The Beveridge Curve: A Survey†

MICHAEL W. L. ELSBY, RYAN MICHAELS, AND DAVID RATNER*.

Important progress has been made in economists’ understanding of the Beveridge curve, from its measurement to its expression in canonical labor market models. Yet enduring puzzles remain. Chief among these are the empirical role of vacancies in the recruitment process; the amplitude, comovement, and persistence of cyclical unemployment–vacancy dynamics; and the sources of lateral shifts in the Beveridge curve. The synthesis of these themes identifies several priorities for ongoing research, including the role of entry costs into vacancy creation in shaping Beveridge dynamics; the cyclicity of search intensity, both off and on the job, and its relation to participation and job-to-job transitions; the theory and measurement of mismatch; and the sources of hysteresis in unemployment flows. (JEL E24, J63, J64)

1. Introduction

The Beveridge curve relation between unemployment and vacancies has emerged in recent decades as a central organizing framework for macroeconomists’ understanding of labor markets. Its conception can be traced back to the formative early work of William Beveridge, who was instrumental in establishing the essential role of vacancies in the determination of unemployment (Beveridge 1944).[1] Seventy years later, the typically inverse relationship between job openings and jobseekers has been shown to have fundamental implications for the efficiency of the matching process that generates employment relationships, and for the nature of shocks that drive fluctuations in the labor market. As a consequence, the Beveridge curve has played a

---

[1] Even in 1944, Beveridge’s emphasis on the role of vacancies as a complement to unemployed jobseekers was longstanding. He had, for example, been an early pioneer of the inception of government-run employment exchanges in Britain in 1910. However, as Yashiv (2008) notes, Beveridge himself did not plot the relation between unemployment and vacancies, although he did offer an empirical analysis of both variables that implied a negative association. The original Beveridge curve plot was instead provided by Dow and Dicks-Mireaux (1958), who also offered an early interpretation based on imbalances in labor supply and demand. Rodenburg (2011) provides an excellent survey on the history of thought on the Beveridge curve.
pivotal role in debates over the functioning of labor markets and has shaped the canonical modern approach to understanding the coexistence and volatility of unemployment and vacancies.

Two key stylized facts have occupied research on the Beveridge curve. First, at cyclical frequencies unemployment and vacancies move in opposite directions, tracing out a negatively inclined Beveridge locus. Second, the position of this locus has shifted periodically in many developed economies, most notably during the persistent rise in European unemployment in the 1980s, and more recently in the wake of the Great Recession in the United States. These facts have been reaffirmed by recent advances in the measurement of vacancies pioneered by the introduction of the Job Openings and Labor Turnover Survey in the United States.

An enduring theme in the analysis of the Beveridge curve has been the interplay of these stylized facts with the development of economic theories of labor market fluctuations. Early work highlighted the negative cyclical comovement between unemployment and vacancies as a hallmark of business cycles that generate broad-based aggregate fluctuations in the demand for labor, as opposed to the reallocation of labor from shrinking to expanding markets (Abraham and Katz 1986). Medium-run shifts in the Beveridge curve were instead identified as a manifestation of changes in the pace of worker reallocation. These can be driven, for example, by demographic shifts associated with the entry of the baby boom generation into the labor market (Abraham 1987; Shimer 2001) or, more alarmingly, declines in the efficiency of the hiring process. The latter argument featured prominently in debate over European unemployment in the 1980s, as well as recently in the United States.

These insights fueled the parallel development of theories of labor market frictions forged by the work of Diamond (1982), Pissarides (1985), and Mortensen and Pissarides (1994). This view of the labor market envisions unfilled vacancies as the counterpart of unemployed workers who are engaged in a search process. The Beveridge curve that emerges is a close relation of the matching technology that brings together jobseekers and job openings. As a result, it is inextricably linked to the processes of job creation and job destruction, and thereby worker flows (Blanchard and Diamond 1989).

A crucial success of this framework is that it identifies fluctuations in vacancy creation as a central determinant of cyclical movements along the Beveridge curve. Declines in vacancy creation slacken the labor market, reducing job openings and raising unemployment, but also depressing job-finding rates faced by the jobless and raising vacancy-filling rates among employers. These qualitative predictions bear a clear resemblance to the paths of labor market stocks and flows during recessions in modern economies, and contributed to the widespread adoption of the Diamond–Mortensen–Pissarides framework as the canonical workhorse model of the aggregate labor market.

But the empirical Beveridge curve has also informed understanding of the quantitative limitations of the canonical model. The amplitude and persistence of fluctuations in vacancies observed in the data significantly dominate that generated by the dynamics of vacancy creation in conventional applications of the standard model (Shimer 2005). Furthermore, persistent adverse shifts in the Beveridge curve relation, such as that currently experienced in the United States, can be accounted for in the model only by an exogenous deterioration in the efficiency of the matching technology.

These shortcomings have generated a reassessment of the key pillars of vacancy creation in the Diamond–Mortensen–Pissarides model—namely, flexible wages,
and free entry—as well as a revived interest in the microeconomic origins of the matching process. The latter have enriched our understanding of Beveridge dynamics by considering the implications of, for example, on-the-job search, and transitions in and out of nonparticipation. By offering a more explicit treatment of match formation among heterogeneous firms and workers, this line of research has also begun to illuminate the microfoundations of matching, and of the Beveridge curve.

The remainder of the survey is organized as follows. In section 2, we summarize challenges and recent developments in the conceptualization and measurement of vacancies and the stylized facts of Beveridge curve movements in the United States and Europe. Section 3 uses the descriptive model of unemployment and vacancy flows developed by Blanchard and Diamond (1989) as a general organizing framework for thinking about Beveridge curve dynamics. Section 4 presents the canonical Diamond–Mortensen–Pissarides model as a special case of the general framework and highlights its qualitative successes and quantitative limitations as an account of unemployment and vacancy fluctuations.

These observations serve as the point of departure for subsequent sections of the paper, which relax in turn several of the assumptions of the canonical model and describe their implications for the Beveridge curve. Sections 5 and 6 turn to vacancy creation and trace out the implications of alternative wage-setting processes and inelastic entry of vacancies. The remaining sections examine departures from the simple matching process implicit in the workhorse model. Section 7 allows for flows in and out of labor force participation; section 8 investigates the role of movements in worker search intensity; section 9 describes the effects of job-to-job flows; and section 10 surveys various sources of heterogeneity in the labor market, in particular notions of mismatch, and the effects of long-term joblessness in shaping the efficiency of the matching process. In each of the ensuing sections, we highlight how these mechanisms inform our interpretation of the recent evolution of the Beveridge curve in the United States.

2. Concepts and Measurement

Much of the vast literature on the Beveridge curve has sought to provide economic explanations for the presence of a relation between unemployment and vacancies. But, before we turn to such explanations, important conceptual and measurement questions must be addressed: What do economists (and survey respondents) mean when they refer to unemployment and vacancies? How are they measured? And, what are the implications of these for the observed relation between measures of unemployment and vacancies?

2.1 Conceptual Issues

Of the two axes of the Beveridge curve, unemployment is relatively well understood. There is a degree of consensus over its measurement, with many statistical authorities now adopting a definition similar to that endorsed by the International Labour Organization. In the United States, for example, the monthly Current Population Survey classifies an individual as unemployed if she did not have a job in the week containing (typically) the twelfth of the month, actively searched for work in the previous four weeks, and is available to start work. Important ambiguities do exist in the interpretation of search intensity in this definition, and thereby between unemployment and nonparticipation (see, for example, Clark and Summers 1979; Abowd and Zellner 1985; Poterba and Summers 1986). Nonetheless, individuals classified as unemployed are much more likely to transition
into employment than those recorded as out of the labor force, suggesting that job seeking is indeed more prevalent among the unemployed (Flinn and Heckman 1983). We shall return to these measurement issues later in the relevant sections of the paper.

By comparison, much less is understood about vacancies. Both conceptual and measurement issues have contributed to this. At a conceptual level, Abraham (1983) proposes a natural definition of a vacancy as unmet labor demand, by analogy to the view of unemployment as unmet labor supply. This parallel suggests that a vacancy indicates the presence of some idle resource—for example, physical or organizational capital—that the firm is seeking to reactivate. Aspects of this analogy are potentially misleading, however.

First, although it usually is straightforward to identify a worker who is not employed, the corresponding idle resource in the case of a vacancy is less conspicuous. This seems particularly likely if the structure of production within an establishment does not admit a simple mapping from unused capacity to unfilled jobs. For example, it would be convenient if a vacancy were simply an empty desk where work orders pile up, but production structures may be more complicated in practice. Second, and relatedly, it is difficult to identify the amount of desired production not being undertaken on account of an unfilled job. For these reasons, the economic meaning of a vacancy—what it signifies in terms of firm activity or the lack thereof—can be unclear. Third, as noted by Wingeard (1966), Frumerman (1978), and Hall (1978), firms are able to recruit for positions in anticipation that they will be open in the future. Such preemptive vacancies need not represent the presence of unmet demand for labor, or foregone output.

2.2 Vacancy Data

Reliable, timely, and comprehensive survey data on vacancies have only recently been made available since the inception of the Job Openings and Labor Turnover Survey (JOLTS) in December 2000, which significantly advanced the measurement of vacancies. The survey covers a sample of approximately 16,000 establishments each month, and collects data on the flows of hires and separations during the month, as well as the stock of job openings at the end of the month. The design of JOLTS has benefited from several decades of research by the Bureau of Labor Statistics into the viability of a vacancy survey. The JOLTS definition of a vacancy that emerged specifies that a position exists for which work could start within thirty days, and for which the employer is actively recruiting from outside the establishment. While the underlying economic content of this definition can remain unclear (for reasons noted above), pilot and validation studies of JOLTS suggest that survey respondents do interpret vacancies in line with the official definition (Ross 1966; BLS 1991; Levin et al. 2000).

---

2 This is also the interpretation that guided early efforts at vacancy measurement. For instance, it was the definition used by the Bureau of Labor Statistics in three significant pilot studies carried out between 1964 and 1990.

3 A 1965 National Industrial Conference Board study in Rochester, New York, found that 10 percent of advertised positions were filled before the current occupant left the firm, as a means to train the new hire (Myers 1966).

4 A pilot study undertaken in the mid-1960s sampled vacancies at the occupation level. Later, in 1969, a question on job openings was included in BLS's labor turnover survey of manufacturing establishments. Occupational detail was then omitted, but with the intent to incorporate it later (Frumerman 1978). However, the question was removed after fewer than five years for budgetary reasons. Subsequently, vacancy surveys were studied periodically, but assessments by the BLS and others questioned the feasibility of accurate vacancy measurement (National Commission on Employment and Unemployment Statistics 1979; Plunkert 1981). JOLTS sidestepped one major concern simply by not pursuing vacancies at the occupation level, which requires larger samples and more intensive survey design.
Prior to the introduction of JOLTS, empirical research on vacancies was forced to rely on limited data for particular states, and especially on proxy evidence from the Conference Board’s Help-Wanted Index. The latter was based on counts of help-wanted ads placed in newspapers in fifty-one large US cities. Given the lack of alternatives, the majority of research on the Beveridge curve in the United States has relied by necessity on the Help-Wanted Index prior to 2000 and the JOLTS vacancy data thereafter. Perhaps surprisingly, Shimer (2005) and Barnichon (2010) observe that these two measures of vacancies comove closely during the short period over which the series overlap, once account is taken of downward trends in newspaper vacancy posting, and the associated rise in online advertising. Indeed, Barnichon uses this property to construct a “composite” vacancy index that splices together these two measures. This composite series is identical to the Help-Wanted Index prior to 1995, and mirrors the JOLTS data subsequent to 2000. To fill in the intervening years, Barnichon notes that the trends in the share of newspaper advertising—implied by either rescaled JOLTS data or the Conference Board’s recently assembled Help-Wanted Online Index—follows a stable downward trend. This trend is used to backcast the composite series between 1995 and 2000.

Figure 1A uses this composite vacancy series to document the reduced-form empirical relationship between unemployment and vacancies in the United States. The figure highlights two key stylized facts of the Beveridge curve relation that have occupied the literature. First, at cyclical frequencies, rises in unemployment during recessions are met with significant declines in vacancies, evocative of the inverse relationship noted originally by Beveridge (1944). Second, clear lateral shifts are apparent at lower frequencies. As stressed by Abraham (1987), a prominent example is the sustained outward shift in the Beveridge curve during the 1970s. More recently, the outward shift in the Beveridge curve in the wake of the Great Recession has loomed large in academic and policy debates (Elsby, Hobijn, and Şahin 2010; Davis, Faberman, and Haltiwanger 2013). The sources and implications of these movements of unemployment and vacancies will act as our point of departure for the remainder of the survey. Before we do so, however, the next subsection considers the extent to which these stylized facts could be a symptom of measurement problems, rather than robust economic phenomena to be explained.

2.3 Measurement Challenges

Abraham (1987) highlights several important shortcomings of the Help-Wanted Index as a measure of vacancies. First, the mapping from the number of newspaper ads to aggregate vacancies is potentially impeded by the fact that one ad may correspond to many vacant positions, and that not all search on behalf of employers is mediated through newspaper ads. Second, the index is sensitive to idiosyncratic trends in the newspaper industry. Abraham highlights the roles of newspaper consolidation and equal employment opportunity laws in inducing an upward trend in the index during the 1960s. More recently, shifts toward online media have induced a steady downward drift in print advertising, as noted above.

In the light of these drawbacks, it is natural to question whether the substantial outward shift in the 1970s in figure 1A is a spurious result of these shortcomings. Abraham (1987) carefully investigated adjustments of the data for occupational composition, employer advertising practices,
and newspaper consolidation. While these factors account for some of the observed movement, Abraham concludes that much of the outward shift in the Beveridge curve in the 1970s was in fact not spurious.

Although the introduction of JOLTS represented a significant advance in the tracking of vacancies, it is not immune from measurement issues. The BLS has often noted that reports from mid-sized employers in particular can be prone to error. These employers are too large to recall exactly job openings from memory, but too small to invest resources in tracking them systematically.\(^6\) Relatedly, a perennial concern has been that vacancies are undercounted (National Commission on Employment and Unemployment Statistics 1979). Recent research pioneered by Davis, Faberman, and Haltiwanger has developed new evidence on this and is relevant to trends in the measured Beveridge curve.

Davis, Faberman, and Haltiwanger’s (2013) investigation of establishment-level observations on hires and vacancies reveals that as much as 40 percent of hires in a typical month are made at establishments that do not report vacancies at the end of the previous month. Based on a formal model of time aggregation, Davis et al. estimate that around two-thirds of such hires can be attributed to vacancies that are posted, and filled, within a given month. The monthly frequency of the JOLTS survey means that such vacancies are missed.

\(^6\)We have benefited from correspondence with John Wohlford of BLS on this topic.
As Davis, Faberman, and Haltiwanger emphasize, however, the remaining one-third of such hires may reflect a deeper, conceptual incompatibility between actual hiring behavior and the notion frequently invoked in the majority of the theoretical literature that hires are mediated entirely through vacancies. For instance, Davis et al. note that many job openings may not be filled via the “active recruiting” specified by the JOLTS vacancy definition. To take an example, employers may engage in “opportunistic” hiring, whereby an employer without job openings pulls forward hiring to recruit an attractive candidate who applies unexpectedly for work.\(^7\) This reiterates the conceptual difficulties associated with defining vacancies noted earlier in this section.

A further concern with the JOLTS vacancy data relates to its sampling frame. Davis, Faberman, and Haltiwanger (2013) and Davis, Faberman, Haltiwanger, and Rucker (2010) emphasize that, by excluding young, fast-growing establishments, the JOLTS sample misses a nontrivial portion of job openings. Indeed, Davis et al. (2010) estimate for the period 2001 to 2006 that the published vacancy series from JOLTS is understated by as much as 8.5 percent as a result.

Both of these measurement issues in the JOLTS could conceivably vary over time, for example with the business cycle, and thereby distort the picture of the Beveridge curve depicted in figure 1A. Evidence on time variation in the prevalence of hires that are not mediated through vacancies, and on the share of vacancies at young establishments, is not yet available. But, Davis, Faberman, and Haltiwanger’s results underscore the value of such research going forward.

### 2.4 Beyond the United States

We close this section with a brief discussion of efforts to measure vacancies outside of the United States. Job vacancy survey data similar to JOLTS have only recently become available for many countries in Europe.\(^8\) For years prior to 2008, European data on vacancies were based on the number of job openings that employers voluntarily registered with government-administered employment agencies. A perennial concern is that such registered vacancies fail to be representative of all job openings. Moreover, methodological inconsistencies can arise from changes in registration procedures and employers’ propensity to register job openings. For instance, in 2001, the U.K. government requested that vacancies be registered with regional offices, rather than more local job centers. This change in procedure was disruptive enough that the estimates were affected and publication of the results was temporarily suspended.

Although these concerns should be borne in mind, available data have revealed such substantial shifts in the Beveridge curve that there is typically thought to be some signal amid the noise. Jackman et al. (1990), for example, plot Beveridge curves for sixteen OECD economies, with registered vacancy

\(^7\)This is also reminiscent of the notion of “gate hiring” noted by Myers (1966), whereby firms can save on the cost of advertising because workers check in to see if there are openings. Interestingly, the 1979 National Commission on Employment and Unemployment Statistics report also cautioned against the requirement of “active recruiting” in the survey definition of a vacancy. More recently, Faberman and Menzio (2010) report that 20 percent of new hires in the 1982 wave of the Employment Opportunity Pilot Project Survey involve no formal vacancy or recruiting time.

\(^8\)These vacancy surveys fulfill the requirements of the European Commission’s Regulation 453, which mandated in 2008 that Member countries administer “business surveys” of vacancies. The definition of a vacancy is very similar to that used in JOLTS, although differences in definitions and coverage still make cross-country comparisons difficult (Kettner and Stops 2008). Only a few countries—namely, Australia, the Netherlands, Sweden, and the United Kingdom—administered vacancy surveys before 2008. The United Kingdom’s push to institute a vacancy survey was partly motivated by the suspension of its “registered” vacancy series, discussed in the text.
data stretching back in some cases to the 1950s. They observe prominent rightward shifts during the 1970s and 1980s, particularly among European economies, suggesting that the European unemployment problem at the time coincided with a deterioration in the Beveridge curve relation. Arguably, these developments do reflect a real change in labor market functioning, measurement issues notwithstanding.

