Why Has the US Economy Recovered So Consistently from Every Recession in the Past 70 Years? *

Robert E. Hall
Hoover Institution and Department of Economics Stanford University
rehall@stanford.edu; stanford.edu/~rehall

Marianna Kudlyak
Federal Reserve Bank of San Francisco and CEPR
marianna.kudlyak@sf.frb.org; sites.google.com/site/mariannakudlyak/

April 4, 2021

Abstract

A remarkable fact about the historical US business cycle is that, after unemployment reached its peak in a recession, and a recovery begins, the annual reduction in the unemployment rate is stable at around one tenth of the current level of unemployment. For example, when the unemployment rate was 7 percent at the beginning of a year, the unemployment rate fell by 0.7 percentage points during the year. The economy seems to have an irresistible force toward restoring full employment. There was high variation in monetary and fiscal policy, and in productivity and labor-force growth during the recoveries, but little variation in the rate of decline of unemployment. We show that the evolution of the labor market involves more than the direct effect of persistent unemployment of job-losers from the recession shock—unemployment during the recovery is elevated for people who did not lose jobs during the recession. We explore models of the labor market’s self-recovery that imply gradual working off of unemployment following a recession shock. We emphasize the feedback from high unemployment to the forces driving job creation. These models also explain why the recovery of market-wide unemployment is so much slower than the rate at which individual unemployed workers find new jobs. The reasons include the fact that the path that individual job-losers follow back to stable employment often includes several brief interim jobs.

JEL: E32, J63, J64.

Keywords: Business cycle, Recovery, Unemployment, Recession

*Prepared for the 2021 NBER Macro Annual, Hall’s research was supported by the Hoover Institution. The opinions expressed are those of the authors and do not reflect those of the Federal Reserve Bank of San Francisco, the Federal Reserve System, or the National Bureau of Economic Research.
Contents

1 Introduction 4

2 Uniform Unemployment Recovery across Recessions 7
   2.1 Inexorable recoveries ............................. 7
   2.2 Estimation methods .............................. 8
      2.2.1 Estimation based on chronologies ................. 9
      2.2.2 Estimation using the hidden Markov approach .......... 10
   2.3 Estimates of the unemployment recovery rate ............... 10
   2.4 Related literature on the statistical properties of unemployment .... 13

3 Job Loss in Recessions 14
   3.1 Layoffs in JOLTS .................................. 15
   3.2 Mass layoffs ...................................... 15
   3.3 Job destruction .................................... 17
   3.4 Displaced workers .................................. 17
   3.5 Comparison of measures of the spike of job loss in a recession .... 19
   3.6 Initial unemployment insurance claims ...................... 21
   3.7 Flow of new permanent layoffs in the Current Population Survey .... 22

4 The Direct Channel from Job Loss to Subsequent Lingering Unemployment 22
   4.1 Information about the subsequent role in unemployment from job displacement 24
   4.2 Application to other measures of job loss .......................... 26
   4.3 Conclusions about the relation between the magnitude of the increase in unemployment following a recession shock and the measures of job loss ...... 26
   4.4 Excess unemployment of new entrants .......................... 28

5 Effective Exit Rate from Unemployment 28
   5.1 Defining and measuring the effective exit rate ................. 28
   5.2 Implications of low effective job-finding rates .................. 30

6 The DMP Model 30
   6.1 Potential driving forces of the DMP model ..................... 33
   6.2 Path of unemployment following a recession in the basic DMP model .... 35
   6.3 Path of unemployment in a model with low effective unemployment exit rate 37
   6.4 Variation of the driving forces over time ..................... 39
1 Introduction

We study data from the labor market during recoveries from past recessions, excluding the nascent recovery from the pandemic recession of 2020. We find that the recovery phase of the US business cycle has invariably been slow but irresistible. Each year, unemployment falls by one tenth of the level of unemployment—say by 0.7 percentage points in a year with 7 percent unemployment. We document this regularity within the two main statistical approaches to business-cycle analysis and measurement — construction of a chronology of turning points and estimation of a Markov regime-switching model.

We note that a well-documented property of the unemployment rate is that unemployment rises rapidly in response to a significant aggregate adverse shock and then gradually recovers to a level of 3 to 5 percent of the labor force. Like fuel prices, unemployment rises like a rocket and falls like a feather. This property was most recently confirmed by Dupraz, Nakamura and Steinsson (2019), with many cites to the earlier literature. We add to another, perhaps surprising, regularity that the speed of the recovery across the last ten recoveries has been remarkably uniform.

We focus on recoveries. Our measurement starts in an economy that has just been hit by an adverse shock that triggered a recession. This paper recognizes that the shocks that propel unemployment sharply upward are heterogeneous. The major recession that began in 1981 is generally viewed as the result of a sharp monetary contraction, while the major recession that began at the end of 2007 got much of its strength from the financial crisis of September 2008. Historical recoveries have been much more homogeneous.

We demonstrate that a steady decline in unemployment following a recession shock occurred reliably in the recoveries in the 70 years we include—years when unemployment has been measured reliably and consistently over time in the Current Population Survey. We point out the puzzle of slow decline of unemployment. Cole and Rogerson (1999) called attention to the puzzle—unemployment declines much more slowly than the measured exit rates from unemployment among individuals would seem to indicate.

We then ask, what accounts for the economy’s consistent, reliable record in recovering from adverse shocks? Our thesis is that the economy has a powerful tendency to recover from serious adverse shocks, but recovery takes time. We show that other potential driving forces do not appear to generate deviations from the downward path of unemployment during recoveries. These forces include monetary and fiscal initiatives and variations in productivity and labor-force growth.

Our results suggest a rather different view of cyclical fluctuations from most current thinking. In our view, a variety of serious shocks can trigger a recession. The financial
crisis of 2008-9 is illustrative. Millions of workers lost jobs they had held for some time and where they had accumulated job-specific human capital. In the following recovery, rebuilding gradually and inevitably occurred. Until the recovery became complete, employment was depressed because the recovery process involves higher unemployment. This unemployment is effectively a reduction in labor supply. In contrast, the dominant current view emphasizes slow recoveries in labor demand and treats the elevated unemployment as the result of a shortfall of demand relative to supply in the labor market.

We emphasize a negative feedback from high unemployment to the job finding rate as a key mechanism behind the slow unemployment recoveries. Our discussion of unemployment is within the framework of Diamond, Mortensen, and Pissarides (DMP). Their model has a well-known but counter-intuitive property—it lacks feedback from unemployment to labor-market tightness. When an adverse shock creates a high volume of unemployment, but the shock subsides so the determinants of tightness return to normal, the legacy of unemployment has no discouraging effect on tightness. Jobs are just as easy to find with unemployment at 10 percent as they are when unemployment is 4 percent. Much of this paper is devoted to studying modifications of the DMP model to alter this property, some of which are based on the ideas in Fujita and Ramey (2007). In the modified model, unemployment is much more persistent, because jobs are hard to find when unemployment is high.

Our view of the labor market has points in common with Pries (2004) and can be seen as responding to the challenge of Cole and Rogerson (1999) to explain why aggregate unemployment recovers much more slowly than does an individual spell of unemployment.

We proceed in the following steps:

First, we show that the recoveries from the past recessions have been reliable and uniform. Second, we study the job loss that occurs when a crisis launches a recession. A spike in job loss is visible in a variety of data sources, measuring layoffs, job destruction, displacement, and unemployment insurance claims. But the spike in job loss in these measures is far smaller than the total increase in the number of unemployed workers.

Third, we ask whether the volume of job losers and their likely speed of finding long-term replacement jobs is enough to explain the long-lasting bulge of total unemployment that is only gradually worked off during even a long recovery like the one that ended in early 2020. We conclude, from data on displaced workers collected every two years in the Current Population Survey, that the number of workers displaced even in the severe recession starting in 2007 was not enough to explain the volume of excess unemployment present in the US economy during the period from 2009 through 2014. Something happened in the
labor market during that period that caused elevated unemployment among workers who were not displaced around 2009. Unemployment proved to be infectious.

Fourth, we examine the puzzle of low recovery speed in the framework of the DMP model. We calculate the effective exit rate from unemployment, which is lower than the exit rate for individuals from one month to the next. Those individual exits are frequently temporary departures from the labor force or short-term jobs, and are then followed by additional spells of unemployment, as described in Hall and Kudlyak (2019). In the DMP equilibrium, the unemployment rate falls more gradually, so it accounts for some of the puzzle of low recovery speed.

We study models that explain slow but sure recoveries through feedback from the level of unemployment to the job-finding rate. We consider the feedback to various driving forces. For example, when unemployment spikes, employers’ costs of recruiting rise. According to standard DMP principles, higher costs of filling vacancies discourage job creation and raises equilibrium unemployment. This model generates a generally slow decline of unemployment during a recovery. We review the extensive literature on the feedback from unemployment to the recovery process that rebuilds employment lost in the earlier crisis.

We describe a wide range of mechanisms that participate in the gradual reduction of unemployment during a recovery. Among the most important of these are

1. A gradual return to the normal mix of unemployment, away from the disproportionate role of hard-to-re-employ workers in the aftermath of a recession
2. Slow but reliable decline in labor-market churn that occurs following a recession
3. Decline back to normal from strict credit standards put in place during a recession
4. Gradual relaxation of adjustment costs for job creation
5. Congestion effects impeding recruiting efforts when unemployment is high

Fifth and finally, we show that recoveries have had a wide variety of accompanying movements of policy instruments and arguably exogenous aggregate driving forces. Unemployment declines smoothly. Movements of possible driving forces, including fiscal and monetary policy, productivity growth, labor-force growth, and variations in stock-market discounts, are irregular.

We conclude that the economy includes a strong internal force toward recovery that operates apart from policy instruments and apart from productivity growth and financial developments revealed in the stock market. After a negative shock, employers gradually find it profitable to hire more aggressively. Unemployment falls as the unemployed are put back
to work. Rather than a pull from expansionary policy, the growth in employment arises from a push toward lower unemployment.

We find that only a small part of the reduction in unemployment following a recession takes the form of the first jobs found by the workers who lost jobs as a direct result of the recession. Additional spells of joblessness occur within that group, and extra unemployment occurs among workers who were not immediate victims, such as people entering the labor market for the first time, after the recession. This induced unemployment gradually returns to normal in the recovery.

This paper is about recoveries from January 1948 through February 2020 and not at all about the recovery that appears to be under way as we write. The pandemic that influenced the labor market starting in late March 2020 created an unprecedented increase in the non-working population—unemployed and out of the labor force. But many of these people had good prospects of recall to their earlier jobs or successful re-entry to the labor force once the pandemic ended. Analysis of unemployment and labor-force participation in the pandemic differs fundamentally from analysis of the periods we consider here.

2 Uniform Unemployment Recovery across Recessions

We study US business-cycle recoveries over the past 70 years. We focus on the unemployment rate. Our key results are (1) the recovery process takes place reliably, regardless of the nature of the shock that causes the preceding economic contraction, and (2) the recovery process is similar in all of the ten past recoveries—unemployment falls by about 0.1 log points per year. The main insights in this section come from the econometric analysis in Hall and Kudlyak (2020).

2.1 Inexorable recoveries

We adopt Romer and Romer’s (2019) conclusion that the unemployment rate, or a measure derived from the unemployment data from the Current Population Survey, is the best available measure of the business cycle.

