quarter ahead forecasts. Specifically, the RMSE of FAVAR 1 and 2 is smaller than that of $BN(f)$ by 7% and that of ARIMA(2,1,0) model by 10%.

Table 3.2 compares the performance of forecasting models to that of SPF and Greenbook forecasts conducted during 1992-2009. The comparison ends in 2009, because the Greenbook forecasts are made public with a 5-year lag. Overall, the RMSE's of VAR, FAVAR the nonlinear steady-state models are smaller than that of median forecasts of SPF in the current quarter and that of Greenbook in the current and next quarter and the current year forecasts.

Among the forecasting models, FAVAR models perform consistently better than other models during 1994-2012. This suggests that the heterogeneity in unemployment exit probabilities captured by the factors could be the important information in forecasting the unemployment rate. Interestingly, the predictability of VAR model becomes stronger when the economy is near or in a recession. During 2000-2002 and 2006-2009, VAR 2 and 3 model outperformed Greenbook forecasts in all of the forecasting horizons, from the current to 4-quarter ahead (Table 3.3). In addition, their RMSE's

[13] Barnichon and Nekarda (2012) also used this approach.

[14] The followings are variables going into the vector x_t in the forecasting models: (1) VAR1: $\ln s_t, \ln f_t$, (2) VAR2: $\ln uic_t, d \ln hwi_t$, (3) VAR3: $\ln uic_t$, (4) VAR4: exit probabilities of those who have been unemployed for 1 month and 2-3 months. (5) FAVAR 1 : the first two factors estimated from the seasonally unadjusted exit probabilities by duration, $\ln uic_t$ and $d \ln hwi_t$, (6) FAVAR 2: the first two factors estimated from the exit probabilities by duration smoothed using 6-month moving average, $\ln uic_t$ and $d \ln hwi_t$ (7) FAVAR 3: the first two factors estimated from the exit probabilities by duration seasonally adjusted with X-12-ARIMA in real time, $\ln uic_t$ and $d \ln hwi_t$.(8) BN(e), BN(f): Barnichon and Nekarda (2012)'a nonlinear steady-state model with expanding and fixed window rolling estimators. (9) BN(ss): the nonlinear steady-state model with constant f_t and s_t.

of 4-quarter-ahead and next-year forecasts are smaller than those of median SPF forecasts. The evidence that the best local forecasting model differs across time suggests that unemployment forecasts could be improved if we adapt our forecasting models to changing economic situations.

3.3 Persistence in Forecasting Performance

In this section, I analyze the persistence in relative out-of-sample performance of unemployment forecasting models. I consider the relative forecasting performance of the entire set of models as well as a small number of top-performing forecasting models.

I characterize the persistence in relative performance in two ways. First, it is measured with the probability that the best model at time τ is different from that of time $\tau + 1$ in forecasting the unemployment rate s-period ahead. Let $P^{s}_{\tau+1|\tau}$ denote the probability. This captures how persistent the performance of the best model is when the information set used for forecasting changes over time. Another way to measure the persistence is the probability that the best model for s-period ahead forecasts made at time τ is the same as the one for $s+1$-period ahead forecasts made at time τ. Let the second probability be $P^{s+1|s}_{\tau}$. This measures how good a forecasting model continues to be as the forecasting horizon expands given the same information set.

Each month t, I make h-month ahead forecasts for $h = 1, 2, 3, ..., 24$ associated with ith model, $\hat{y}^{i}_{t+h|t}$, and then make average of three monthly forecasts as follows

$$\hat{z}^{i}_{t+s} = \frac{\hat{y}^{i}_{t+(3s-2)|t} + \hat{y}^{i}_{t+(3s-1)|t} + \hat{y}^{i}_{t+3s|t}}{3}$$

for $s = 1, 2, 3, ..., 8$. The s-period ahead performance of the ith forecasting model at time t is measured through the loss function, $L^{i}_{t+s|t} = L(z_{t+s}, \hat{z}^{i}_{t+s|t})$ where z_{t+s} is the data. I assume squared forecast error (SFE) loss:

$$L(z_{t+h}, \hat{z}^i_{t+s|t}) = (z_{t+h} - \hat{z}^i_{t+s|t})^2 \tag{3.4}$$

I rank $L^i_{t+s|t}$ for all i. Let m^s_{t+s} be the index of forecasting model of which s-period ahead forecasts made at time t yields the smallest SFE computed at time $t+s$,

$$m^s_{t+s} = \arg\min_i (L^i_{t+s|t}). \tag{3.5}$$

It follows that $P^s_{\tau+1|\tau}$ and $P^{s+1|s}_{\tau}$ can be written as follows

$$\begin{aligned} P^s_{\tau+1|\tau} &= \sum_{t=1}^{T-1} \frac{\mathbf{1}(m^s_{t+s} \neq m^s_{t+s+1})}{T-1} \\ P^{s+1|s}_{\tau} &= \sum_{t=1}^{T} \frac{\mathbf{1}(m^s_{t+s} \neq m^{s+1}_{t+s+1})}{T}, \end{aligned}$$

where $\mathbf{1}(\cdot)$ is an indicator function and T is the number of forecasts made.

First I examine the relative performance of the entire set of 161 models. Table 3.4 reports $P^s_{\tau+1|\tau}$ computed from the forecasts of the 161 models. There are two interesting facts. First, the best forecast model changes very frequently among 161 forecasting models. Over the forecasting horizons, $s = 1, 2, 3, ..., 8$, the probabilities that the best forecast model in month t changes next month in $t+1$ are between 0.87 and 0.99 on average. Second, the best forecast model switches more frequently in the near-term forecasts than it does in the forecasts of longer-horizons. The probability that the best model in forecasting the unemployment rate 1-period ahead ($s = 1$) in month t changes next month is 0.99, but that in 8-period ahead ($s = 8$) forecasting is 0.87. This implies that the performance of the best forecast model is more persistent when we try to forecast the unemployment rate over longer horizons. Third, this pattern gets stronger during recessions. During the Great Recession, the best model in 1-period ahead forecasting changes every month. However, the probability that the best model in 8-period-ahead forecasting in month t changes next month is 0.75, much lower than the average probability of the entire sample period. This suggests that the performance of best long-term forecast model becomes more persistent during economic downturns.

Table 3.5 reports $P_{\tau}^{s+1|s}$,the probability that the best model for s-period ahead forecasts made in a month is not the best one for $s+1$-period ahead forecast generated in the same month. Three patterns are noticeable, which are quite similar to those in Table 3.4. First, the best model for s-period ahead forecasts is often the best one in $(s+1)$-period ahead forecasting. Over the forecasting horizons, $s = 1, 2, 3, ..., 7$, the probability that the best model in forecasting the unemployment rate s-period ahead is not the best one for $(s+1)$-period ahead forecasting is between 0.67 and 0.91 on average. Second, the performance of the best forecast model gets more persistent as s becomes larger. The probability that the best model for 1-period-ahead forecasts is not the best one for 2-period-ahead forecasts is 0.91, while that for $s = 7$ is not the best one for $s = 8$ is 0.67. Third, this feature becomes more dramatic during economic downturns. Recently, between 2007 and 2012, the probabilities all decrease over the forecasting horizons. However, the difference between the short and long horizons becomes larger. The probability that the best model of 7-period-ahead forecasts is not the best one for 8-period-ahead forecasts is 0.38.