Hobijn and Şahin (2013) update the OECD data on registered vacancies used by Jackman et al. to document recent movements in Beveridge curves across countries. Figure 2 presents these data for a selection of European countries to give a sense of the diversity of experiences across these economies. As in the US case illustrated in figure 1A, Beveridge curves across countries typically exhibit negative comovement between unemployment and vacancies at higher frequencies, combined with periodic lateral shifts in the relation at lower frequencies. The patterns for France and the United Kingdom are archetypal of the European unemployment problem of the 1980s. Both economies experienced significant rises in unemployment over this period accompanied by only small declines in vacancy rates. By contrast, Sweden, like other Nordic countries, was spared from the pathologies of labor market outcomes experienced in the 1980s in other European economies, and exhibited a relatively stable Beveridge curve over this period. With the onset of a financial crisis in 1990, however, Sweden’s Beveridge curve shifts out substantially to the right. Finally the evolution of the Beveridge curve in the Netherlands mirrors the US experience, with significant cyclical loops around a relatively stable downward-sloped locus.

3. A Descriptive Model of Beveridge Curve Dynamics

To organize ideas for the remainder of the paper, it is useful to consider a descriptive model of the joint dynamics of unemployment and vacancies in the spirit of Blanchard and Diamond (1989). The key ingredient of this model is a system of differential equations that governs the evolution of the number of unemployed workers $U$ and the stock of vacancies $V$ over time. These dynamics can be traced to the underlying worker and job flows that shape the processes of the hiring and separation of workers to and from jobs. A particularly simple, and often used example is the following description of Beveridge curve dynamics,

$$\frac{dU}{dt} = \lambda (L - U) - m(U, V), \quad \text{and}$$

$$\frac{dV}{dt} = \gamma - m(U, V).$$

Here, $\lambda$ is the rate at which employed workers flow from employment to unemployment, $L$ is the labor force, and $m(U, V)$ is a matching function that describes the flow of hires generated by $U$ unemployed workers and $V$ unfilled vacancies. $\gamma$ denotes the flow of new vacancies.

Clearly, the differential system in equation (1) necessarily misses some potentially important aspects of unemployment and vacancy dynamics: it abstracts from labor force entry and exit; it omits job-to-job flows; it abstracts from heterogeneity and segmentation in the labor market across sectors; and so on. These issues will be the subjects of subsequent sections of the paper. Nonetheless, equation (1) is a useful starting point for organizing thoughts on the Beveridge curve and its dynamics. To begin, it captures an interpretation of the Beveridge curve that has pervaded much of modern research on the topic, namely that it is the locus of
unemployment and vacancy combinations consistent with a stable level of unemployment, $\frac{dU}{dt} = 0$. From equation (1), this implies that

\begin{equation}
\frac{dU}{dt} = 0 : \lambda(L - U) = m(U, V).
\end{equation}

Under this interpretation, the Beveridge curve is a close cousin of the matching...
function $m(U, V)$ that has been the subject of another JEL survey by Petrongolo and Pissarides (2001). A lesson from their analysis is that estimates of the matching function broadly support the notion that it exhibits constant returns to scale. Under this restriction, the Beveridge curve can be expressed as

\[
\frac{dU}{dt} = 0 : \lambda = m \left( \frac{u}{1-u}, \frac{v}{1-u} \right),
\]

where $u \equiv U/L$ and $v \equiv V/L$ are the associated unemployment and vacancy rates.

This simple description of the Beveridge curve has immediate implications for the empirical relationship between vacancies and unemployment. In particular, under a stable matching function relation between hires and searchers, the Beveridge curve is predicted to take the form of a downward sloping locus in $(u, v)$ space for a constant rate of job loss $\lambda$. This prediction resembles the inverse empirical relationship traced out over the business cycle in figures 1 and 2. In addition, equation (3) suggests that changes in the rate of job loss $\lambda$ induce shifts in this equilibrium locus. Increases in $\lambda$ raise the inflow into unemployment for a given level of vacancies, shifting out the Beveridge curve relation.

A virtue of this predicted link between the flow of workers into the unemployment and the Beveridge curve is that it can be confronted with available data on unemployment flows. Shimer (2012) uses data from the Current Population Survey on employment and unemployment by duration, together with the law of motion for the stock of unemployed workers in (1), to infer the inflow rate $\lambda$. Panel A of figure 3 plots the time series of this measure from 1948 to 2013. As can be seen, the unemployment inflow rate displays substantial variation at both high and low frequencies. Over the business cycle, the inflow rate is prominently countercyclical, rising at the onset of recessions, and receding quickly in subsequent recoveries. Over the medium term, $\lambda$ rose during the 1970s and early 1980s, and has trended downward since then. Shimer (2001) has argued that these trend movements in the degree of labor market churning can be traced to the entry of the baby boom generation into the labor force.

It is possible to use these data to perform a counterfactual exercise that can be applied to this simple model, as well as to several extensions considered in the remainder of the paper. To this end, we note that the vertical shift in the Beveridge curve induced by changes in the unemployment inflow rate $\lambda$ can be obtained by differentiating (3), holding $u$ fixed:

\[
\frac{d \ln v}{d \ln \lambda} \bigg|_{u, \dot{u}=0} = \frac{1}{1 - \alpha},
\]

where $\alpha$ denotes the elasticity of the matching function with respect to unemployment. It follows that we can sketch out a $\lambda$-constant unemployment-vacancy locus by subtracting the shifts implied by equation (4) from

---

10 A notable omission from this formulation is the flow of workers who are recalled by the same firm from which they separated. Fujita and Moscarini (2013) use SIPP data to show that 45 percent of unemployment spells end in recalls. They argue that recalled workers (typically) are not active searchers. In that case, the Beveridge curve would steepen in $(u, v)$ space, principally because a decline in vacancies no longer affects the unemployment outflow rates of the temporarily laid off. This effect may also have a cyclical component, as the temporary layoff share tends to be high in recessions. These rises in temporary layoffs tend to be especially short-lived, however.

11 For recent evidence emphasizing these dynamics, see Elsby, Michaels, and Solon (2009), Fujita and Ramey (2009), and Yashiv (2007). An abundant earlier literature also emphasized these stylized facts, including Perry (1972), Marston (1976), and Blanchard and Diamond (1990).
the observed path of vacancies. To be exact, starting from some initial date 0, the counterfactual vacancy series can be computed as \( \tilde{v}_t = v_t (\lambda_t / \lambda_0)^{-1/(1-\alpha)} \). For the initial condition, \( \lambda_0 \), we typically will use the first available data point, and for the matching
elasticity $\alpha$, we use 0.5 in light of evidence from Petrongolo and Pissarides (2001).\footnote{Consistent with our assumption of a constant matching elasticity, Petrongolo and Pissarides (2001) survey evidence suggesting that a Cobb–Douglas matching function describes the data well. In addition, their summary of leading estimates of $\partial \ln m / \partial \ln u = \alpha$ suggests that $\alpha$ lies in the range of 0.4 to 0.7.}

Note that the thought experiment here is not to claim that the path $(u_t, \tilde{v}_t)$ is the equilibrium path that would be realized in the absence of movements in the unemployment inflow rate. Rather, it is sketching the implied counterfactual shape of the Beveridge curve locus in (3), which is one input into that equilibrium.

Figure 1B reveals that this exercise provides an instructive interpretation of the shifts in the realized Beveridge curve documented in figure 1A. Two stark observations emerge. First, consistent with the trend rise and fall in the measured inflow rate during the 1970s and 1980s shown in figure 3A, the outward shift in the realized Beveridge curve over that period vanishes in the counterfactual locus. Low-frequency movements in rates of worker reallocation are able to account for much of the historical Beveridge curve shifts in the 1970s and 1980s.\footnote{The OECD (2014, chapter 1) summarizes estimates of similar constant inflow-rate counterfactual Beveridge curves for a large set of countries.}

By contrast, the second message of figure 1B is that the outward shift in the Beveridge curve since the late 2000s is thrown into even sharper relief once account is taken for the path of unemployment inflows. While this outward shift is discernible in the realized $(u, v)$ dynamics in figure 1A, the shift is all the more extreme, and conspicuous relative to the past, in the counterfactual locus in figure 1B. The reason can be traced to the path of the unemployment inflow rate. As emphasized in figure 3A, $\lambda$ recovered quickly in the wake of the Great Recession, and since then appears to have continued on its downward prerecession trend. The outward shift in the realized Beveridge curve since the Great Recession has therefore occurred despite recent movements in worker reallocation. Viewed through the lens of equation (3), this recent shift can be rationalized only if the efficiency of the matching technology $m(u, v)$ has deteriorated since the Great Recession.

These observations highlight the usefulness of the above interpretation of the Beveridge curve. But the picture painted by equation (3) is also necessarily simplified, abstracting as it does from the richer detail of labor markets. In later sections of the paper, we will enrich the model to account for flows of workers in and out of the labor force, between jobs, and so on. For now, however, we address a more fundamental issue: that equation (3) is just one of the equilibrium conditions that determine the joint dynamics of unemployment and vacancies. To close the model, one must take a stand on the determination of vacancies. The leading approach to this question was pioneered by the work of Diamond, Mortensen, and Pissarides on search and matching models.

### 4. The Canonical Search Model as a Benchmark

The simplest and most widely used model in the search and matching tradition was first formulated in Pissarides (1985). Unemployment in the steady state of this model is governed by equation (2) above, so that all the implications of that relationship noted thus far apply. What the model provides in addition is a theory of the flow of new vacancies, $\gamma$, which closes the system of equations in (1). Since this theory has become the canonical approach to understanding the joint dynamics of unemployment and vacancies, and because it is a useful benchmark point of comparison for what follows, we review its structure here.
4.1 The Pissarides (1985) Model

Production is organized in worker–firm pairs. A filled job produces a flow output $p$, and pays a wage $w$. As above, jobs are filled via a constant-returns matching technology $m(u, v)$, and are subsequently destroyed at rate $\lambda$. Unemployed workers receive a flow payoff $b < p$ that reflects the value of leisure, home production, and unemployment benefits, and find jobs at a rate $m(u, v)/v \equiv f(\theta)$. Under constant returns, the job-finding rate can be expressed simply in terms of the ratio of vacancies to unemployment $\theta \equiv v/u$, commonly referred to as labor market tightness. With time discount rate $r$, then, the values to a worker of unemployment $\mathcal{V}$, and employment $\mathcal{W}$, are given by

$$\begin{align*}
\mathcal{V} &= b + f(\theta)(\mathcal{W} - \mathcal{U}), \\
\mathcal{W} &= w + \lambda(\mathcal{U} - \mathcal{V}).
\end{align*}$$

From the firm’s perspective, a vacancy is subject to a flow cost $c$ for the period that it remains unfilled, and is filled at rate $m(u, v)/v \equiv q(\theta)$. Thus, the values to the firm of a vacant job $\mathcal{V}$, and a filled job $\mathcal{J}$, can be expressed as

$$\begin{align*}
\mathcal{V} &= -c + q(\theta)(\mathcal{J} - \mathcal{V}), \\
\mathcal{J} &= p - w + \lambda(\mathcal{U} - \mathcal{J}).
\end{align*}$$

Pissarides (1985) invokes two assumptions to close the model. The first is that there is free entry into vacancy creation. As a result, vacancies are opened or closed until the value of a vacancy is driven to zero, $\mathcal{V} = 0$. Free entry thus implies that

$$\mathcal{J} = \frac{p - w}{r + \lambda} = \frac{c}{q(\theta)}.$$

That is, the value of a filled job must equal the expected cost of a vacancy. The second key assumption of Pissarides (1985) is that wages are determined according to an asymmetric Nash bargain. The presence of search frictions generates a surplus from continuing employment relationships, as it is costly for both firm and worker to find alternative matches. Nash bargaining imposes that this surplus is split in fixed proportions so that

$$(1 - \beta)(\mathcal{W} - \mathcal{U}) = \beta(\mathcal{J} - \mathcal{V}),$$

where $\beta$ indexes the bargaining power of the worker. The wage equation that emerges from these assumptions has the form

$$w = \beta(p + c \theta) + (1 - \beta)b.$$
ray and the downward-sloped Beveridge curve in equation (2) determines equilibrium vacancies and unemployment.

4.2 Qualitative Implications

This simple theory provides a rich set of implications for the nature of the Beveridge curve and its dynamics, many of which can be confronted with available data. In particular, the theory delivers clear predictions for the movements in unemployment and vacancies implied by three types of disturbances: shifts in the productivity of labor $p$; changes in the pace of job destruction $\lambda$; and movements in matching efficiency embodied in $m(u, v)$.

Figure 4 sketches the implied responses of the labor market to these shocks. Panel A depicts the effects of increases in the job destruction rate, or decreases in matching efficiency. These mirror the effects already noted above in equation (3): they shift the Beveridge curve outward. The model of vacancy creation adds a further nuance to the story, however. Increases in $\lambda$ also reduce the present value of a match; and declines in match efficiency also reduce the rate at which vacancies can be filled $q(\theta)$, increasing recruitment costs $c/q(\theta)$. Both reduce job creation incentives. As Panel A of figure 4 demonstrates, this additionally pivots the job creation condition clockwise. It is possible to show, however, that reasonable parameterizations of the model imply that the outward shift in the Beveridge curve dominates, so that vacancies and unemployment comove.

Notes: The figure sketches out the qualitative dynamics of unemployment and vacancies implied by the Pissarides (1985) model. Panel A depicts the effects of an increase in the rate of job destruction ($\lambda$ in the notation of section 4) or a decrease in match efficiency (a reduction in $m(u, v)$ for any given $u$ and $v$). Panel B depicts the effects of a reduction in aggregate labor productivity ($p$ in the notation of section 4). The arrows describe the implied dynamics. Paths without intervening arrows connote jump dynamics; paths with intervening arrows connote persistent dynamics.
positively in steady state. Figure 4A is drawn to reflect this.\(^{15}\)

Panel B of figure 4 traces out the effects of a negative shift in the productivity of labor, \(p\). It is clear from the job creation condition (9) that a decline in labor productivity reduces the equilibrium vacancy–unemployment ratio, and pivots the job creation schedule clockwise. Intuitively, reductions in productivity are not offset fully by declines in wages, and so the labor market must slacken to restore incentives to create vacancies consistent with free entry. In steady state, then, such productivity shocks trace out a negative relationship between vacancies and unemployment along a stable Beveridge curve. This qualitative implication of the standard search model thus provides a potential explanation of the prominent inverse comovement between unemployment and vacancies observed at business cycle frequencies in figure 1.

Interestingly, this interpretation parallels the early insights of Abraham and Katz (1986). They used vacancy data from the Help-Wanted Index as a means to tease out the sources of cyclical fluctuations in the United States, with particular emphasis on two channels: aggregate demand and sectoral shifts (Lilien 1982). Their intuition can be understood in the context of the above discussion of figure 4. Economywide declines in labor demand have the implication that vacancies decline as unemployment rises. In contrast, higher rates of turnover (in some sectors) can increase returns to vacancy creation (in other sectors) as more unemployed workers become available. Based on the dominant negative cyclical comovement of unemployment and vacancies, Abraham and Katz argue that broad-based reductions in labor demand are the proximate determinant of recessions in the United States.\(^{16}\)

A distinctive feature of the above framework is its emphasis on the flows associated with workers moving in and out of unemployment, and of vacancies as they are created and filled. Accordingly, the model’s predictions for Beveridge curve dynamics impart parallel implications for these flows. In particular, as more unemployed workers search among fewer vacancies during recessions, the job-finding rate \(f(\theta)\) is predicted to fall, and the rate at which firms are able to fill their vacancies \(q(\theta)\) to rise.

These predictions are mirrored in the data. Figure 3B plots a measure of the rate at which unemployed workers flow out of the unemployment pool. This is the counterpart to Shimer’s (2012) measure of the inflow rate \(\lambda\) in figure 3A. Consistent with the implications of the model, the unemployment outflow rate is clearly procyclical, falling prominently in downturns and recovering in expansions.\(^{17}\) Likewise, figure 3C depicts a measure of the vacancy-filling rate inferred from the Job Openings and Labor Turnover Survey by Davis, Faberman, and Haltiwanger (2013). As in the model,

\(^{15}\)To see this, fix the unemployment rate \(u\), and recall from equation (4) that the implied Beveridge curve elasticity of vacancies \(v\) to the job destruction rate \(\lambda\) will be on the order of two under a conventional matching elasticity of \(\alpha = 0.5\). By contrast, from (9), a 1 percent rise in \(\lambda\) induces a decline in vacancies of \(1/\alpha[\alpha(r + \lambda) + \beta f(\theta)]\) percent, for given \(u\). For observed values of \(\lambda\) and \(f\), the latter approaches two only if \(\beta\) is in the neighborhood of zero. It follows that, for conventional calibrations, the decline in the expected length of matches is not enough to offset the stimulus to job creation that derives from an increase in the number of job searchers.

\(^{16}\)Following Blanchard and Diamond (1990), this discussion uses changes in \(\lambda\) in a one-sector model as a shortcut for considering the effect of a general rise in “turbulence” across sectors. Later, in section 10, we will consider explicitly a multisector model of search and matching.

\(^{17}\)These properties have also been noted in the large literature on worker flows cited in footnote 11. The outflow rate from unemployment is the sum of the job-finding rate—the rate at which workers transition from unemployment to employment—and the rate at which unemployed workers exit the labor force. We will see in section 7 that direct measures of the job-finding rate also are strongly procyclical, like the overall outflow rate.
this measure displays dynamics that are the approximate mirror image of the unemployment outflow rate in figure 3B, rising sharply in recessions and subsiding in booms.