Figure 1 displays the log of the unemployment rate during the 10 recoveries since 1948, with the recession spells of sharply rising unemployment left blank. Throughout the paper we are excluding the incomplete recovery from the pandemic recession that started in 2020.

The key fact about recoveries is apparent in the figure: Unemployment declines smoothly but slowly throughout most recoveries most of the time, at close to the same proportional rate. In the log plot, the recoveries appear as impressively close to straight lines.
2.2 Estimation methods

We now turn to estimating the rate of decline of unemployment during recoveries. We treat log-unemployment as the sum of a latent trend component and a latent stationary component capturing survey sampling errors and other deviations from the trend. In a crisis, the trend is fairly sharply upward. During a recovery, the trend is modestly downward. Our objective is to measure the central tendency and dispersion of the rate of decline of the latent systematic component of the monthly change of log-unemployment rate during recoveries.

We consider the general class of models

\[ f(u_t) = x_t + \epsilon_t, \]

where \( f(\cdot) \) is a monotonic transformation, \( x_t \) is the systematic trend component capturing the business cycle, and \( \epsilon_t \) is the random unsystematic component, taken to be uncorrelated with \( x_t \). The model for \( x_t \) describes the random arrival of peaks and troughs with linear paths connecting these turning points. We use the log transformation because it fits the data quite a bit better than a model where the level of unemployment is the dependent variable. In Figure 1 and later in Figure 2, the implied linearity of the recovery paths of log unemployment is quite clearly confirmed.

The literature has focused on two general classes of specifications for the systematic component. One is chronology-based and proceeds by assigning turning points—dates when recessions end and recoveries begin and dates when recoveries end and recessions begin.
Chronologies are available from published sources, notably the National Bureau of Economic Research, which identifies monthly dates of turning points in a latent measure called economic activity. They can be created for a particular time series, such as the unemployment rate, as an exercise in human pattern recognition. And chronologies can be created by algorithms, such as the one described in Dupraz et al. (2019).

Given a chronology, one can approximate the systematic component $x_t$ by interpolating between the turning points and measuring the noise $\epsilon_t$ as the residual between $\log u_t$ and $x_t$. The systematic trend component $x_t$ is a smooth function of $t$. We take it to be a straight line between the turning points of the series, so $x_t$ has equal increments over time, between the turning points. Overall, the trend component is a linear spline.

The other class of models focuses on *regime switching*, where the systematic component modeled as a statistical time series that obeys one model in contractions and another in recessions. Hamilton (1989) launched the econometric literature on Markov-switching models in this class.

A somewhat related literature focuses on the asymmetry of recoveries—long and gentle—and contractions or recessions—brief and fierce. This research generally focuses on the third moment of unemployment levels or changes as the measure of asymmetry. We will mention some of the results from this literature but do not believe that it competes in statistical power with the methods we use for the issue we investigate—the uniformity of the slope of recoveries in unemployment.

### 2.2.1 Estimation based on chronologies

We consider three monthly business-cycle chronologies:

1. NBER: The chronology maintained by the National Bureau of Economic Research identifying turning points in economic activity, as described in detail at NBER.org.

2. DNS: The chronology produced by the DNS algorithm based on US unemployment from January 1948 through February 2020, with size parameter 1.5.

3. HK: The chronology produced by the authors based on observed business cycle peaks and troughs, using the same data as DNS.

In general, DNS and HK are similar to one another and differ from the NBER one (see Hall and Kudlyak (2020) for the chronologies). The reason is that DNS and HK are chronologies for unemployment alone, while NBER is a chronology for latent economic activity. For the dates in the table starting with January 1980, NBER.org has published explanations of the various indicators that form the basis for the determination of the dates.
We measure the recovery rate as the average annual decline in log-unemployment over the entire set of recoveries and over individual recoveries. Although the mean is the natural estimator irrespective of the time-series process of the shocks $\epsilon$, the process does matter for the standard error of the sample mean.

Given a recovery running from an initial high point of unemployment, which we number as zero, to the following low point, which we number as $T$, our model for a single recovery is

$$12(\log u_T - \log u_0) = -\beta T + \epsilon_T.$$  \hspace{1cm} (2)

We include the 12 so that the recovery rate $\beta$ is in log points per year. We use the estimator

$$\hat{\beta} = -\frac{12(\log u_T - \log u_0)}{T}.$$  \hspace{1cm} (3)

We use a modified bootstrap procedure, described in our other paper, to approximate the sampling distribution of $\hat{\beta}$.

**2.2.2 Estimation using the hidden Markov approach**

Our second approach to modeling business cycles posits the same basic cyclical structure,

$$\log u_t = x_t + \epsilon_t,$$  \hspace{1cm} (4)

where $x_t$ is unobserved, but hypothesized to be piecewise linear, with switching between positively and negatively sloped segments at random, according to a Markov process. Under the assumption that the disturbance is a random walk, $\Delta \epsilon_t = \eta_t$, with $\eta_t$ being white noise, the model becomes

$$\Delta \log u_t = \beta_i + \eta_t.$$  \hspace{1cm} (5)

The monthly increment, $\beta_i$, $i \in \{1, 2\}$, shifts back and forth between $i = 1$ for recessions and $i = 2$ for recoveries. We focus on $\beta_2$, the log-decline in unemployment during recoveries. James Hamilton pioneered the econometric analysis of this class of models. He derived the likelihood function in a computationally convenient form (Hamilton (1989)). Additional details about the estimation are provided in Hall and Kudlyak (2020).

**2.3 Estimates of the unemployment recovery rate**

Table 1 shows the statistical results for both approaches. The upper panel displays results for the entire sample, starting in 1949, and the lower panel displays the same statistics for the post-1959 sample. It shows the estimates of the two key statistics in this study: the annual recovery rate in log points, $\beta$, and the standard deviation of $\beta_r$ across the recoveries, indexed by $r$. The upper panel covers the entire sample period, from October 1949 through
February 2020. The left three columns show the recovery rate $\beta$ in log points per year using the chronology approach, along with the standard deviation of the recovery rate, both with bootstrap standard errors.

For the NBER chronology with the full sample, the decline rate pooled across recoveries is estimated as 0.087. Recovery rates for the DNS and HK chronologies are similar to each other and are well above the NBER level, at 0.129 and 0.132. The DNS and HK chronologies, constructed from unemployment alone, are more successful at capturing the movements of unemployment during recoveries because they are better synchronized with the actual movements. Of course, DNS and HK would be correspondingly poorer at tracking economic activity, the concept behind the NBER chronology.

We illustrate the interpretation of the annual decline figures in the table with an example from the recovery rates based on the NBER chronology. Consider the situation just after a severe recession, with the unemployment rate starting at 10 percent. The expected unem-

Table 1: Statistical Results

<table>
<thead>
<tr>
<th></th>
<th>NBER</th>
<th>Dupraz-Nakamura-Steinsson</th>
<th>Hall-Kudlyak</th>
<th>Hidden Markov</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Full sample</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Annual recovery rate, log points</td>
<td>0.087</td>
<td>0.132</td>
<td>0.129</td>
<td>0.066</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.017)</td>
<td>(0.017)</td>
<td>(0.042)</td>
<td></td>
</tr>
<tr>
<td>Bootstrap standard error</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Information matrix standard error</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard deviation of recovery rate across recoveries</td>
<td>0.076</td>
<td>0.084</td>
<td>0.084</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.119)</td>
<td>(0.117)</td>
<td>(0.115)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bootstrap standard error</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference between HK estimate and hidden-Markov estimate</td>
<td></td>
<td></td>
<td></td>
<td>0.066</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.015)</td>
</tr>
<tr>
<td><strong>After 1959</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Annual recovery rate, log points</td>
<td>0.067</td>
<td>0.106</td>
<td>0.103</td>
<td>0.070</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.039)</td>
<td></td>
</tr>
<tr>
<td>Bootstrap standard error</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Information matrix standard error</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.014)</td>
</tr>
<tr>
<td>Standard deviation of recovery rate across recoveries</td>
<td>0.025</td>
<td>0.011</td>
<td>0.016</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.038)</td>
<td>(0.039)</td>
<td>(0.038)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bootstrap standard error</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference between HK estimate and hidden-Markov estimate</td>
<td></td>
<td></td>
<td></td>
<td>0.036</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.012)</td>
</tr>
</tbody>
</table>
ployment rate a year later is $10 \exp(-0.087) = 9.2$ percent. With the higher recovery rate based on DNS, the rate a year later would be $8.8$ percent, essentially the same as with the HK chronology. According to the DNS rate, starting from 6 percent, the unemployment rate a year later would be $5.3$ percent.

The recovery rates for the sample running from May 1961 through February 2020, shown in the lower panel of Table 1, are somewhat lower, because, as shown in Figure 3, the first three recoveries in the full sample had substantially higher recovery rates than did any of the later recoveries.

The column of Table 1 labeled *Hidden Markov* shows results from estimating the hidden Markov model using the full sample and the sample starting in 1961. For the full sample, the estimated recovery rate is $0.066$, well below the earlier results for the chronology-based estimates, especially in the case of DNS and HK. Estimates for the post-1960 sample are closer, at $0.070$ for hidden-Markov and $0.103$ for DNS. The NBER estimate concurs with the hidden-Markov one, but we have good reason to believe that it is downward biased relative to estimates that are based only on the unemployment rate, which include hidden-Markov.

In the lower right-hand corner of both panels of Table 1, we report the difference between the HK variant of the chronology-based estimate of the pooled recovery rate and the hidden-Markov-based estimate. Below each estimate, we give the bootstrap standard error of the difference, based on the difference in 500 replications of the bootstrap simulations. In both cases, we reject the hypothesis that the difference arises from sampling variation alone. The rejection is stronger for the full sample than for the later sample.

One reason for the disagreement between the two estimators is that the theory of the application of the hidden-Markov setup to our problem requires the assumption that the disturbance is a random walk, whereas it is actually an AR(1) process with coefficient somewhat less than one. The disagreement between the information-matrix estimate of the hidden-Markov standard error and the bootstrap may also result from this conflict.

The chronology-based estimator can be considered an application of Bayesian thinking, in that it imposes prior beliefs about the process. The posterior may involve a higher implied value of the recovery rate because the prior belief pushes the likelihood in that direction.

The two statistical approaches agree that the basic structure that they share is supported in the data. It really is true and not the result of an optical illusion that unemployment rises rapidly and then declines slowly and uniformly over the business cycle. Figure 2 shows how well the fitted value from the HK estimation with different recovery rates for each recovery (those shown in Figure 3) fits the data. The $R^2$ of the fit is $0.91$.

Figure 3 shows the results of estimating each of the 10 recoveries in the HK chronology separately. The separate rates from 1961 onward cluster close to $0.10$. Over that 60-year
period, with 7 recessions and recoveries, some mild and two quite severe, the recovery rates are remarkably similar. Table 1 summarizes this finding in terms of the standard deviations across either the 10 recoveries of the full sample or the 7 recoveries of the later sample.

