This paper proposes a new method to estimate the unemployment gap (the actual unemployment rate minus the efficient rate). While lowering unemployment puts more people into work, it forces firms to post more vacancies and devote more resources to recruiting. This unemployment-vacancy tradeoff, governed by the Beveridge curve, determines the efficient unemployment rate. Accordingly, the unemployment gap can be measured from three sufficient statistics: the elasticity of the Beveridge curve, cost of recruiting, and social cost of unemployment. In the United States the unemployment gap is countercyclical, reaching 1.5–6.5 percentage points in slumps. Thus the US labor market appears inefficient—especially inefficiently slack in slumps.
1. Introduction

Perhaps the most important statistic in the design of stabilization policy is the unemployment gap: the difference between the actual and the socially efficient unemployment rate.\(^1\) Once the world emerges from its coronavirus quarantine and economic activity resumes, policymakers will need to assess the unemployment gap. The gap will indicate how much inefficient slack remains within their economies, and the extent to which stabilization policies need to be brought to bear as a result.

Two measures of the unemployment gap are commonly used (Crump et al. 2019); but neither incorporates an appropriate estimate of the efficient unemployment rate. The first is the difference between actual unemployment and its secular trend. This measure is flawed because in most models trend unemployment is not efficient (Pissarides 2000, chap. 8). The second is the difference between actual unemployment and the non-accelerating inflation rate of unemployment (NAIRU), obtained by estimating a Phillips curve. This measure is flawed because the NAIRU never indicates labor-market efficiency (Rogerson 1997). Thus, although these two measures are easy to use, they lack a theoretical foundation.

This paper proposes a new measure of the unemployment gap that builds upon the theory of efficiency in modern labor-market models (Pissarides 2000, chap. 8). These models feature both unemployed workers and job vacancies, each associated with welfare costs: more unemployment means fewer people at work so less output; more vacancies mean more work effort devoted to recruiting and also less output. Furthermore, these models feature a Beveridge curve, so unemployment and vacancies cannot be simultaneously reduced: less unemployment requires more vacancies, and fewer vacancies create more unemployment. Our analysis resolves this unemployment-vacancy tradeoff, characterizing the efficiency point on the Beveridge curve.

At the same time, we strive to develop an unemployment-gap formula that is usable for policy work. To that end, we adhere to the sufficient-statistic method from public economics (Chetty 2009). A first advantage is that our formula requires little theoretical structure. It applies to any labor market with a Beveridge curve, irrespective of the structure of the labor market, production, preferences, wage setting, and shocks; the model’s relevant properties are captured

---

\(^1\)In practice, many governments are mandated to reduce the unemployment gap to zero. In the United States, the 1978 Humphrey-Hawkins Full Employment Act mandates that the government maintains the economy at “full employment” using monetary and fiscal policy. Full employment should not be interpreted as no unemployment, which is physically impossible, but rather as a socially efficient amount of unemployment; therefore, the mandate of US policymakers can be interpreted as closing the unemployment gap. Macroeconomic theory also shows that many stabilization policies should move in tandem with the unemployment gap. Such policies include labor-market subsidies and taxes (Pissarides 2000, chap. 9); monetary policy (Michaillat and Saez 2014); public expenditure (Michaillat and Saez 2019); and short-time work (Giupponi and Landais 2018).
by the formula’s sufficient statistics. Since the Beveridge curve appears in many models and has been observed in many countries (Elsby, Michaels, and Ratner 2015), the formula applies broadly. A second advantage is that the sufficient statistics are estimable, so the formula is simple and transparent to apply.

We begin by solving the problem of a social planner who must allocate labor between production, recruiting, and unemployment subject to the Beveridge-curve constraint. The solution gives the efficient unemployment rate, which we express as a function of actual unemployment and vacancy rates, and three sufficient statistics: the elasticity of the Beveridge curve, cost of recruiting, and social cost of unemployment.

Next, we compute the efficient unemployment rate in the United States. Between 1951 and 2019 the efficient unemployment rate averaged 4.2%. It started around 3% in the 1950s, steadily climbed to reach about 6% in the mid-1980s, fell to 4% in 1990, and remained between 3% and 4% until 2019. These variations are caused by shifts of the Beveridge curve.

Since the efficient unemployment rate is slow-moving while the actual unemployment rate is countercyclical, the unemployment gap is countercyclical. We infer that the US unemployment gap is almost never zero: the US labor market does not operate efficiently. In fact the unemployment gap is generally positive, averaging 1.6 percentage points over 1951–2019, so the US labor market is generally inefficiently slack. The unemployment gap is especially high in slumps, reaching for instance 5 percentage points in 1982 and 6.5 points in 2010: unsurprisingly, inefficiencies are exacerbated in slumps.

The most uncertain statistic in our formula is the social cost of unemployment, which corresponds to one minus the social value of nonwork. Our midrange estimate for the social value of nonwork is 0.25, meaning that the value from home production and recreation during unemployment replaces 25% of the marginal product of labor. For other plausible estimates of the social value of nonwork, the efficient unemployment rate does not change too much: from a low-end estimate of 0 to a high-end estimate of 0.5, the average efficient unemployment rate only increases by 1.5 percentage points. In contrast, some macroeconomic studies argue that the social value of nonwork is almost 1. Under this calibration the efficient unemployment rate is so high that unemployment is always inefficiently low—even during the Great Recession. This result seems implausible, suggesting that such calibration understates the social cost of unemployment.