4.3 Quantitative Limitations

Despite these qualitative successes, however, the standard search model faces challenges in explaining three crucial quantitative features of observed Beveridge curve dynamics—the amplitude, comovement and persistence of unemployment and vacancy fluctuations. This point, originally conceived in the influential work of Shimer (2005), has since been the topic of a burgeoning literature expertly surveyed by Mortensen and Nagypál (2007).

Two implications of the standard model are key in this regard. First, the job creation condition (9) implies that the vacancy–unemployment ratio $\theta$ is a jump variable, adjusting immediately to contemporaneous changes in the parameters of the model. Thus, the job creation condition holds both in and out of steady state. Since the dynamics of unemployment implied by the Beveridge curve are naturally slow moving, and the vacancy–unemployment ratio moves instantaneously in response to shocks, it follows that vacancies themselves are a jump variable in the model. Intuitively, this is a corollary of the assumption of free entry: vacancies enter and exit instantaneously to enforce zero expected profits. The dynamics illustrated in figure 4 are drawn to reflect this.

The second insight is that, by virtue of these jump dynamics, the amplitude of the response of the vacancy–unemployment ratio to changes in labor productivity can therefore be deduced simply by total differentiation of the job creation condition (9). A particularly useful result is that the job creation response to changes in labor productivity takes the simple form:

$$\frac{d \ln \theta}{d \ln p} = \frac{r + \lambda + \beta f(\theta)}{\alpha(r + \lambda) + \beta f(\theta) 1 - \left(b/p\right)}.$$  

Thus, the magnitude of the implied rotation of the job creation condition in figure 4 can be gauged from the size of the elasticity in equation (10). Note that, with a conventional matching elasticity around $\alpha = 0.5$, the first ratio on the right-hand side of (10) is bounded between 1 and 2. By contrast, $1/(1 - b/p)$ can potentially be very large depending on the size of the flow surplus to employment relationships, as measured by the distance between the flow payoff from unemployment $b$ and the flow productivity of a job $p$. In the presence of a small surplus, slight changes in the productivity of labor can exhaust that surplus, so that incentives to create vacancies evaporate. Consequently, the amplitude of the negative cyclical comovement between unemployment and vacancies implied by the model hinges crucially on the surplus to employment relationships.

Figure 5 provides an illustration of the implications of the standard search model for the amplitude, comovement, and propagation of unemployment and vacancy dynamics. Panel A highlights the first two of these. Following Shimer (2005), it plots the relation between unemployment and vacancies implied by simulations in which actual realized shocks to $p$ and $\lambda$ since 1948 are fed through a conventional calibration of the model (see note to figure 5). These shocks are inferred respectively from the detrended paths of output-per-worker from the Bureau of Labor Statistics’s Labor Productivity and Costs program, and the unemployment inflow rate in figure 3A. In the light of equation (10),

---

18 In other words, as stressed in Pissarides (1985), there is no backward-looking component to job creation. Hence, for given $p$ and $\lambda$, (9) holds at all times. Shimer (2005) extends the Pissarides (1985) model to include shock processes for $p$ and $\lambda$ but argues that the comparative statics based on (9), and reported in equation (10), continue to serve as a good guide to the model dynamics.
we implement this exercise for a range of values of $b/p$, from 0.4 (as implemented by Shimer 2005), to 0.8 (an intermediate value), and 0.95 (similar to the calibration advocated by Hagedorn and Manovskii 2008).

Figure 5A reveals that the amplitude and comovement of the implied $(u, v)$ dynamics are two sides of the same coin. As noted in figure 4A, shocks to $\lambda$ in isolation induce a positive Beveridge curve correlation. To counteract this tendency, it is necessary for job creation to respond sufficiently—that is, for the amplitude of the response of the vacancy–unemployment ratio to productivity shocks be large enough. As emphasized by equation (10), this in turn requires that the match surplus be small. Consistent with this, figure 5A demonstrates that the comovement between $u$ and $v$ turns from positive to negligible to negative as $b/p$ rises from 0.4 to 0.8 to 0.95 and the amplitude of fluctuations successively expands. Thus, the canonical search model requires a relatively small surplus to account for the amplitude and comovement of $(u, v)$ variation, as in Hagedorn and Manovskii (2008).

Panel B of figure 5 turns attention to the dynamic properties of the model.
highlighting the tension faced by the standard model in replicating the persistence of unemployment, and especially vacancy dynamics. This is illustrated by the following exercise: for every recession since 1953, we identify start dates for each associated ramp up in unemployment and trace out the cumulative log deviation in the unemployment and vacancy rates over the subsequent twelve quarters relative to their prerecession levels. This is implemented symmetrically in both detrended empirical data, and data generated from simulations of the model with \( b/p \) set to 0.95—the parameterization that is able to account for the amplitude of empirical fluctuations. Figure 5B plots the unweighted average of these paths across recessions.

The empirical dynamics of both unemployment and vacancies in figure 5B are quite persistent, recovering slowly from their recessionary peaks and remaining on average around thirty log points away from their prerecession values even three years subsequent to the initial rise in unemployment. To place this in perspective, the average duration of an NBER-dated recession since 1953 has been around four quarters. Thus, unemployment remains high, and vacancies low, long after a recession is deemed officially to have ended. By contrast, the model-implied trajectories in figure 5B fall far short of the persistence of empirical Beveridge curve dynamics. The implied jump dynamics of vacancies noted earlier are manifested in figure 5B by a rapid decline and rebound in implied vacancies, which return to prerecession values after only nine quarters. Due to the natural persistence in the law of motion for unemployment in (1), the model-implied path of the unemployment rate is more persistent, but only slightly so, recovering much faster than its empirical analog.¹⁹

A simple analysis of the time series behavior of unemployment and vacancies thus suggests that both variables—vacancies in particular—exhibit dynamics that are much more persistent than implied by standard search models. This analysis did not attempt to identify formally the timing and dynamics of shocks to productivity and job destruction, nor their causal effects on Beveridge curve dynamics, however. While the identification of these causal processes is not straightforward, more formal analyses that have estimated such impulse responses of aggregate unemployment and vacancies under different assumptions support the impression of figure 5B (see, for example, Blanchard and Diamond 1989; Fujita and Ramey 2007; Hagedorn and Manovskii 2011). Intuitively, the comparatively high-frequency dynamics of labor productivity and job loss in the data are in large part unimpaired by modern attempts to purge these measures of potential endogenous feedback effects. On this basis, the standard search model thus appears to have insufficient internal propagation mechanisms to account for the persistence of unemployment and vacancies.²⁰

### 4.4 Summary, and a Look Ahead

We have presented an apparatus for understanding Beveridge curve dynamics that has become a standard reference point. This framework has the ability to account for some of the behavior of the empirical Beveridge

\[
\dot{u} = - (\lambda + f(\theta))(u - u^*), \text{ where } u^* = \lambda/(\lambda + f(\theta)) \text{ is the flow steady-state unemployment rate. Thus, the rate of convergence of unemployment to its steady state is given by } \lambda + f(\theta). \text{ Since the unemployment outflow rate } f \text{ is estimated to be on the order of 0.5 at a monthly frequency in the United States, the dynamics of } u \text{ are fast, closing around half of any deviation from flow steady state in a month.}
\]

²⁰It is possible that there exist other unobserved, and more persistent, shocks that drive the persistent dynamics of unemployment and vacancies. However, the literature has not reached a consensus on any such alternatives.

¹⁹Formally, the reason for this can be understood by expressing the law of motion for unemployment as
The outward shift in the relationship witnessed from the mid-1970s to mid-1980s can be accounted for by low-frequency movements in the degree of reallocation in the labor market. In addition, the canonical search model of Pissarides (1985) is able to provide a qualitative account of the negative comovement between unemployment and vacancies over the business cycle if shocks to aggregate labor productivity are present.

However, the model also faces several challenges that will be the point of departure for the remainder of the paper. First, we saw that it is difficult to account quantitatively for the amplitude, comovement, and persistence of unemployment and vacancies. In sections 5 and 6, we revisit the two pillars that close the benchmark model, namely, flexible Nash bargaining over wages and free entry. We consider alternatives to the benchmark model's treatment of these issues, with an eye toward each alternative's implications for the comovement of $u$ and $v$.

Second, although movements in the rate of job loss are able to account for some of the historical shifts in the Beveridge curve, we saw that the recent outward shift in the locus remains a particularly stark outlier. Within the benchmark search model, such a shift can be accounted for only by an unexplained deterioration in the functioning of matching markets—that is, a decline in matching efficiency. To reiterate this point, panels B and C of figure 3 demonstrate that a significant part of recent deterioration in unemployment exit and vacancy-filling flows also cannot be explained without appealing to a decline in match efficiency. Accordingly, the ensuing sections explore richer specifications of the matching process that respond to many of the abstractions maintained in the simple search model: we reflect on the effects of allowing for flows of workers in and out of the labor force, search intensity on behalf of firms and workers, job-to-job flows, heterogeneity in workers and jobs, and so on. In each case, we will apply these enriched models to recent US data on Beveridge curve dynamics.

5. Wage Setting

A feature of the canonical search model that has been the subject of particular scrutiny is its assumption over wage setting. As noted by Hall (2005b), the Nash surplus-sharing rule that gives rise to the wage equation (8) is but one of a continuum of feasible, and privately efficient, wage outcomes. There is no reason a priori why the rents generated by search frictions cannot be split arbitrarily. It therefore makes sense to consider alternative wage setting protocols.

It has long been understood that the flexibility of wages has an important bearing on the volatility of the demand for labor. As we shall see, the search model is no exception. Wage setting has important implications for the amplitude and comovement of unemployment and vacancy fluctuations summarized in the Beveridge curve. In particular, it useful to consider the implications of deviations from Nash surplus sharing in the simple canonical model outlined above. Wages in that model affect the Beveridge curve entirely through their effects on job creation incentives. The effects of wage flexibility on Beveridge dynamics can therefore be traced through by reconsidering the volatility of job creation under more general wage

21 Another channel through which wage setting can differ from the standard model is to relax the assumption of ex post wage bargaining. Wage-posting models instead allow firms to commit to a wage contract upfront as a means to attract workers. Thus, an interesting possibility in this environment is that wages can be used as a more purposive instrument of recruiting, as an alternative to vacancies. However, simple models in this vein without on-the-job search typically invoke free entry into vacancy creation, so that it is dominant to use vacancies as the primary recruitment tool, just as in the standard search model (see, for example, Schaal 2012). We will see in later sections that matters are different once free entry is relaxed (Kaas and Kircher forthcoming), or if workers are able to search on the job (Menzio and Shi 2011).
adjustment. Using the free entry condition (7), it is possible to generalize equation (10) to show that the volatility of the vacancy–unemployment ratio $\theta$ in response to a change in labor productivity is given by

$$\frac{d \ln \theta}{d \ln p} = \frac{1}{\alpha} \frac{1 - \varepsilon_w (w/p)}{1 - (w/p)}.$$  

Here, $\varepsilon_w = \frac{d \ln w}{d \ln p}$ is the elasticity of real wages with respect to a change in productivity. Note that $\varepsilon_w$ represents a total derivative, in the sense that it incorporates not only the direct response of wages to $p$, but also any indirect responses, for example via changes in $\theta$.

This simple observation nests several cases noted in recent literature. The first is the influential insight of Hall (2005b). Motivated by the indeterminacy of wages in the presence of ex post rents, he shows how a simple modification of the canonical search model in which wages are assumed to be completely rigid, $\varepsilon_w = 0$, need not imply a violation of bilateral efficiency. Intuitively, search frictions drive a wedge between the reservation wages of a firm and its worker, since it is costly for both to find alternative partners. It is therefore possible for wages to remain constant within a match without inducing an inefficient separation. Thus, as Hall (2005a) notes, this form of wage rigidity does not raise unemployment volatility by triggering a “burst” of layoffs into unemployment. Rather, it amplifies the volatility of job creation in equation (11), and thereby the negative comovement of unemployment and wages, alleviating one of the shortcomings of the canonical model.

As emphasized by Mortensen and Nagypál (2007) and Kenman (2010), the effect of wage rigidity in all these accounts is especially pronounced if the wage is also high relative to the product of a match—that is if $w/p$ is close to one. This echoes the message of Hagedorn and Manovskii (2008) that the canonical model delivers greater amplitude if employment relationships generate smaller rents for the employer. It is therefore natural to question whether the rents required to generate sufficient amplitude also are realistic. In a richer environment with idiosyncratic heterogeneity and endogenous job destruction, Elsby and Michaels (2013) demonstrate that conventional models with constant marginal returns to labor face a tension: the counterpart of a surplus small enough to generate realistic amplitude is the presence of excessive worker turnover. Elsby and Michaels show that this tension is resolved in a “large firm” model in which the demand for labor slopes downward. In that setting, job creation decisions are informed by the marginal surplus, which can remain small even in the presence of significant inframarginal rents that shape worker turnover among incumbent workers.

Michaillat (2012) enriches this point by noting that the combination of rigid wages and diminishing marginal returns to labor implies the presence of some job rationing in equilibrium: even in the absence of search frictions, some unemployment would remain. This observation in turn delivers an array of interesting implications. First, as indicated by equation (11), such a model is able to generate large, negatively correlated swings in unemployment and vacancies over
the cycle, as seen in the observed Beveridge curve. In addition, these swings are dominated by movements in unemployment due to job rationing, as opposed to frictions.

Given the theoretical importance of rigid wages in generating the amplitude and comovement of unemployment and vacancies, a parallel empirical literature has sought to devise measures of the relevant flexibility of wages. Following Bewley (1999) and Shimer (2004b), this literature has concentrated on the subtle but crucial role of rigidity in wages for newly hired workers. Since vacancy creation is a marginal decision, it is informed by a comparison of productivity and wages for a marginal hire. If the wages of the latter decline sufficiently in recessions, job creation incentives will be upheld.

Supporting the notion that wage rigidities of this form exacerbate labor market fluctuations, Bewley (1999) reports interview evidence that morale considerations constrain employers’ ability to cut wages, and that the pay of new hires and incumbent workers are tied together by internal equity considerations. By contrast, subsequent research based on survey data has emphasized the flexibility of entry wages. Pissarides (2009) surveys an older literature that documented the significant procyclicality of wage changes, especially among job changers (Bils 1985; Solon, Barsky, and Parker 1994). Likewise, Haefke, Sonntag, and van Rens (2013) find that the average wages among workers hired from nonemployment in the Current Population Survey are about as procyclical as labor productivity.

As Gertler and Trigari note, however, such evidence need not rule out the possibility that wages within particular jobs are rigid, which is what matters for job creation, and thereby the volatility of labor market fluctuations. The reason is that the allocation of workers across jobs may change over the cycle. If workers move up the job ladder in expansions—a process known as cyclical upgrading—individuals’ wages may rise even if wages within jobs are rigid. Recent research has attempted to control for such effects by tracking the wages of particular jobs within firms using employer–employee matched data. In Portuguese data, Martins, Solon, and Thomas (2012) and Carneiro, Guimarães, and Portugal (2012) find entry wages to be about as procyclical as employment. Stüber (2013) applies similar methods to German data and finds comparable results. Thus, the limited evidence available thus far is suggestive of substantial flexibility in new hires’ real wages.

The accumulation of further evidence on the cyclical behavior of entry wages is therefore a priority.

Kudlyak (2013) stresses that the path of wages subsequent to initial hiring is also crucial to the cyclical volatility of hiring. Intuitively, if workers hired in recessions are locked in to a persistently low wage trajectory, job creation incentives will be upheld in downturns. By contrast, if wage concessions among those hired in recessions dissipate over the course of an employment relationship—for example, due to internal equity considerations, as Bewley (1999) suggests—job creation incentives will be more prominently cyclical. Kudlyak distills this logic into the notion of a user cost of labor that summarizes the opportunity cost of delaying hiring decisions, and presents estimates to suggest necessarily informative on wage rigidity. In a simple model without frictions, for example, firms set the real wage equal to the marginal product of labor, which typically will be closely tied to average labor productivity regardless of the rigidity of wages.

Rogerson and Shimer (2011) note another reason why the comovement of real wages and labor productivity is not necessarily informative on wage rigidity. In a simple model without frictions, for example, firms set the real wage equal to the marginal product of labor, which typically will be closely tied to average labor productivity regardless of the rigidity of wages. See Solon, Whatley, and Stevens (1997), McLaughlin and Bils (2001), and the references therein.

22 Rogerson and Shimer (2011) note another reason why the comovement of real wages and labor productivity is not necessarily informative on wage rigidity. In a simple model without frictions, for example, firms set the real wage equal to the marginal product of labor, which typically will be closely tied to average labor productivity regardless of the rigidity of wages.

23 See Solon, Whatley, and Stevens (1997), McLaughlin and Bils (2001), and the references therein.

24 Haefke, Sonntag, and van Rens (2013) do note, however, that theoretically a small amount of rigidity in real wages will amplify labor market fluctuations if the firm’s surplus is small, as equation (11) emphasizes.
that the latter is very procyclical. Obtaining further understanding of the path of wages subsequent to initial recruitment is thus crucial for future work.

6. **Costly Entry**

Figure 5B and its surrounding discussion noted the inability of the standard search model to generate sluggish unemployment and vacancy dynamics. Recall that the failure of the model on this dimension can be traced to its assumption of free entry into vacancy creation. By rendering the entry of new vacancies infinitely elastic with respect to the return to vacancy creation, free entry induces jump dynamics in the vacancy–unemployment ratio \( \theta \), and thereby in the job-finding and vacancy-filling rates, \( f(\theta) \) and \( q(\theta) \), as well as the vacancy stock \( v \) itself.

All of these predictions stand in direct contrast to the persistent dynamics observed in the data.