The standard deviations across the individual recoveries over the full sample, shown in the upper panel in Table 1, are all around 0.08 with standard errors around 0.12. As Figure 3 shows, the first three recoveries have rather higher recovery rates than the later seven. The hypothesis of equality, that is, a standard deviation of zero, is easily not rejected, but the confidence interval is fairly wide, for all three chronologies. On the other hand, the standard deviations of the seven later recoveries have point estimates of 0.025 or under confirming the visual evidence of the figure. And the confidence intervals, constructed with standard errors of 0.39 or 0.38, are tight. These results nail down the primary thesis of our study—the uniformity of recovery rates over the past 60 years.

2.4 Related literature on the statistical properties of unemployment

The asymmetry in unemployment dynamics, where unemployment rises faster in contractions than it falls during expansions, has been found in a long line of research. Neftci (1984) used a finite state Markov process and found that jumps are sudden and declines are slow. Sichel (1993), Andolfatto (1997), Kim and Nelson (1999), and Sinclair (2009) confirmed that finding. Dupraz et al. (2019) find that the increase in unemployment during a contraction
forecasts the amplitude of the subsequent expansion one-for-one, while the fall in unemployment during an expansion has no explanatory power for the size of the next contraction. They develop a search model of the business cycles in which the source of asymmetry is downward nominal wage rigidity. Ferraro (2017) also finds that recoveries are less dramatic than recessions. Following much of the literature on this subject, the paper examines the asymmetry of the distribution of log differences, specifically, the third moment of that distribution. The paper is intended to demonstrate the asymmetry. By contrast, we focus mainly on the uniformity of the decline in unemployment in recoveries. We work within the statistical literature of regime changes, while the third-moment literature works within single-regime statistics.

3 Job Loss in Recessions

In this section, we examine the job loss in recessions. Data on layoffs, job destruction, and long-term worker displacement show the substantial but short-lived spikes of job loss in recessions. The worker-level data from the CPS and data on the initial unemployment insurance claims show the initial substantial spike and a subsequent lingering of the elevated job loss.

We consider a number of measures of job loss:

- Layoffs, the flow of workers whose jobs ended at the initiative of employers.
• Job destruction, the amount of employment decline among establishments with shrinking employment.

• Worker displacement, job loss among workers with at least three years of tenure at the lost job.

• Unemployment insurance claims.

3.1 Layoffs in JOLTS

Figure 4 shows data on layoffs from the Job Openings and Labor Turnover Survey. A layoff occurs when an employer terminates a worker without prejudice, typically because continuing employment has become unprofitable. Most layoffs occur without any definite promise to rehire, but explicitly temporary layoffs are an important part of layoffs. On average, 20 million workers lose their jobs each year in normal times. A substantial but short-lived burst of above-normal layoffs occurred soon after the financial crisis in the fall of 2008.

3.2 Mass layoffs

Mass layoff occurs when a relatively large number of firm’s employees lose jobs. Such events often involve high-tenured workers who tend to suffer prolonged periods of joblessness following job loss (Jacobson, LaLonde and Sullivan (1993), Davis and von Wachter (2011)).
Figure 5: Extended Mass Layoffs, in Thousands of Initial Claimants per Year

Note: Data from the Mass Layoffs Statistics program from the BLS.

The Mass Layoffs Statistics program from the BLS tracks the effects of major job cutbacks using data from state unemployment insurance databases. A mass layoff is defined as 50 or more initial claims for unemployment insurance benefits being filed against an employer during a 5-week period. These employers are contacted by the state agency to determine whether the separations lasted more than 30 days. Such events are termed extended mass layoffs. The BLS obtains information on the total number of workers separated during the extended mass layoffs, including the workers who do not file for unemployment insurance, and the reasons for these separations according to the employer. These layoffs involve both people subject to recall and those who are terminated. The program ran from 1995 to the first quarter of 2013.

Figure 5 shows the number of initial claimants from extended mass layoffs. The number hovers around a million in normal times but spikes during recessions. A decline in business demand and financial difficulties are the main reasons cited behind the spikes. In 2009, extended mass layoffs spiked to 2.4 million.

Another source of data on mass layoffs is the Worker Adjustment and Retraining Notification Act (WARN), which requires employers to provide notice 60 days in advance of covered plant closings, covered mass layoffs, or sale of business that result in an employment loss. Employers are covered by WARN if they have 100 or more employees, not counting
employees who have worked less than 6 months in the last 12 months and not counting employees who work an average of less than 20 hours a week. The term employment loss means (1) an employment termination, other than a discharge for cause, voluntary departure, or retirement; (2) a layoff exceeding 6 months; or (3) a reduction in an employee’s hours of work of more than 50 percent in each month of any 6-month period. A plant closing occurs if an employment site will be shut down, and the shutdown will result in an employment loss for 50 or more employees during any 30-day period. A mass layoff occurs without a plant closing if the layoff results in an employment loss at the employment site during any 30-day period for 500 or more employees, or for 50-499 employees if they make up at least 33 percent of the employer’s active workforce. Under certain circumstances, smaller employment losses also trigger notification requirements.

The WARN data over an extended period of time are publicly available for many states. Figure 6 shows the number of layoffs for Alabama, Michigan, and Washington, as examples. The data show clear spikes in layoffs in 2009. For Alabama and Washington the figure shows layoffs sorted by the effective date. For Michigan, we have information about the date of the WARN notice but not about the effective date of the layoff, so we sort the layoffs by the expected effective date, which is the notice date plus two months.

3.3 Job destruction

The Business Dynamics Statistics data report job destruction. This measure is defined as the sum of all establishment-level reductions in employment. Davis and Haltiwanger (1992) and Davis, Faberman and Haltiwanger (2013) proposed job destruction as a measure of separations and validated the definition through study of the microdata from JOLTS. Although an employer could accomplish a reduction in employment by cutting back hiring and relying on normal attrition, in fact, almost all employment reductions take the form of separations. When an adverse shock hits the economy, separations jump and quits fall. Figure 7 shows data from the BDS on job destruction. It shows a considerable bulge of job destruction immediately after the financial crisis.

3.4 Displaced workers

Displaced workers are defined as those 20 years old and over who have worked for their employers for 3 or more years at the time of displacement, who lost or left jobs because their plants or companies closed or moved, because there was insufficient work for them to do, or because their positions or shifts were abolished. These are job losses among workers with
Figure 6: Mass Layoffs, by State

Note: Data from the layoff notices under the Worker Adjustment and Retraining Notification Act.
substantial tenure, in contrast to layoffs measured in JOLTS. These are called long-tenured displaced workers.

Table 2 shows the findings of the displaced workers supplement to the CPS taken in January of even-numbered years starting in 2002. The survey inquires about current unemployment and displacement in the year ended the month before, and one and two years earlier.

The design of the displaced workers supplement to the CPS poses an interesting challenge to inference about the time path of long-term displacements and the path of unemployment following displacement. Figure 8 shows an initial attempt. The annual estimates satisfy the overlapping three-year sums and are informed by the timing of layoffs and job destruction within each three-year span. The figure also shows a counterfactual path of displacements, which eliminates the two recession spikes present in the actual data.

3.5 Comparison of measures of the spike of job loss in a recession

Figure 9 compares the estimated long-term worker displacement counts to the tabulations of layoffs, job destruction, and extended mass layoffs. Although the normal level of displacement is far below the levels of layoffs or job destruction, the increase in displacements at
<table>
<thead>
<tr>
<th>Survey in January of</th>
<th>Displacement occurring in calendar years</th>
<th>Number of displaced workers</th>
<th>Unemployed at time of survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>1999 2000 2001</td>
<td>3,969</td>
<td>841</td>
</tr>
<tr>
<td>2004</td>
<td>2001 2002 2003</td>
<td>5,329</td>
<td>1,076</td>
</tr>
<tr>
<td>2006</td>
<td>2003 2004 2005</td>
<td>3,815</td>
<td>511</td>
</tr>
<tr>
<td>2008</td>
<td>2005 2006 2007</td>
<td>3,641</td>
<td>655</td>
</tr>
<tr>
<td>2010</td>
<td>2007 2008 2009</td>
<td>6,938</td>
<td>2,505</td>
</tr>
<tr>
<td>2012</td>
<td>2009 2010 2011</td>
<td>6,121</td>
<td>1,634</td>
</tr>
<tr>
<td>2014</td>
<td>2011 2012 2013</td>
<td>4,292</td>
<td>893</td>
</tr>
<tr>
<td>2016</td>
<td>2013 2014 2015</td>
<td>3,191</td>
<td>507</td>
</tr>
<tr>
<td>2018</td>
<td>2015 2016 2017</td>
<td>2,981</td>
<td>429</td>
</tr>
</tbody>
</table>

Table 2: Total Long-Tenured Displaced Workers and the Number of Unemployed at the Time of the Survey, in Thousands

Note: Data from the Worker Displacement Supplement to the Current Population Survey.

Figure 8: Estimated Annual Displacements, Actual and Counterfactual
the outset of the two recessions is an important fraction of the increases for layoffs and job destruction.

Figure 10 shows excess job loss associated with the 2009 recession by four measures of job loss, together with excess unemployment. We measure excess job loss as the job loss in excess of the average job loss just before and just after the year of the job loss spike. We measure excess unemployment, as unemployment in excess unemployment in 2007. All four job loss measures show a substantial but short-lived spike. Unemployment shows a substantial increase and slow return to its pre-recession level.

### 3.6 Initial unemployment insurance claims

Figure 11 shows the initial unemployment insurance claims. In contrast to layoffs (Figure 4) but similarly to unemployment, during recessions the initial UI claims go up like a rock and go down as a feather.

Why is there a discrepancy between the number of layoffs and the initial UI claims? One factor is that not all eligible unemployed individuals claim the benefits. Building on Blank and Card (1991), Auray, Fuller and Lkhagvasuren (2019) find that from 1989 through 2012, the take-up rate averaged 77 percent. Research shows that the number of those who are eligible but do not claim benefits increases in recession and declines in recoveries (see Fuller, Ravikumar and Zhang (2012) and Auray et al. (2019)). Thus, fluctuations in take-up rates
Figure 10: Excess Job Loss in 2009 and Excess Unemployment, in Thousands of Workers

goes in the wrong direction as an explanation of the discrepancy between layoffs and the initial UI claims.

3.7 Flow of new permanent layoffs in the Current Population Survey

Figure 12 shows unemployment involving permanent job loss by duration, from the CPS. Layoffs with duration of 5 weeks or less is a flow of new layoffs. The flow spikes at the onset of recessions and declines only slowly afterwards.

4 The Direct Channel from Job Loss to Subsequent Lingering Unemployment

We consider the hypothesis that excess job loss directly accounts for the spike and the subsequent long slow decline of excess unemployment. We call it the direct-channel hypothesis. According to this hypothesis, the extra individuals who become unemployed because of the recession shock follow a path similar to those found in research such as Jacobson et al. (1993) and Davis and von Wachter (2011) that tracks the post-displacement paths of workers who lose their jobs from layoffs. These paths often include multiple spells of unemployment.
Figure 11: Initial Unemployment Insurance Claims, in Thousands

Figure 12: Unemployment due to Permanent Job Loss, by Duration

Note: Data from the Current Population Survey.
4.1 Information about the subsequent role in unemployment from job displacement

The CPS survey supplement measuring job displacement contains the crucial information about lingering unemployment among job-losers in the years following job loss.