2. Beveridgean labor market

We introduce the labor-market model used to compute the unemployment gap. The main ingredient is a Beveridge curve—a negative relation between unemployment and vacancies.
2.1. Beveridge curve

We consider a labor market with both unemployed workers and job vacancies. The unemployment rate $u$ is the number of unemployed workers divided by size of the labor force. The vacancy rate $v$ is the number of vacancies divided by size of the labor force. Labor-market tightness is number of vacancies per unemployed worker: $\theta = v/u$. The employment rate is the number of employed workers divided by size of the labor force: $n = 1 - u$.

Unemployment and vacancy rates are related by a Beveridge curve: the vacancy rate is a strictly decreasing and convex function of the unemployment rate, $v(u)$. Many labor-market models feature a Beveridge curve and so are nested into our framework (Elsby, Michaels, and Ratner 2015). Importantly, models build around matching functions exhibit a Beveridge curve (Petrongolo and Pissarides 2001, eq. (12)). This category includes the canonical Diamond-Mortensen-Pissarides model (Pissarides 2000, chap. 1); but also its variants with rigid wages (Hall 2005a; Hall and Milgrom 2008), large firms (Cahuc, Marque, and Wasmer 2008; Elsby and Michaels 2013), and job rationing (Michaillat 2012). Even models without a matching function may feature a Beveridge curve: for instance, models of mismatch (Shimer 2007) and of stock-flow matching (Ebrahimy and Shimer 2010).

Our framework applies to any country that exhibits a Beveridge curve. There are many such countries (Jackman, Pissarides, and Savouri 1990; Nickell et al. 2002; Elsby, Michaels, and Ratner 2015), including the United States (Blanchard and Diamond 1989; Diamond and Sahin 2015; Elsby, Michaels, and Ratner 2015). As an illustration, we construct the US Beveridge curve. For the unemployment rate, we use the standard measure constructed by the Bureau of Labor Statistics (BLS) from the Current Population Survey (CPS) (figure 1, panel A). For the vacancy rate, we use two different sources because there is no continuous national vacancy series over the period. For 1951–2000, we use the vacancy proxy constructed by Barnichon (2010). Barnichon starts from the help-wanted advertising index constructed by the Conference Board—a proxy for vacancies proposed by Abraham (1987) that has become standard (Shimer 2005, p. 29). He then corrects the Conference Board index, which is based on newspaper advertisements, to take into account the shift from print advertising to online advertising after 1995. Finally, he rescales the index into vacancies, and divides the vacancy number by the size of the labor force to obtain a vacancy rate.

---

2In many matching models, unemployment follows a law of motion, and the Beveridge curve is defined as the locus of unemployment and vacancies consistent with a steady level of unemployment. If unemployment converged slowly to steady state, unemployment and vacancies would not always be on the Beveridge curve. However, Pissarides (2009a, p. 236) notes that “Perhaps surprisingly at first, but on reflection not so surprisingly, we get a good approximation to the dynamics of unemployment if we treat unemployment as if it were always on the Beveridge curve.” The reason is that labor-market flows are large so after a shock, unemployment adjusts rapidly to its new steady-state level, where inflows into unemployment equal outflows from unemployment (Pissarides 1986; Hall 2005b; Pissarides 2009b; Elsby, Michaels, and Solon 2009; Shimer 2012).
Panel A: The unemployment rate is constructed by the BLS from the CPS. Panel B: For 1951–2000, the vacancy rate is constructed by Barnichon (2010) from the Conference Board help-wanted advertising index; for 2001–2019, the vacancy rate is the number of job openings measured by the BLS in JOLTS, divided by the civilian labor force constructed by the BLS from the CPS. All unemployment and vacancy rates are quarterly averages of seasonally adjusted monthly series. The shaded areas represent recessions, as identified by the National Bureau of Economic Research (NBER). Panels C and D display scatterplots of the unemployment and vacancy rates (from panels A and B) using a logarithmic scale. Panel C depicts the 1951–1989 period, and panel D the 1990–2019 period. The subperiods during which the Beveridge curve is stable are depicted in color: 1951Q1–1959Q2, 1959Q4–1971Q1, 1971Q3–1975Q1, 1975Q3–1987Q3, 1990Q1–1999Q1, 2001Q1–2009Q3, and 2010Q1–2019Q4. Quarters during which the Beveridge curve shifts are depicted in gray.
For 2001–2019, we obtain the vacancy rate from the number of job openings measured by the BLS in the Job Opening and Labor Turnover Survey (JOLTS), divided by the civilian labor force constructed by the BLS from the CPS. We then splice the Barnichon and JOLTS series to obtain a vacancy rate for 1951–2019 (figure 1, panel B).

The Beveridge curve appears in scatterplots of the unemployment and vacancy rates (panels C and D of figure 1; for readability we separately plot the 1951–1989 and 1990–2019 periods). The Beveridge curve was stable in seven subperiods, during which unemployment and vacancies moved up and down along a clearly defined curve: 1951Q1–1959Q2, 1959Q4–1971Q1, 1971Q3–1975Q1, 1975Q3–1987Q3, 1990Q1–1999Q1, 2001Q1–2009Q3, and 2010Q1–2019Q4. At the end of each of the first three subperiods, the Beveridge curve shifted outward. After the 1975Q3–1987Q3 and 1990Q1–1999Q1 subperiods, the Beveridge curve shifted back inward. Finally, after the 2001Q1–2009Q3 subperiod, the Beveridge curve shifted back outward.