In the light of this, it is natural to postulate that a more accurate account of Beveridge curve dynamics might be delivered by entertaining some departure from free entry. Interestingly, although the majority of modern applications of search theory have embraced the assumption of free entry, this is not reflected in some of the pioneering early work on the Beveridge curve. Blanchard and Diamond (1989), for example, consider an environment in which the stock of potential jobs is completely inelastic. This model therefore encapsulates the polar opposite case to that implicit in standard search models with free entry. Similarly, the seminal contribution of Diamond (1982) envisioned entry as the outcome of a process in which the returns to vacancy creation are set against stochastic startup costs.

After characterizing Beveridge curve dynamics in this basic setting, we further show how costly entry activates channels that are suppressed in conventional extensions of the standard search model. A prominent example that has been a focus of recent debate over the Beveridge curve is the notion of recruiting intensity emphasized by Davis, Faberman, and Haltiwanger (2013).

6.1 **Costly Entry and Beveridge Dynamics**

Entry of new vacancies corresponds to the flow \( \gamma \) in the law of motion for vacancies in equation (1). The returns to posting an active vacancy are given by the value of a vacancy, \( \mathcal{V} \) in the above notation. In the standard search model, free entry renders \( \gamma \) infinitely elastic with respect to \( \mathcal{V} \), with the implication that \( \mathcal{V} = 0 \) in equilibrium. It follows that a general way of relaxing the assumption of free entry is to allow the elasticity of the flow of new vacancies with respect to the value of an active vacancy, denoted \( \varepsilon_{\gamma} \), to be positive but finite. This result emerges, for example, in the early work of Diamond (1982), where potential entrants face a stochastic setup cost. In this case, entrepreneurs take a draw each period from a distribution \( G \) of startup costs, choosing to form vacancies only if the cost lies below the benefit \( \mathcal{V} \), so that the flow of new vacancies is given by \( \gamma = G(\mathcal{V}) \).

The implications of costly entry for Beveridge curve dynamics can be gleaned by inserting this alternative model of the flow of new vacancies into the law of motion for vacancies in equation (1) above. After the startup cost is paid, recruitment and production proceed just as in the standard model, and so this alternative entry process does not
alter the steady state\footnote{Of course, the point of this subsection is that the transition dynamics will depart from the jump dynamics of the standard model. Thus, the Bellman equations will diverge out of steady state.} Bellman equations (5) and (6). It follows that the value of an active vacancy $V$ can generically be expressed as a function of the productivity of labor $p$, the rate of job loss $\lambda$, and the vacancy–unemployment ratio $\theta$. Thus, the steady-state locus for vacancies takes the form:

\begin{equation}
\frac{dv}{dt} = 0 : \gamma(V(p, \lambda, \theta)) = m(u, v).
\end{equation}

A number of important implications follow. First, consistent with the motivation for analyzing costly entry, vacancy dynamics become persistent away from the free-entry limit, a point first noted by Blanchard and Diamond (1989). If vacancies lie above their steady-state level implied by (12), for example, outflows from the vacancy stock $m(u, v)$ will rise, and inflows $\gamma(V(p, \lambda, \theta))$ will fall as a tighter labor market reduces returns to vacancy creation $V$. Deviations of vacancies from steady state are thus closed incrementally over time. In this sense, relaxing free entry aids the model’s implications for Beveridge curve dynamics\footnote{An implicit assumption in the forgoing analysis is that exit of vacancies responds inelastically to changes in vacancy creation incentives. That is, there exists some (sufficiently large) cost to withdraw a vacancy such that (small) aggregate productivity shocks never induce a firm to do so.}

The presence of costly entry need not imply a large loss of amplitude in equilibrium unemployment and vacancy fluctuations, however. Costly entry also has important implications for the slope of the steady-state vacancy locus,

\begin{equation}
\frac{d \ln v}{d \ln p}
|_{u \ln p}
\end{equation}

In the free entry limit, as $\epsilon_{\gamma\lambda}$ goes to infinity, equation (14) recovers the implication of the standard search model noted above, namely that the (logarithmic) slope of the job creation condition approaches one, tracing out an upward-sloping ray from the origin in $(u, v)$ space. But, matters are potentially quite different away from this limit. Indeed, in the polar opposite case of completely inelastic entry, $\epsilon_{\gamma\lambda} = 0$, the slope of the $dv/dt = 0$ locus is equal to $-\alpha/(1 - \alpha)$, approximately \textit{minus} one under conventional parameterizations of the matching elasticity $\alpha$. Thus, the steady-state locus can become \textit{downward sloping}. Between these polar extremes, the slope of the locus is monotonically increasing in $\epsilon_{\gamma\lambda}$.
Entertaining the possibility of costly entry thus uncovers vacancy dynamics that are veiled in the free entry limit. In that limit, aggregate volatility is dampened by the jump dynamics of vacancies: as vacancies jump down and the market slackens in the wake of a recessionary shock, firms’ recruitment and wage costs also jump downward—it becomes easier and cheaper to fill vacancies—and in turn this maintains incentives to create jobs in a recession. Away from this limit, however, the slope of the locus in equation (14) reflects a delicate counterbalance. With sluggish vacancies, the flow of new vacancies responds only partially to job creation incentives. As a result, a rise in job searchers does not induce rapid entry of vacancies but instead partially depletes the outstanding vacancy stock, which slackens the market and depresses vacancies still further. This can lead to the negatively sloped steady-state locus noted above. Intuitively, in the absence of an infinitely elastic reserve of potential vacancies, new vacancies γ can “dry up.” It is in this sense that unemployment and vacancy dynamics become intertwined in the presence of costly entry.

Costly entry has more fundamental implications for the effects of shocks to job destruction λ. This is the central theme of Coles and Moghaddasi Kelishomi (2011), who show that unemployment and vacancy dynamics in this case become mutually reinforcing, both amplifying and propagating shocks to λ, inducing Beveridge curve dynamics that mirror those seen in the data. Figure 6A illustrates the idea. Because costly entry reduces the slope of the steady-state vacancy locus—because the replenishment of new vacancies is stemmed as the labor market slackens—the rise in unemployment induced by increased rates of job destruction can deplete the vacancy stock. Moreover, this sets in motion an adverse chain of events whereby the depletion of vacancies exacerbates the job-finding prospects of the unemployed, further raising unemployment and reducing vacancies.

To pursue this point further, figure 6C illustrates the results of numerical simulation of the model of Coles and Moghaddasi Kelishomi (2011). Figure 6C is the costly entry counterpart of figure 5B for the standard model, with one exception: we feed productivity shocks is comparable with the free entry case.

Formally, equations (13) and (14) imply that the equilibrium response of unemployment, \( d\ln u/d\ln p \), equals its free entry counterpart multiplied by \( \varepsilon_{\gamma V}/R(1 - \alpha)u \). While the latter increases in absolute value as the elasticity of entry \( \varepsilon_{\gamma V} \) rises, a conventional unemployment rate \( u \approx 0.06 \) and matching elasticity \( \alpha \approx 0.5 \) imply very small differences, provided \( \varepsilon_{\gamma V} \) is not too small.

We focus on the case in Coles and Moghaddasi Kelishomi (2011) with entry costs, and do not consider their “time-to-build” extension, whereby entrepreneurs must wait to form a vacancy after the startup cost is paid. We retain Coles and Moghaddasi Kelishomi’s calibration in all respects except that we focus on an environment where the average entry cost is large enough to arrest the response of vacancies to changes in unemployment, yielding a downward-sloped job creation locus, as in figure 6A. To match the average unemployment rate in the data, we adjust matching efficiency as needed.
Panel A. Rise in job destruction

Panel B. Decline in labor productivity

Panel C. Response to job destruction shocks

Figure 6. Beveridge Curve Dynamics with Costly Entry

Notes: Panels A and B sketch out the qualitative dynamics of unemployment and vacancies implied by relaxing free entry in a standard search model. Panel A depicts the effects of an increase in the rate of job destruction ($\lambda$ in the notation of section 4) or a decrease in match efficiency (a reduction in $m(u,v)$ for any given $u$ and $v$). Panel B depicts the effects of a reduction in aggregate labor productivity ($p$ in the notation of section 4). The solid lines in panel C are the same as in panel B of figure 5: the series for unemployment and vacancies are log differences relative to the start dates over NBER recessions. The dashed lines are simulations of the model in Coles and Moghaddasi Kelishomi (2011) based on the following parameter values, largely from their baseline calibration, but set at a weekly frequency: the discount factor is 0.999; the matching elasticity is set to $\alpha = 0.6$; worker bargaining power is set to 0.6; the flow payoff from unemployment $b = 0.7$; vacancy posting costs are set to zero; average $\lambda$ is 0.0085 and shocks to $\lambda$ are calculated as in figure 5, linearly interpolated to the weekly frequency. We assume that the elasticity of the flow of new vacancies with respect to the value of a vacancy, $\varepsilon_v \gamma$, is one eighth.
through the model shocks to job destruction $\lambda$ only, and suppress shocks to labor productivity $p$, to underscore the diametrically different dynamics implied by inelastic entry on this dimension. Specifically, entrepreneurs in the model are confronted with the (detrended) observed time series of $\lambda$ over the post-war period. We isolate the model-implied responses of vacancies and unemployment around each NBER-dated recession since 1953, average across them, and compare these theoretical responses to their empirical analogues. The results are very much in line with the intuition gleaned from figure 6A, replicating the persistence and amplitude of both unemployment and vacancies quite closely.

Departing from free entry thus turns on its head one of the key implications of standard search model. In that model, figure 4B demonstrated that movements in labor productivity provide the most promising potential reconciliation of the negative cyclical comovement of unemployment and vacancies seen in the data. The standard model suffers from important drawbacks in its ability to replicate the amplitude and persistence of empirical Beveridge curve dynamics, however. In stark contrast, a model with costly entry identifies job destruction shocks as a leading potential explanation for cyclical fluctuations in the labor market. A tantalizing aspect of this alternative account is that it provides a natural amplification and propagation mechanism that is able to explain the bulk of the volatility and persistence of Beveridge curve fluctuations.

6.2 Recruitment Intensity

An uncomfortable feature of the models of vacancy creation considered in the preceding sections is that firms are rather passive agents—they post vacancies and wait. In reality, firms take a more active role in recruiting, using tools in addition to vacancy posting to aid the matching process, such as greater recruiting effort, relaxation of hiring standards, higher wage offers, and so on (Barron, Berger, and Black 1997). Influential recent work by Davis, Faberman, and Haltiwanger (2013) has further revealed evidence that such recruitment intensity has a prominent cyclical component, falling significantly in the Great Recession.

Perhaps surprisingly, conventional extensions of the standard search model that impose free entry predict that recruitment intensity will be invariant to aggregate fluctuations (see chapter 5 of Pissarides 2000). That this neutrality result rests on the premise of free entry underscores the restrictiveness of that assumption for labor market dynamics noted above. By the same token, it also suggests that recruiting intensity can be revived as an active margin in the more general case in which entry is inelastic to some degree. We will see that allowing for recruiting intensity in the presence of inelastic entry has interesting implications for Beveridge curve dynamics.

To begin, note that the simplest interpretation of recruitment intensity is that it augments the stock of vacancies in the matching function:

\begin{equation}
    m = m(u, av),
\end{equation}

where $a$ represents average recruiting effort per vacancy. In this way, movements in recruitment intensity $a$ will leave an imprint on the observed Beveridge curve relation between $u$ and $v$. The augmented matching function in (15) implies that the number of

\[29\text{Alternative explanations involve recruiting tools that do not operate through the matching technology. In Kaas and Kircher (forthcoming), firms face a convex cost of vacancy posting so that wage posting becomes a valued recruiting tool. Firms that wish to grow post higher wages as a means to attract more workers and therefore fill their vacancies at a faster rate. Faberman and Nagypál (2008) present a model of on-the-job search in which more productive, growing firms are more likely to be able to compete away employed job seekers, raising their vacancy filling rates.}\]
hires per unit of recruiting effort is given by $q(a\theta) = m/av = m(1/\theta, 1)$, where $\theta$ denotes the vacancy–unemployment ratio $v/u$, as before. Thus, an individual firm expending $a_i$ units of recruitment effort will fill its vacancy with probability $a_i q(a\theta)$: greater effort pays off by increasing the probability of filling a job. If the cost of exerting effort is summarized by an increasing, convex function $c(a_i)$, a vacant firm solves the following decision problem:

$$r V = \max_{a_i} \{-c(a_i) + a_i q(a\theta)(J - V)\}.$$  

The symmetry of this simple problem implies that all firms choose the same recruitment effort $a$ that solves the first-order condition $c'(a) = q(a\theta)(J - V)$. Firms trade off the increased marginal costs of recruiting intensity with its returns, which are given by the marginal increase in the likelihood of filling a vacancy times the expected value of such a filled job. Combining the above first-order condition and the expressions for $J$ and $V$ in equation (6), one can show that, for a given stock of vacancies, the steady-state optimal search intensity has an intuitive form: $a$ is increasing in the productivity of filled jobs ($p$ in our earlier notation), and decreasing in both the rate of job loss ($\lambda$) and the vacancy–unemployment ratio ($\theta$).

The reason conventional models imply that recruiting intensity has no bearing on labor market equilibrium is because the stock of vacancies is not taken as given in the standard search model. Instead, it is determined by free entry, forcing the returns to vacancy creation $V$ to zero in equilibrium. Imposing this and the first-order condition for optimal recruitment intensity implies that the steady-state value of a vacancy becomes

$$r V = -c(a) + ac'(a)$$

$$= 0 \Rightarrow \frac{c'(a)a}{c(a)} = 1.$$  

Under free entry, the elasticity of the recruiting cost with respect to effort is set to unity in all states. The level of recruiting intensity $a$ that solves this unit-elasticity condition is thus invariant to aggregate fluctuations.\(^{30}\) The intuition is that all adjustments occur at the extensive margin under free entry, because that margin is infinitely elastic. For the same reason, the only role played by search intensity is to minimize the cost per unit of (constant) effort, $c(a)/a$.

This observation suggests that there is considerable merit to investigating the role of recruitment intensity in the presence of inelastic entry, as in the model sketched earlier in this section. This setting will resemble that environment, with one exception: responses of recruitment intensity to shocks will feed back into the matching technology itself.

To see how, it is useful to return to the laws of motion for unemployment and vacancies:

$$\frac{du}{dt} = \lambda(1 - u) - m(u, av), \quad \text{and}$$

$$\frac{dv}{dt} = \gamma(V(p, \lambda, \theta)) - m(u, av).$$

Consider first the steady-state unemployment condition that sets $du/dt = 0$. Standard search models with free entry imply that this relation is independent of movements in labor productivity, and shifts outward in $(u, v)$ space when job loss is higher. Allowing for costly entry and endogenous recruitment effort changes these predictions in two ways. First, reductions in labor productivity will now also induce outward shifts in this locus, by reducing recruiting intensity. Second, outward shifts originating from rises in job loss are now amplified, as increases in $\lambda$ further depress recruitment intensity. Intuitively, declines in recruiting intensity imply fewer

\[^{30}\text{It is simple to confirm that a unique positive level of } a \text{ will solve (17) provided } c(0) > 0.\]
matches for any given \((u, v)\) pair, and provide a source of procyclical fluctuations in measured matching efficiency. Formally, the vertical shifts in the locus induced by changes in labor productivity and job loss are given by

\[
\left. \frac{d \ln v}{d \ln p} \right|_{u,u^\prime} = -\varepsilon_{ap} < 0, \quad \text{and} \quad \left. \frac{d \ln v}{d \ln \lambda} \right|_{u,u^\prime} = \frac{1}{1 - \alpha} - \varepsilon_{a\lambda} > 0,
\]

where, as in prior notation, \(\varepsilon_{ap} > 0\) and \(\varepsilon_{a\lambda} < 0\) respectively denote the elasticities of recruitment effort with respect to labor productivity and the rate of job loss. As we noted in section 6.1, and as illustrated in figure 6A, increases in the number of unemployed searchers associated with such outward shifts in the \(du/dt = 0\) locus can deplete vacancies faster than new vacancies are created under inelastic entry. As a result, shifts in recruitment intensity can amplify labor market fluctuations in Beveridge space.

Variable recruiting intensity also influences the steady-state vacancy condition. In particular, as we saw in section 6.1, the slope of the \(dv/dt = 0\) locus can change dramatically in the presence of costly entry, and have important implications for Beveridge dynamics. Strikingly, recruitment intensity also amplifies these effects. Specifically, the slope of the steady-state vacancy locus takes the form:[31]

\[
\left. \frac{d \ln v}{d \ln u} \right|_{\theta = 0} = -\frac{\alpha + \varepsilon_{\gamma} \varepsilon_{\gamma \theta} (1 - \alpha) \varepsilon_{a\theta}}{1 - \alpha - \varepsilon_{\gamma} \varepsilon_{\gamma \theta} + (1 - \alpha) \varepsilon_{a\theta}}.
\]

Since a lower \(\theta\) suppresses recruiting and wage costs, it also induces an increase in recruiting intensity, ceteris paribus. Thus \(\varepsilon_{a\theta} < 0\). It follows that the steady-state vacancy locus becomes more negatively sloped, magnifying the negative comovement between unemployment and vacancies induced by the outward shifts in the \(du/dt = 0\) locus noted above. These predictions suggest there is significant promise for future research that pursues an account of labor market fluctuations based on movements in recruiting intensity.

Recently, Davis, Faberman, and Haltiwanger (2013) have provided evidence on the empirical importance of recruitment intensity and its cyclical variation. Examining the establishment-level data on hiring and vacancies underlying the Job Openings and Labor Turnover Survey, they identify a simple but compelling fact: fast-growing employers fill a greater share of their vacancies. In particular, they uncover a strikingly tight cross-sectional relationship between an establishment’s propensity to fill a vacancy \(q_{it}\) and its hiring, \(m_{it}\).

\[
\ln q_{it} = \chi_0 + \chi_1 \ln m_{it} + d_t + e_{it},
\]

where the time effects \(d_t\) collect all aggregate determinants of job-filling propensities, such as \(q(ab)\) in the model above, and the error term \(e_{it}\) captures idiosyncratic variation in vacancy-filling rates across establishments. Davis, Faberman, and Haltiwanger emphasize that their finding of \(\chi_1 > 0\) challenges the implications of conventional models in the search tradition, in which all employers face the same probability of filling a vacancy, regardless of their desired change in employment.[32]

31 Note that, from the envelope theorem, the responsiveness of the value of a vacancy \(v\) to \(\{p, \lambda, \theta\}\) will resemble that in the free entry case.