We fit a simple time-series regression with the biennial data for unemployment in January of even-numbered years of workers suffering displacements in the previous three years as the left-hand variable and three lagged values of the estimated displacement counts as right-hand variables, along with a constant. The relation takes the form

\[ u_t = f_1(D_{t-1}) + f_2(D_{t-2}) + f_3(D_{t-3}) \]  \hspace{1cm} (6)

We linearize as

\[ u_t = \alpha + \beta_1 D_{t-1} + \beta_2 D_{t-2} + \beta_3 D_{t-3} \]  \hspace{1cm} (7)

If market tightness were constant over time, \( \beta_1 \) would be the unemployment rate among workers who suffered displacement within the past year, \( \beta_2 \) one to two years ago, and \( \beta_3 \) two to three years ago. The design of the survey prevents learning about unemployment among people displaced more than 3 years ago. However, job-finding rates are lower in the same years that displacements are high, so \( f(D) \) is a convex function of \( D \). This property implies that the intercept \( \alpha \) should be negative and the coefficients should be greater than the unemployment rates.

Table 3 shows the regression results. The good fit suggests that the imputation of annual timing for the displacements is reasonably successful. The fact that the 3rd-year coefficient is somewhat larger than the 2nd-year one is within sampling variation, but may also reflect the fact that a worker with displacement 3 years ago also suffered an earlier displacement as well. In addition, there may be a stronger convexity effect for the 3rd-year displacement. The negative intercept confirms the expectation of a convex relation between displacements and later unemployment.

Our first use of the regression is to impute unemployment of workers suffering displacements in the previous three years in January of the odd-numbered years when the supplement to the CPS was not performed. Figure 13 shows the fitted values from the regression for the years 2002 through 2018, in red, along with the actual unemployment counts for the even-numbered years when the supplement to the CPS occurs, in blue.

Our second use of the regression results is to calculate how much lower displacement-related unemployment would have been absent the spikes of displacement in the two recessions. Figure 14 shows the results of this counterfactual and compares displacement-related unemployment from the two recessions to overall unemployment in January in the years

24
Table 3: Regression Results for the Relation between Lagged Displacements and Current Unemployment of Workers Suffering those Displacements

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Interpretation</th>
<th>Coefficient</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha )</td>
<td>Intercept</td>
<td>-991</td>
<td>(144)</td>
</tr>
<tr>
<td>( \theta_1 )</td>
<td>Effect of prior year’s displacements</td>
<td>0.76</td>
<td>(0.07)</td>
</tr>
<tr>
<td>( \theta_2 )</td>
<td>Effect of displacement 1 to 2 years ago</td>
<td>0.21</td>
<td>(0.08)</td>
</tr>
<tr>
<td>( \theta_3 )</td>
<td>Effect of displacement 2 to 3 years ago</td>
<td>0.37</td>
<td>(0.06)</td>
</tr>
</tbody>
</table>

\( R^2 \) 0.991 \\
\( \sigma \) 116

Figure 13: Actual and Fitted Unemployment Counts of Workers Suffering Displacements in the Previous Three Years, in Thousands
Figure 14: Total Unemployment and the Component Related to Bursts of Displacement Associated with Two Recessions, in Thousands

since 2001. The rise after the recession is material relative to the overall increase in unemployment following the recession that began in 2001, but is a small part of the large increase in unemployment following the financial crisis.

4.2 Application to other measures of job loss

We use the estimates from Table 3 to calculate excess unemployment from excess job loss by the four measures shown in Figure 10. Figure 15 shows the unemployment resulting from excess job loss in 2009 and total excess unemployment.

Figure 16 shows the contribution of unemployment from excess job loss in 2009 to the cumulative excess unemployment during the 2007-09 recession.

4.3 Conclusions about the relation between the magnitude of the increase in unemployment following a recession shock and the measures of job loss

An unambiguous spike in regular and mass layoffs, job destruction, and displacement, and mass layoffs accompanies the shock that marks a recession. Figure 15 shows that the spike in layoffs more than fully accounts for the spike in unemployment in 2009 but cannot account for all of the excess unemployment afterwards. That is, the excess job loss accounts for the magnitude of the initial increase in unemployment, but not its persistence. The persistence
Figure 15: Unemployment from Excess Job Loss in 2009 and Total Excess Unemployment, by Four Measures of Job Loss, in Thousands of Workers

Figure 16: Contribution of Excess Job Loss to Cumulative Excess Unemployment, by Four Measures of Job Loss, in Thousands of Worker-Years
is too large to be explained as reflecting only the personal experiences of the extra job-losers dating from the spike. The direct channel is only part of the story of persistent high unemployment after the crisis. This conclusion is reinforced by the rise in unemployment among new entrants to the labor force, as we show below.

The results show that the spike of job loss during recessions induces a downstream effect on unemployment. Some of subsequent unemployment comes from the unemployed who suffered the original displacement event and are circling through short employment spells, and the rest of subsequent elevated unemployment appears to come from new job loss, not associated with the original job loss.

4.4 Excess unemployment of new entrants

By definition, new entrants to the labor force are not the victims of job loss events. A bulge of unemployment of new entrants following an adverse shock indicates that either (1) unemployment is infectious, or (2) the bulge of unemployment arises from a decline in the incentives to create jobs. Figure 17 shows that new-entrant unemployment nearly doubled after the financial crisis. This finding rules out the hypothesis that the sole cause of lingering unemployment following the crisis was the slow absorption of workers who suffered job loss from the crisis. The direct channel cannot be the only link between a crisis and its subsequent gradual recovery.

5 Effective Exit Rate from Unemployment

5.1 Defining and measuring the effective exit rate

From the data on unemployment in the displaced workers supplement, we can estimate what we call the effective exit rate from unemployment, denoted $f_t$. We know the number of people in the survey who were displaced in the prior three years and who are currently unemployed. We also have estimates of the number of people displaced in each of those years. The effective exit rate is based on the assumption that the probability of being unemployed $\ell$ months following a displacement is the product of the monthly exit rates from the time of displacement up to the survey. Here we adopt the perspective originated in Krueger, Cramer and Cho (2014), and expanded in Hall and Schulhofer-Wohl (2018) and Hall and Kudlyak (2019), that the typical path from initial unemployment to current labor-market activity often involves a mixture of spells of short jobs, time out of the labor force, and unemployment. As Krueger and co-authors showed, the probability of being unemployed a year later conditional on starting unemployed is much higher than would be expected from
the monthly probability of unemployment ending raised to the 12th power. Our calculation here extends the calculation by two additional years, as we exploit the 3-year look-back in the displaced workers supplement of the CPS.

The implied relation between the observed number of people unemployed in the January survey of month $t$ is

$$U_t = \sum_{\ell} \prod_i (1 - f_{t-i}) N_{\ell}. \quad (8)$$

We parameterize as

$$f_\tau = a - b u_\tau. \quad (9)$$

The parameter $b$ is the negative sensitivity of the effective exit rate to the standard national unemployment rate $u$. Not surprisingly, it turn out to be essentially 1. We estimate $a$ and $b$ by minimizing the sum of squared residuals of the actual values to the implied values of $U_t$. The estimated value is $b = 1.00$ and the monthly effective exit rates range from 0.042 in 2010 to 0.099 in 2018. By contrast, the monthly exit rate is around 0.5.
5.2 Implications of low effective job-finding rates

This subsection provides evidence that recessions are followed by long periods of high but not continuous unemployment among those who lost jobs in the recession. During the long re-employment process, the unemployed often circle among unemployment, out of the labor force, and short-term jobs.

Figure 18 shows unemployment by reason as a share of the labor force, except for labor force new entrants. The figure shows that recession involve not only an increase in unemployment from permanent and temporary layoffs but also due to completion of temporary jobs and labor force re-entry. This points towards an elevated number of individuals taking temporary jobs and circling between unemployment and spells out of the labor force.

When a crisis causes a spike in unemployment, there is a shift away from stable jobs and toward brief jobs in the working-age population. This shift gradually subsides during the recovery. To demonstrate this phenomenon, we study the 8-month CPS activity paths, as in Hall and Kudlyak (2019). We define short employment spells as those lasting one or two months. These are the spells that are preceded and succeeded by unemployment or out of the labor force. We define an individual to have stable employment if employed in all 8 reported months. We calculate the average number of short employment spells among the of CPS respondents of working age. We also calculate the average number of respondents in stable employment. We create an index of the shift toward short jobs as the difference between first and second of these calculations. Figure 19 shows the short-spell index starting in 1976 for four demographic groups. The indexes jump upward in recessions and gradually declines during the ensuing recovery for all four groups.

6 The DMP Model

This section describes the basic DMP model and discusses potential mechanisms and driving forces in the model that may help explain the slow but steady recovery of unemployment following a spike in job loss.

Here and in the rest of the paper, we refer to driving forces, which are variables taken as exogenous to the labor market, in the sense that we do not consider that actors in the labor market can influence the variables. We also refer to endogenous variables, notably unemployment, determined within the labor market. Models have parameters, taken as constants in derivations, but often later treated as variables, either as time series or as functions of other variables. Thus matching efficiency \( \mu \) is a parameter of the matching function, but it can change over time, or it can be a function of unemployment, in which case we will always write it \( \mu(u) \).
Figure 18: Unemployment by Reason, as Share of Labor Force

Note: Data from the Current Population Survey.
Figure 19: Indexes of the Duration of Employment Spells

Note: Authors calculations using data from the Current Population Survey.
The DMP class of models treats the level of unemployment as a state variable. At
the outset, in our application of the model, unemployment has a high value left behind by a
recession. The model traces the movements of unemployment for a decade, aiming to emulate
the slow but steady decline of unemployment documented earlier in this paper. We discuss
potential driving forces of the model that generate persistence. We confirm the well-known
failure of the simple DMP model to account for the slow pace of recovery without invoking
persistent movements of driving forces.

6.1 Potential driving forces of the DMP model

Our discussion of the model will follow the principle that the only flow in the labor market
that is sensitive to tightness is from unemployment to employment. To keep the exposition
compact for now, we take the size of the labor force to be a constant. Thus we neglect
variations in flows into and out of the labor market. Another important simplification is
that we omit on-the-job search. A large and rich literature deals with this topic.

The labor market operates on the principles of random search. The tightness of the
market, denoted $\theta$, is measured by the ratio of vacancies to job-seekers. We normalize
the labor force at 1 and measure vacancies as the ratio to the labor force, denoted $v$, and
unemployment as the unemployment rate, denoted $u$, so

$$\theta = \frac{v}{u}. \quad (10)$$

An aggregate constant-returns matching function $m(u, v)$ gives the flow of job matches
as a function of the inputs—the stock of searchers, $u$, and the stock of vacancies $v$. We
parameterize $m(u, v)$ as a Cobb-Douglas function $m(u, v) = \mu u^{1/2} v^{1/2}$. The parameter $\mu$ is
matching efficiency.

The job-finding rate is the number of matches per job-seeker per month,

$$f = \mu \frac{u^{1/2} v^{1/2}}{u} = \mu \theta^{1/2}. \quad (11)$$

Thus, tightness determines the monthly job-finding rate, an increasing function. Similarly
tightness determines the monthly job-filling rate, $q = \mu \theta^{-1/2}$. The latter is the number of
jobs filled by holding a vacancy open for a month, so it can be greater than 1. The job-filling
rate decreases with tightness.