Plotted on a logarithmic scale, all the branches of the Beveridge curve are almost linear, so each branch is isoelastic. A central statistic to measure the unemployment gap will be the Beveridge elasticity:

**Definition 1.** The Beveridge elasticity is the elasticity of the vacancy rate with respect to the unemployment rate along the Beveridge curve, normalized to be positive:

\[
\epsilon = -\frac{d \ln(v(u))}{d \ln(u)}.
\]

In a Diamond-Mortensen-Pissarides model with Cobb-Douglas matching function \( m(u, v) = \mu u^\alpha v^{1-\alpha} \), the Beveridge elasticity is closely related to the matching elasticity \( \alpha \). In that model the Beveridge curve is obtained by equating inflows into unemployment with outflows from unemployment: \( s \cdot (1 - u) = m(u, v) \), where \( s \) is the job-separation rate, so that

\[
v(u) = \left[ \frac{s(1 - u)}{\mu u^\alpha} \right]^{1/(1-\alpha)}
\]

(Shimer 2005, p. 36). Hence, the Beveridge elasticity is

\[
\epsilon = \frac{1}{1 - \alpha} \left( \alpha + \frac{u}{1 - u} \right).
\]

### 2.2. Social welfare

The Beveridge curve determines the tradeoff between unemployment and vacancies. This tradeoff is central to the welfare analysis because both unemployment and vacancies enter the welfare
We assume that social welfare is given by a function \( W(n, u, v) \), where \( n \) is the employment rate, \( u \) is the unemployment rate, and \( v \) is the vacancy rate. The function \( W \) is differentiable and strictly increasing in \( n \). The effects of unemployment and vacancies on welfare are given by the following two statistics, which play a key role in calculating the unemployment gap:

**Definition 2.** The social value of nonwork is the marginal rate of substitution between unemployment and employment in the welfare function:

\[
\zeta = \frac{\partial W/\partial u}{\partial W/\partial n} < 1.
\]

The social cost of unemployment is \( 1 - \zeta > 0 \).

**Definition 3.** The recruiting cost is minus the marginal rate of substitution between vacancies and employment in the welfare function:

\[
\kappa = -\frac{\partial W/\partial v}{\partial W/\partial n} > 0.
\]

Employed workers contribute to social welfare through market production. Unemployed workers contribute to social welfare through home production and recreation (Aguiar, Hurst, and Karabarbounis 2013); this contribution is diminished if people suffer psychic pain from being unemployed (Brand 2015). The social value of nonwork \( \zeta \) measures the marginal contribution of unemployed workers relative to that of employed workers. Since \( \zeta \) is a social concept, it does not include monetary transfers received by unemployed workers from the government or others.

Unemployment is socially costly because unemployed workers’ contribute less to welfare than employed workers (\( \zeta < 1 \)). The social cost of unemployment \( 1 - \zeta \) measures the social loss of having a person unemployed rather than employed. Such loss comprises foregone market production and psychological pain of being unemployed, net of the value home production and recreation when unemployed.

Vacancies enter the welfare function because the recruiting activity required to fill vacancies diverts labor and other resources away from market production. The recruiting cost \( \kappa \) measures the resources absorbed by maintaining a vacancy, expressed in terms of labor. It is normalized to be positive.

Since the labor force is divided between employed and unemployed workers, the employment rate satisfies \( n = 1 - u \), and social welfare can be written as a function of unemployment and
vacancy rates:

\[(u, v) \mapsto W(1 - u, u, v).\]

The assumptions that \(\zeta < 1\) and \(\kappa > 0\) imply that the function (2) is strictly decreasing in \(u\) and \(v\)—which captures the social costs of unemployment and vacancies. Moreover, to ensure that the social planner’s problem is well behaved, we assume that this function is quasiconcave.

In the Diamond-Mortensen-Pissarides model the statistics \(\zeta\) and \(\kappa\) take a simple form. Let \(L\) be the size of the labor force; \(p\) the productivity of employed workers; \(z < p\) the productivity of unemployed workers; and \(p \times c\) the resource cost of posting a vacancy (Pissarides 2000, chap. 1). The utility function is linear, so the welfare function is

\[W(n, u, v) = (pn + zu - pc) L.\]

Accordingly, the social value of nonwork and recruiting cost are

\[
(3) \quad \zeta = \frac{z}{p} \quad \text{and} \quad \kappa = c.
\]

3. Efficient unemployment rate and unemployment gap

In a Beveridgean labor market, we define efficiency as the solution to the problem of a social planner who is subject to the Beveridge-curve constraint:

**Definition 4.** The efficient unemployment and vacancy rates, denoted \(u^*\) and \(v^*\), maximize social welfare (2) subject to the Beveridge-curve constraint \(v = v(u)\). The efficient labor-market tightness is \(\theta^* = u^* / v^*\), and the unemployment gap is \(u - u^*\).

We now represent labor-market efficiency in a Beveridge diagram, and then derive a sufficient-statistic formula for the efficient unemployment rate and unemployment gap.

3.1. Representation in a Beveridge diagram

The Beveridge diagram, with unemployment rate on the \(x\)-axis and vacancy rate on the \(y\)-axis, is depicted in figure 2, panel A. In the diagram the Beveridge curve is downward-sloping and convex; it gives the locus of unemployment and vacancy rates that are feasible in the economy. The diagram also features an isowelfare curve: the locus of unemployment and vacancy rates such that social welfare (2) remains constant at some level (the equivalent of an indifference curve for a utility function or an isoquant for a production function). Since (2) is decreasing in
unemployment and vacancies, all the points inside the isowelfare curve yield higher welfare, so the green area delineated by the isowelfare curve is an upper contour set of (2). Since the function (2) is quasiconcave, its upper contour sets are convex, which implies that the isowelfare curve must be concave.