Similar patterns are found when we consider monthly forecasting horizons up to 24 months, as seen in Table 3.6.

So far the persistence is measured with the monthly probability that the best forecasting model switches. In practice, we are interested in how well the best model with the smallest squared error for s-period-ahead forecasts evaluated in month t (model m_t^s) would perform in forecasting the unemployment rate in longer horizons. Table 3.7 shows the relative performance of model m_t^s in forecasting the unemployment rate n quarters ahead for $n = 0, 1, 2, 3$ and 4. We can first observe that m_t^6 model showed the highest average ranks over n across $s = 1, 2, ...8$. Among the 161 models, model m_t^6 is ranked 53 on average in 4-quarter-ahead forecasting. The relative performance of models m_t^6 become stronger particularly during the most recent recession. Between

December 2007 and June 2009, the average rank of m_t^6 models is 21 in 2-quarter-ahead forecasting and 28 in 4-quarter ahead forecasting. This implies that the relative performance of the best out-of-sample forecasting model is persistent in forecasting the unemployment rate in longer horizons.

Would the relative performance of forecasting models also exhibit substantial persistence if we only consider a small set of better-performing models instead of including hundreds of models in the forecast combination? If it does, it will save our effort to estimate a large number of models to improve forecasts. To verify this, I choose 8 models based on their RMSE of out-of-sample forecasts and compute $P_{\tau+1|\tau}^s$, $P_\tau^{s+1|s}$ and the relative out-of-sample performance of model m_t^s. The selected models are the nonlinear steady-state model with expanding-window estimation ($BN(e)$), VAR 1, 2, 3 and 4 model, FAVAR 1, 2 and 3 model in Table 1-3.

Table 3.8 and 3.9 report $P_{\tau+1|\tau}^s$ and $P_\tau^{s+1|s}$ for the 8 models, respectively. Since we consider fewer models, the best forecast model changes less frequently. The persistence features in the relative performance observed in the entire set of models also appear in the case of 8 models. The best forecast model switches more frequently in the near-term forecasts than it does in the forecasts of longer-horizons. This again suggests that the performance of the best forecast model is more persistent when we try to forecast the unemployment rate over longer horizons. In addition, the performance of the best forecast model becomes more persistent as s becomes larger. During the recession, the relative persistence gets smaller throughout 's'.

Table 3.10 shows the average ranks for n-quarter ahead forecasts generated by model m_t^s among the 8 models. Across s, model m_t^s performs slightly better than the average. Since we consider a small number of top-performing models, the persistence of relative performance measured with the average ranks for n-quarter-ahead forecasts is less dramatic than that observed from the entire set of models. Nonetheless, the pattern is robust that the relative performance of the best out-of-sample forecasting model is persistent in forecasting the unemployment rate in longer horizons and the persistence is stronger during the Great Recession.

The takeaway of this exercise is summarized as follows. First, as the forecasting horizon gets longer, the persistence in the relative performance of best forecasting model gets higher. Even though the information set changes, the best forecasting model in the current period is more likely to be the best one in the next period. In addition, the best model in forecasting unemployment s periods ahead is more likely to be the best one for $(s + 1)$-period ahead forecasts. Second, the persistence of model performance has a counter-cyclical feature. The persistence of both types becomes greater around economic recessions. Third, the persistence in the relative performance is preserved in out-of-sample forecasts of longer horizons. Most of all, these features are robust although we consider fewer top-performing models. Considering that the unemployment rate is hard to forecast near the economic downturns due to model instability, these findings suggest that we might be able to improve out-of-sample forecasts of unemployment rate by taking advantage of the persistence in relative performance of forecasting models with a small number of forecasting models.

3.4 Dynamic Model Adaptation

3.4.1 Description

We saw that there is the substantial persistence in the performance of best forecasting models in longer horizons and that the persistence is counter-cyclical. Aiolfi and Timmermann (2006) observed that persistence in forecasting performance is likely to be a key determinant of the optimal degree of averaging across models, with less averaging being required the more persistent the relative performance is. This suggests that choosing the recent top forecasting model may improve out-of-sample forecasts. In this section, I propose a new strategy of forecasting combination, *Dynamic Model Adaptation,* which is to use the model of which the recent performance of out-of-sample forecasts is the best among the candidate models for forecasting.

Dynamic Model Adaptation is composed of two steps. First, at time t we calculate the realized squared error of s-period ahead forecasts of model i made at time $t-s$. Let $L^i_{t|t-s}$ be the squared forecast error. We choose the model whose $L^i_{t|t-s}$ is the smallest among the forecasting models,

$$m^s_t = \arg\min_i (L^i_{t|t-s}),$$

where m^s_t is the index for the best model. Second, we use the model m^s_t to make n-quarter ahead out-of-sample forecasts, $\hat{y}^{m^s_t}_{t+n|t}$. I will call s the switching criteria and the s that produces the best forecasts the optimal switching criteria.

Dynamic Model Adaptation is comparable to the performance based weighting by Stock and Watson (2004) and the information criterion methods (AIC and BIC) by Pesaran, Pick and Timmermann (2010). The proposed method is similar to the two existing methods in that the combination of forecasts is based on the relative performance of forecasting models and that the time-variation in the relative performance is taken into account.

The difference is as follows. In the performance based weighting, the history of forecasting performance is reflected on the combination weight on individual model and thus non-negative weights are allocated to the forecasting models other than the best model.[15] Meanwhile, in Dynamic Model Adaptation only the most recent out-of-sample performance is considered in the model selection procedure and only the best model is used to make out-of-sample forecasts. The proposed approach is also differentiated from Pesaran et al.(2012). Their information criterion method is based on the in-sample performance of forecasts made from rolling regressions, while the performance of out-of-sample forecasts is the model selection criterion in the method.

[15]The weight on each forecasting model is calculated from the sum of discounted squared errors realized up to time t.

3.4.2 Empirical Evaluation

The proposed method is applied to the selected 8 models as well as the entire set of 161 models.

In Table 3.11, the performance of Dynamic Model Adaptation with the optimal switching criteria s for forecasting the unemployment rate is reported and compared with that of existing combination approaches. I consider the equal weighted combination (simple averaging), median, least squares combination (LS), Bayesian combination (BS) and performance based weighting scheme (PW).[16] There are three important features about the empirical performance of proposed method.

First, for the forecasts made during 1994-2012, I found that Dynamic Model Adaptation with the 8 models with $s = 3$ works better than other combination methods in forecasting the unemployment rate from the current to 4 quarters ahead and the average unemployment rate of the current as well as next year (Table 3.11).[17] For the forecasts made during 1994-2009, it performs better than Greenbook forecasts in out-of-sample forecasting throughout the forecasting horizons (Table 3.12). The RMSE's of proposed method are smaller than those of equal weighted combination, the most popular method, by 10% in the 2-quarter ahead forecast and in the 4-quarter ahead forecast, respectively.

Second, it turns out that the proposed method with $s = 6$ performs better than other model averaging methods, if we take the entire set of forecasting models into account. The striking feature is that including more models in the forecast combination does not improve out-of-sample forecasts. The RMSE's are close to those for the 8 models in size and even slightly larger in the 3 and 4-quarter-ahead and the next year forecasts as shown in Table 11 and 12.