We let $P$ be the present value of a newly hired worker’s productivity and $W$ be the present
value of their wage. The difference $P - W$ is the net benefit to the firm from hiring a new
worker. For simplicity, we call $W$ the wage but we mean the present value the worker earns
from the job, as of the time of hiring. The model operates in an environment of certainty,
so expectation operators are omitted.
The flow cost of recruiting is \( \kappa \). Recruiting satisfies the zero-profit condition,

\[
\kappa = \mu \theta^{-1/2} (P - W)
\]  

(12)

This condition pins down tightness:

\[
\theta = \left( \frac{\mu}{\kappa} \frac{P - W}{\kappa} \right)^2
\]  

(13)

Tightness is increasing in matching efficiency \( \mu \), increasing in productivity \( P \), decreasing in the wage, \( W \), and decreasing in the recruiting cost, \( \kappa \).

The law of motion of unemployment describes the rate of change of unemployment as the net of inflows from separations and outflows from job-finding:

\[
\dot{u} = (1 - u) s - u \mu \theta^{1/2}.
\]  

(14)

The parameter \( s \) is the separation rate into unemployment. Its reciprocal, \( 1/s \), is the expected duration of a job. In addition to its visible role in the law of motion, the separation rate is one of the determinants of the present values \( P \) and \( W \).

Separations matter in two ways. First, a shock that hits the economy just before the starting time of the model results in a pulse of separations and an elevated unemployment rate at the outset of the time span of the model. We do not model the shocks or recessions in general—they are simply the source of a legacy of unemployment when the model swings into action. Second, the separation rate \( s \) describes the flow of separations during the recovery. The separation rate controls the inflow to unemployment and the job-finding rate controls the outflow. Along a realistic recovery path, the two flows are almost equal—unemployment declines quite slowly. For now, we take the separation rate to be an exogenous constant. Later we will consider its role as a driving force.

The job-finding rate, \( \mu \theta^{1/2} \), controls the flow out of unemployment. Its reciprocal is the expected duration of a spell of unemployment. Note that matching efficiency \( \mu \) appears in both the tightness equation and in the law of motion for unemployment.

The potential driving forces of the model are

- productivity \( P \),
- the wage, \( W \),
- the flow cost of a vacancy, \( \kappa \),
- matching efficiency, \( \mu \),
- the separation rate, \( s \).
With these specified as constants, time series, or functions, the model is a first-order differential equation in the single state variable, \( u \).

We consider the path of the economy immediately after a shock has left unemployment at an elevated level. The model then evolves according to its law of motion. In the cases we consider, it converges to a stationary state because the driving forces approach constant levels.

### 6.2 Path of unemployment following a recession in the basic DMP model

We start by describing a basic DMP model, with constant productivity, separation rate, matching efficiency and vacancy cost. We also assume that the wage is constant even though the unemployment rate declines over time. This assumption mirrors the behavior of the canonical DMP model of Mortensen and Pissarides (1994). In that model, the wage is the endogenous result of bargaining. Under our assumptions of constant \( P \) applied to that model, the bargained wage is constant, even though unemployment follows a path that starts above its stationary value and converges over time to its stationary value. Because the wage is constant, all of the driving forces are constant and labor-market tightness is constant during the recovery.

In the data, tightness rises during recoveries. Modifying the DMP model to match this key fact is one of the main topics of recent theoretical work on the model and of this paper.

We parametrize the model to resemble the economy in early 2020. Time is monthly. The stationary unemployment rate is \( u^* = 0.035 \), separation rate \( s = 0.018 \), and matching efficiency \( \mu = 0.5 \).

Figure 20 describes the behavior of the model in a phase diagram, with unemployment \( u \) on the horizontal axis and tightness \( \theta \) on the vertical axis (see Pissarides (2000), Figure 1.3, p. 30). Equation (13) determines tightness. It describes hypothetical combinations of tightness and unemployment. It is the horizontal line in the phase diagram because in the basic model tightness is a constant, not a function of unemployment.

The downward-sloping curve in the phase diagram is the locus of stationary values of unemployment, derived from the law of motion by setting \( \dot{u} \) to zero. It is

\[
\theta = \left( \frac{1 - u \ s}{u \ \mu} \right)^2.
\]  

We call this the “\( \dot{u} = 0 \) curve” throughout the paper. It appears in all of our phase diagrams. It only changes when we consider different values of matching efficiency \( \mu \) or the separation rate \( s \). Points above and to the right of the \( \dot{u} = 0 \) curve have declining
The labor market begins at the right end of the horizontal line, with high unemployment. The market moves to the left, along that line, as unemployment falls but tightness remains the same. The market approaches the left end of the line where unemployment becomes constant at its stationary level.

unemployment, while points down and to the left have rising unemployment. At high unemployment rates, a given job-finding rate generates a higher outflow from unemployment because the rate applies to more people. With lower unemployment, constancy of unemployment along the locus requires a higher job-finding rate and thus higher tightness. Or, to put it another way, higher tightness means a higher rate of growth of unemployment at a given unemployment rate, and so a lower level of unemployment to achieve constancy.

All the combinations of unemployment and tightness consistent with the model will be on the horizontal line labeled $\theta$, tightness. After a recession creates a legacy of high unemployment, the economy starts at the right end of that line. As the recovery proceeds, unemployment moves horizontally to the left according to equation (14).

Figure 21 shows the path of unemployment implied by the simple model with constant driving forces, along with the actual path starting from the peak in 2009. The dots show the progress by month. In the first month, unemployment falls substantially. As the unemployment rate falls during the recovery, the steps become smaller. The model economy closes most of the gap in just three or four months. The model’s recovery is far speedier than actuality, the point made emphatically by Cole and Rogerson (1999). And the path is more convex, away from the nearly linear path of actual unemployment.
All of our model solutions in this paper are effectively exact, that is, not based on any approximation, and using double-precision arithmetic. Petrosky-Nadeau (2014) demonstrated the importance of accurate solutions of DMP models, arising from the substantial concavity of the matching function, which impairs the accuracy of approximation by log-linearization.

6.3 Path of unemployment in a model with low effective unemployment exit rate

We know that the 50-percent per month transition rate from unemployment to employment in the basic DMP model greatly overstates the actual exit rate from unemployment (see section 5). Figure 22 shows the phase diagram and Figure 23 shows the model’s unemployment path together with the actual path, with the lower effective exit rate from unemployment of 0.1 per month in the range estimated in Section 5. This alteration substantially delays the recovery but not nearly enough to match the actual path of unemployment. The phase diagram with lower exit rate is same as the earlier one, so we do not repeat it.

We conclude that using the estimated effective unemployment exit rate of 0.10 makes an important contribution to matching the model’s unemployment path but cannot be a full resolution of the high-persistence puzzle.

Pries (2004) builds a DMP model that explains the high persistence of unemployment as the result of recurrent spells of unemployment following a shock. Once a job match is
Figure 22: Phase Diagram with Low Effective Unemployment Exit Rate

Figure 23: Recovery Path of Unemployment with Estimated Effective Exit Rate from Unemployment, and Actual Unemployment, 2000 to 2020
made, the parties are at risk of an adverse productivity realization that reveals that the match should end and the worker should return to the labor market. In normal times, most matches will have become known to be reliable and no longer at risk of being found unproductive. At random, a cloud may form over the labor market that calls into question the earlier belief that a match is good—the parties need to receive a new signal of reliability for a fraction of the existing matches. Some of the matches end immediately and the others are exposed to the possibility that they will be found to be unproductive from a later draw of productivity. If the aggregate shock simply knocks out some of the existing matches, the model generates little persistence—the victims of the shock regain reliable employment almost as quickly as they would without the learning-by-experience feature of the model (see Pries’s Figure 3). The broader version of the shock, which induces the parties to wait to determine who are the job losers, makes the effect of the shock realistically persistent.

6.4 Variation of the driving forces over time

So far we have discussed the model’s dynamics when the driving forces are constant. The basic model generates speedy recovery of unemployment. Tightness is fixed.

In this subsection, we consider evidence of time variation in the driving forces. Where possible, we use evidence from the direct empirical counterparts of the driving forces. Otherwise, we infer variation in a combination of driving forces using the model.

The model has five driving forces—productivity, $P$, the wage, $W$, the flow cost of a vacancy, $\kappa$, matching efficiency, $\mu$, and the separation rate, $s$.

A phase diagram is not a good way to describe the effect of a changing driving force, because that change involves a shift of the $\theta$-function over time. If a recovery occurs because of higher values of $\mu$ or $D$, that function shifts upward and tightness rises correspondingly. We discuss potential shifts in driving forces with this point in mind, without using phase diagrams. At this point, we do not take a stand on whether a driving force is exogenous to the labor market, arising from policy or financial developments or the like, or whether it is endogenous within the labor market, arising from intrinsic features of the labor market that generate persistent recoveries.

We infer variation in matching efficiency from the matching function and data on hires, vacancies, and unemployment. It is the ratio of the hiring flow to weighted matching inputs:

$$\mu = \frac{H}{U^{1/2}V^{1/2}}.$$  \hfill (16)

Tightness in equation (13) can be expressed as the square of the product of matching efficiency, $\mu$, and the compound force, $D = (P - W)/\kappa$. The separate elements of the compound force—productivity, wage, and the vacancy cost—are not easily identified, as a
Figure 24: Matching Efficiency and the Compound Driving Force Embodying Productivity, Wage, and Vacancy Cost

practical matter. Time series on productivity and wages are available, but not in a form that yields accurate measures of the difference, $P - W$. In a typical calibration of the DMP model, the match surplus is around 6 percent of $P$ and the part of that accruing to the employer is around 3 percent. Measuring $P - W$ as the difference between two series, each with measurement errors little correlated with one another, produces results that are highly unreliable. Data on recruiting costs are scarce—the relevant costs are only those incurred by an employer prior to making a wage bargain.

We infer the compound driving force, $D$, as the part of the overall movement of the square root of $\theta$ not attributable to $\mu$, using equation (13). This turns out to be the ratio of vacancies to hires:

$$D = \frac{\theta^{1/2}}{\mu} = \frac{V}{H}. \hspace{1cm} (17)$$

Figure 24 shows the results. The labor market tightened during the long expansion starting in 2009 from a combination of rising matching efficiency $\mu$, and rising value of the compound job-creation incentive $D$.

The remaining driving force is the separation rate into unemployment, $s$. Figure 25 shows the time variation in the separation rate constructed from the CPS data. The separation rate increases sharply in recessions and declines slowly in recoveries. The separation rate contributes negatively to tightness, so its decline in recoveries contributes to the gradual decline in unemployment.
Figure 25: Transition Rate from Employment to Unemployment, Monthly

Note: Authors' calculations using data from the Current Population Survey, monthly seasonally adjusted.
7 Exogenous Driving Forces of Recoveries

Next we explore modifications of the basic DMP model with time-varying driving forces that are exogenous to the labor market. They would be endogenous in a full general-equilibrium model.

7.1 Profitability of hiring a worker

The financial incentive to create a job, \( P - W \), is at the center of the DMP class of unemployment models. If \( P - W \) remains at a low value after the crisis and only gradually trends upward during the recovery, tightness will gradually rise from a low value back to normal, and unemployment will gradually decline.