The efficient unemployment and vacancy rates can easily be found in the Beveridge diagram. First, they have to lie on the Beveridge curve. Second, since both unemployment and vacancies impose a welfare cost, they must lie on the isowelfare curve that is as close to the origin as possible. The closest that the isowelfare curve can be while remaining in contact with the Beveridge curve is at the tangency point with the Beveridge curve. This is where the efficient unemployment and vacancy rates are found. The efficient labor-market tightness is also visible on the diagram: it is the slope of the origin line going through the tangency point.

As with indifference curves and isoquants, the slope of the isowelfare curve is minus the marginal rate of substitution between unemployment and vacancies in the welfare function (2):

$$-\frac{(\partial W/\partial u) - (\partial W/\partial n)}{\partial W/\partial v} = -\frac{1 - (\partial W/\partial u)/(\partial W/\partial n)}{-(\partial W/\partial v)/(\partial W/\partial n)} = \frac{1 - \zeta}{\kappa}.$$  

The efficient unemployment rate is found at the point where the Beveridge curve, with slope $v'(u)$, is tangent to isowelfare curve, with slope $-(1 - \zeta)/\kappa$. This yields a first result:

**Proposition 1.** In a Beveridge diagram, efficiency is achieved at the point where the Beveridge curve is tangent to an isowelfare curve. Hence, the efficient unemployment rate is implicitly defined by

$$v'(u) = \frac{1 - \zeta}{\kappa},$$

where $\zeta < 1$ is the social value of nonwork and $\kappa > 0$ the recruiting cost.

Formula (4) simply says that when the labor market operates efficiently, welfare costs and benefits from moving one worker from employment to unemployment are equalized. When one worker moves from employment to unemployment, the reduction in welfare is the social cost of unemployment, $1 - \zeta$. Having one more unemployed worker also means having $-v'(u) > 0$ fewer vacancies. Each vacancy reduces welfare by the recruiting cost, $\kappa$, so welfare improves by $-v'(u)\kappa$ through the reduction in recruiting activity. When welfare costs and benefits are equalized, we have $1 - \zeta = -v'(u)\kappa$, which is equivalent to (4).

Of course, there is no guarantee that the labor market operates efficiently (figure 2, panel B). The labor market may be above the efficiency point, where unemployment is too low, vacancies are too high, and the unemployment gap is negative. This situation corresponds to a boom. It
may also be below the efficiency point, where unemployment is too high, vacancies are too low, and the unemployment gap is positive. This situation corresponds to a slump. As slumps and booms are inefficient, they lie on a worse isowelfare curve than the efficiency point.

3.2. Formula with sufficient statistics

We rework the efficiency condition (4) to obtain an explicit expression for the unemployment gap. We begin by introducing the Beveridge elasticity: \( \epsilon = -(u/v)v'(u) \) so \( \epsilon \theta = -v'(u) \). With this result, we can re-express (4) as \( \theta = (1 - \zeta)/(\kappa \epsilon) \).

In panel B of figure 2, we also see that any point on the Beveridge curve above the efficiency point has \( -v'(u) > (1 - \zeta)/\kappa \), and any point below it has \( -v'(u) < (1 - \zeta)/\kappa \). Using again \( \epsilon \theta = -v'(u) \), we infer that tightness is inefficiently high whenever \( \theta > (1 - \zeta)/(\kappa \epsilon) \); and tightness is inefficiently low whenever \( \theta < (1 - \zeta)/(\kappa \epsilon) \).

Hence we can assess the efficiency of labor-market tightness from three sufficient statistics:

**Proposition 2.** Consider a point on the Beveridge curve with labor-market tightness \( \theta \), Beveridge elasticity \( \epsilon \), recruiting cost \( \kappa \), and social value of nonwork \( \zeta \). Then tightness is inefficiently high if \( \theta > (1 - \zeta)/(\kappa \epsilon) \), inefficiently low if \( \theta < (1 - \zeta)/(\kappa \epsilon) \), and efficient if

\[
\theta = \frac{1 - \zeta}{\kappa \epsilon}.
\]

Formula (5) can be seen as a reformulation of the well-known Hosios (1990) condition. Hosios resolves the unemployment-vacancy tradeoff and derives a condition to ensure efficiency when wages are determined by Nash bargaining. The condition is that workers’ bargaining power equals the matching elasticity. While the Hosios condition has had a tremendous theoretical impact, its practical impact has been more limited. In our view this is due to two limitations, which we aim to address. First, measuring bargaining power is challenging (Pissarides 2000, p. 229). In contrast, formula (5) involves labor-market statistics that are estimable. Second, Nash bargaining poorly describes wage-setting over the business cycle (Shimer 2005; Hall 2005a; Jager et al. 2018). In contrast, our formula applies to any model with a Beveridge curve, irrespective of how wages are set.

Since the statistics \( \epsilon, \kappa \), and \( \zeta \) generally depend on tightness \( \theta \), formula (5) characterizes the efficient tightness only implicitly. This limitation is typical of the sufficient-statistic approach.