32 In standard search models there are no “firms”—the presence of constant returns in production implies no well-defined notion of size. However, several papers have studied “large-firm” versions of these models that preserve the property that job-filling rates are orthogonal to idiosyncratic fluctuations in desired employment growth (see Acemoglu and Hawkins forthcoming; and Elsby and Michaels 2013).
To see how this relation between filling rates and hires can be interpreted as a sign of recruitment intensity, it is useful to return to the simple theoretical framework above. There, \( q_{it} = a_{it} q(a\theta) \) and \( m_{it} = a_{it} q(a\theta) v_{it} \), so that the relationship between the employer-level job-filling and hires rates in equation (21) will reflect recruitment intensity through two channels. First, recruitment intensity \( a_{it} \) is mechanically a common component of both \( q_{it} \) and \( m_{it} \). Second, an establishment’s vacancies may also be correlated with recruitment intensity. As an example of the latter, note from equation (17) that the returns to vacancy creation \( Y \) can be expressed as a direct, increasing function of recruitment intensity. In that model, then, a positive correlation between the idiosyncratic components of \( q_{it} \) and \( v_{it} \) is embedded in estimates of equation (21) and is a further sign of variable recruitment intensity.

33 A difference between this model and the empirical analysis of Davis et al. (2013) is that the latter relates the job-filling rate to the hiring rate, which divides hires by the establishment’s size. By contrast, the framework of this section relates the job-filling rate to the level of hires. While these are related, it is possible, for example, that recruiting intensity is particularly high at small, but fast-growing establishments whose absolute number of hires is not that large. Characterizing the mapping between hiring and job-filling rates requires a more explicit theory of establishment-level dynamics, which pin down employer size. The latter is beyond the scope of this survey.

34 This reflects the fact that recruitment intensity and vacancies are assumed to be complements in the matching technology. It is conceivable that some degree of substitutability could also arise between these two margins.

35 Davis, Faberman, and Haltiwanger (2013) also provide a careful investigation of potential confounds to this interpretation. They find little evidence of increasing returns to vacancy posting at the microeconomic level. A possibility that is harder to rule out is the presence of idiosyncratic variation in matching efficiency, which would induce a correlation between filling rates and hiring that is unrelated to recruitment intensity. A source of reassurance on this dimension is that Davis, Faberman, and Haltiwanger find little difference in their estimates if they control for time-invariant differences using establishment fixed effects. It is considerably more difficult to isolate, and control for, variation within establishment in matching efficiency.

An interesting corollary of this interpretation is that recruitment intensity is likely to have a cyclical dimension. If the response of employers’ recruitment intensity to an aggregate expansion mirrors the response to idiosyncratic expansions in equation (21), Davis, Faberman, and Haltiwanger (2013) note that \( x_1 \Delta E[\ln m_{it}] \) can be interpreted as an index of aggregate recruitment intensity. Figure 7A depicts the time series of this measure over the period since JOLTS data were made available in December 2000. By this measure, recruiting intensity fell prominently during the Great Recession, raising the provocative possibility that this variation might account for recent movements in the Beveridge curve.

To address this question, we again conduct a simple accounting exercise. From the law of motion for unemployment in (18), the vertical shift in the Beveridge curve attributable to recruiting intensity is simply

\[
\frac{d \ln v}{d \ln a} \bigg|_{u, it=0} = -1.
\]

Figure 7B accordingly constructs a counterfactual Beveridge curve that traces out the path of vacancies implied by holding recruiting intensity constant and taking as given the path of the unemployment rate, as we have done in previous sections. The decline in recruiting effort during the Great Recession is manifested in a prominent inward shift of the counterfactual Beveridge curve: if recruiting effort had not fallen, many fewer vacancies would have been needed to support the observed levels of unemployment. Thus, cyclical movements in recruiting intensity leave a clear impression on the

36 Following Davis et al. (2013), in figure 7A, we construct the actual index using aggregate time series data by computing \( x_1 \Delta E[\ln m_{it}] \), rather than the object in the main text. The two will differ by Jensen’s inequality. Also, notwithstanding the concern raised in footnote 34, we use the hiring rate, again following Davis et al.
Beveridge curve and appear to have contributed significantly to the rise in unemployment in the Great Recession. An additional implication of figure 7B, however, is that these significant fluctuations in recruiting effort are not persistent enough to account for the sustained outward shift of the Beveridge curve in the wake of the recession: the “loop” in the counterfactual Beveridge curve is left largely intact.

7. The Participation Margin

The analysis thus far has inferred Beveridge curve dynamics from a model of worker flows between just two labor market states: unemployment and employment. A common motivation for this abstraction is the observation that the stock of labor force participants $L$ is much less cyclical than, for instance, employment or unemployment. In what follows, however, we shall see that worker flows between each of these states and nonparticipation are nontrivial, and some display important cyclical components.

To understand how participation flows shape the dynamics of unemployment, it is necessary to amend the dynamic system in equation (1) to consider the joint dynamics of nonparticipation $N$.

\[
\frac{dU}{dt} = \pi_{EU}(L - U) + \pi_{NU}N - \pi_{UN}U - m(U, V),
\]

Notes: Panel A depicts an index of recruiting intensity calculated in the following way. We take the total non-farm hires from JOLTS and divide by CPS employment to derive the hiring rate. We follow Davis, Faberman, and Haltiwanger (2013), who report an elasticity of recruiting intensity with respect to the hiring rate of 0.82. Recruiting intensity is then normalized to equal 1 in 2008 Q1. The counterfactual Beveridge curve in panel B is estimated by netting off the shifts in vacancies implied by changes in recruiting intensity as measured in panel A.
and

\[ \frac{dN}{dt} = \pi_{EN}(L - U) + \pi_{UN}U - \pi_{NU}N - \pi_{NE}N. \]

Here, the \( \pi_{ij} \)s denote the flow hazard rates associated with worker transitions from state \( i \) to state \( j \) for \( i, j \in \{E, U, N\} \).

Three aspects of equation (23) are worth noting. First, flows in and out of unemployment now may be associated either with job loss/finding or with labor force entry/exit. These were implicitly conflated in the above two-state unemployment analysis. Second, the job-finding rate \( \pi_{UE} = m(U, V)/U = m(1, V/U) \) ties the dynamics of \( U \) and \( N \) with those of vacancies, and thereby the Beveridge curve. Finally, while we allow for the presence of direct hiring from nonparticipation in this framework, we assume for now that such hires are not mediated through the matching function. We return to this in more detail later in this section.

Figure 8 depicts the time series of the empirical analogues of the flow transition rates \( \pi_{ij} \) for the United States. Since the early work of Perry (1972), Marston (1976), and Blanchard and Diamond (1990), these flows conventionally have been estimated using the two-thirds or so of individuals in the Current Population Survey whose responses can be linked longitudinally across consecutive months. Figure 8 plots quarterly averages of these monthly measures of the flow transition probabilities from 1967 to 2013.38

In addition to reaffirming the marked countercyclicality of the job loss rate \( \pi_{EU} \), and the strong procyclicality of the job finding rate \( \pi_{UE} \), figure 8 reveals that the flow transition rates between unemployment and non-participation also exhibit significant cyclical components. Moreover, these fluctuations at the participation margin clearly serve to contribute to rises in unemployment in times of recession: the rate at which nonparticipants enter unemployment \( \pi_{NU} \) rises, and the rate at which unemployed workers exit the labor force \( \pi_{UN} \) declines, in all recessions in the sample period. Several authors have emphasized that these forces contribute importantly to unemployment cyclicality.39

The cyclicity of these participation flows will also leave an imprint on the realized Beveridge curve. In particular, Petrongolo and Pissarides (2008) show that the dynamic system in equation (23) implies a steady-state unemployment rate that can be written as:

\[ u = \frac{\pi_{EU} + \epsilon}{\pi_{EU} + \epsilon + \frac{m(1, V/U)}{\pi_{UE}}} + \xi, \]

where

\[ \epsilon \equiv \frac{\pi_{EN}}{\pi_{NU} + \pi_{NE}} \] and \[ \xi \equiv \frac{\pi_{NE}}{\pi_{NU} + \pi_{NE}}. \]

Note that this mirrors equation (2) above, except that the additional terms \( \epsilon \) and \( \xi \)

38 Data for the period since 1990 are published on the Bureau of Labor Statistics’ website (see http://bts.gov/webapps/legacy/cpsflowstab.htm). Data prior to 1990 have been tabulated by Joe Ritter and made available by Hoyt Bleakley. Shimer (2012) also has computed measures of these flows using Current Population Survey microdata from 1976 to 2007 (see https://sites.google.com/site/robertshimer/research/flows).

39 See Shimer (2012); Elsby, Hobijn, and Şahin (2015); Smith (2011); Elsby, Smith, and Wadsworth (2011); Barnichon and Figura (2013a); Kudlyak and Schwartzman (2012). Elsby, Hobijn, and Şahin (2015), for example, provide a metric to infer the contributions of participation flows using successive local linear approximations to equation (23). They find that the cyclicity of \( \pi_{NU} \) and \( \pi_{UN} \) in particular explain around one third of the variance in unemployment in the United States.
Figure 8. Gross Worker Flow Transition Probabilities among Employment, Unemployment, and Nonparticipation
denote respectively unemployment inflows and outflows associated with entry to and exit from the labor force. One interpretation of $\epsilon$, for instance, is that it measures the rate at which workers enter unemployment by flowing from employment to nonparticipation ($\pi_{EN}$) and then, conditional on leaving nonparticipation, transition to unemployment ($\pi_{NU}/(\pi_{NU} + \pi_{NE})$). A similar interpretation applies to $\xi$—it captures indirect flows from unemployment to employment via nonparticipation. Thus $\pi_{EU} + \epsilon$ can usefully be thought of as a composite inflow rate into unemployment, and likewise $\pi_{UE} + \xi$ a composite outflow rate.

Figure 9A depicts the implied fluctuations in the participation-related flows, $\epsilon$ and $\xi$ since 2000. During the Great Recession, inflows to unemployment by flowing from employment to nonparticipation ($\pi_{EN}$) and then, conditional on leaving nonparticipation, transition to unemployment ($\pi_{NU}/(\pi_{NU} + \pi_{NE})$). A similar interpretation applies to $\xi$—it captures indirect flows from unemployment to employment via nonparticipation. Thus $\pi_{EU} + \epsilon$ can usefully be thought of as a composite inflow rate into unemployment, and likewise $\pi_{UE} + \xi$ a composite outflow rate.

Figure 9A depicts the implied fluctuations in the participation-related flows, $\epsilon$ and $\xi$ since 2000. During the Great Recession, inflows to unemployment by flowing from employment to nonparticipation ($\pi_{EN}$) and then, conditional on leaving nonparticipation, transition to unemployment ($\pi_{NU}/(\pi_{NU} + \pi_{NE})$). A similar interpretation applies to $\xi$—it captures indirect flows from unemployment to employment via nonparticipation. Thus $\pi_{EU} + \epsilon$ can usefully be thought of as a composite inflow rate into unemployment, and likewise $\pi_{UE} + \xi$ a composite outflow rate.

by about one-third in the recession, reflecting in large part declines in flows from unemployment to nonparticipation, $\pi_{UN}$.

7.1 Participation Flows and the Beveridge Curve

The reinterpretation of the Beveridge curve in equation (24) allows one to implement a counterfactual exercise similar to those used in prior sections of this survey. The thought experiment here is to infer the vertical—that is, the $u$-constant—shift in the Beveridge curve implied by the participation flows embedded in $\epsilon$ and $\xi$. Total differentiation of (24) reveals that the log-change in vacancies associated with flows between unemployment and nonparticipation takes the form:

$$
\frac{d \ln v}{du, \dot{\bar{u}} = 0} = \frac{1}{1 - \alpha} \left[ \frac{\omega_\epsilon}{1 - \omega_\xi} d \ln \epsilon - \frac{\omega_\xi}{1 - \omega_\xi} d \ln \xi \right],
$$

Figure 9. The Participation Margin and the Beveridge Curve

**Notes:** Panel A is calculated from the gross flows data depicted in figure 8 according to the equations for $\xi$ and $\epsilon$ given in equation (24) in section 7. The counterfactual Beveridge curve in panel B is estimated by netting off the shifts in vacancies implied by changes in $\xi$ and $\epsilon$ as measured in panel A.
where, as before, $\alpha$ denotes the elasticity of the matching function, $\omega_\epsilon \equiv \epsilon / (\pi_{EU} + \epsilon)$ and $\omega_\xi \equiv \xi / (\pi_{UE} + \xi)$. Thus, as happens during recessions, both the increased inflows into unemployment via nonparticipation $\epsilon$, as well as declines in exits from unemployment via nonparticipation $\xi$, will raise the level of vacancies needed to support a given rate of unemployment. Accordingly, the realized Beveridge curve will shift up. Equivalently, if the influence of the participation-related flows were removed, the implied counterfactual Beveridge curve would lie below its realized position, as we will see momentarily.

The counterfactual Beveridge curve implied by holding constant the participation flows $\epsilon$ and $\xi$ is illustrated in figure 9B. Two key features stand out. First, the observation noted earlier that movements in participation flows must contribute to fluctuations in unemployment finds its expression in $(u, v)$ space in the form of a significant flattening of the realized Beveridge curve, relative to its counterfactual. Thus, participation flows alter substantially the empirical relationship between vacancies and unemployment, echoing the conclusions of Elsby, Hobijn, and Şahin (2015) and Kudlyak and Schwartzman (2012). Second, figure 9B provides a perspective on whether the participation margin can account for the outward shift in the Beveridge curve that has emerged since the end of the Great Recession. Indeed, the outward shift is noticeably more modest in the counterfactual path—by the second quarter of 2013, for example, the vertical outward shift is about 0.7 percentage points in the realized Beveridge curve, but only around 0.4 points if participation flows are held constant. Thus, recent movements in participation flows are able to account for a nontrivial fraction of the recent outward shift in the Beveridge curve.

As we noted earlier in this section, however, an implicit assumption underlying the exercise in figure 9B is that flows from nonparticipation to employment are not mediated through the matching function. This rules out the possibility that hires from nonparticipation come at the expense of hires from the unemployment pool, which in turn could have important implications for the Beveridge curve. If vacancies are chased not only by the unemployed, but also by some segment of nonparticipants, the Beveridge curve locus will depend on the degree of search of the nonparticipants. This congestion effect has been emphasized recently by Diamond (2013) as a possible contributor to the outward shift in the Beveridge curve since the Great Recession. As demonstrated in figure 8, although $\pi_{NE}$ is clearly procyclical, it is not as procyclical as the job finding rate from unemployment $\pi_{UE}$.

Under certain conditions, however, our preceding results survive largely intact even after incorporating nonparticipation into the matching technology. Imagine, for example, that a fraction $\eta$ of nonparticipants search for employment, so that the

---

41 It is also possible to break the contributions of $\xi$ and $\epsilon$ into constituent components associated with each of the participation flows. As we hinted earlier, doing so suggests that the flow from nonparticipation to unemployment has been particularly influential in both the rise and persistence of unemployment during the Great Recession.

42 Of course, if search behavior is recorded accurately in the CPS and individuals are sampled continuously, such direct N-to-E flows should theoretically be nonexistent, except in cases such as school-to-work transitions. Studies that have attempted to adjust for classification errors in recorded labor market status, and for missing intervening unemployment spells between surveys (known as time aggregation bias), have found that the levels of adjusted N-to-E flows are approximately halved, but remain procyclical (Elsby, Hobijn, and Şahin 2015).
matching function becomes \( m(U + \eta N, V) \), and the probability of any searcher contacting a vacancy is \( m(1, V/(U + \eta N)) \). As Diamond (2013) stresses, an increase in \( \eta \) implies that there are fewer vacancies to go around, diminishing the job-finding rate of unemployed workers. Yet increases in \( \eta \) have an additional effect in this simple environment: for a given flow hazard \( \pi_{NU} \), higher search intensity also implies that a nonparticipant is more likely to find work before a transition to unemployment occurs. In other words, search among nonparticipants reduces not only outflows of unemployed workers, but also inflows into unemployment via nonparticipation—\( \epsilon \) in our notation. We find that these two effects approximately balance out with respect to the observed outward shift of the Beveridge curve.

7.2 Understanding the Participation Margin

The preceding analysis has highlighted the potential importance of participation flows as a proximate determinant of the Beveridge curve, taking as given the path of those flows. A complete understanding of the participation margin, however, requires an understanding of why participation flows vary as they do over the cycle. At first blush, the cyclical variation in flows between unemployment and nonparticipation can appear distinctly puzzling: it is natural to imagine, for example, that the search incentives of nonemployed workers would deteriorate in times of recession in which the returns to search are low—the so-called discouraged worker effect. By contrast, figure 8 suggests an opposite message: the unemployed appear less likely to exit the labor force during downturns, and nonparticipants are more likely to initiate a job search.

This puzzle shares a counterpart in much recent theorizing on the participation margin. The point can be illustrated in a simple extension of the baseline search model in the spirit of Garibaldi and Wasmer (2005) and Haefke and Reiter (2011). The key difference is that nonemployed individuals may now choose whether or not to search for employment. For example, if individuals are subject to shocks to their market productivity \( x \) over time, the value of participating in the labor market also will fluctuate with those shocks.