A large fraction of the DMP literature takes \( P \) as an exogenous time-series process. Unless \( W \) changes by the same amount as \( P \), an unlikely configuration, \( P - W \) inherits some of the movements in \( P \) and so tightness will have some volatility on that account. Productivity has essentially no correlation with tightness, so this channel cannot explain the cyclical movements of tightness and unemployment, with persistent recoveries (Hall (2017)).

A straightforward way to increase volatility in tightness relative to the volatility in \( P \) is through a rigid wage, \( W \). The wage is the result of a bargaining process and is generally endogenous, although a fixed wage is an interesting special case. Wage-bargaining models different from the Nash bargain of the canonical DMP model can deliver realistic volatility but do not generally add to persistence of recoveries. Shimer (2005) found that the movements in \( P - W \) induced by movements in \( P \) and thus in unemployment, were tiny, in the DMP model with Nash bargaining under plausible assumptions about parameter values. Numerous subsequent papers altered the original model to boost its response to productivity.

A different and more fundamental shortcoming of the hypothesis that productivity drives unemployment is the lack of any correlation of measured productivity and unemployment. Some authors, notably Gertler and Trigari (2009), have built models where a bargained wage has a persistent effect because it becomes a fixed wage for a period of time. This setup is a cousin of Calvo (1983) pricing in product markets.

Literature finds little support for the relevant wage rigidity in the data. Bils (1994) finds that wages of newly hired workers are more cyclical than wages of incumbent workers.

Kudlyak (2014) estimates the cyclicality of the user cost of labor, which is the wage statistics that takes into account the present discounted value of wages and is relevant for job creation decisions of the firms. She finds that the user cost of labor is more pro-cyclical than the average wage and, more importantly, than the wages of newly hired workers (see also Basu and House (2016)). Kudlyak quantitatively examines the implications of the esti-
mated cyclicity of the user cost for the basic DMP model. First, she finds that the data lack relevant wage rigidity to amplify the unemployment volatility in response to productivity shocks in the model. More importantly, the basic DMP model cannot simultaneously generate empirical elasticities of the vacancy–unemployment ratio and the user cost of labor—both are too volatile for the basic model’s job creation equation to hold. She concludes that the job creation equation in the basic model is missing an element that makes it unprofitable to hire in recessions but this element is not wages because labor in recessions is cheap.

Our conclusion is that a gradual improvement in the value of the marginal revenue product of labor that results in a realistic upward trend in $P - W$ is a candidate explanation for the gradual decline in unemployment. However, the empirical support for that mechanism is weak. In particular, there is no systematic improvement in productivity above trend in recoveries.

### 7.2 Financial sources of rising $P - W$ in recoveries

Attention has turned in the DMP literature to financial factors as driving forces for unemployment (Hall (2017), Kilic and Wachter (2018), Kehoe, Lopez, Midrigan and Pastorino (2020)). These papers observe that $P - W$ is the discounted value of the future cash flow to the employer, along with other discounted flows in more elaborate models, and thus are sensitive to fluctuations in discount rates. These fluctuations are substantial, according to financial principles set forth in Campbell and Shiller (1988) that are widely accepted in financial economics today.

Discount effects operating through $P - W$ will be present if $W$ moves in proportion to $P$ (because necessarily $W < P$) but will be more powerful if $W$’s movement is less than $P$’s—that is, if wages are somewhat sticky relative to the present value of future productivity. Wage stickiness is fully consistent with DMP principles, provided it is not so severe as to dictate a wage outside the bargaining set of the worker and the employer, which would destroy a match despite its joint value to the parties (Hall (2005b)).

Spikes in general financial discounts coincide with spikes in unemployment, and so are logical candidates to be the source of high unemployment from recessions. But declines in discounts are not nearly persistent enough to account for the lengthy recoveries observed in unemployment.

One potential source of persistent financial effects is a crisis-induced cut in the availability of credit, which raises discount rates for credit-dependent firms and thereby cuts $P - W$. As the availability of credit gradually returns back to normal, unemployment also returns to normal. Dromel, Kolakez and Lehmann (2010) pursue this approach to explaining the high persistence of unemployment.
Figure 26: Scaled Index of Loan Availability Compared to the Unemployment Rate

We provide suggestive evidence of the persistent influence of credit conditions on unemployment in recoveries. To measure the availability of credit, we use data from the Federal Reserve Board’s Survey of Senior Loan Officers. Respondents in the survey answer in terms of tightening and easing of commercial loan standards. We cumulate these answers using the statistical model in Hall (2011) to form an index of loan availability. The scale of the index is arbitrary. In Figure 26, we scale it to have the same standard deviation as the observed compound driving force $D_t$ and compare the scaled index to $D_t$. The two variables move closely together. The slow relaxation of lending standards matches the slow decline in the driving force $D_t$.

7.3 Other exogenous factors

Petrosky-Nadeau (2014) introduces external financing of vacancy costs in frictional credit markets. The easing of financing constraints during an expansion as firms accumulate net worth reduces the opportunity cost for resources allocated to job creation. Agency-related credit frictions endogenously generate persistence in the dynamics of labor-market tightness.

Garin (2015) studies the effects of changes in collateral requirements on the cyclical properties of unemployment and job creation. In the model, borrowing limits are linked to the firm’s physical capital stock. Financial frictions arise from an imperfect enforcement contract. Financial frictions in the form of borrowing constraints create a wedge in the job
creation equation as in Petrosky-Nadeau’s paper. To the extent the constraint is binding, the marginal cost of hiring an employee increases.

Under borrowing constraints, productivity shocks move the incentive of firms to post vacancies in the same direction as they move the borrowing constraint. That is, in Garin’s model, periods characterized by low realizations of TFP are also periods in which the firms are less eager to increase their hiring. Garin argues that fluctuations in collateral requirements is main driving force behind movements in the vacancy creation costs in the model. His paper endogenizes an effect from credit conditions similar to the one we discussed earlier that we treated as exogenous. See also Dromel et al. (2010).

Gavazza, Mongey and Violante (2018) develop a model where firms adjust their recruiting effort in response to movements in labor market tightness.

8 Endogenous Mechanisms Implying a Slow Downward Trend in Unemployment in Recoveries

The high persistence of economic activity in general, and unemployment in particular, has puzzled macroeconomists for decades. Finding explanations of endogenous fluctuations in the labor market or other markets has been a goal of many generations of researchers. Our particular interest is labor-market mechanisms operating to generate the observed pattern of reliable but slow recovery of the unemployment from the high levels experienced in recessions.

We study the situation immediately after a major shock has left a legacy of high unemployment. In the basic model, tightness is determined by the equation,

$$\theta = \left( \mu \frac{P - W}{\kappa} \right)^2,$$

which excludes any influence of unemployment except through the driving forces. Even if half the labor force is unemployed, jobs are not hard to find, as long as the driving forces are at normal levels. Instead, the volume of vacancies created by employers is at a high enough level to bring $\theta$, the vacancy/unemployment ratio, to a normal level. The supply of vacancies is perfectly elastic.

With instantaneous response of tightness to restored normal driving forces, unemployment returns to normal fairly quickly. Recoveries are unrealistically speedy. A number of interesting contributions to the DMP literature, mostly recent, alter the model to mimic the high persistence of unemployment.

One appealing notion in the quest for persistence is that the legacy of high unemployment from a recession creates congestion in the labor market—the high levels of vacancies hypoth-
esized by the canonical DMP model are impractical because employers would interfere with each other just as additional cars joining a crowded highway slow down all of the traffic.

### 8.1 Unemployment influences tightness $\theta$

We consider a class of DMP-type models in which the unemployment rate influences labor-market tightness $\theta$. We will review an extensive literature that deals with this modification of the basic DMP model. In this class, the $\theta$ function is not a horizontal line as in Figure 20 but rather slopes downward in the unemployment-tightness diagram.

Unemployment plays the role of a *mediating force* in this class of models. It is endogenous, but transmits movements in somewhat the same way as a driving force.

The $\theta$ function becomes

$$\theta(u) = (\mu D)^2 - \gamma(u),$$

where $\gamma$ is increasing in $u$. We recover the $\gamma$ function from the observed relation between unemployment and tightness during the expansion from 2009 to 2020, shown as the blue line in Figure 27. The fact that the line fits a smooth curve, except for small transitory deviations, supports the hypothesis that a functional relationship exists between $u$ and $\theta$. 

---

**Figure 27:** Phase Diagram for the DMP Model with Negative Dependence of tightness on unemployment
As before, the $\dot{u}$ curve in the figure traces out the relation between unemployment and tightness such that the change in unemployment is zero. It is

$$\theta = \left(\frac{1 - us}{u \cdot \mu}\right)^2,$$

(20)

using the values of the parameters $s$ and $\mu$ just discussed.

A recovery involves a gradual movement along the $\theta$ function starting at the lower right and moving toward the stationary point with unemployment $u^*$. Figure 27 implies that recoveries will proceed slowly, because the two curves are close to each other. In October 2009, at the beginning of the recovery, high unemployment discouraged tightness to a small fraction of its normal value. Job creation proceeded only enough to lower unemployment slowly.

Though congestion externalities may play a role in understanding unemployment persistence, it is important to recognize that concavity of the empirical matching function captures the congestion externality to the extent it is reflected in the actual level of vacancies.

Progress in this area either (1) revises the technology to increase the effective concavity in vacancies, or (2) invokes adverse forces that counteract the incipient high level of vacancies early in recoveries. The resulting modified DMP model would have a lower elasticity of supply of vacancies and higher the persistence of unemployment in comparison to the canonical DMP model.

### 8.2 Higher vacancy costs early in the recovery

The cost of maintaining vacancies, $\kappa$, can be a channel of gradual decline in unemployment in recoveries. The idea is that elevated unemployment in a recession raises $\kappa$ by creating congestion. In turn, as discussed earlier, higher $\kappa$ slows down the decline in unemployment. We consider mechanisms to describe this feedback from unemployment to tightness as we described earlier, where the $\theta$-equation involves an offset $\gamma(u)$ to the driving force $D$ depending negatively on unemployment.

It is clear that the story of the phase diagram only works if the effect embodied in $\gamma$ is reasonably strong—enough to twist the curve clockwise from flat to downward sloping, and lying close to the $\dot{u} = 0$ curve.

One way to incorporate the idea that there is feedback from unemployment to tightness is to make the cost of maintaining a vacancy an increasing function of unemployment. This alteration of the basic DMP model eliminates the disconnect of unemployment from the determination of tightness. As we noted earlier, in the basic model this property arises from the perfectly elastic supply of vacancies at cost $\kappa$, a fixed parameter of the model.
If expanding vacancies involves increasing marginal cost, making $\kappa$, the marginal cost of maintaining a vacancy, an increasing function of unemployment, the elasticity of supply of vacancies will be finite. An extensive literature on adjustment costs in vacancy creation has pursued this point.