---

\(^3\) As a result the bargaining power is usually simply calibrated to 0.5 (den Haan, Ramey, and Watson 2000; Pissarides 2000; Gertler and Trigari 2009) or to the matching elasticity (Mortensen and Pissarides 1994; Shimer 2005; Costain and Reiter 2008).
Panel A depicts an efficient labor market in the Beveridge diagram. The statistic $\kappa$ is the recruiting cost, $\zeta$ is the social value of nonwork, and $\nu'(u)$ is the slope of the Beveridge curve. Panel B describes inefficient labor markets. In slumps, unemployment is inefficiently high, vacancies are inefficiently low, and the unemployment gap is positive; in booms, unemployment is inefficiently low, vacancies are inefficiently high, and the unemployment gap is negative. Panels C–F are special cases of panel A: assumption 1 holds, so the isowelfare curve is linear with slope $-(1 - \zeta)/\kappa$, and the Beveridge curve is isoelastic with elasticity $\epsilon$. Panel C shows that the efficient unemployment rate increases when the recruiting cost increases. Panel D shows that the efficient unemployment rate increases when the social value of nonwork increases. Panel E shows that the efficient unemployment rate increases when the Beveridge elasticity increases (keeping welfare constant). Panel F shows that the efficient unemployment rate increases when the Beveridge curve shifts outward (keeping the Beveridge elasticity constant).
(Chetty 2009), but it complicates the task of computing the unemployment gap. To address it, we use a workaround developed by Kleven (2018):

**Assumption 1.** The Beveridge elasticity ($\epsilon$), recruiting cost ($\kappa$), and social value of nonwork ($\zeta$) do not depend on the unemployment and vacancy rates.

How realistic is this assumption? Panels C and D in figure 1 suggest that the Beveridge curve is isoelastic, so the assumption on the Beveridge elasticity seems valid in US data. We do not have comparable evidence on the recruiting cost and social value of nonwork, but at least in the Diamond-Mortensen-Pissarides model, these two statistics are independent of the unemployment and vacancy rates (equation (3)).

Under assumption 1, we obtain a simple formula for the unemployment gap. The assumption implies that the Beveridge curve is isoelastic:

$$v(u) = v_0 \cdot u^{-\epsilon},$$

where the parameter $v_0 > 0$ determines the location of the curve. On the Beveridge curve, tightness is related to unemployment by

$$\theta = \frac{v(u)}{u} = v_0 \cdot u^{-(1+\epsilon)}$$

and

$$\theta^* = v_0 \cdot (u^*)^{-(1+\epsilon)}.$$

We can therefore link the unemployment gap to the tightness gap:

$$\frac{u^*}{u} = \left( \frac{\theta}{\theta^*} \right)^{1/(1+\epsilon)}.$$

Under assumption 1, (5) gives $\theta^* = (1 - \zeta)/(\kappa \epsilon)$, which yields the following proposition:

**Proposition 3.** Under assumption 1, the efficient unemployment rate and unemployment gap can be measured from current unemployment rate ($u$), vacancy rate ($v$), Beveridge elasticity ($\epsilon$), recruiting cost ($\kappa$), and social value of nonwork ($\zeta$). The efficient unemployment rate satisfies

$$u^* = \left( \frac{\kappa \epsilon}{1 - \zeta} \cdot \frac{v}{u} \right)^{1/(1+\epsilon)} u,$$

from which the unemployment gap $u - u^*$ follows.

---

4The Beveridge elasticity does depend on the unemployment rate in the Diamond-Mortensen-Pissarides model (equation (1)). However, as the unemployment rate is an order of magnitude smaller than the matching elasticity $\alpha$, the Beveridge elasticity is approximately constant and equal to $\alpha/(1 - \alpha)$. 

11
The proposition gives an explicit formula for the unemployment gap, expressed in terms of observable sufficient statistics. The formula is valid in any Beveridgean labor-market model, irrespective of the structure of the labor market, production, preferences, and wage setting. Another advantage of the formula is that we do not need to keep track of all the shocks disturbing the labor market—shocks to productivity, wages, labor-force participation, matching function, job separations, preferences, etc. The sufficient statistics are all we need to observe.\footnote{Without assumption 1, we could still obtain a formula for the unemployment gap, but it would require three additional statistics: the elasticities of $\epsilon$, $\kappa$, and $\zeta$ with respect to the unemployment rate (Kleven 2018).}

A secondary benefit of assumption 1 is to simplify the shapes of the curves in the Beveridge diagram: the Beveridge curve becomes isoelastic, and the isowelfare curve becomes linear with slope $-(1 - \zeta)/\kappa$. The simplification yields additional results:

**Proposition 4.** Under assumption 1, the following comparative statics hold:

- An increase in recruiting cost raises the efficient unemployment rate and lowers the efficient tightness.

- An increase in the social value of nonwork raises the efficient unemployment rate and lowers the efficient tightness.

- A compensated increase in the Beveridge elasticity (increase in $\epsilon$ keeping welfare constant) raises the efficient unemployment rate and lowers the efficient tightness.

- An outward shift of the Beveridge curve (increase in $v_0$) raises the efficient unemployment rate but does not affect the efficient tightness.

The intuition is conveyed by the diagrams in figure 2. When recruiting is more costly or unemployment is less costly, the unemployment-vacancy tradeoff becomes more favorable to unemployment, and the efficient unemployment rate increases (panels C and D). A compensated increase in the Beveridge elasticity (analogous to a compensated price increase in the context of Hicksian demand) is an increase in $\epsilon$ compensated by a change in $v_0$ so that the new Beveridge curve remains tangent to the same isowelfare curve (panel E). Such a change steepens the Beveridge curve: a rise in unemployment triggers a larger drop in vacancies, so the unemployment-vacancy tradeoff is more favorable to unemployment, and the efficient unemployment rate increases. Finally, the outward shift of the Beveridge curve does not change its elasticity so the efficient tightness remains the same. Since the Beveridge curve is further out, however, the efficient unemployment rate is higher.
We apply formula (6) to measure the unemployment gap in the United States over the 1951–2019 period. The first step is to estimate the sufficient statistics: Beveridge elasticity, recruiting cost, and social value of nonwork.