Suppose for simplicity that shocks to \( x \) arrive at Poisson rate \( \delta \) and are distributed according to a distribution function \( H(x) \), with lower support such that the value of working always dominates the value of nonparticipation. Then, the participation decision can be summarized by the following value functions for unemployment \( U \), and nonparticipation \( N \):

\[
(b + f(\theta)(W(x) - U(x)))
\]

\[
+ \delta \int \max \{U(x'), N(x')\} dH(x'),
\]

As noted by Veracierto (2011), this model suggests a simple measure of \( \eta \). Since \( \pi_{UE} = m(1, V/(U + \eta N)) \) and \( \pi_{NE} = \eta m(1, V/(U + \eta N)) \), it follows that \( \eta = \pi_{NE}/\pi_{UE} \). The gross flows depicted in figure 8 suggest this is strongly countercyclical, since \( \pi_{UE} \) falls far more than \( \pi_{NE} \) in recessions. Veracierto uses this observation to undertake a different exercise. He solves for the equilibrium paths of unemployment and vacancies within a Mortensen and Pissarides model, and uses these to infer aggregate matching efficiency.

Although Garibaldi and Wasmer (2005) and Haefke and Reiter (2011) also include shocks to home productivity, it is possible to recast the setting described here to incorporate these shocks. What is important is the relative productivity at home versus at work.

The latter assumption allows us to focus on implications for worker flows between unemployment and nonparticipation. The results of Garibaldi and Wasmer (2005) suggest the qualitative patterns we describe hold in a more general environment.
\((r + \delta)N(x)\)

\[= z + \delta \int \max\{U(x'), N(x')\} dH(x').\]

Here, \(z\) is the flow payoff to individuals out of the labor force, and the flow payoff from unemployment \(b\) can be thought of as net of the costs of searching (including any foregone leisure or home production due to search activity). It follows that an individual will choose to engage in search provided their market productivity \(x\) exceeds a threshold \(S\) that equates the opportunity cost of search to the returns to search, \(z - b = f(\theta)(W(S) - U(S))\). The implied hazard rates for worker flows between unemployment and nonparticipation are therefore:

\[
\pi_{UN} = \delta H(S), \quad \text{and} \quad \pi_{NU} = \delta(1 - H(S)).
\]

The implied effects of a recession are simple to characterize. One can verify that the worker’s surplus, \(W(x) - U(x)\), is increasing in individual productivity \(x\), and will decline for each \(x\) following a reduction in aggregate productivity \(p\). In addition, a slackening of the labor market will depress the job finding probability, \(f(\theta)\). Both will lead to an increase in \(S\). Thus, by reducing the returns to search, a business cycle slump will imply that only workers with higher \(x\)s will participate. The result is an exodus from unemployment into nonparticipation, raising \(\pi_{UN}\) and depressing \(\pi_{NU}\) in recessions, precisely the opposite of what is shown in figure 8.

Examples of this result abound in the literature. The early insights of Tripier (2004) and Veracierto (2008) highlight that models similar to that considered above can fail in a rather spectacular way: the unemployment rate can become procyclical as labor force entrants flood the market in response to increases in the value of work, generating a positively sloped Beveridge curve. Although Ebell (2011) and Shimer (2012) demonstrate that this particular pathology can be surmounted by generating an offsetting flood of searchers into employment in expansions—echoing the (un)employment volatility puzzle noted by Shimer (2005)—models in this vein share the same counterfactual implications with respect to worker flows: labor force entry rate is predicted to be pro-cyclical and labor force exit countercyclical.

So, what might explain the seemingly perverse cyclicality of participation flows? Recent literature has begun to make progress on this puzzle. The most promising avenue highlights the importance of worker heterogeneity, and associated composition effects over the business cycle, for understanding the participation margin. Elsby, Hobijn, and Sahin (2015) emphasize the role of cyclical shifts in the labor force attachment of the unemployed. The onset of recessions is accompanied by a wave of job loss, inducing an inflow into unemployment of attached workers with a low propensity to exit the labor force. This generates a procyclical movement in the labor force exit rate \(\pi_{UN}\). In complementary work, Barnichon and Figura (2013a) document that rises in marginally attached nonparticipants who

---

46 Explanations that receive weak support include the roles of extensions in unemployment insurance during recessions in deterring labor force exit (Rothstein 2011; Farber and Valletta 2013), and the added worker effect in inducing labor force entry (Elsby, Hobijn, and Sahin 2015). There is modest evidence that part of the rise in \(\pi_{NU}\) during recessions is a symptom of a blurred distinction between unemployment and nonparticipation. Elsby et al. (2011) document that many individuals identified as labor force entrants in longitudinal data subsequently report substantial search durations. Elsby, Hobijn, and Sahin (2013) find that adjustments for classification errors between unemployment and nonparticipation shade down the countercyclicality of labor force entry somewhat. Nonetheless, prominent cyclicality remains even in the adjusted flows. Performing the same exercise in figure 9A using adjusted worker flows reveals similar results.
report wanting a job, individuals more likely to flow into unemployment, have contributed to the rise in labor force entry in the recent recession.

These empirical findings dovetail with recent theoretical work that has sought to unify the theory of labor supply with unemployment determination. Krusell et al. (2012) generate compositional shifts that resemble those noted above from a rich distribution of worker heterogeneity driven by idiosyncratic productivity shocks, labor market frictions, and an endogenous wealth distribution arising from borrowing constraints. For example, increased rates of job loss and reduced rates of job finding after a recessionary shock imply that the average unemployed worker is further from the participation boundary—that is, more attached to the labor market—reducing the rate of labor force exit from unemployment.

8. Search Intensity

The focus of the previous section was on the extensive margin of the participation decision—that is, the decision to search or not. But it is natural to expect that the intensive margin of participation—search intensity—will also respond to changes in the economic environment. Furthermore, given the inherent ambiguity of the distinction between unemployment and nonparticipation, there is considerable merit in viewing search as a more continuous decision. The importance of search intensity for movements in the Beveridge curve derives from its potential role in raising the number of hires generated from a given number of searchers.

The literature on search intensity shares important parallels with the decision to participate, as one would expect, given their close conceptual link. A leading example is the extension of the canonical search model presented in chapter 5 of Pissarides (2000). There, aggregate worker search intensity $s$ augments the stock of unemployed workers in the matching function $m(s, v)$. Under constant returns, the number of hires per unit of worker search effort is therefore given by $m(1, \theta/s) \equiv f(\theta/s)$. An individual unemployed worker exerting search effort $s_i$ trades off an increasing and convex cost $c(s_i)$ with increases in the rate at which they find a job, $s_i f(\theta/s)$, and the associated surplus from work $W - U$. Optimal search effort equates the marginal costs and benefits of this trade-off. Given the symmetry of the problem, all workers choose the same search effort $s$ that solves

$$c'(s) = f(\theta/s)(W - U).$$

It is straightforward to show that search intensity is predicted to rise with the vacancy–unemployment ratio $\theta = v/u$—that is, $s$ is predicted to be procyclical.

8.1 Evidence on Search Intensity

Although direct measures of search intensity remain in short supply, the limited evidence available challenges this theoretical prediction, again echoing what we saw at the extensive margin. Early work by Shimer (2004a) used information from the Current Population Survey on the average number of search methods used by unemployed workers as a proxy for search intensity, noting that it rose significantly in the 2001 recession. Recent research has found similar patterns using newly available data from the American Time Use Survey to measure time allocated to search activity among the unemployed. While each unemployed worker spends, on average, just thirty to forty minutes per day on search, this number has risen over the course of the Great Recession (Krueger and Mueller 2010; Aguiar, Hurst, and Karabarbounis 2013; Mukoyama, Patterson, and Şahin 2014). In particular, Mukoyama, Patterson, and Şahin show that time devoted to search is strongly correlated in the cross-section with the
number of search methods, lending credence to Shimer’s estimates, and remains countercyclical even after controlling for compositional differences of the unemployed in recession.\footnote{These results recently have been challenged by Gomme and Lkhagvasuren (forthcoming). Using ATUS data, they find that recessions are accompanied by compositional shifts toward jobseekers that spend more time searching. Once account is taken of such biases, Gomme and Lkhagvasuren find that time spent on job search is procyclical. DeLouch and Kurt (2013) and Timen (2014) also report evidence for procyclical search effort.}

As Shimer (2004a) and Mukoyama, Patterson, and Şahin (2014) stress, conventional models like the one outlined above fail because they implicitly assume a complementarity between individual search intensity and labor market tightness in job-finding rates. Although these authors have explored alternative matching functions that allow for some substitutability in this technology, there is no consensus on the form of that technology. Consequently, it is not straightforward to provide a quantitative sense of the effect of search intensity on the Beveridge curve as we have done in previous sections of the paper. What does seem clear, however, is that the persistent rise in measured search intensity during and after recessions will contribute to a prolonged inward shift in the Beveridge curve—as workers search harder on average, fewer vacancies are needed to maintain unemployment at a given level. Since this prediction runs opposite to what has been observed since the Great Recession, a tentative conclusion is that recent trends in the Beveridge curve have emerged despite, rather than because of, movements in search intensity.

8.2 The Role of Unemployment Insurance

The conclusion that search intensity among the unemployed tends to rise in recessions is perhaps surprising in light of the fact that unemployment insurance (UI) benefits are routinely extended during recessions in the United States. In the simple model underlying equation (28), for example, unemployment benefits reduce a worker’s anticipated surplus from a job, \( W - U \), thereby discouraging search intensity.

Consistent with this, microeconometric evidence for the United States and Europe consistently finds that longer-duration benefits extend unemployment spells. However, leading estimates of these disincentive effects are statistically significant, but small. Recent research by Rothstein (2011), for example, finds that the substantial rise in the duration of UI benefits that accompanied the Great Recession can explain less than half of a percentage point of the 5.5 percentage point rise in the unemployment rate, a point reaffirmed by Farber and Valletta (2013).\footnote{Using cross-state variation in the duration of UI benefits, earlier studies in this field report estimates that are somewhat larger (Moffitt 1985; Katz and Meyer 1990; Meyer 1990; Card and Levine 2000; and Krueger and Meyer 2002). As Card and Levine (2000) note, however, many of the larger estimates in this range are likely overstated because many states extend UI benefits as a mechanical response to poor job-finding prospects.}

Moreover, these studies further suggest that recent UI extensions have operated mainly to reduce exits from unemployment to non-participation, rather than to employment. Analysis based on European data—see, for instance, Hunt (1995), Lalive (2008), and Schmieder, von Wachter, and Bender (2012)—report estimates largely in line with US results.\footnote{An exception is van Ours and Vodopivec (2006), who study Slovenian data. Other studies are somewhat difficult to compare with those noted in the main text. Bover, Arellano, and Bentolila (2002), for instance, estimate that the receipt of benefits reduces the exit rate from unemployment by 50 percent. However, they do not observe benefit levels or duration, so one cannot re-express their result in terms common to the rest of the literature. For a survey of earlier analyses, see the review in this journal by Atkinson and Micklewright (1991).}

The potential disincentive effects of unemployment insurance have been emphasized in the macroeconomic literature on the role of labor market institutions in...
shaping the dynamics of unemployment and vacancies. This literature has often appealed to differences in the generosity of UI to account for differing labor market experiences across countries. For example, Jackman et al. (1990) highlight the potential of greater benefit durations to account for higher unemployment rates, and Beveridge curves that lie to the right of those in countries with lower benefit durations. Nickell and Layard (1999) and Nickell et al. (2002) also argue that differences in UI durations across countries have clear predictive power for differences in unemployment rates.

Several caveats must be borne in mind when assessing this evidence, however. First, a drawback of such cross-country analyses is that it is difficult to rule out the possibility that unobserved characteristics of different economies drive the correlation between benefit durations and labor market outcomes. Second, it is possible that any such macroeconomic effects are conceptually distinct from the search intensity effects noted in the microeconometric literature. For example, the standard search model outlined in section 4 implies that more generous unemployment benefits will reduce the job creation incentives of firms by raising the wage demands of workers. In stark contrast to a decline in search intensity, these general equilibrium effects will not be manifested in a long-run outward shift in the Beveridge curve, but rather will mirror the effects of a reduction in labor demand, generating the negative comovement between unemployment and vacancies noted in figure 4B. Such macroeconomic effects, therefore, are unable to account for sustained disparities in Beveridge curves across countries.

9. Job-to-Job Transitions

In his presidential address to the American Economic Association, Tobin (1972) challenged the common abstraction from direct job-to-job transitions implicit in the previous analysis, and the related assumption that all worker turnover is mediated through the unemployment pool. The limited evidence available even at that time suggested that many quits instead transition immediately to another employer, with no intervening unemployment spell. Additional evidence amassed since then confirms that employer-to-employer (E-to-E) transitions are indeed abundant in the US labor market, and display systematic fluctuations over the business cycle. Using information on these transitions available in the Current Population Survey since its redesign in 1994, Fallick and Fleischman (2004) document that E-to-E transitions comprise around one-third of all hires, and fell markedly in the 2001 recession (see also Nagypál 2008).

51 Hagedorn et al. (2013) test the claim that the macroeconomic effects of UI extensions in the United States substantially depressed job creation in the Great Recession. They exploit cross-state differences in UI by comparing labor market outcomes across counties on opposite sides of a state border. The identifying assumption is that the county in a low-UI state can serve as the counterfactual for the county in a high-UI state. They find large effects: UI extensions during the recession account for more than 3 percentage points of the increase in the unemployment rate. At the time of writing, there remains some disagreement over these results (see Hall 2013).

52 Shortly after Tobin’s address, the work of Mattila (1974) summarized a range of early evidence highlighting the empirical importance of job-to-job quits. See also Akerlof, Rose, and Yellen (1988).
confirms that \( E \rightarrow E \) transitions also fell dramatically in the 2008 recession.\(^{53}\)

From the perspective of the Beveridge curve relation between unemployed searchers and vacancies, it is not immediately obvious how such job-to-job flows will play a role, since they bypass the unemployment pool. In what follows, we highlight three possible channels. First, the presence of employed searchers may crowd out the employment prospects of unemployed searchers. Second, if vacancies are posted in response to the search efforts of both unemployed and employed searchers, we will see that this can affect the amplitude and dynamics of the vacancy dimension of the Beveridge curve. Finally, influential recent work has emphasized that, even if the labor markets faced by unemployed and employed searchers are segmented, the presence of a \textit{job ladder} that links these labor markets yields important implications for Beveridge dynamics.

To begin, then, we first consider the implications of on-the-job search from the perspective of a standard random matching search model in which all job seekers, both employed and unemployed, search in the same market.\(^{54}\) The primary effect of this extension is that the stock of employed searchers, which we shall denote \( s_e \), now enters the matching technology that determines the number of hires in the economy, \( m(u + s_e, v) \). Under random matching and constant returns, all searchers will meet a vacancy at rate \( m(1, v/(u + s_e)) \equiv f(\sigma \theta) \), where \( \sigma = u/(u + s_e) \) denotes the share of searchers that is unemployed and, as before, \( \theta \) denotes the vacancy–unemployment ratio \( \epsilon/u \). An immediate implication is that changes in the composition of searchers embodied in \( \sigma \) will alter the dynamics of the law of motion for unemployment. To be precise, the evolution equation becomes

\[
\frac{du}{dt} = \lambda (1 - u) - f(\sigma \theta) u.
\]

This reveals that increases in the share of employed searchers \textit{crowd out} the job finding prospects of the unemployed for a given vacancy–unemployment ratio by depressing \( \sigma \).

There is, however, another channel through which on-the-job search can influence the job-finding rate, namely, by stimulating vacancy creation. To see this, note that the rate at which unfilled vacancies meet a searcher is given by \( m((u + s_e)/v, 1) \equiv q(\sigma \theta) \). It follows that, for a given vacancy–unemployment ratio \( \theta \), a greater number of employed searchers will \textit{raise} incentives for vacancy creation.\(^{55}\) Intuitively, the presence of an additional pool of available searchers raises the rate at which vacancies can be filled, which reduces the costs of recruitment.\(^{55}\) If we maintain the assumption of free entry from the canonical search model, so that the returns to vacancy creation \( \nu \) must satisfy the zero-profit condition \( \nu = 0 \), the resultant job creation condition will specify that augmented labor market tightness \( \sigma \theta \) will be pinned down in equilibrium, at a level determined by the state of the aggregate labor market.

\(^{53}\)It is difficult to distinguish a clear cyclical pattern over the short sample available in the CPS. Indeed, Fallick and Fleischman caution that \( E \rightarrow E \) transitions did not rise noticeably in the 1990s expansion. Other data sources developed since shed more light on this period. Using administrative data from unemployment insurance records, Bjelland et al. (2011) find that the ramp-up in \( E \rightarrow E \) transitions in the 1990s was concentrated between 1992 and 1995, a period not covered in full by the CPS.

\(^{54}\)Pissarides (1994) and chapter 4 of Pissarides (2000) extend the baseline search model to incorporate on-the-job search. Fujita and Ramey (2012) provide a detailed quantitative analysis of the model.

\(^{55}\)Conversely, an increase in the number of unemployed searchers slackens the market and reduces the rate at which the \textit{employed} contact vacancies. This implies a lower quit rate and, hence, longer-lived matches, which also raises the return to vacancy creation.
Figure 10. Job-to-Job Transitions and the Beveridge Curve

Notes: The employer-to-employer transition probability in panel A is estimated by Fallick and Fleischman based on updates of their 2004 manuscript. Implied on-the-job search intensity is inferred under the assumption of a common matching market (see text). The counterfactual Beveridge curve is estimated by netting off the shifts in vacancies implied by changes in on-the-job search intensity as measured in panel C.
The steady-state loci that shape the Beveridge curve will thus resemble those depicted in figure 10B. As before, recessionary declines in labor productivity \( p \) will rotate the equilibrium vacancy locus clockwise. But in addition, the latter locus will also shift down, amplifying the decline in vacancies. This latter shift reflects the presence of employed search. Since vacancies are posted, in part, to attract employed searchers, a slacker market now means that there are fewer vacancies for any level of \( u \). An increase in job destruction \( \lambda \) has the same qualitative effect on the vacancy locus in addition, of course, to shifting out the steady-state unemployment locus. The additional implications of the presence of on-the-job search turn crucially on the direction of cyclical movements in the search effort of employed workers \( s_e \), however.