Third, Dynamic Model Adaptation with the 8 models performs particularly bet-

[16] I used the implementation by Stock and Watson (2004) for Bayesian combination and performance based weighting. I followed Granger and Ramanthan (1984) to compute least squares combination. The details on other combination methods are found in Genre, Kenny, Meyler and Timmermann (2013).

[17] I cannot use the forecasts of the first two years, since I compare out-of-sample forecasting up to 2 years to implement the proposed method.

ter during economic recessions. Between the 1st quarter of 2000 and 4th quarter of 2002 and between the 1st quarter of 2006 and 4th quarter of 2009, the RMSE of Dynamic Model Adaptation with the 8 models is smaller than that with 161 models by 9% and than that of equal weighted combination by 16% in the next-year forecasts (Table 3.13). In addition, it performs better than the Greenbook forecasts throughout the forecasting horizons and produces more accurate forecasts of the average unemployment rate next year than the median forecast of SPF.

Meanwhile, during economic booms with less fluctuations in the unemployment rate, the equal weighted combination with the 161 models is the best in forecasting the unemployment rate from the current to 4 quarters ahead and the average unemployment rate of the current as well as next year. It is strictly better than the median forecasts of SPF and the simple averaging with the 8 models (Table 3.14). It is notable that the significant forecasting improvement comes from additional models that we consider in the equal weighted combination.

Previous research on forecasting combination has found that it is hard to beat the simple averaging with many models. This phenomenon is often dubbed ”forecasting combination puzzle”. In case of forecasting the unemployment rate, the equal weighted combination with a large number of models is a powerful tool during booms, but its performance deteriorates substantially during recessions. However, the predictive power of Dynamic Model Adaptation with a small number of top-performing models is particularly strong during economic downturns. The overall improvement of the proposed method in out-of-sample forecasts and its superior prediction surrounding economic recessions imply that we can effectively and easily mitigate the model instability problem in forecasting the unemployment rate with ease. This further suggests the possibility that the forecasting combination strategies might have to differ depending on the nature of time-series which we try to forecast.

3.5 Conclusion

The paper proposes a new forecast combination strategy, Dynamic Model Adaptation, to tackle the model instability problem in forecasting the unemployment rate. I find that with a small number of forecasting models we can improve the out-of-sample forecasts of short as well as long horizons by continuously switching forecasting models to the best one in terms of the recent performance of out-of-sample forecasts. Dynamic Model Adaptation produces more accurate forecasts than existing forecasting combination methods. Its predictability is particularly stronger than other methods during economic recessions. The proposed method performs better than the Greenbook forecasts.

The source of improvement is twofold. First, forecasters can use better performing forecasting models relatively quickly by continuously updating the forecasting model to changing economic situations. Second, the persistence in the relative performance of unemployment forecasting models, particularly the top one, enables us to take advantage of the recent best forecasting model for out-of-sample forecasts.

I considered forecasting the unemployment rate in real time as empirical application. Whether the proposed method will work well in forecasting different time-series is definitely a question to be answered. More fundamentally, it is an open question where the persistence of relative performance of forecasting models comes from and whether it is legitimate to use the same model combination strategies for forecasting time-series with different dynamic nature. These can be interesting future research topics.

3.6 Acknowledgements

Chapter 3 is in preparation for submission.

3.7 Tables

Table 3.1: Performance of selected forecasting models (1994–2012)

Root-mean-squared-error(percentage points). Calculated from the monthly forecasts made during 1994–2012 that share a common information set with the historical SPF forecasts. $t+0$ denotes the current quarter at the time of forecasts and $t+n$ denotes n quarters from the current quarter t for $n = 1, 2, 3, 4$. $Y+0$ denotes the current year at the time of the forecast and $Y+1$ denotes next year. The last column shows the regressors (x_t) that are included in the VAR forecasting models. $p_t^1, p_t^{7.+}$: the exit probabilities of those who have been unemployed for 1 month and longer than 6 months, 12 month moving average. f_{1t}^n, f_{2t}^n : the first two factors estimated using principal components from seasonally unadjusted exit probabilities by duration. f_{1t}^{6m}, f_{2t}^{6m} : the first two factors estimated from exit probabilities by duration smoothed using 6 month moving average. f_{1t}^s, f_{2t}^s : the first two factors estimated from exit probabilities by duration seasonally adjusted with X-12-ARIMA in real time. BN(e), BN(f): Barnichon and Nekarda (2012)'a nonlinear steady-state model with expanding and fixed window rolling estimators. BN(ss): the nonlinear steady-state model with constant f_t and s_t.

		Forecast Horizon (quarters)						
Model	t+0	t+1	t+2	t+3	t+4	Y+0	Y+1	x_t
ARIMA (2,0,1)	0.15	0.45	0.70	0.95	1.16	0.32	1.14	-
ARIMA (2,1,0)	0.14	0.43	0.66	0.91	1.12	0.28	1.11	-
VAR1	0.12	0.36	0.57	0.81	1.03	0.25	1.03	$\ln s_t, \ln f_t$
VAR2	0.12	0.35	0.56	0.80	1.02	0.23	1.03	$\ln s_t, \ln f_t,$ $\ln uic_t, d\ln hwi_t$
VAR3	0.12	0.36	0.57	0.82	1.04	0.25	1.04	$\ln uic_t$
VAR4	0.12	0.36	0.58	0.84	1.08	0.25	1.09	$p_t^1, p_t^{7.+},$ $\ln uic_t, d\ln hwi_t$
FAVAR 1	0.12	0.35	0.55	0.79	1.01	0.23	1.02	$f_{1t}^n, f_{2t}^n,$ $\ln uic_t, d\ln hwi_t$
FAVAR 2	0.12	0.35	0.55	0.79	1.01	0.23	1.00	$f_{1t}^{6m}, f_{2t}^{6m},$ $\ln uic_t, d\ln hwi_t$
FAVAR 3	0.12	0.35	0.55	0.80	1.02	0.23	1.02	$f_{1t}^s, f_{2t}^s,$ $\ln uic_t, d\ln hwi_t$
BN(e)	0.12	0.37	0.58	0.86	1.09	0.27	1.10	-
BN(f)	0.11	0.36	0.57	0.85	1.09	0.25	1.10	-
BN(ss)	0.18	0.50	0.69	0.94	1.12	0.35	1.12	-
Median SPF	0.14	0.32	0.49	0.71	0.92	0.19	0.93	-

Table 3.2: Performance of selected forecasting models (1994–2009)

Root-mean-squared-error, (percentage points) Calculated from the monthly forecasts made during 1994–2009 that share a common information set with the historical Greenbook and SPF forecasts.