4.1. Beveridge elasticity (ε)

We estimate the Beveridge elasticity by OLS regression of log vacancy rate on log unemployment rate. Since the Beveridge curve shifts over time, we separately estimate the elasticity on the seven subperiods during which the Beveridge curve is stable: 1951Q1–1959Q2, 1959Q4–1971Q1, 1971Q3–1975Q1, 1975Q3–1987Q3, 1990Q1–1999Q1, 2001Q1–2009Q3, and 2010Q1–2019Q4. One such regression is illustrated in figure 3, panel A; the results on each subsample are displayed in figure 3, panel B.

We find that from 1951 to 2019, the Beveridge elasticity fluctuated between 0.81 and 1.24, averaging 1.03. The Beveridge elasticity steadily increased from 0.92 in the 1950s to 1.24 in the 1980s, before dropping back to 0.92 in 1990, and dropping further to 0.81 in 2010. The elasticities are precisely estimated: the standard errors vary between 0.02 and 0.10. And the fit of the regressions is good: the $R^2$ vary between 0.90 and 0.97. The good fit confirms that unemployment and vacancies travel on tightly defined, isoelastic curves.

Our estimates of the Beveridge elasticity are consistent with the estimates of the matching elasticity obtained by the empirical literature studying the matching function. A midrange estimate of the matching elasticity is $\alpha = 0.5$ (Petrongolo and Pissarides 2001). The average unemployment rate over the 1951–2019 period is $u = 5.8\%$. Combining these values with the expression of the Beveridge elasticity in a Diamond-Mortensen-Pissarides model, given by (1), we obtain $\epsilon = [0.5 + 0.058/(1 - 0.058)]/(1 - 0.5) = 1.12$. This value is close to our average elasticity of $\epsilon = 1.03$.

4.2. Recruiting cost (κ)

To estimate the recruiting cost we use the 1997 National Employer Survey. In the survey, the Census Bureau asked 4,500 establishments about their recruiting process, and found that firms spend 2.5% of their labor costs on recruiting (Villena Roldan 2010). This means that in 1997, $\kappa v = 2.5\% \times (1 - u)$. In 1997, the average vacancy rate is 3.3% and the average unemployment rate is 4.9% (figure 1). Hence, we find that the recruiting cost in 1997 is $\kappa = 2.5\% \times (1 - 4.9\%)/3.3\% = 0.72$. Since there is no other comprehensive measure of recruiting cost in the United States, we assume that the recruiting cost remains at its 1997 value over the entire 1951–2019 period.
4.3. Social value of nonwork (ζ)

To estimate the social value of nonwork, we rely on the work of Borgschulte and Martorell (2018) and Mas and Pallais (2019).

Using military administrative data covering 1993–2004, Borgschulte and Martorell study how servicemembers’ reenlistment choice is influenced by unemployment. This choice allows them to estimate the dollar value of the utility loss caused by higher unemployment during the transition to civilian life, and compare it to the earnings loss caused by higher unemployment. They find that during unemployment home production, recreation, and public benefits offset between 13% and 35% of lost earnings.

Using a large field experiment, Mas and Pallais study how unemployed job applicants choose between randomized wage-hour bundles. These choices imply that the value of home production and recreation during unemployment amounts to 58% of predicted earnings.

Next, we translate these estimates into social values of nonwork. The first step is to express the estimates relative to the marginal product of labor rather than earnings. The marginal product of labor is higher than earnings for several reasons. First, the wage paid by firms is usually lower than the marginal product of labor. In a matching model the wedge amounts to the share of workers allocated to recruiting, so the marginal product is about 3% higher than the wage (Landais, Michaillat, and Saez 2018, eq. (1)). In a monopsony model, the wedge depends on the elasticity of the labor supply, and the marginal product may be 25% higher than the wage (Mas and Pallais 2019, p. 121). Second, the wage received by workers is lower than that paid by firms because of employer-side payroll taxes, which amount to 7.7%. Third, Mas and Pallais discount predicted earnings by 6% to capture the wage penalty incurred by workers who recently lost their jobs; we undo the discounting because the penalty does not seem to apply to the marginal product of labor (Davis and von Wachter 2011). To conclude, to obtain a marginal product of labor, Borgschulte and Martorell’s earnings have to be adjusted by a factor between 1.03×1.077 = 1.11 and 1.25×1.077 = 1.35, and Mas and Pallais’s earnings by a factor between 1.03×1.077×1.06 = 1.18 and 1.25×1.077×1.06 = 1.43. Accordingly, to obtain a value relative to the marginal product of labor, Borgschulte and Martorell’s estimates have to be adjusted by a factor between 1/1.35 = 0.74 and 1/1.11 = 0.90, and Mas and Pallais’s estimates by a factor between 1/1.43 = 0.70 and 1/1.18 = 0.85.

The second step only applies to Borgschulte and Martorell’s estimates, from which we subtract the value of public benefits. All servicemembers are eligible to unemployment insurance (UI). Chodorow-Reich and Karabarbounis (2016, pp. 1585–1586) find that UI benefits amount to 21.5% of the marginal product of labor. But this quantity has to be reduced for several reasons: the UI takeup rate is only 65%; UI benefits and consumption are taxed, imposing a factor of 0.83; the disutility from filing for benefits imposes a factor of 0.47; and UI benefits expire, imposing a factor of 0.83. In
sum, the average value of UI benefits to servicemembers is $21.5 \times 0.65 \times 0.83 \times 0.47 \times 0.83 = 5\%$ of the marginal product of labor. Like all unemployed workers, servicemembers are also eligible to other public benefits, which Chodorow-Reich and Karabarbounis quantify at $2\%$ of the marginal product of labor. Hence, to account for benefits, we subtract $5\% + 2\% = 7\%$ of the marginal product of labor from Borgschulte and Martorell’s estimates.