Conventional extensions of the standard search model to incorporate on-the-job search capture a substitution effect whereby \( s_e \) responds to the returns to search (Pissarides 1994; Pissarides 2000, chapter 4). Fujita and Ramey (2012) note that these effects generate additional volatility in vacancies, relieving some of the deficiencies of the standard search model. The logic can be traced through using the arrows in figure 10B. Declines in search effort \( s_e \) among employees in recessions ease congestion in the search market, moderating the outward shift in the \( du/dt = 0 \) locus. This attenuates the rise in \( u \), but also curtails what would be a counterfactual recessionary increase in vacancies. In addition to this, potential employers lose a segment of searchers in the market following declines in \( s_e \), reducing vacancy-filling rates for a given vacancy–unemployment ratio. This diminishes vacancy creation incentives and shifts downward the job creation ray in figure 10B, thereby amplifying the rise in \( u \) and fall in \( v \). These dynamics have, then, partially offsetting effects on unemployment, but yield an unambiguously more precipitous decline in vacancies. As Fujita and Ramey (2012) stress, a particular benefit of this result is that this extended model is more adept at matching the negative comovement between unemployment and vacancies seen in the data, which was a failing of the standard model in section 4.

If the search effort of the employed \( s_e \) is procyclical, then, the presence of on-the-job search does much to resolve some of the enduring puzzles of the observed Beveridge curve relation. It is natural to ask, therefore, whether the search effort of the employed is indeed procyclical. Unfortunately, comparatively little research has addressed this important question, likely because available data on the issue are limited. In what follows, we suggest and review a couple of shreds of evidence.

First, it is important to note that the procyclicality of Fallick and Fleischman’s (2004) estimated job-to-job transition probability

\[ 57 \text{These offsetting effects on unemployment can in fact cancel exactly. In Pissarides (2000), the response of steady-state unemployment to a change in } p \text{ is unaffected by the change in } s_e. \text{ Fujita and Ramey (2012) confirm that this also serves as a good guide to out-of-steady-state dynamics. Interestingly, Nagypál (2008) shows that on-the-job search can amplify the effect of shocks on unemployment if firms must pay to train new hires. The reason is that hires originating from employment are expected to last longer—they are consummated only if the match is superior to the worker’s previous match. Accordingly, the training cost can be amortized over a longer period, making employed searchers more attractive. Since a vacancy is more likely to match with an unemployed worker in recessions, this attenuates the expected payoff from job creation and drives further increases in unemployment.} \]
illustrated in figure 10A is not sufficient to imply that \( s_e \) itself is procyclical, since the employer-to-employer rate is a conflation of search effort and the job-finding rate per searcher. Interestingly, however, the above framework suggests that a measure of \( s_e \) can be inferred using precisely this logic. In particular, note that the number of employer-to-employer transitions is given simply by the on-the-job searchers, \( s_e \), who contact a vacancy. The contact rate, for employed and unemployed searchers alike, is \( f(\sigma \theta) \). Hence, the job-finding rate of the unemployed and the employer-to-employer transition rate are given respectively by

\[
(30) \quad \pi_{UE} = f(\sigma \theta), \quad \text{and} \quad \pi_{EE'} = \frac{s_e f(\sigma \theta)}{1 - u}.
\]

Viewing the data through the lens of a simple model in which employed and unemployed job seekers search in the same market therefore suggests a simple indirect measure of the search effort of the employed given by

\[ s_e = (1 - u) \pi_{EE'}/\pi_{UE}. \]

Figure 10C depicts the time series of this measure of \( s_e \) using Fallick and Fleischman’s estimates. In contrast to the implications of the on-the-job search model summarized in chapter 4 of Pissarides (2000), this measure of \( s_e \), which is suggested by the model itself, is instead modestly countercyclical. Intuitively, the decline in the job-to-job transition rate in recessions is dominated by declines in the job-finding rate. The only way to rationalize this under the assumption of a common labor market is if effective search on the job increases in recessions.

While these estimates are only suggestive, it is worth noting that they have important implications. The presence of countercyclical search behavior among the employed suggests that on-the-job search decisions are shaped by an income effect, perhaps reflecting a deterioration in average match quality during recessions, and a related desire to search for better jobs (Barlevy 2002). In addition, the implications of on-the-job search for Beveridge dynamics are quite different. The arrows in figure 10B again illustrate the point. Instead of amplifying the response of vacancies, a countercyclical \( s_e \) contributes further to outward shifts in the Beveridge curve. Although this limits the usefulness of job-to-job flows as means to understanding the amplitude of labor market fluctuations, it does present the possibility that some of the lateral shift in the Beveridge curve seen since the Great Recession has its origins in the crowding out of unemployed searchers by the increased search efforts of the employed.

As in our analysis of the role of other extensions to the canonical search model, the potential effects of on-the-job search on the Beveridge curve can be quantified using a simple counterfactual exercise. Proceeding as before, differentiation of the law of motion for unemployment (29) suggests that the vertical shift in the Beveridge curve induced by

\[ \text{60One can estimate job search by employed workers more directly using the American Time Use Survey. We find that the average minutes of search by the employed are in fact relatively low in the aggregate expansion of 2003–07, and have been higher since then. But the incidence of search is surprisingly small, given the extent of employer-to-employer transitions. There could be good reasons such search activity is underreported, though. For instance, respondents are asked to report their “primary” activity during each hour at work. Even if they do search online job postings while at work, this is unlikely to be their primary activity.} \]

\[ \text{58In this benchmark model, all new matches are endowed with maximum productivity, so all contacts result in a match (Pissarides 2000).} \]

\[ \text{59For internal consistency, to infer } s_e \text{ we use Fallick and Fleischman’s (2004) estimate of the unemployment-to-employment transition probability to measure } f \text{ in equation (30). Similar results emerge if one uses instead Shimer’s (2012) simple two-state measure of the outflow rate from unemployment.} \]
variable search intensity of the employed equals

\[
\left. \frac{d \ln v}{d \ln s_e} \right|_{u, \bar{u} = 0} = 1 - \sigma.
\]

Figure 10D uses the latter to illustrate the counterfactual path of vacancies implied by holding constant on-the-job search intensity using the measure of \(s_e\) from figure 10C. This exercise suggests that the countercyclicality of this measure of \(s_e\) contributes somewhat to the rise in unemployment in recessions. Holding \(s_e\) constant shifts the counterfactual Beveridge relation inward, relative to its realized counterpart—fewer vacancies are needed to support a given level of unemployment. Similarly, this approach suggests that movements in on-the-job search are able to account for a fraction of the observed outward shift in the \((u, v)\) locus. A comparison of vacancy stocks in the first quarters of 2009 and 2012 is particularly instructive, since the unemployment rate was approximately 8.25 percent on both of these dates. The vertical shift in the observed Beveridge curve over this period amounts to 0.55 percentage points, or a little over 30 percent. Its counterfactual analogue is 0.4 percentage points, suggesting that this view of on-the-job search can account for around one-quarter of the realized shift in the Beveridge curve.

The goal of this exercise is not to claim that the measure of employed search intensity inferred above or its related counterfactual exercise is definitive. Rather, we wish to highlight the potential importance of gaining better measures and understanding of on-the-job search.

In the remainder of this section, we summarize some influential recent research that has identified additional channels through which job-to-job flows interact with unemployment and vacancy dynamics. A particularly important assumption underlying the above analysis is that both employed and unemployed job seekers search in the same market. In contrast to this is the model of directed search in Menzio and Shi (2011), in which unemployed and employed searchers sort endogenously into different submarkets by trading off the probability of finding work quickly with the value of the job. The unemployed sort into submarkets where the probability of finding work is high, and so accept low-value jobs. The employed wait longer for better jobs that dominate their current match.

Interestingly, Menzio and Shi’s (2011) results echo the implication described earlier that on-the-job search generates additional procyclicality in vacancies. After a positive productivity shock, vacancies rise on impact. As unemployment falls, however, the cost of recruiting the unemployed ramps up, and vacancies posted for the unemployed reverse course, reminiscent of the dynamics of the standard model depicted in figure 4B. However, the key difference is that newly hired workers are inclined to search on the job in this model. As a result, vacancies continue to rise in submarkets in which the employed search to match with workers who wish to transition to more productive matches. On net, aggregate vacancies rise much more than in the canonical search model without on-the-job search, and this is associated with a gradual ramp up in job-to-job flows.

Several recent papers have also captured this kind of job-ladder dynamic. In Krause and Lubik (2010), there are two types of jobs, bad (low-productivity) and good (high-productivity). In expansions, there is a gradual transition of workers toward their comparative-advantage jobs, which draws in vacancies and amplifies and propagates the
initial shock. Moscarini and Postel-Vinay (2013) develop a model in which this job-ladder dynamic emerges endogenously from wage posting decisions in a rich model of firm dynamics. Large, higher-productivity firms “poach” workers from smaller firms as an aggregate expansion matures. As theirs is a model with a well-defined notion of firm size, the authors can use this model to engage evidence on the dynamics of the firm size distribution and cyclical fluctuations across size classes (see Moscarini and Postel-Vinay forthcoming).

10. Heterogeneity

We have seen that an influential interpretation of the Beveridge curve is that it is a close relation of the matching function. As emphasized by Petrongolo and Pissarides (2001), the usefulness of the matching function as an abstraction is hinged on the existence and stability of an aggregate relation between the number of hires and the numbers of unemployed workers and vacant jobs. Such an abstraction becomes strained, however, when evidence emerges for a breakdown in that relationship. Viewed through the lens of the canonical search and matching model discussed in section 4, outward shifts in the Beveridge curve would be interpreted as reductions in matching efficiency. This in turn has raised the question of what exactly matching efficiency might be, and stimulated a reconsideration of the microfoundations of the matching function.

A key assumption maintained in the foregoing analysis is that unemployed job seekers face homogeneous employment prospects. A consequence is that the analytical framework considered thus far is incapable of assessing the effects of shifts in the composition of unemployment for the Beveridge curve. In this section, we consider two prominent sources of worker heterogeneity. The first, mismatch, refers to a horizontal notion of heterogeneity associated with the possible misallocation of workers across jobs. The second captures a particular form of vertical heterogeneity, known as duration dependence, in which employment prospects deteriorate with the length of an unemployment spell.

Both sources of heterogeneity have risen to the forefront of academic and policy debate in the wake of persistent rises in unemployment, such as those witnessed in European economies in the 1980s, and more recently during the lackluster recovery of the US labor market following the Great Recession. Both episodes were accompanied by changes in the structure of labor demand—for example, away from manufacturing in the 1980s, and away from construction and financial services in the recent recession—as well as large rises in the duration of unemployment. If there are frictions that prevent workers from reallocating across sectors, or reduce the search effectiveness of the long-term unemployed, these forces will naturally raise the level of unemployment for a given stock of vacancies, shifting out the Beveridge curve.

This reasoning can be formalized in a simple analytical framework. In particular, if we index subgroups of the labor market by \( j \), the law of motion for unemployment can be expressed as

\[
\frac{du}{dt} = \lambda (1 - u) - \Omega fu,
\]

where \( \Omega = \sum_j \omega_j p_j \), \( \omega_j = u_j / u \) is the unemployment share of group \( j \), and \( p_j \) is

\[
the job-finding rate of group \( j \), relative to an aggregate measure, \( f \). Although the exact interpretation of the latter will vary with the application, one can broadly conceive of \( f \) as the (unique) job-finding rate that would emerge if heterogeneity were suppressed. Note that changes in the unemployment pool are isomorphic to shifts in matching efficiency, providing a potential microfoundation for those shifts.

In this way, differences across groups in job-finding rates can interact with changes in the composition of the unemployment pool to induce shifts in the observed Beveridge curve. Specifically, assuming the elasticity of \( f \) with respect to vacancies is a constant and denoting it by \((1 - \alpha)\), the vertical shift in the Beveridge curve induced by movements in \( \Omega \) is simply:

\[
\frac{d \ln v}{d \ln \Omega} \bigg|_{u, v, \dot{u} = 0} = -\frac{1}{1 - \alpha}.
\]

In what follows, we use disaggregated data to infer measures of the compositional shifts associated with the sectoral and duration structure of unemployment embodied in \( \Omega \) and use equation (33) to trace out their implications for recent trends in the Beveridge curve.

10.1 Mismatch

To infer a measure of mismatch, it is necessary to take a stand on the allocation of workers across labor markets that would emerge in the absence of reallocation frictions. This is a particularly thorny problem, requiring an articulation of the equilibrium of an unobserved counterfactual economy. Building on the early work of Jackman and Roper (1987), \( \check{\text{S}}\text{ahin et al. (2014)} \) have proposed one fruitful way forward using simple models of labor market frictions, and successive extensions thereof, to infer proxies for mismatch. The ensuing discussion draws heavily from their insights.

An important motivation for the matching function is that it provides a reduced-form representation of heterogeneities in labor supply and demand that, in the presence of frictions, slow down the rate at which suitable employment relationships can be found. Thus, in contrast to a framework in which workers and firms participate in a single matching market, \( \check{\text{S}}\text{ahin et al. (2014)} \) envision the labor market as a collection of distinct submarkets differentiated by geography, industry, and occupation, each of which mirrors the standard search and matching model described in section 4. Thus, within each market \( j \), the presence of matching frictions slows down the rate at which workers searching in \( j \) are able to find jobs. To transit across markets, however, the worker must pay an additional reallocation cost. This cost is intended to represent a composite of potential sources—building a network of contacts, learning a new trade, moving locations, and so on. The key is that \( \check{\text{S}}\text{ahin et al. (2014)} \) propose a measure of the distortion of labor market outcomes implied by whatever frictions operate across sectors.

To convey the basic idea, in what follows we consider a simple special case of this environment in which labor markets are symmetric—they possess the same matching technology, \( m_j = m(u_j, v_j) \), and are equally productive—and where the vacancy stock is taken as given. Absent reallocation frictions, a social planner in this economy would wish to distribute unemployed workers across

\[62\text{This is in fact how we specify the aggregate object, } f, \text{ in all applications.}\]

\[63\text{In terms of the exact, underlying frictions, } \text{Herz and van Rens (2011)} \text{ argue that slow (real) wage adjustment is critical. Intuitively, mismatch can arise if wages in tight markets cannot adjust to absorb new workers. Herz and van Rens find that wages in some states do seem “too high,” in that they fail to track profits as prescribed by a Nash sharing rule. Their estimates of mismatch are no higher than those in } \check{\text{S}}\text{ahin et al. (2014).}\]
markets in such a way as to maximize the total number of hires. If the matching function is concave, the allocation that achieves this goal equalizes the marginal product of unemployed searchers with respect to hires, ∂m/∂u, across markets. Under constant returns, the optimal allocation takes a particularly simple form, namely that vacancy–unemployment ratios, and thereby job-finding rates, be equalized across sectors, vj/uj = θ* and fj = f*. Deviations from this benchmark, manifested in cross-sectoral dispersion in job finding rates, provide one interpretation of mismatch, and a potential foundation for decreases in observed matching efficiency.

As made clear earlier in this section, the effects of mismatch on the Beveridge curve are embodied in Ω in equation (32). In this context, Ω reflects the aggregate effects of deviations in the sectoral job finding rates from their optimal counterpart, Ω = ∑j ωj(fj/f*). A virtue of the conceptual framework articulated by Şahin et al. (2014) is that a measure of f* presents itself. To see this, one can imagine the planner as observing the aggregate stock of unemployment, u, and vacancies, v, that prevail at the start of any one period and redistributing workers across markets to equalize tightness, θj = θ*. Since the observed ratio, v/u, is a weighted average of the θj’s, it follows that v/u = θ*.

Therefore, the optimal job-finding rate is simply f* = m(1, v/u).64

Şahin et al. (2014) demonstrate how this framework can usefully be extended to allow for real distinctions between sectors. For instance, they incorporate sectoral differences in matching and production technologies, as well as endogenous vacancy creation. Because their results suggest that these extensions do not significantly alter implied trends in mismatch, we draw out implications for the Beveridge curve using the simple model outlined above.65

To illustrate the implications of this view of mismatch, figure 11A plots the implied path of Ω under the interpretation of a labor market as an industry. The figure is constructed using JOLTS data on vacancies for seventeen nonfarm industries matched with CPS data on the number unemployed that report that their last job was in one of these industries.

To compute Ω, all that remains is to specify the matching technology. Şahin et al. (2014) adopt a conventional Cobb–Douglas specification and report that a matching elasticity of α = 0.5 is representative of their estimated industry matching functions.

Figure 11A suggests that effective matching efficiency Ω fell by around five log points during the Great Recession as a result of increased cross-industry mismatch.66 However, this measure also suggests that the rise in mismatch subsides relatively quickly, returning to its prerecession level within two years of the NBER-dated trough. As noted, Şahin et al. (2014) find similar time variation in more elaborate indexes of mismatch that account for heterogeneity in matching and production.

65 In these extensions, the constrained-efficient level of tightness can no longer simply be inferred from the aggregate time series. For instance, if there were differences in matching efficiency across sectors, the planner would wish to know these in order to decide the allocation.

66 This measures industry unemployment by counting up the number of workers whose last job was in that industry. Yet in the model, what matters is the industry in which a worker now searches—uj is the number of workers who search in industry j. See Şahin et al. (2014) for an attempt to test the sensitivity of the analysis to this point.

67 Examination of the sources of the increase in dispersion during 2008 to 2009 suggest that it is due especially to a relative decline in vacancy–unemployment ratios in durable goods manufacturing and transportation, and a comparable relative tightening of the labor markets in health care and government.
technologies, endogenous vacancy creation, and proprietary data on help-wanted ads that allow a richer definition of labor markets according to industry, occupation, and geography. They conclude that rises in the cross-sectional dispersion of labor market outcomes have been modest and mostly a cyclical (as opposed to structural) phenomenon in the Great Recession.