	Forecast Horizon (quarters)						
Model	t+0	t+1	t+2	t+3	t+4	Y+0	Y+1
(Univariate Model)							
ARIMA (2,0,1)	0.15	0.47	0.74	1.02	1.24	0.33	1.22
ARIMA (2,1,0)	0.14	0.44	0.68	0.94	1.16	0.30	1.15
(VAR Model)							
VAR1	0.12	0.37	0.59	0.87	1.11	0.25	1.11
VAR2	0.12	0.36	0.58	0.85	1.10	0.23	1.09
VAR3	0.12	0.37	0.59	0.87	1.11	0.25	1.11
VAR4	0.12	0.37	0.60	0.88	1.13	0.24	1.11
(FAVAR Model)							
FAVAR1	0.12	0.36	0.59	0.86	1.10	0.24	1.10
FAVAR 2	0.12	0.37	0.60	0.87	1.12	0.24	1.11
FAVAR 3	0.12	0.37	0.60	0.89	1.14	0.25	1.14
(Non-linear Model)							
BN(e)	0.11	0.36	0.59	0.88	1.13	0.24	1.12
BN(f)	0.11	0.35	0.58	0.87	1.13	0.23	1.13
BN(ss)	0.17	0.49	0.69	0.95	1.15	0.34	1.15
(Professional Forecasts)							
Median SPF	0.13	0.32	0.52	0.76	0.98	0.20	1.00
Greenbook	0.18	0.37	0.56	0.78	1.00	0.25	0.99

Table 3.3: Performance of selected forecasting models during recessions

Root-mean-squared-error (percentage points). Calculated from the monthly forecasts made during 2000-2002 and 2006-2009 that share a common information set with the SPF forecasts.

	Forecast Horizon (quarters)						
Model	t+0	t+1	t+2	t+3	t+4	Y+0	Y+1
(Univariate Model)							
ARIMA (2,0,1)	0.20	0.69	1.12	1.55	1.90	0.46	1.85
ARIMA (2,1,0)	0.19	0.63	1.05	1.47	1.83	0.43	1.82
(VAR Model)							
VAR1	0.15	0.49	0.79	1.18	1.51	0.30	1.48
VAR2	0.14	0.48	0.78	1.16	1.48	0.27	1.45
VAR3	0.15	0.48	0.78	1.16	1.48	0.30	1.45
VAR4	0.15	0.52	0.88	1.34	1.75	0.32	1.72
(FAVAR Model)							
FAVAR1	0.14	0.51	0.84	1.25	1.59	0.31	1.59
FAVAR 2	0.15	0.52	0.89	1.33	1.71	0.32	1.68
FAVAR 3	0.15	0.52	0.88	1.31	1.68	0.33	1.66
(Non-linear Model)							
BN(e)	0.14	0.52	0.88	1.34	1.75	0.32	1.72
BN(f)	0.14	0.49	0.85	1.33	1.74	0.30	1.74
BN(ss)	0.19	0.63	0.96	1.38	1.73	0.40	1.73
(Professional Forecasts)							
Median SPF	0.14	0.44	0.75	1.14	1.50	0.27	1.52
Greenbook	0.20	0.49	0.81	1.15	1.50	0.31	1.50

Table 3.4: Probability of the best s-period ahead forecasting model at τ to change at $\tau + 1$ (161 forecasts)

Calculated from the monthly forecasts made during 1992–2012

	Forecast Horizon							
s	1	2	3	4	5	6	7	8
All	0.99	0.96	0.95	0.94	0.90	0.88	0.86	0.87
1998-2001	1.00	0.96	0.90	0.90	0.81	0.79	0.79	0.81
2007-2012	1.00	0.90	0.93	0.89	0.88	0.78	0.74	0.75

Table 3.5: Probability of the best s-period ahead forecasting model being different from the best $s+1$-period ahead forecasting model at time τ (161 forecasts)

Calculated from the monthly forecasts made during 1992–2012

	Forecast Horizon						
s	1	2	3	4	5	6	7
All	0.91	0.85	0.84	0.81	0.74	0.70	0.67
1998-2001	0.96	0.88	0.75	0.71	0.71	0.58	0.58
2007-2012	0.81	0.76	0.67	0.61	0.50	0.46	0.38

Table 3.6: Probability of the change in the best forecasting model (161 forecasts)

Calculated from the monthly forecasts made during 1992–2012

h	Next month			Next horizon		
	All	1998-2001	2007-2012	All	1998-2001	2007-2012
1m	1.00	1.00	1.00	-	-	-
2m	0.98	1.00	0.97	0.88	0.92	0.83
3m	0.95	0.96	0.94	0.81	0.85	0.74
6m	0.95	0.94	0.93	0.81	0.77	0.68
9m	0.95	0.90	0.93	0.82	0.71	0.68
12m	0.88	0.88	0.82	0.74	0.69	0.56
15m	0.90	0.79	0.83	0.75	0.73	0.46
18m	0.86	0.75	0.72	0.66	0.54	0.43
19m	0.87	0.79	0.78	0.65	0.54	0.31
20m	0.87	0.77	0.75	0.68	0.58	0.38
21m	0.86	0.81	0.74	0.64	0.58	0.33
22m	0.86	0.81	0.69	0.61	0.48	0.31
23m	0.85	0.83	0.69	0.62	0.52	0.33
24m	0.85	0.79	0.72	0.62	0.54	0.33

Table 3.7: Relative out-of-sample performance of model m_t^s (161 forecasts)

Ranks out of 161 models. Calculated from the monthly forecasts made during 1992–2012

Model	Forecast Horizon (quarters) t+0	t+1	t+2	t+3	t+4
($s = 1$)					
Average	80	85	76	69	69
Dec.2007-Jun.2009	85	95	88	64	58
($s = 2$)					
Average	77	72	70	71	72
Dec.2007-Jun.2009	51	52	46	54	49
($s = 3$)					
Average	70	71	65	66	65
Dec.2007-Jun.2009	44	37	39	43	38
($s = 4$)					
Average	71	65	62	60	58
Dec.2007-Jun.2009	43	29	28	33	31
($s = 5$)					
Average	70	68	64	61	60
Dec.2007-Jun.2009	36	42	42	39	39
($s = 6$)					
Average	63	62	57	54	53
Dec.2007-Jun.2009	38	26	21	29	28
($s = 7$)					
Average	67	68	60	56	55
Dec.2007-Jun.2009	42	38	35	44	44
($s = 8$)					
Average	66	68	62	60	57
Dec.2007-Jun.2009	52	47	45	48	46

Table 3.8: Probability of the best s-period ahead forecasting model at τ to change at $\tau + 1$ (8 models)

Calculated from the monthly forecasts made during 1992–2012

	Forecast Horizon							
s	1	2	3	4	5	6	7	8
All	0.77	0.75	0.67	0.58	0.53	0.51	0.51	0.43
1998-2001	0.75	0.79	0.79	0.69	0.58	0.52	0.54	0.48
2007-2012	0.67	0.63	0.58	0.49	0.43	0.44	0.49	0.42

Table 3.9: Probability of the best s-period ahead forecasting model being different from the best $s+1$-period ahead forecasting model at time τ (8 models)

Calculated from the monthly forecasts made during 1992–2012

	Forecast Horizon						
s	1	2	3	4	5	6	7
All	0.62	0.56	0.49	0.38	0.30	0.26	0.25
1998-2001	0.52	0.56	0.54	0.38	0.31	0.19	0.25
2007-2012	0.39	0.47	0.42	0.18	0.17	0.28	0.28

Table 3.10: Relative out-of-sample performance of model m_t^s (8 models)

Ranks out of 161 models, Calculated from the monthly forecasts made during 1992–2012