Combining these two steps, we find that Mas and Pallais’s estimates imply a social value of nonwork between $0.58 \times 0.70 = 0.41$ and $0.58 \times 0.85 = 0.49$; and that Borgschulte and Martorell’s estimates imply a social value of nonwork between $(0.13 \times 0.74) - 0.07 = 0.03$ and $(0.35 \times 0.90) - 0.07 = 0.25$. The range of plausible values for the social value of nonwork therefore is $0–0.5$; we set the statistic to its midrange value, $\zeta = 0.25$.

We also keep the social value of nonwork constant over the business cycle. In some models, the productivities of unemployed and employed workers do not move in tandem over the business cycle, generating fluctuations in the social value of nonwork (for instance in the Diamond-Mortensen-Pissarides model, as shown by (3)). However, Chodorow-Reich and Karabarbounis (2016, pp. 1599–1604) find no evidence of such fluctuations in US data. Instead, they establish that the utility derived by unemployed workers from recreation and home production moves proportionally to labor productivity—which implies that $\zeta$ is acyclical. The social value of nonwork could also exhibit medium-run fluctuations, but we omit them by lack of evidence.

### 4.4. Unemployment gap

We now use the estimates of the Beveridge elasticity, recruiting cost, and social value of nonwork, as well as the unemployment and vacancy rates from figure 1, to measure the unemployment gap in the United States between 1951 and 2019.

We begin by computing the efficient unemployment rate from formula (6) (figure 3, panel C). It hovered around 3% in the 1950s, rose to 4% in the 1960s, and climbed to reach 5.9% in 1980—a level it maintained until 1986. The steady increase of the efficient unemployment rate between 1951 to 1986 was caused by two factors: a steady increase of the Beveridge elasticity, from 0.92 in 1951 to 1.24 in 1986 (figure 3, panel B), and a steady outward shift of the Beveridge curve (figure 1, panel C). Then, in 1987–1990, the efficient unemployment rate sharply declined to reach 4%. The decline was caused both by a drop of the Beveridge elasticity from 1.24 to 0.91 (figure 3, panel B), and by an inward shift of the Beveridge curve (figure 1, panels C–D). The efficient unemployment rate then remained stable through the 1990s, 2000s, and 2010s, hovering between 3.3% and 4%.

Interestingly the efficient unemployment rate did not increase in the aftermath of the Great Recession—despite the outward shift of the Beveridge curve (figure 1, panel D). This is because the Beveridge curve also became flatter after 2009: the Beveridge elasticity fell from 0.97 to 0.81.
Panels A and B: We estimate the Beveridge elasticity $\epsilon$ with OLS regressions of log vacancy rate on log unemployment rate (from figure 1, panels A and B) for seven subperiods: 1951Q1–1959Q2, 1959Q4–1971Q1, 1971Q3–1975Q1, 1975Q3–1987Q3, 1990Q1–1999Q1, 2001Q1–2009Q3, and 2010Q1–2019Q4. Panel A illustrates, as an example, the estimation of $\epsilon$ for 2010Q1–2019Q4; the slope of the regression line gives $\epsilon = 0.81$. Panel B depicts the estimates of $\epsilon$ for all subperiods. Panel C: The actual unemployment rate comes from figure 1, panel A. The efficient unemployment rate is computed from (6) with $\kappa = 0.72$, $\zeta = 0.25$, $\epsilon$ from panel B, and the unemployment and vacancy rates from figure 1. Panel D: Actual and efficient unemployment rates come from panel C. CBO unemployment rate is the long-term natural rate of unemployment constructed by the Congressional Budget Office. NAIRU and trend unemployment rates come from Crump et al. (2019, fig. 8B). The shaded areas represent recessions, as identified by the NBER.
(figure 3, panel B). The flattening offset the outward shift, leaving the efficient unemployment rate almost unchanged by the recession.

Measuring the distance between the actual and the efficient unemployment rate, we obtain the unemployment gap (figure 3, panel C). Between 1951 and 2019, the unemployment rate averages 5.8% and the efficient unemployment rate average 4.2%, so the unemployment gap averages 1.6 percentage points. The unemployment gap is sharply countercyclical. It is close to zero at business-cycle peaks: sometimes negative (−0.4 percentage points in 1953 and −0.7 points in 1969); sometimes zero (in 1973, 1979, and 2019); and sometimes positive (for instance 0.4 points in 2000 and 0.9 points in 2007). The effect of wartime mobilization on the labor market is visible: the unemployment gap was only negative in 1951–1953, during the Korea war, and in 1965–1970, at the peak of the Vietnam war. And the unemployment gap is highly positive at business-cycle troughs: for instance, 5.0 points in 1982, 3.8 points in 1992, and 6.5 points in 2009. Unsurprisingly, the largest unemployment gap occurred in the wake of the Great Recession.

Assumption 1 is required to obtain precise values of the unemployment gap (proposition 3). Yet, even without the assumption, the graph in panel C of figure 3 continues to be informative: it indicates whether unemployment is inefficiently high or low. Whenever unemployment is above the efficiency line, we know from (6) that $v/u < (1 - \zeta)/\kappa\epsilon$, which indicates that unemployment is inefficiently high (proposition 2). Conversely, whenever unemployment is below the efficiency line, we know that unemployment is inefficiently low.