Adding adjustment costs to the vacancy-creation technology would be a natural way to tame the behavior of vacancies at the beginning of recoveries. Fujita and Ramey (2007) is the pioneering analysis of adjustment costs in vacancy creation. Their model makes the marginal cost of creation rise in proportion to the rate that employers raise vacancy creation. The feedback generates higher persistence of unemployment. When unemployment is high and employers are creating vacancies at higher rates, $\kappa$ is high, offsetting some of the incentive to hire and slowing the recovery. Creation costs induce firms to smooth the adjustment of new openings following a shock, leading the stock of vacancies to react sluggishly. Fujita and Ramey’s modification of an otherwise standard DMP model eliminates the counter-intuitive property of that model, that vacancies are a jump variable that increases by the full amount of the increase in unemployment following a crisis. Their model makes vacancies a state variable obeying an adjustment process. During that process, the decline in unemployment from a high initial level is slower.

Ferraro (2017) confirms the conclusion in Fujita and Ramey (2007) that the perfect elasticity of supply of vacancies assumed in the canonical DMP model is unrealistic. Within its third-moment framework, the paper argues that upward-sloping supply results in more realistic performance of the DMP model, notably in its ability to match slow recoveries. Shao and Silos (2013) present a model in which firms face sunk costs to enter the production process.

The effect of these models in extending unemployment persistence operates through the marginal cost of adjustment, which enters the model through the parameter $\kappa$, now reinterpreted as the derivative of a convex adjustment cost function. The analysis in the previous section considers the response to higher values of $\kappa$ in the compound driving force $D = (P - W)/\kappa$.

Coles and Kelishomi (2018) test and reject the assumption of the canonical DMP model that tightness is orthogonal to unemployment. They conclude that the vacancy creation process is less than infinitely elastic. Leduc and Liu (2020) develop a model where recruiting intensity varies with labor-market tightness in a way that makes unemployment more persistent in recoveries.
8.3 Models of the adverse effects of unemployment on the recruiting process

One line of modeling to support the proposition that higher unemployment raises recruiting costs is the following: Employers have a choice between costly screening of applicants or hiring without screening. In normal times, most employers do not screen, because most applicants self-select to be well matched to the jobs being filled. In times of higher unemployment, self-selection breaks down and employers invest in screening prior to negotiating terms with qualified applicants. The effective cost $\kappa$ of maintaining a vacancy rises and the labor market slackens rapidly. As time passes, conditions gradually reduce $\kappa$ (this is the challenging part) and unemployment begins to decline. The process gains momentum as the pool of job-seekers begins to increase its self-selection. Unemployment gradually declines along the path described earlier in this paper.

This mechanism was considered in Hall (1990) and Hall (2005a). The cost of evaluation per hire depends on the fraction of applicants who are qualified for the job. Applicants may be better informed about their qualifications than are employers. If incentives induce self-selection by job-seekers, so that they apply mainly for jobs where they are qualified, friction and thus unemployment will be low. Self-selection is strongest in markets where unemployment is low and jobs are easy to find. Because of this positive feedback, the equilibrium in a market with self-selection is fragile—unemployment is sensitive to its determinants. Self-selection provides a mechanism for amplification of small changes in the determinants of unemployment.

Moscarini (2001) builds a model in which slack labor markets involve elevated levels of what he calls “excess worker reallocation.” In the model, workers with specialized skills search selectively and contact few vacancies, where they are likely to be hired; while workers with weak comparative advantage apply to any vacancy, driven by the low anticipated acceptance rate. The latter workers produce movements across job types, both job-to-job and through unemployment–excess worker reallocation. He finds that in a tight labor market, comparative advantage dominates waiting costs and excess worker reallocation is lower and matches are more successful. Vacancy costs are correspondingly lower for employers.

Engbom (2021) develops a theory of a positive relation between recruiting cost and unemployment in the presence of on-the-job search, based on the observation that the unemployed apply for many jobs that are unlikely to be a good fit, compared to the employed. This makes it harder for recruiting firms to determine who is a good fit for a job when unemployment is high. Separation shocks are the central driver of unemployment fluctuations in the model. A higher separation rate influences job creation through two channels. First, new matches are expected to be briefer, which discourages job creation. Second, job creation is affected
by a composition effect, as the pool of potential hires shifts toward the unemployed. If application behavior differs little between the unemployed and employed, the composition effect encourages job creation because the unemployed can be hired at a low wage. If, on the other hand, differences in application behavior are large, the composition effect discourages vacancy creation because the unemployed applicants raise firms’ cost of recruiting. Using the Survey of Consumer Expectations, Engbom estimates that the unemployed send over 10 times as many applications as the employed, but are less than half as likely to move per application. The model interprets this as the unemployed being less selective in terms of what jobs they apply to. The results are not driven by firms receiving a greater quantity of applications in recessions, but that the quality of the pool of applicants falls in recessions as the unemployed are less selective in their application decision. In Engbom’s model, by raising the cost of recruiting, a short-lived increase in the separation rate has a persistent, negative impact on the job finding rate, accounting for a large share of its empirical volatility and persistence. Because the increase in the separation rate is not particularly persistent, it does not account for much of the persistence in the job finding rate. Crucial for propagation is the gradual shift in the pool of potential hires toward the unemployed and poorly matched applicants. Because such workers are less selective in their application decisions, this composition effect serves to discourage job creation.

Engbom provides direct evidence that firms receive more applications and spend more time on recruiting per vacancy when unemployment is high. He estimates in the Earnings and Opportunities Pilot Project across 28 US locations for 1978–1981 that one percent higher unemployment is associated with firms receiving 1.34 percent more applications per vacancy and spending 0.90 percent more hours on recruiting per vacancy, while filling a vacancy 0.76 percent faster. Because applications per vacancy and hours spent on recruiting per vacancy rise more than the fill rate, these estimates imply that applications per hire and time per hire rise with the unemployment rate.

### 8.4 Externalities from recruitment selection

Congestion externalities infect the participants in random-search models. As the level of unemployment goes up and the mix of unemployment worsens, recruitment selection imposes externalities on the unemployed and on employers. Firms post fewer vacancies and set higher thresholds for hiring.

Gautier (2002) focuses on an externality in the labour market which is caused by non-sequential search. He develops a model in which unemployment is caused by selection delays of employers. The individual probability to get a job offer is increasing in the number of applications while the aggregate hiring rate is decreasing in the average amount of appli-
cations per applicant. In the decentralized equilibrium, screening information is lost when a worker is found unsuitable for a particular job and the next firm has to spend screening time on this worker again. The externality can be reduced by institutions which specialize in screening large groups of workers such as temporary employment agencies, labor offices and head hunters, which only screen workers a single time before they allocate them to suitable jobs.

Villena-Roldan (2012) studies the implications of employer search with recruiting selection. When unemployment is high and it is hard to find a job, employers have a chance of picking better applicants given a particular distribution of the unemployed. However, vacancy posting and screening generate externalities that have general equilibrium effects on the composition of the unemployment. By screening harder, firms lower the productivity of the unemployment pool and make it harder for other employers to find good workers. On the other hand, if more vacancies are posted, the expected number of applicants decreases so that firms receive fewer applications which limits their capability of hiring the best workers. For this reason, the average productivity of the unemployment pool increases, with positive consequences for all employers. If vacancy posting activity is low, a firm’s individual posting decision becomes less attractive because the unemployment pool may consist mostly of low productivity workers.

Molavi (2018) show that a decline in the quality of the pool of applicants leads firms to post fewer vacancies and to set a higher threshold for the signal of a worker’s ability needed to hire the worker. When the pool becomes more adversely selected, these two effects conspire to depress the job-finding rates of all workers irrespective of their ability.

Fishman, Parker and Straub (2020) develop a dynamic model of credit markets in which lending standards and the quality of potential borrowers are endogenous. Lending standards set privately by the banks have negative externalities and are dynamic strategic complements—tighter screening worsens the future pool of borrowers for all banks and increases their incentives to screen in the future. Lending standards can amplify and prolong temporary downturns, affecting lending volume, credit spreads, and default rates. In the model, when markets recover, they may do so only slowly, a phenomenon the authors call “slow thawing.” This line of thought may apply to labor markets.

Lockwood (1991) develops a setup where employers may administer a test. Then employers also consider unemployment duration as informative about how many times the job-seeker has flunked previous tests. When unemployment is higher, this problem worsens, creating a congestion externality.
8.5 Matching efficiency

Hall and Schulhofer-Wohl (2018) study job-finding rates and match efficiency in CPS data broken down by multiple categories based on the personal circumstances of working-age individuals. These include individuals who are unemployed for various reasons, those currently employed, and those out of the labor force who are and are not interested in working. These categories are further broken down by the duration of unemployment to date in the cases of the unemployed. All categories have positive job-finding rates, ranging from high values for most of the unemployed to quite low values for those not interested in working.

The paper estimates job-finding rates by category, as functions of a single measure of labor-market tightness, the duration of vacancies as reported in JOLTS as the ratio of the stock of vacancies to the flow of hires. The coefficient for each category is a measure of matching efficiency—it is the component of the job-finding rate not accounted for by cyclical movements in overall tightness.

The primary finding of the paper is that matching efficiency does not track the business cycle—efficiency at the category level has moderate trends but little cycle. Neither the plunge in matching efficiency in recessions nor the gradual growth in recoveries could reasonably be considered as an exogenous driving force. Rather, our finding displayed in Figure 24 of strong cyclical shifts in efficiency reflects major cyclical changes in the composition of unemployment.

When unemployment is high at the beginning of a recovery, a larger fraction of the unemployed belong to low-matching-efficiency groups, notably losers of permanent jobs. The DMP model has the property that unemployment is higher in markets for workers with lower matching efficiency. Matching efficiency declines in recessions as the pool of the unemployed shifts towards workers with low job finding rates (Sahin, Song, Topa and Violante (2014), Hornstein and Kudlyak (2016), Hall and Schulhofer-Wohl (2018)). Then aggregate matching efficiency gradually improves during recoveries. This behavior contributes to the gradual downward glide of unemployment in recoveries.

A full treatment of this effect would make each category of unemployment a separate state variable of the model, but the basic effect can be modeled in the framework of this paper, with a single state variable.

8.6 Endogenous impaired profitability when unemployment is high

Other models have incorporated the property that feedback from unemployment causes $P - W$ to decline when unemployment is high.
One way for such a feedback is when $P$ declines more than $W$ in response to higher unemployment. Ljungqvist and Sargent (1998) propose a model in which workers accumulate skills on the job and lose skills during unemployment, while their non-employment option remains unchanged. In turbulent economic times, the loss of skills is faster and the decline in productivity is greater.

Eeckhout and Lindenlaub (2019) observe that, in a recession, the composition of job-seekers shifts toward the unemployed and away from on-the-job searchers. Incentives for job creation are diminished because the productivity of job-seekers is lower.

Mercan, Schoefer and Sedlacek (2020) propose a model in which newly hired workers are imperfect substitutes for seasoned workers. In their model, a greater share of the unemployed among the potential new hires in the recessions discourages job creation and helps explain the persistence of aggregate unemployment following an adverse shock.

### 8.7 Gradual decline of the separation rate as a function of the unemployment rate

The separation rate $s$ is another parameter of the DMP model that could contribute to the explanation of the slow recovery of unemployment. An elevated separation rate shifts the stationary locus in the DMP phase diagram to the right. A gradual decline in the separation rate results in a gradual decline in unemployment.

Our data on the separation rate support that account of unemployment persistence. The separation rate spikes in a recession, then slowly returns to its normal level.