	Forecast Horizon (quarters)					
Model	t+0	t+1	t+2	t+3	t+4	t+5
$(s=1)$						
Average	5.5	5.7	3.7	3.6	4.2	4.5
Dec.2007-Jun.2009	5.9	6.1	3.5	3.1	3.7	4.0
$(s=2)$						
Average	4.7	4.5	4.5	4.3	4.4	4.5
Dec.2007-Jun.2009	4.7	4.5	4.3	4.1	4.5	4.5
$(s=3)$						
Average	4.3	4.5	4.4	4.0	4.0	3.7
Dec.2007-Jun.2009	3.7	3.6	3.4	3.0	3.3	3.2
$(s=4)$						
Average	4.3	4.1	4.0	3.7	3.6	3.3
Dec.2007-Jun.2009	4.0	3.5	3.3	3.2	2.9	2.9
$(s=5)$						
Average	4.5	4.3	4.1	4.0	3.7	3.6
Dec.2007-Jun.2009	3.6	3.8	3.9	3.8	3.6	3.7
$(s=6)$						
Average	4.5	4.1	3.8	3.7	3.5	3.5
Dec.2007-Jun.2009	4.1	3.7	3.1	3.3	3.2	3.1
$(s=7)$						
Average	4.6	4.4	3.8	3.7	3.6	3.5
Dec.2007-Jun.2009	4.5	3.9	3.1	3.2	3.5	3.4
$(s=8)$						
Average	4.1	4.2	4.1	3.9	3.8	3.6
Dec.2007-Jun.2009	4.2	3.7	3.6	3.8	3.7	3.8

Table 3.11: Comparing Unemployment Forecasts: Dynamic Model Adaptation versus other methods, 1994-2012

Root-mean-squared error (percentage points). Calculated from the monthly forecasts made during 1994–2012 that share a common information set with the historical SPF forecasts. $t+0$ denotes the current quarter at the time of forecasts and $t+n$ denotes n quarters from time t (quarter) for $n=1,2,3,4$. Likewise, $Y+0$ denotes the current year at the time of the forecast and $Y+1$ denotes next year. For the least squares (LS) combination and Bayesian combination (BS), the monthly forecasts are calculated over 1999-2012.

		Forecast Horizon (quarters)						
	Method	t+0	t+1	t+2	t+3	t+4	Y+0	Y+1
8 models	DMA(s=3)	0.12	0.36	0.53	0.75	0.94	0.23	0.93
	Average	0.12	0.39	0.59	0.84	1.04	0.26	1.04
	Median	0.12	0.38	0.58	0.84	1.04	0.26	1.05
	LS	0.19	0.43	0.80	1.22	1.59	0.32	1.57
	BS	0.25	0.43	0.67	0.94	1.17	0.23	1.16
	PW	0.12	0.36	0.56	0.81	1.03	0.24	1.03
161 models	DMA(s=6)	0.12	0.34	0.53	0.76	0.98	0.23	0.98
	Average	0.12	0.37	0.58	0.82	1.03	0.24	1.02
	Median	0.12	0.36	0.57	0.81	1.03	0.24	1.02
	LS	0.14	0.49	0.82	1.20	1.53	0.34	1.52
	BS	0.25	0.46	0.70	0.97	1.20	0.26	1.18
	PW	0.13	0.38	0.58	0.82	1.02	0.24	1.01
SPF	Median	0.14	0.32	0.49	0.71	0.92	0.19	0.93

Table 3.12: Comparing Unemployment Forecasts to Greenbook and Survey of Professional Forecasters, 1994-2009

Root-mean-squared error (percentage points). Calculated from the monthly forecasts made during 1994–2009 that share a common information set with the historical Greenbook and SPF forecasts. $t+0$ denotes the current quarter at the time of forecasts and $t+n$ denotes n quarters from time t (quarter) for $n = 1, 2, 3, 4$. Likewise, $Y+0$ denotes the current year at the time of the forecast and $Y+1$ denotes next year. For the least squares (LS) combination and Bayesian combination (BS), the monthly forecasts are calculated over 1999-2009

		Forecast Horizon (quarters)						
	Method	t+0	t+1	t+2	t+3	t+4	Y+0	Y+1
8 models	DMA(s=3)	0.12	0.36	0.55	0.79	0.99	0.22	0.97
	Average	0.12	0.39	0.62	0.89	1.12	0.26	1.12
	Median	0.12	0.38	0.61	0.88	1.12	0.25	1.11
	LS	0.14	0.44	0.90	1.35	1.76	0.35	1.74
	BS	0.27	0.43	0.70	0.98	1.22	0.25	1.20
	PW	0.12	0.37	0.59	0.86	1.10	0.24	1.09
161 models	DMA(s=6)	0.11	0.33	0.55	0.80	1.03	0.23	1.04
	Average	0.12	0.39	0.62	0.89	1.11	0.26	1.11
	Median	0.12	0.37	0.60	0.87	1.11	0.25	1.10
	LS	0.14	0.52	0.90	1.32	1.68	0.38	1.67
	BS	0.25	0.45	0.72	1.00	1.24	0.27	1.22
	PW	0.13	0.39	0.62	0.88	1.10	0.26	1.09
SPF	Median	0.13	0.32	0.52	0.76	0.98	0.20	1.00
Greenbook		0.18	0.37	0.56	0.78	1.00	0.25	0.99

Table 3.13: Comparing Unemployment Forecasts (Economic recessions)

Root-mean-squared error (percentage points). Calculated from the monthly forecasts made during 2000-2002 and 2007-2009 that share a common information set with the SPF forecasts. $t+0$ denotes the current quarter at the time of forecasts and $t+n$ denotes n quarters from time t (quarter) for $n = 1, 2, 3, 4$. Likewise, $Y+0$ denotes the current year at the time of the forecast and $Y+1$ denotes next year. For the least squares (LS) combination and Bayesian combination (BS), the monthly forecasts are conducted during 1999-2012.

		Forecast Horizon (quarters)						
	Method	t+0	t+1	t+2	t+3	t+4	Y+0	Y+1
8 models	DMA(s=3)	0.14	0.48	0.78	1.16	1.49	0.28	1.47
	Average	0.16	0.56	0.94	1.38	1.76	0.36	1.75
	Median	0.15	0.55	0.93	1.37	1.75	0.35	1.74
	LS	0.17	0.58	1.21	1.83	2.39	0.46	2.36
	BS	0.27	0.47	0.85	1.24	1.54	0.29	1.53
	PW	0.15	0.53	0.88	1.29	1.63	0.33	1.60
161 models	DMA(s=6)	0.14	0.48	0.82	1.23	1.60	0.32	1.61
	Average	0.16	0.56	0.94	1.36	1.72	0.36	1.70
	Median	0.16	0.54	0.90	1.34	1.71	0.35	1.68
	LS	0.18	0.69	1.22	1.80	2.28	0.51	2.27
	BS	0.25	0.52	0.88	1.26	1.56	0.32	1.55
	PW	0.17	0.57	0.94	1.35	1.70	0.36	1.67
SPF	Median	0.14	0.44	0.75	1.14	1.50	0.27	1.52
Greenbook		0.20	0.49	0.81	1.15	1.50	0.31	1.50

Table 3.14: Comparing Unemployment Forecasts (Economic booms)

Root-mean-squared error (percentage points). Calculated from the monthly forecasts made during 1994–2009 that share a common information set with the historical Greenbook and SPF forecasts. $t + 0$ denotes the current quarter at the time of forecasts and $t+n$ denotes n quarters from time t (quarter) for $n = 1, 2, 3, 4$. Likewise, $Y+0$ denotes the current year at the time of the forecast and $Y + 1$ denotes next year. For the least squares (LS) combination and Bayesian combination (BS), the monthly forecasts are conducted during 1999-2009.