In sum, the US unemployment rate is generally inefficiently high—an inefficiency exacerbated in slumps. These results have implications for policymaking and macroeconomic modeling. First, because the unemployment spikes in recessions are inefficient, it is warranted to deploy fiscal and monetary policy in bad times to reduce unemployment. Second, given that the unemployment rate is almost always inefficient, and sometimes markedly so, it might not be productive to insist upon modeling the labor market as efficient—whether it is by assuming that the Hosios condition holds or by focusing on a competitive-search equilibrium.

Finally, to provide some context, we compare our measure of the unemployment gap to three other measures: the differences between actual unemployment and its trend, the NAIRU, and the Congressional Budget Office’s natural rate of unemployment (figure 1, panel D). Trend unemployment and the NAIRU are constructed by Crump et al. (2019, fig. 8B) using state-of-the-art techniques. The CBO’s natural rate of unemployment—which features prominently in policy discussions—is constructed by blending trend and NAIRU considerations (Shackleton 2018, app. B). The main similarity is that all four measures are countercyclical. This is because efficient, trend, NAIRU, and CBO unemployment rates are slow-moving while the actual unemployment rate is countercyclical. The main difference is that our measure is higher than the three others;
the average difference over the period is 1.3–1.8 percentage points. This is because the efficient unemployment rate is lower than trend, NAIRU, and CBO unemployment rates.

4.5. Alternative calibrations of the social value of nonwork

Among the three statistics in the unemployment-gap formula, the least understood is the social value of nonwork. Given this uncertainty, we consider alternative calibrations of the social value of nonwork and explore their impact on the efficient unemployment rate (figure 4).

First, we consider an estimate at the low end of the range of plausible values: \( \zeta = 0 \). Under this calibration, the efficient unemployment rate follows the same pattern as under the baseline calibration (\( \zeta = 0.25 \)) but is on average 0.6 percentage point lower (panel A).

Next, we consider an estimate at the high end of the range of plausible values: \( \zeta = 0.5 \), which is consistent with the estimates from the macro-labor literature (Chodorow-Reich and Karabarbounis 2016, eq. (30)). The efficient unemployment rate follows again the same pattern as under the baseline calibration, but it is on average 0.9 percentage point higher (panel B).

These different calibrations show that around our baseline calibration of \( \zeta = 0.25 \), the efficient unemployment rate is fairly insensitive to the precise value of \( \zeta \). For any \( \zeta \) in the 0–0.5 range, the efficient unemployment rate remains contained in a band whose average width is less than 1.5 percentage point. This is reassuring as the range of plausible values for \( \zeta \) is quite broad.

Our baseline calibration implies that the social value of unemployment is much lower than labor productivity; in contrast, some macro-labor studies argue that the two are very close. A well-known calibration, due to Hagedorn and Manovskii (2008), is \( \zeta = 0.96 \). Such a calibration has a drastic impact: it pushes the efficient unemployment rate above 13%, and sometimes as high as 22%, with an average value of 17.5% (panel C). Under this calibration, unemployment is always inefficiently low, even at the peak of the Great Recession, which seems implausible.

Finally, to provide further perspective, we compute the social value of nonwork that arises under the assumption that US unemployment is efficient at all time. Proposition 2 shows that this value is

\[
(7) \quad \zeta^* = 1 - \kappa \epsilon \theta,
\]

where \( \theta \) is the actual labor-market tightness. To sustain efficiency, the social value of nonwork would need to be immensely countercyclical, as low as −0.16 in booms and as high as 0.90 during the Great Recession (panel D). Then recessions would merely be vacations.
Figure 4. Efficient unemployment rate in the United States for different social values of nonwork

Panels A–C reproduce panel C of figure 3, and add the efficient unemployment rate computed from (6) and different social values of nonwork: \( \zeta = 0 \) in panel A, \( \zeta = 0.5 \) in panel B, and \( \zeta = 0.96 \) in panel C. Panel D displays the social value of nonwork that would ensure that the observed unemployment rate is efficient at all time; the value is obtained from (7). The shaded areas represent recessions, as identified by the NBER. The dark gray line in panels A–C represents the actual unemployment rate.
5. Conclusion

This paper develops a new method to measure the unemployment gap—the difference between the actual and the socially efficient unemployment rate. We consider a labor-market model with only one structural element: a Beveridge curve relating unemployment and vacancies. This framework covers many modern labor-market models, including the Diamond-Mortensen-Pissarides model. We show that the unemployment gap can simply be measured from current unemployment and vacancy rates, and three sufficient statistics: the elasticity of the Beveridge curve, cost of recruiting, and social cost of unemployment.

We apply our unemployment-gap formula to the United States, 1951–2019. We find that the US unemployment gap is countercyclical: the gap is close to zero in booms, slightly positive or negative; it is highly positive in slumps. We infer that the US unemployment rate is generally inefficiently high, and such inefficiency worsens in slumps. Historically, it was therefore warranted to activate fiscal and monetary policy in bad times to reduce unemployment.

To cement our estimates of the unemployment gap, it would be invaluable to obtain more evidence on the three statistics in the formula. The Beveridge elasticity could be estimated more finely by using more sophisticated econometric techniques or by imposing more structure on the labor market; Ahn and Crane (2020) take a step in that direction. The recruiting cost could be measured better in the future by adding a new question into JOLTS, asking firms to report the number of man-hours devoted to recruiting in addition to the number of vacancies. Last, additional estimates of the social cost of unemployment could be obtained from natural and field experiments, following the methodology in Borgschulte and Martorell (2018) and Mas and Pallais (2019).

References