This conclusion has its parallel in terms of its implications for the Beveridge curve. Figure 11B uses equation (33) to trace out a counterfactual path of vacancies implied by the realized path of unemployment, but holding constant $\Omega$. Consistent with the message of figure 11A and the conclusions of Şahin et al. (2014), removing the recessionary rise in measured mismatch shifts the Beveridge curve inward somewhat during this period, but this effect largely vanishes thereafter as mismatch subsides. Importantly, essentially none of the outward shift in the Beveridge curve that has emerged since the Great Recession can be explained by a degradation in matching efficiency associated with increased dispersion in cross-sectoral labor market outcomes. The simple reason is that apparent declines in match efficiency, as indicated by the shortfalls in job-finding and vacancy-filling rates in figure 3, have been broad based across sectors of the economy (Barnichon et al. 2012).

The theoretical framework underlying Şahin et al. (2014) has the advantage of suggesting a straightforward strategy for measuring mismatch. However, it is not the only conceptualization of mismatch proposed in the literature. In the remainder of this subsection, we briefly survey alternative approaches. While these do not lend themselves to simple accounting exercises like that above, quantitative analysis of these models suggests that mismatch can additionally provide important insights on the
negative comovement of unemployment and vacancies over the cycle.

Shimer (2007) envisions a set of distinct labor markets, as in the above framework, except that worker mobility across markets is assumed to follow an exogenous process and each labor market is taken to be competitive. It follows that markets exhibit either excess demand, in which there exist vacant jobs but no unemployment, or excess supply, whereby all jobs are filled and a fraction of workers remains unemployed. Unemployment is therefore the outcome of such mismatch. The lone endogenous variable in the model is the aggregate number of jobs, which is determined by a free entry condition. Through this channel, reductions in aggregate labor demand reduce aggregate job creation, increasing the number of markets with unemployed workers, and reducing the number of markets with vacancies. The result is greater mismatch. But note that the economy’s response traces out a downward-sloped locus in \((u, v)\) space. Thus, endogenous shifts in mismatch are associated with movements along a stable Beveridge curve locus, rather than movements of the locus itself. As Shimer emphasizes, the model-implied Beveridge curve bears a remarkable resemblance to observed \((u, v)\) dynamics in the United States prior to the Great Recession.

As with any parsimonious model, however, there are drawbacks. Importantly, the only means by which the model can generate shifts in the Beveridge locus is via changes in the numbers of workers and jobs per labor market. Intuitively, joining markets together lowers the likelihood of mismatch, as some markets with unemployed workers will be joined with markets with vacant jobs. Thus, outward shifts in the Beveridge curve, such as that witnessed since the Great Recession in figure 1, correspond to a world in which labor markets appear “smaller.” A fruitful avenue of future research in this vein, then, would be to model the boundaries of labor markets, and thereby the forces that can induce shifts in apparent match efficiency.

More recent work has enriched the theory of mismatch by introducing some notion of comparative advantage. This feature is omitted in the Şahin et al. (2014) framework, as workers inherit the productivity of the sector to which they move. Relaxing this assumption, several recent papers have allowed that workers either accumulate skills specific to a job (and thus find it costly to leave) or ex ante are a better “fit” with some jobs.

Carrillo-Tudela and Visschers (2013) study an environment in which mismatch emerges endogenously because workers can move out of their (potentially long-held) occupations only at a cost. They highlight a natural interaction between worker mobility and the business cycle: if increases in aggregate productivity during expansions raise the returns to a good worker–occupation match, reallocation of workers across markets will be procyclical, a prediction mirrored in data on occupational transitions. Carrillo-Tudela and Visschers show that this procyclical reallocation in turn generates a feedback effect on job creation: employers post even more vacancies in expansions to take advantage of the improvement in the quality of the searching pool. In this way, endogenous worker mobility interacts with job creation decisions to magnify the negative comovement of unemployment and vacancies along the Beveridge curve.

Complementary to this is work that considers permanent two-sided firm and worker heterogeneity. In these settings, relatively straightforward measures of mismatch obtain only in restrictive cases. Eeckhout

---

68 The implicit premise of Shimer’s model is that the costs of reallocation are large, so that endogenous mobility is not a first-order concern at business cycle frequencies. Accordingly, he assumes that jobs and workers are randomly allocated across sectors. Intuitively, the distribution of workers across sectors reflects idiosyncratic factors that are orthogonal to the business cycle.
and Kircher (2010) show that the prescription that it is efficient to equate labor market tightness across markets holds only in very special cases—for example, in settings with positive assortative matching in which worker and firm types are symmetric in production and matching technologies, and where the type distributions also are symmetric.

Ongoing work by Lise and Robin (2013) uses a model of permanent two-sided heterogeneity to infer the cyclicity of mismatch. In their setting, there exists a firm that is most suitable for each worker, but search frictions and aggregate productivity shocks prevent all workers from matching with their ideal partner. This model therefore captures a richer notion of mismatch. Recessions have two effects on mismatch. First, during downturns, new hires from unemployment accept jobs only if the match is more suitable, reducing mismatch. Second, employed workers move more slowly toward their ideal job as the job-finding rate falls, raising mismatch. On net, the authors find that aggregate mismatch is endogenously procyclical, in contrast to the results of Şahin et al. (2014) noted above. This richer environment therefore highlights an important distinction to notions of mismatch: if the allocations of workers and firms observed in the economy reflect the endogenous sorting of workers and firms according to productivity, the gains from reallocation become more intricate, and may even decline during recessions.

10.2 Duration Dependence

Historically, shifts in Beveridge curves often have coincided with substantial rises in long-term unemployment. This was apparent in the European unemployment problem of the 1980s (Layard, Nickell, and Jackman 1991), and has been mirrored more recently in the US experience since the Great Recession. Of course, it is natural that a worsening of the Beveridge relation between unemployment and vacancies will be accompanied by increased duration of jobless spells: if matching efficiency deteriorates, this will depress job finding prospects of the unemployed.

But it is also possible that declines in match efficiency are themselves aggravated by rises in the duration of unemployment. For example, if reemployment rates decline with the length of an unemployment spell, shifts in the duration structure toward longer spells will generate fewer hires for a given level of vacancies at the aggregate level. The economy will appear as if matching efficiency has declined. This is the possibility we consider in this subsection.

At first blush, there is ample evidence in the United States to suggest that average exit rates from unemployment decline with duration. \[ \text{Figure 12A} \] illustrates this point using the longitudinally linked Current Population Survey microdata. Specifically, it plots the monthly probability of exiting unemployment for five duration bins, averaged over the third quarter of 2010 to the second quarter of 2011 (which captures the state of the labor market shortly after the recession ended). The outflow probability declines by 30 percent as duration progresses from less than five weeks to between five and fourteen weeks. Those with duration exceeding twenty-seven weeks are around half as likely to escape from unemployment than those with durations less than five weeks.

69 Interestingly, the shift in the US Beveridge curve in the 1970s was not accompanied by as dramatic a rise in mean duration. Instead, the rise in unemployment inflows is able to account for the entire shift during the 1970s.

70 As surveyed by Machin and Manning (1999), evidence on the relationship between average unemployment exit rates and duration outside of the United States is remarkably inconclusive. For example, a negative relation between exit rates and duration is typically found in the United Kingdom, there is little evidence for a relation in France, and estimates for Germany and Spain differ significantly in sign across studies. Likewise, Elsby, Hobijn, and Şahin (2013) find limited evidence for a relation between exit rates and duration among European economies.
Panel A. Negative duration dependence in exit rates

Panel B. Long-term unemployment and implied match efficiency

Panel C. Counterfactual Beveridge curve

Figure 12. Duration Dependence and the Beveridge Curve

Notes: Panel A plots the average unemployment exit probability by duration class for the period 2010Q3 to 2011Q2 estimated from matched monthly Current Population Survey data. The matching follows the procedure described in Nekarda (2009). The long-term unemployment share is the fraction unemployed at durations 27 weeks or greater calculated from the same matched CPS data as in panel A. The counterfactual Beveridge curve is estimated by netting off the shifts in vacancies implied by changes in duration as measured in panel A.
It has long been understood, however, that such negative duration dependence in average outflow rates need not imply that spell length has a causal effect on the propensity to exit unemployment, what is referred to as “true” duration dependence. Instead, it can reflect a process of dynamic selection whereby workers with high job-finding rates exit unemployment quickly, leaving behind job seekers who have low job-finding rates (Salant 1977). This distinction has a bearing on the interpretation of the effects of duration dependence on the Beveridge curve. If duration dependence reflects unobserved heterogeneity, declines in average exit rates by spell length are themselves the outcome of compositional shifts across unobserved worker types. In this context, it is not appropriate to attribute the effects of such compositional shifts to duration.

One example of compelling evidence against duration dependence comes from data on recalls, whereby an unemployed worker returns to her original employer. A consistent empirical finding is that negative dependence seems confined in household survey data to workers who are waiting to be recalled (Katz 1986; and Fujita and Moscarini 2013). Since the employers know these workers, and since the workers are returning to jobs they have been doing all along, it is hard to see why such duration dependence should be attributed to a decline in the worker’s productivity over her unemployment spell. Rather, it is tempting to conclude that negative dependence simply reflects unobserved (to the econometrician) heterogeneity.

This evidence notwithstanding, the literature has not settled on a clear consensus regarding the source of duration dependence.72 Reacting to conflicting findings, recent efforts have applied randomized control trial methods in an attempt to sharpen identification. Kroft, Lange, and Notowidigdo (2013) examine the effect of unemployment duration on callback rates using fictitious job applications in which duration is manipulated exogenously. Kroft et al. find that eight months of unemployment reduce the probability of callback by 3 percentage points, which amounts to a 45 percent reduction relative to short-duration (one-month) applicants. Eriksson and Rooth (2014) and Ghayad (2013) also run field experiments of this form, finding evidence of negative dependence. While the mapping from callback rates to job-finding propensities is not straightforward, this evidence suggests there is some merit in investigating the implications of negative duration dependence under the assumption that it is causal.

Following Barnichon and Figura (2013b), and Kroft et al. (2013), we use the apparatus in equations (32) and (33) to infer the implications of shifts in the duration structure for the Beveridge curve. To implement this exercise, however, it is useful to consider briefly a theoretical setting in which $\Omega$ can be measured and interpreted. The simplest such setting follows a long literature in assuming workers’ skills depreciate with the duration of a jobless spell,74 and that workers of all duration classes search in the same matching market. The latter implies that unemployed workers of all durations contact vacancies at

---

71 See in particular the early work of Kaitz (1970), Lancaster (1979), and Nickell (1979).
72 Using different approaches to control for unobserved heterogeneity, US studies variously have concluded that duration dependence is negative (Imbens and Lynch 2006; white men in van den Berg and van Ours 1996; absent black job seekers in van den Berg and van Ours 1996). Buffet (2010), and positive (Heckman and Singer 1984; and Meyer 1990); and positive (Heckman and Singer 1984; and Meyer 1990).
73 This quasi-experimental evidence provides a compelling reason to believe that employers discriminate against the long-term unemployed. But even so, there might not be equilibrium duration dependence if job searchers adjust their search strategies accordingly. For instance, they could “lower their standards” by accepting lower-paying offers.
the same rate $\phi$ determined by the aggregate vacancy–unemployment ratio $\theta = \psi/h$, as in the standard search model. Upon contacting an employer, however, a worker’s duration is revealed and a decision is made on whether to consummate a match. By linking expected productivity to unemployment duration, the probability of consummation $p_d$, and hence job-finding rates $f_d = \phi(\theta)p_d$, will decline in duration $d$. Thus, movements in the rates at which the unemployed of all durations contact vacancies $\phi(\theta)$, and the relative job-finding rates embodied in $p_d$ are decoupled.

In this setting, the role of duration dependence in shaping match efficiency can be inferred from a simple measure in the spirit of equation (33) above. In particular, the average exit rate from unemployment can be written as $\Omega \phi(\theta)$, where $\Omega = \sum_d \omega_dp_d$ and the $\omega_d$s are the unemployment shares of duration group $d$. Many of the elements of $\Omega$ can be measured using data from the Current Population Survey. The unemployment shares $\omega_d$ can be computed from the cross section, and outflow rates by duration $f_d = \phi(\theta)p_d$ can be estimated from the underlying microdata matched across consecutive months. Because the contact rate $\phi(\theta)$ is not observed, however, we make the identifying assumption that the very short-term unemployed match upon contact with probability one, $p_1 \equiv 1$. Accordingly, we can interpret $p_d$ for all other duration classes as their job-finding rates relative to the finding rate of the lowest-duration searchers, and the aggregate effects of shifts in the duration structure are summarized by $\Omega = \sum_d \omega_d(p_d/p_1)$.

The time series for this measure of $\Omega$ is depicted in figure 12B. The results are stark. Viewed through the lens of the simple model above, matching efficiency declined by almost one quarter around the Great Recession as a result of changes in the duration structure interacted with negative duration dependence. Moreover, the path for matching efficiency since the recent recession has been somewhat subdued.

These observations are echoed in the counterfactual Beveridge curve displayed in figure 12C. Recall that the interpretation of this exercise is that it sketches out the $\Omega$-constant path of vacancies implied by the observed path of unemployment. Two features stand out. First, the counterfactual Beveridge curve is, on average, shifted inward the

---

75 The literature has questioned whether workers continue to search through the full stock of vacancies during the course of their unemployment spell. Instead, the notion of stock-flow matching characterizes the hiring process as one where jobseekers wait for a suitable match among the incoming flow of vacancies and vacant jobs remain unfilled as employers await a suitable employee among the flow of new searchers. Coles and Smith (1998) apply this notion of stock-flow matching to the UK labor market. Consistent with this alternative view of the hiring process, they find that the exit rates of jobseekers that have been unemployed for at least a month are more closely associated with the flow of new vacancies than the stock. Nonetheless, Ebrahimy and Shimer (2010) show that a stable Beveridge curve relationship in the stocks of $u$ and $v$ can emerge from a model of stock-flow matching.

76 An example of how this decoupling can fail is Lockwood’s (1991) model of adverse selection. As Lockwood stresses, when duration serves as a noisy signal of worker quality, employers should recognize that even good workers exit unemployment at a lower rate in recessions, shifting the Beveridge curve inward.

77 It should be noted, however, that there are important inconsistencies between reported durations in the cross-section, and the path of the duration structure implied by the longitudinally linked microdata on worker flows by duration (Elsby et al. 2011; Rothstein 2011). Almost half of measured inflows into unemployment in the longitudinal data report durations exceeding four weeks, contributing to the mass at higher durations in the cross-sectional duration distribution. This discrepancy is almost entirely associated with inflows into unemployment of individuals classified as out of labor force in the previous month.

78 One difference to the preceding analysis is that, in this case, $1 - \alpha$ corresponds to the elasticity of the exit rate of the lowest-duration job searcher with respect to labor market tightness $\theta$. Accordingly, we estimate $1 - \alpha$ by regressing the log of the short-duration exit rate on the log of the vacancy-unemployment ratio.

79 The data begin in 1994, which was the year of the redesign of the Current Population Survey.
origin: for any given vacancy stock, duration dependence accounts for a portion of unemployment in any state. Second, and more striking, after accounting for the persistent decline in estimated matching efficiency \( \Omega \) since the Great Recession, the outward shift in the Beveridge curve essentially vanishes.

The message of this exercise is that it is possible to reconcile much of the recent behavior of the Beveridge curve in the United States by appealing to a view of the labor market in which skills are lost during unemployment spells. A few qualifications should be noted, however. On the one hand, the above exercise attributes the entirety of the observed duration dependence in average exit rates to skill depreciation. Although the recent evidence of Kroft et al. (2013) is consistent with the notion that higher durations are associated with some loss of (perceived) skill, it seems unlikely that this drives all of the observed duration dependence. On the other hand, indirect effects of skill depreciation may amplify its influence on duration. For instance, a worker may withdraw search effort if skill loss diminishes the returns from effort (by reducing wage offers). This feedback onto search effort would amplify duration dependence. For example, in his discussion of evidence presented by Krueger and Mueller (2011) that search effort declines over the course of unemployment spells, Davis (2011) stresses that this will, in turn, interact with the increase in long-duration unemployment to lower the aggregate job-finding rate.

11. Conclusions

Important progress has been made in economists’ understanding of the dynamics of unemployment and vacancies summarized in the Beveridge curve. Our survey began with a general description of the dynamics of vacancies and unemployment devised in the early work of Blanchard and Diamond (1989). We showed that this framework provides a tractable lens through which movements along the Beveridge curve, and of the locus itself, inform the sources of cyclical disturbances, as well as long-run trends in the functioning of the labor market. The canonical Diamond–Mortensen–Pissarides search and matching model provides one formalization of job creation within this framework, and has yielded important insights into the essential role of fluctuations in vacancy creation in generating observed Beveridge dynamics.

Yet substantial gaps in our understanding remain. Chief among these are the amplitude, comovement, and persistence of cyclical variation in unemployment and vacancies, and the sources of the seeming shifts in the efficiency of the matching process implied by lateral movements of the Beveridge curve, such as that witnessed recently in the United States. The general framework that we use can be amended to incorporate a range of models of vacancy creation and matching assumptions that are omitted in the baseline search model. Doing so identifies several areas in which additional research seems especially fruitful. These include the role of wage determination and entry into vacancy creation on the volatility and sluggishness of vacancy dynamics; the cyclicality of labor market flows across the participation threshold; the apparent countercyclicality of search effort, both off and on the job; and the sources and consequences of duration dependence. Further theoretical and empirical work is needed on the mechanisms through which each of these channels affects the Beveridge curve. We are hopeful that future work will provide just that.

---

80 This feedback would not be picked up in the audit studies of, for instance, Kroft et al. (2013), which estimate callback conditional on application. This effect instead suggests that fewer long-duration searchers would bother to apply in the first place.
References


Davis, Steven J., R. Jason Faberman, John C.
Eriksson, Stefan, and Dan-Olof Rooth. 2014. “Do


Kocherlakota, Narayana. 2010. “Inside the FOMC.” President’s Speech, Marquette, August 17.


Moscarini, Giuseppe, and Fabien Postel-Vinay. Forthcoming “Did the Job Ladder Fail after the Great Recession?” Journal of Labor Economics.