		Forecast Horizon (quarters)						
	Method	t+0	t+1	t+2	t+3	t+4	Y+0	Y+1
8 models	DMA(s=3)	0.11	0.31	0.39	0.46	0.52	0.23	0.51
	Average	0.10	0.27	0.30	0.34	0.34	0.18	0.36
	Median	0.10	0.27	0.31	0.37	0.39	0.19	0.42
	LS	0.19	0.26	0.25	0.37	0.45	0.14	0.46
	BS	0.24	0.39	0.50	0.65	0.81	0.18	0.78
	PW	0.10	0.26	0.32	0.41	0.49	0.18	0.51
161 models	DMA(s=6)	0.11	0.26	0.29	0.36	0.45	0.12	0.42
	Average	0.10	0.24	0.26	0.29	0.32	0.11	0.30
	Median	0.10	0.24	0.28	0.33	0.38	0.14	0.38
	LS	0.13	0.25	0.27	0.35	0.48	0.10	0.46
	BS	0.25	0.42	0.54	0.68	0.84	0.20	0.81
	PW	0.10	0.25	0.27	0.32	0.36	0.10	0.35
SPF	Median	0.14	0.25	0.28	0.31	0.39	0.10	0.35

3.8 Appendix

Alternative variables of labor market flows for VAR model

Let $p_t^1, p_t^{2.3}, p_t^{4.6}$ and $p_t^{7.+}$ be the exit probability of those who have been unemployed for 1 month, 2-3 months, 4-6 months and longer than 6 months. The exit probabilities are calculated from the following:

$$\begin{aligned} p_t^1 &= 1 - \frac{U_t^2}{U_{t-1}^1} \\ p_t^{2.3} &= 1 - \frac{U_t^3 + U_t^4}{U_{t-1}^2 + U_{t-1}^3} \\ p_t^{4.6} &= 1 - \frac{U_t^5 + U_t^6 + U_t^7}{U_{t-1}^4 + U_{t-1}^5 + U_{t-1}^6} \\ p_t^{7.+} &= 1 - \frac{U_t^{7.+} - U_t^7}{U_{t-1}^{7.+}}, \end{aligned}$$

where U_t^x denotes the number unemployed for x months and $U_t^{7.+}$ is the number unemployed for longer than 26 weeks. The numbers are constructed using CPS micro data. Seasonally unadjusted CPS microdata publicly available at the NBER website (http://www.nber.org/data/cps_basic.html).

References

Ahn, Hie Joo and James D. Hamilton (2014), “Heterogeneity and Unemployment Dynamics,” Working paper, University of California, San Diego.

Aaronson, Daniel, Bhashkar Mazumder, and Shani Schechter (2010). ”What is behind the rise in long-term unemployment,” *Economic Perspectives,* QII:28-51.

Abbring, Jaap H., Gerard J. van den Berg, and Jan C. van Ours (2001). ”Business Cycles and Compositional Variation in U.S. Unemployment,” *Journal of Business and Economic Statistics,* 19(4):436–48.

Abbring, Jaap H., Gerard J. van den Berg, and Jan C. van Ours (2002). ”The Anatomy of Unemployment Dynamics,” *European Economic Review,* 46(10):1785-1824.

Abraham, Katharine G. and Robert Shimer (2002). ”Changes in Unemployment Duration and Labor Force Attachment,” in Alan Krueger and Robert Solow, eds., *The Roaring Nineties,* pp. 367-420. Russell Sage Foundation.

Acemoglu, Daron (1995), “Public Policy in a Model of Long-term Unemployment,” *Economica* 62:161-178.

Aiolfi, M. and C.A. Favero, 2005. ”Model Uncertainty, Thick Modeling and the Predictability of Stock Returns,” Journal of Forecasting, 24(4):233–254.

Aiolfi, Marco & Timmermann, Allan, 2006.”Persistence in forecasting performance and conditional combination strategies,” Journal of Econometrics, Elsevier, vol. 135(1-2), pages 31-53.

Bachmann, Ronald, and Mathias Sinning (2012). ”Decomposing the Ins and Outs of Cyclical Unemployment.” Working paper, Australian National University.

Baker, Michael (1992). ”Unemployment Duration: Compositional Effects and Cyclical Variability,” *American Economic Review*, 82(1):313–21.

Baker, Michael amd Melino, Angelo (2000). ”Duration dependence and non-parametric heterogeneity: A Monte Carlo study,” *Journal of Econometrics*, Elsevier, vol. 96(2), pages 357-393, June.

Barnichon, Regis & Christopher J. Nekarda, 2012. ”The Ins and Outs of Forecasting Unemployment: Using Labor Force Flows to Forecast the Labor Market,” Brookings Papers on Economic Activity, Economic Studies Program, The Brookings Institution, vol. 45(2 (Fall)), pages 83-131.

Barnichon, Regis and Andrew Figura (2013). ”Labor Market Heterogeneity and the Aggregate Matching Function,” Working paper, Universitat Pompeu Fabra.

Bates, J.M. and C.W.J. Granger, 1969. ”The Combination of Forecasts,” Operations Research Quarterly 20:451-468.

Baumeister, Christiane and Gert Peersman, ”Time-Varying Effects of Oil Supply Shocks on the U.S. Economy” American Economic Journal: Macroeconomics, 5(4), October 2013, p. 1-29

Bednarzik, Robert (1983). ”Layoffs and Permanent Job Losses: Workers' Traits and Cyclical Patterns,” *Monthly Labor Review*, 1983-09.

Billio, Monica & Roberto Casarin & Francesco Ravazzolo & Herman K. van Dijk, 2012. ”Combining predictive densities using Bayesian filtering with applications to US economic data,” Working Papers 2012_16, Department of Economics, University of Venice Ca' Foscari.

Billio, Monica & Roberto Casarin & Francesco Ravazzolo & Herman K. van Dijk., 2012. ”Combination schemes for turning point predictions,” The Quarterly Review of Economics and Finance, Elsevier, vol. 52(4), pages 402-412.

Blanchard, Olivier J. and Lawrence H. Summers (1987). ”Hysteresis in Unemployment,” *European Economic Review*, 31(1-2):288–295.

Bleakley, Hoyt, Ann E. Ferris, and Jeffrey C. Fuhrer (1999). ”New Data on Worker Flows During Business Cycle,” *New England Economic Review*, July/August:49–76, 1999.

Botosaru, Irene (2011). ”A Duration Model with Dynamic Unobserved Heterogeneity,” TSE Working Papers, 11-262, Toulouse School of Economics.

Buthler, Richard J. and James B McDonald (1986). ”Trends in Unemployment Duration Data,” *The Review of Economics and Statistics*, 68(4):545–57.

Clements, Michael P. & David F. Hendry, 2004, ”Pooling of forecasts,” Econometrics Journal, Royal Economic Society, vol. 7(1), pages 1-31, 06.

Darby, Michael R., John C. Haltiwanger, and Mark W. Plant (1986). ”The Ins and Outs of Unemployment: The Ins Win,” NBER Working Paper # 1997.