Research in the VAR framework provides evidence for the importance of job loss in understanding unemployment dynamics. Fujita (2011), using structural VARs with sign restrictions, finds that when unemployment increases and vacancies drop, the separation rate and gross separations rise quickly and remain persistently high. Barnichon (2012) finds that the separation rate accounts for about 40 percent of unemployment’s variance and contributes to about 60 percent of unemployment steepness asymmetry—the observation that unemployment increases faster than it decreases. Using wavelets analysis, Portugal and Rua (2020) find that the job separation rate accounts for, on average, 53% of the unemployment variability at short-run fluctuations during recessions.

Fujita and Ramey (2012) find that the DMP model with endogenous separation can generate realistic volatility and productivity responsiveness of the separation rate and worker flows. It, however, fails to generate sufficient volatility of the job finding rate.
8.8 Conclusions about the variants of the DMP model with adverse feedback through unemployment

The simple DMP model with a job-finding rate calibrated to the high observed rate of monthly unemployment-to-job transitions generates a severely unrealistically rapid recovery, as Cole and Rogerson (1999) observed early in the development of the model. Using an effective unemployment exit rate inferred from the 3-year histories of displaced workers improves the fit but still does not match the data over the 10 year span in our example.

Our main focus is the mechanisms in which driving forces endogenously impair tightness when unemployment is high. In contrast, the basic model has completely constant driving forces, so the dynamics come entirely from the law of motion and there is no movement of tightness.

The mechanisms of the slow downward trend in unemployment in recoveries based on feedback from unemployment that we discuss involve, when unemployment is high:

1. higher recruiting costs early in the recovery,
2. congestion in recruitment,
3. externalities from recruitment selection,
4. lower matching efficiency,
5. impaired profitability of new matches, and
6. persistently higher separation rates

9 Other Forces Operating during Recoveries

The DMP model provides a disciplined framework for studying the issues considered in this paper. But a great deal of business-cycle thinking occurs outside the DMP framework. In this section we examine the behavior of policy instruments and other potential driving forces without trying to determine how the might operate through the DMP model.

Next we take a look at a variety of macro variables that may be involved in recoveries. These are policy instruments—government spending and monetary policy—and influences that might be considered exogenous determinants—productivity, labor-force growth, and the stock market.

We use the NBER business-cycle chronology, so that our timing results are measured over the general business cycle, rather than a cycle pertaining specifically to unemployment.
Figure 28: Real Government Purchases of Goods and Services during Cyclical Recoveries, as the Ratio to Potential GDP, Quarterly

### 9.1 Fiscal and monetary policy

**Government purchases.** Figure 28 displays consolidated government purchases of goods and services divided by the CBO’s potential GDP series. The dates of peaks in the business cycle appear along the bottom—not the peaks in the purchases series itself. Essentially all macroeconomic models agree that an increase in government purchases stimulates output. The figure shows that purchases in the first recovery, 1949 through 1953, grew rapidly because of the Korean War. The Reagan military buildup in the 1980s also accounted for rising purchases relative to potential GDP in that recovery—in all other recoveries, even the one in the 1960s containing the Vietnam war, purchases failed to keep up with potential GDP. The conclusion with respect to those, notably including the most recent recovery, is that fiscal policy taking the form of deliberate expansion of purchases—such as the American Recovery and Reinvestment Act—provided stimulus when the economy was weak. As the economy recovered, the stimulus was withdrawn. This influence was much greater, relative to potential GDP, in the current recovery, compared to earlier recoveries.

**Government transfers.** The US has large and effective countercyclical government transfer programs and practices. Figure 29 shows the history of dollar benefits in terms of our unemployment recovery chronology. We standardize the data by dividing by nominal disposable income. Some of the countercyclical pattern arises from automatic stabilizers—programs that enroll more dependents in bad times—and some from discretionary expansion
of programs and creation of new ones—such as extending unemployment insurance benefits to cover more weeks.

The figure shows that there is a good deal of heterogeneity across the recoveries.

**Monetary policy.** The central instrument of monetary policy in the US is the Federal Reserve’s policy interest rate. The standard way to state its effect as an instrument is to define it as the margin of the economy’s natural or equilibrium short interest rate over the policy rate. To expand, the Fed depresses the policy rate and increases the margin. And to contract, the Fed raises the policy rate above the natural rate to drive the margin negative. Laubach and Williams (2003) is a widely used estimate of the natural short rate.

Figure 30 shows the expansionary margin of interest-rate policy, according to Laubach and Williams. The Fed has chosen net expansion in four expansions and net contraction in two. In the recovery from the 2007-09 recession, the Fed has chosen substantial expansion, almost as much as in the recovery of second half of the 1970s. Oddly, the late 1970s were a period of high and rising inflation, so the Fed was failing in its duty to lean against the wind.

As with the other policy instruments, we find heterogeneity in the setting of the Fed’s interest-rate margin during the recoveries of the past 70 years.
9.2 Other forces during recoveries

Financial discounts. Forces other than macroeconomic policy may influence unemployment declines during recoveries. For example, a recent literature has described a relation between financial discounts and unemployment. See Hall (2017) in the context of the aggregate labor market and Kilic and Wachter (2018) and Kehoe et al. (2020) in general equilibrium. These papers consider DMP-type models of unemployment and events that alter economy-wide discount rates, thus changing the job-value, which is the present value of the contribution of a newly hired worker net of the wage paid to the worker. Discounts sometimes jump upward almost discontinuously, as they did immediately after the Lehman bankruptcy in 2008. The job value represents the incentive to recruiting. When it declines, the labor market slackens and unemployment rises. In the recovery phase, falling discounts raise the job value and unemployment falls.

According to principles of modern finance elucidated in Campbell and Shiller (1988), discount rates for risky future cash payouts are equal to the expected rates of returns associated with those payouts. In a recovery, the stock market rises, the price/dividend ratio rises, and expected rates of return decline. According to the literature linking financial events to the labor market, unemployment declines back to normal. Figure 31 shows the history of the ratio for recoveries since 1949. The ratio rose dramatically during the recovery of the 1990s. It fell substantially during the financial crisis in 2008 and 2009, but recovered during 2010,
when unemployment was still rising. Its relation to the business-cycle chronology in earlier years is less apparent. A rising price/dividend ratio is sometimes important for a recovery, but does not explain the reliability of US business-cycle recoveries.

**Productivity growth.** Another aggregate influence of unquestioned importance for GDP growth is productivity growth. If the topic of this paper were real GDP growth in recoveries, productivity would receive top billing. But the relation of productivity growth to the gradual rise of economic activity in recoveries is ambiguous and may well be small. Figure 32 shows that productivity level. The productivity growth tended to be high in recoveries through the 1980s, had a small comeback in the recovery starting in 2003, and had a spectacular shortfall in the recovery from the 2007-09 recession. Overall, productivity growth tended to be irregular in recoveries.

**Variations in labor-force growth.** The DMP model of Mortensen and Pissarides (1994) has a constant labor force. Extensions to endogenous participation may involve positive or negative co-movements of participation and unemployment. Figure 33 shows that the participation rate grew during the years up to 1990 when the rising rate for women was a key factor for overall participation (to achieve a basic adjustment for demographic influences, the data refer to ages 25 through 54). In the the recovery from the 2020 recession, participation was essentially unchanged. In the recovery from the 2007-09 recession, participation declined.
Figure 32: Total Factor Productivity during Unemployment Recoveries, Quarterly

Figure 33: Labor-Force Participation Rate during Business-Cycle Recoveries
9.3 Discussion of policies and other forces operating during recoveries

Based on this evidence, we conclude that the economy includes a strong internal force toward recovery that operates apart from policy instruments and from financial developments or productivity growth. Policymakers understand this point and withdraw expansionary policies as the internal force does its job.

We should make it clear that optimal policy that resulted in uniform growth of economic activity might be quite irregular as it fights off disturbances, so the irregularity of instruments is not conclusive evidence of the irrelevance of policy. That is, the evidence of stable outcomes and unstable policy instruments is also consistent with the view that policymakers understand the workings of the economy well and deploy the instruments to deliver stable outcomes.

Our tentative conclusion that policy has little impact on unemployment still leaves room for effective policy to prevent or moderate recessions.

10 Inference about the Mechanisms of Unemployment Recoveries

We have discussed a variety of mechanisms propelling the consistent recovery of unemployment after recessions result in spikes of unemployment. In those based on the DMP model, recoveries are endogenous—there is a natural force causing job-seekers to match up with available jobs and thus to lower unemployment. Our discussion showed that recoveries from high unemployment is rapid in the basic DMP model, so there is no doubt that recoveries occur, and the main question is why the actual recovery process is so slow. Our Figure 27 demonstrates that a model with negative feedback from unemployment to tightness provides an internally consistent version of the DMP model with much slower, but equally reliable, recoveries. We do not claim that we have ruled out alternative mechanisms.

We have also investigated a range of recovery mechanisms exogenous to the labor market. Our main tool is to study correlations. For example, we find that credit conditions tend to ease systematically following a recession, so we entertain that variable as a potential exogenous driving force. And we reject the hypothesis that unemployment recovers because productivity grows faster than average in recoveries because that is not what we find in the data.

Our tentative conclusions do not rise to the level of firmly established causal inference. Much more remains to be done.
11 Concluding Remarks

Why has the US economy recovered so consistently from every recession in the past 70 years? Our answer is that the labor market operates according to the principles of Diamond, Mortensen, and Pissarides, with one major new element: unemployment itself inhibits the rebuilding process that follows a recession that has caused a spike in unemployment. Strong negative feedback results in slow removal of excess unemployment.

Our view of the recovery of the US economy from a recessionary shock differs from the standard view. In the standard view, unemployment is high following a recessionary shock because there is a shortfall of demand. As time passes, demand recovers and unemployment returns to normal. Under the standard view, the reliable persistence of unemployment during recoveries arises from persistence in demand.

In our view, unemployment remains high after a recession and declines only gradually, because of frictions in rebuilding employment. These frictions impede both the individuals who lost jobs from the recession and those who did not, but found it more difficult to navigate the labor market.

According to this view of the labor market, the average level of unemployment depends on the frequency and severity of recessionary shocks. The natural rate of unemployment is not immutable. Instability arising from monetary policy shocks prior to the 1990s and financial shocks since then tended to elevate average unemployment, while long stretches of stability in the 1990s and 2010s demonstrated that the economy could achieve unemployment around 3.5 percent.
References


Gavazza, Alessandro, Simon Mongey, and Giovanni L. Violante, “Aggregate Recruiting

Gertler, Mark and Antonella Trigari, “Unemployment Fluctuations with Staggered Nash

Hall, Robert E., “High and Low Unemployment Equilibria, Self-Selection, and Screening in
the Labor Market,” February 1990.


____ and Marianna Kudlyak, “Job-Finding and Job-Losing: A Comprehensive Model of
Heterogeneous Individual Labor-Market Dynamics,” Working Paper 2019. Hoover In-
stitute.


____ and Sam Schulhofer-Wohl, “Measuring Job-Finding Rates and Matching Efficiency


Hornstein, Andreas and Marianna Kudlyak, “Estimating Matching Efficiency with Variable

Jacobson, Louis S., Robert J. LaLonde, and Daniel G. Sullivan, “Earnings Losses of Dis-

Kehoe, Patrick, Pierlauro Lopez, Virgiliu Midrigan, and Elena Pastorino, “Asset Prices and