Davis, Steven J., John Haltiwanger and Scott Schuh (1996). ”Small Business and Job Creation: Dissecting the Myth and Reassessing the Facts,” *Small Business Economics*, Springer, vol. 8(4), pages 297-315, August.

Davis, Steven J., R. Jason Faberman, and John Haltiwanger (2006). ”The Flow Approach to Labor Markets: New Data Sources and Micro-Macro Links,” *Journal of Economic Perspectives*, 20(3):3–26.

Davis, Steven J. and Till von Wachter (2011). ”Recessions and the Costs of Job Loss,” *Brookings Papers on Economic Activity*, Economic Studies Program, The Brookings Institution, vol. 43(2 (Fall)), pages 1-72.

Davis, Steven J., R. Jason Faberman and John C. Haltiwanger (2013). ”The Establishment-Level Behavior of Vacancies and Hiring,” The Quarterly Journal of Economics, Oxford University Press, vol. 128(2), pages 581-622.

Den Haan, Wouter J. (2000). “The comovement between output and prices,” *Journal of Monetary Economics* 46:3-30.

Diebold, F.X. and P. Pauly, 1990. ”The Use of Prior Information in Forecast Combination,” International Journal of Forecasting 6:503-508.

Diebold, Francis X. & Pauly, Peter, 1990. ”The use of prior information in forecast combination,” International Journal of Forecasting, Elsevier, vol. 6(4):503-508.

Elliott, Graham & Timmermann, Allan, 2004. ”Optimal forecast combinations under general loss functions and forecast error distributions,” Journal of Econometrics, Elsevier, vol. 122(1):47-79.

Elliott, Graham & Allan Timmermann, 2005. ”Optimal Forecast Combination Under Regime Switching ,” International Economic Review, Department of Economics, University of Pennsylvania and Osaka University Institute of Social and Economic Research Association, vol. 46(4), pages 1081-1102, November.

Elliott, Graham & Gargano, Antonio & Timmermann, Allan, 2013.”Complete subset regressions,” Journal of Econometrics, Elsevier, vol. 177(2), pages 357-373.

Elsby, Michael W., Bart Hobijn, and Aysegul Sahin (2010). ”The Labor Market in the Great Recession,” *Brookings Papers on Economic Activity*, 41(1):1-69.

Elsby, Michael W., Bart Hobijn, and Aysegul Sahin (2013). ”On the Importance of the Participation Margin for Market Fluctuations,” Working Paper Series, 2013-05, Federal Reserve Bank of San Francisco.

Elsby, Michael W. L., Ryan Michaels, and Gary Solon (2009). ”The Ins and Outs of Cyclical Unemployment,” *American Economic Journal: Macroeconomics*, 1(1): 84-110.

Estevao, Marcello and Christopher Smith, ”Skill mismatches and Unemployment in the United States,” January 2013. International Monetary Fund.

Estevao, Marcello and Evridiki Tsounta, ”Has the Great Recession Raised U.S. Structural Unemployment?,” May 2011. International Monetary Fund.

Eriksson, Stefan, and Dan-Olof Rooth (2014). ”Do Employers Use Unemployment as a Sorting Criterion When Hiring? Evidence from a Field Experiment,” *American Economic Review* 104(3): 1014-1039.

Faberman, R. Jason and Kudlyak, Marianna (2014). ”The Intensity of Job Search and Search Duration,” Working Paper 14-12, Federal Reserve Bank of Richmond.

Farber, Henry S. and Robert G. Valletta (2013). ”Do Extended Unemployment Benefits Lengthen Unemployment Spells? Evidence from Recent Cycles in the U.S. Labor Market,” NBER Working Paper No. 19048.

Forni, Mario & Hallin, Marc & Lippi, Marco & Reichlin, Lucrezia, 2005. ”The Generalized Dynamic Factor Model: One-Sided Estimation and Forecasting,” Journal of the American Statistical Association, American Statistical Association, vol. 100, pages 830-840, September.

Forni, Mario & Hallin, Marc & Lippi, Marco & Zaffaroni, Paolo, 2015. ”Dynamic factor models with infinite-dimensional factor spaces: One-sided representations,” Journal of Econometrics, Elsevier, vol. 185(2), pages 359-371.

Fujita, Shigeru and Giuseppe Moscarini (2013). ”Recall and Unemployment,” NBER Working Papers, w19640.

Fujita, Shigeru and Garey Ramey (2006). ”The cyclicality of job loss and hiring,” Working Papers 06-17, Federal Reserve Bank of Philadelphia.

Fujita, Shigeru and Garey Ramey (2009). ”The Cyclicality of Separation and Job Finding Rates,” *International Economic Review*, 50(2):415-430.

Genre, Véronique & Kenny, Geoff & Meyler, Aidan & Timmermann, Allan, 2013. ”Combining expert forecasts: Can anything beat the simple average?,” International Journal of Forecasting, Elsevier, vol. 29(1), pages 108-121.

Hall, Robert E.(2005). ”Job Loss, Job Finding and Unemployment in the U.S. Economy over the Past Fifty Years,” in Mark Gertler and Kenneth Rogoff, eds., *NBER Macroeconomics Annual*, pp. 101-137. Cambridge, MA: MIT Press.

Hall, Robert E.(2014). ”Quantifying the Lasting Harm to the U.S. Economy from the Financial Crisis,” in Jonathan Parker and Michael Woodford, eds., *NBER Macroeconomics Annual.*

Hall, Robert E. and Sam Schulhofer-Wohl (2014). ”Measuring Matching Efficiency with Heterogeneous Jobseekers,” working paper, Stanford University.

Hagedorn, Marcus, Fatih Karahan, Iourii Manovskii and Kurt Mitman (2013). ”Unemployment Benefits and Unemployment in the Great Recession: The Role of Macro Effects,” NBER Working Papers 19499, National Bureau of Economic Research, Inc.

Hamilton, James D. (1988). ”A Neoclassical Model of Unemployment and the Business Cycle,” *Journal of Political Economy*, 96(3):593-617.

Hamilton, James D. (1994a). *Time Series Analysis.* Princeton: Princeton University Press.

Hamilton, James D. (1994b). ”State-Space Models,” in Robert F. Engle and Daniel L. McFadden, eds., *Handbook of Econometrics, Volume IV,* Chapter 50, pp. 3039-3080. Amsterdam: Elsevier.

Hamilton, James D., 2005. ”What's real about the business cycle?,” Review, Federal Reserve Bank of St. Louis, issue Jul, pages 435-452.

Haltiwanger, John and Mark Plant (1987). ”Alternative Measures of Slackness in the Labor Market and their Relationship to Wage and Price Inflation,” in Kevin Lang and Jonathan S. Leonard, eds., *Unemployment and the Structure of Labor Markets.* Oxford, England: Basil Blackwell Ltd.

Heckman, James J. and Burton Singer (1984). ”Econometric duration analysis,” *Journal of Econometrics*, Elsevier, vol. 24(1-2), pages 63-132.

Heckman, James J. and Burton Singer (1984). ”The Identifiability of the Proportional Hazard Model,” *Review of Economic Studies*, 51(2):231-41.

Herz, Benedikt and Thijs van Rens, ”Structural Unemployment,” July 2011. Universitat Pompeu Fabra.

Hornstein, Andreas (2012). ”Accounting for Unemployment: the Long and Short of it,” FRB Richmond Working Paper, No.12-07, Federeal Reserve Bank of Richmond.

Jaimovich, Nir and Henry E. Siu.(2012) ”The Trend is the Cycle: Job Polarization and Jobless Recoveries,” NBER Working Papers, w18334.

Jordà, Òscar, 2005. ”Estimation and Inference of Impulse Responses by Local Projections,” American Economic Review, American Economic Association, vol. 95(1), pages 161-182, March.

Kaitz, Hyman B. (1970). ”Analyzing the Length of Spells of Unemployment,” *Monthly Labor Review*, 93(11):11–20.

Katz, Lawrence F. (1986). ”Layoffs, Recall and the Duration of Unemployment,” NBER Working Papers, w1825.

Katz, Lawrence F. and Bruce D. Meyer (1990a). ”The Impact of the Potential Duration of Unemployment Benefits on the Duration of Unemployment,” *Journal of Public Economics* 41:45-72.

Katz, Lawrence F. and Bruce D. Meyer (1990b). ”Unemployment Insurance, Recall Expectations, and Unemployment Outcomes,” *The Quarterly Journal of Economics*, 105(4):973-1002.

Kerchkoffs, C., C. De Neubourg and F. (1994). ”The Determinants of Unemployment and Job Search Duration in the Netherlands,” *De Economist* 142(1):21-42.

Koop, G. and Korobilis, D., 2012. “Forecasting Inflation Using Dynamic Model Averaging”, International Economic Review, 53, pp. 867-886.

Kroft, Kory, Fabian Lange, and Matthew J. Notowidigdo (2012). ”Duration Dependence and Labor Market Conditions: Theory and Evidence from a Field Experiment”, NBER Working Papers, w18387.

Kroft, Kory, Fabian Lange, and Matthew J. Notowidigdo, and Lawrence F. Katz (2013). ”Long-Term Unemployment and the Great Recession: The Role of Composition, Duration Dependence, and Non-Participation,” Working Paper, University of Chicago.

Krueger, Alan B. and Andreas I. Mueller (2010). ”Job search and unemployment insurance: New evidence from time use data,” *Journal of Public Economics*, Elsevier, vol. 94(3-4), pages 298-307, April.

Krueger, Alan B. and Andreas I. Mueller (2011). ”Job Search and Job Finding in a Period of Mass Unemployment: Evidence from High-Frequency Longitudinal Data,” Working Papers 1295, Princeton University, Department of Economics, Center for Economic Policy Studies..

Krueger, Alan B. and Andreas I. Mueller (2012). ”The Lot Of The Unemployed: A Time Use Perspective,” *Journal of the European Economic Association, European Economic Association*, vol. 10(4), pages 765-794, 08.

Krueger, Alan B., Judd Carmer and David Cho (2014), ”Are the Long-Term Unemployed on the Margins of the Labor Market?” *Brookings Panel on Economic Activity,* March 20–21, 2014

Lazear, Edward P. and James R. Spletzer (2012). ”Hiring, Churn and the Business Cycle”, *American Economic Review*, 102(3):575-79.

Ljungqvist, Lars and Thomas J. Sargent (1998). ”The European Unemployment Dilemma,” *Journal of Political Economy*, University of Chicago Press, vol. 106(3), pages 514-550, June.

Marcellino, Massimiliano & Stock, James H. & Watson, Mark W., 2006. ”A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series,” Journal of Econometrics, Elsevier, vol. 135(1-2), pages 499-526.

Meyer, Bruce D, 1990. ”Unemployment Insurance and Unemployment Spells,” *Econometrica*, Econometric Society, Econometric Society, vol. 58(4), pages 757-82, July.

Montgomery, Alan R., & Victor Zarnowitz & Ruey S. Tsay & George C. Tiao. 1998. ”Forecasting the U.S. Unemployment Rate,” Journal of the American Statistical Association, 1998, 93(442), pp. 478.

Oyer, Paul. 2006. ”Initial Labor Market Conditions and Long-Term Outcomes for Economists.” Journal of Economic Perspectives, 20(3): 143-160.

Pesaran, M. Hashem & Pick, Andreas & Timmermann, Allan, 2011. ”Variable selection, estimation and inference for multi-period forecasting problems,” Journal of Econometrics, Elsevier, vol. 164(1), pages 173-187, September.

Rothstein, Jesse (2011). ”Unemployment Insurance and Job Search in the Great Recession,” *Brookings Papers on Economic Activity* 43(2):143-213.

Sahin, Aysegul, Joseph Song, Giorgio Topa, and Giovanni L. Violante, ”Mismatch Unemployment,” Working Paper 18265, National Bureau of Economic Research, August 2012.

Schwarz, Gideon (1978), “Estimating the Dimension of a Model,” *Annals of Statistics* 6:461-464.

Sider, Hal (1985), ”Unemployment Duration and Incidence: 1968-1982,” *American Economic Review*, 75(3):461-72.

Shimer, Robert (2012). ”Reassessing the Ins and Outs of Unemployment,” *Review of Economic Dynamics*, 15(2):127-148.

Stock, James H & Watson, Mark W, 1996. ”Evidence on Structural Instability in Macroeconomic Time Series Relations,” Journal of Business & Economic Statistics, American Statistical Association, vol. 14(1), pages 11-30, January.

Stock J.H. & Watson M.W., 2002. ”Forecasting Using Principal Components From a Large Number of Predictors,” Journal of the American Statistical Association, American Statistical Association, vol. 97, pages 1167-1179, December.

Stock, James H & Watson, Mark W, 2004. ”Combination forecasts of output growth in a seven-country data set,” Journal of Forecasting, John Wiley & Sons, Ltd., vol. 23(6), pages 405-430.

Stock, James H. & Watson, Mark W., 2014. ”Estimating turning points using large data sets,” Journal of Econometrics, Elsevier, vol. 178(P2), pages 368-381.

Swanson, Norman R & White, Halbert, 1995. ”A Model-Selection Approach to Assessing the Information in the Term Structure Using Linear Models and Artificial Neural Networks,” Journal of Business & Economic Statistics, American Statistical Association, vol. 13(3), pages 265-75, July.

Valletta, Rob (2013). ”Long-term Unemployment: What do we know,” *FRBSF Economic Letters*, 2013-03. Federal Reserve Bank of San Francisco.

Van den Berg, Gerard J. (2001). ”Duration models: specification, identification and multiple durations,” *Handbook of Econometrics*, edition 1, volume 5, chapter 55, pages 3381-3460, Elsevier.

Van den Berg, Gerald J., and Jan C van Ours (1994). ”Unemployment Dynamics and Duration Dependence in France, the Netherlands and the United Kingdom,” *Economic Journal* 104:432-443.

Van den Berg, Gerald J., and Jan C van Ours (1996). ”Unemployment Dynamics and Duration Dependence,” *Journal of Labor Economics*, 14(1):100-125.

Van den Berg, Gerard J., and van der Klaauw, Bas, (2001). ”Combining micro and macro unemployment duration data,” *Journal of Econometrics*, Elsevier, vol. 102(2), pages 271-309, June.

White, Halbert (1982). ”Maximum Likelihood Estimation of Misspecified Models,” *Econometrica* 50:1-25.

Whittaker, Julie M. and Katelin P. Isaacs (2014). ”Unemployment Insurance: Programs and Benefits,”*Congressional Research Service Report for Congress*, 7-5700.